



Universiteit
Leiden
The Netherlands

Kwaliteit en verantwoording in het hoger onderwijs: indicatoren, processen en beleidsalternatieven: eindrapport

Beerkens, M.; Groeneweg, S.L.

Citation

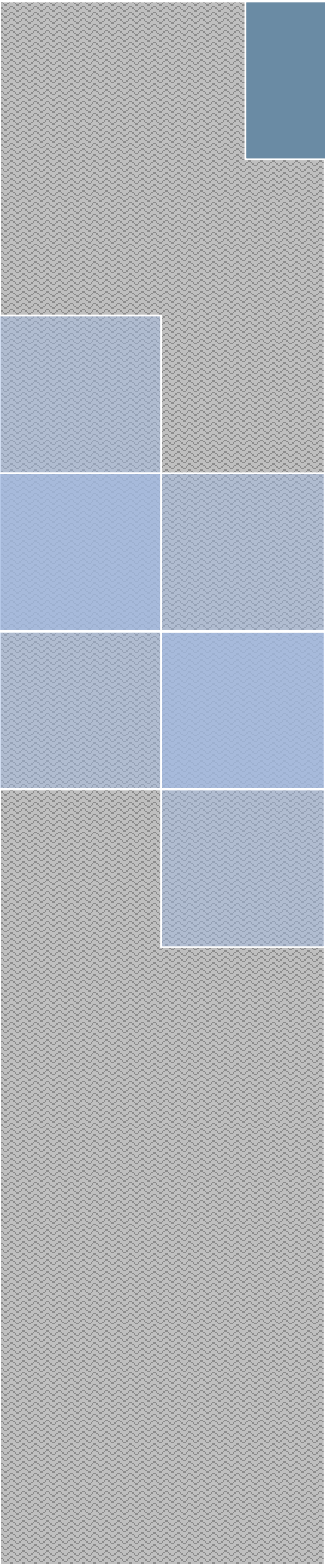
Beerkens, M., & Groeneweg, S. L. (2022). *Kwaliteit en verantwoording in het hoger onderwijs: indicatoren, processen en beleidsalternatieven: eindrapport*. Den Haag: NRO. Retrieved from <https://hdl.handle.net/1887/3484703>

Version: Publisher's Version

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/3484703>

Note: To cite this publication please use the final published version (if applicable).



KWALITEIT EN VERANTWOORDING IN HET HOGER ONDERWIJS:

INDICATOREN, PROCESSEN EN
BELEIDSLTERNATIEVEN

EEN OVERZICHTSSTUDIE

MAARJA BEERKENS
SVENNE GROENEWEG



Inhoudsopgave

Samenvatting	2
1. Inleiding	6
2. Kwaliteit en kwaliteitszorg	7
2.1 Wat is kwaliteit?	7
2.2 Wat is kwaliteitszorg?	9
2.3 Een publiek belang in kwaliteitszorg	10
3. Methodologie	12
3.1 Literatuurstudie	12
3.2 Interviews	13
3.3 Case studies	14
4. Databronnen en indicatoren	15
4.1 Verschillende soorten databronnen	15
4.2 Verschillende soorten indicatoren	20
4.3 Het gebruik en selectie van indicatoren: betrouwbaarheid, validiteit en haalbaarheid	27
4.4 Indicatoren in Nederland	30
5. Interne processen	34
5.1 Algemeen beeld van kwaliteitsmanagementprocessen: eenheid in verscheidenheid	34
5.2 Specifieke managementaanpakken	35
5.3 Interne processen in Nederland	41
5.4 Het gebruik en vormgeving van processen: helderheid, transparantie en efficiëntie	42
6. Internationale case studies	46
6.1 Vier alternatieven voor het beoordelen van opleidingskwaliteit	46
6.2 Casus 1: Zelfaccreditatie in Noorwegen	49
6.3 Casus 2: Publiek toegankelijke <i>scoreboards</i> en rapportkaarten in Engeland	54
7. Nieuwe ontwikkelingen in het openbaar bestuur en de regelgeving	58
7.1 Nieuwe benaderingen van regelgeving	59
7.2 Lessen voor de beoordeling van opleidingskwaliteit	64
8. Conclusie	66
Literatuurlijst	68

Samenvatting

Achtergrond

De afgelopen 25 jaar hebben kwaliteitsborgingssystemen in het hoger onderwijs in Europa aanzienlijke vooruitgang geboekt. De oude kwaliteitsborgingsmodellen beginnen hun aantrekkingskracht te verliezen, en er is groeiende belangstelling voor alternatieven die aansluiten op de volwassen, geavanceerde kwaliteitsborgingssystemen en het veranderende hoger onderwijslandschap. In deze nieuwe fase van kwaliteitszorg wordt externe controle op de kwaliteit van opleidingen vaak vervangen door institutionele audits die de nadruk leggen op de eigen verantwoordelijkheid van onderwijsinstellingen en hun bekwaamheid om kwaliteit te borgen. Tegelijkertijd is het waarborgen van kwaliteit op opleidingsniveau nog steeds van publiek belang. Dit rapport gaat in op dit dilemma. Op basis van academische literatuur en andere bronnen wordt getracht een antwoord te geven op de vraag: *“Hoe kan de kwaliteit op opleidingsniveau beoordeeld en bewaakt worden in deze nieuwe fase van kwaliteitszorg?”*. Het rapport biedt een overzichtsstudie met als voornaamste doel een inventaris op te maken van de huidige omvangrijke, maar gefragmenteerde kennis op dit gebied.

De studie combineert drie elementen. In het literatuuronderzoek wordt een synthese gecreëerd van bestaande kennis over mogelijke kwaliteitsindicatoren in het hoger onderwijs en over processen die worden gebruikt om de kwaliteit op opleidingsniveau te waarborgen. Interviews met kwaliteitszorgmedewerkers bij twintig Nederlandse hoger onderwijsinstellingen geven een beeld van het gebruik van indicatoren en interne processen bij Nederlandse hoger onderwijsinstellingen. Tot slot wordt ingegaan op twee internationale voorbeelden waar opleidingskwaliteit op alternatieve wijze wordt gemonitord.

Kwaliteitsindicatoren op opleidingsniveau

De literatuur over kwaliteitsindicatoren is omvangrijk en er zijn erg veel mogelijke indicatoren voorgesteld. In werkelijkheid, echter, wordt in alle landen een betrekkelijk beperkte en homogene set indicatoren gebruikt voor een regelmatig kwaliteitsmonitoring en voor vergelijkingsdoeleinden:

- Directe outputindicatoren, zoals het percentage voortijdig studieverlaters, de tijd die nodig is om een diploma te behalen, en de arbeidsmarkresultaten van afgestudeerden;
- Studenttevredenheidsindicatoren, gebaseerd op nationale studenten-enquêtes en enquêtes onder afgestudeerden waarin de tevredenheid over diverse aspecten van de studie wordt aangegeven;
- Procesindicatoren, zoals het gebruik van verschillende onderwijstechnieken en maatstaven voor de betrokkenheid van studenten;
- Inputindicatoren, zoals de student-stafratio, kwalificaties van docenten en investeringen in faciliteiten.

Kwaliteitsindicatoren worden regelmatig zwaar bekritiseerd omdat zij niet de werkelijke ‘kwaliteit’ van het leerproces meten, maar ook omdat er zorgen bestaan over de betrouwbaarheid, relevantie, ongelijkheid en

disfunctionele effecten. Er zijn meer verfijnde maatstaven ontwikkeld om bijvoorbeeld metingen bij te stellen aan de hand van (socio-economische) kenmerken van instromende studenten, of om de toegenomen kennis te meten, maar dergelijke maatstaven zijn vaak onbetaalbaar, arbeidsintensief en ingrijpend, en zij hebben de neiging om hun intuïtieve eenvoud te verliezen. Bij de selectie van indicatoren lijkt de beschikbaarheid van gegevens de meest bepalende factor te zijn. Het is duidelijk dat overheidsinitiatieven om de informatie te verzamelen en te presenteren een groot effect hebben op welke gegevens zowel extern als intern worden gebruikt.

Kwaliteitsprocessen op opleidingsniveau

Veel hoger onderwijsinstellingen hebben uitgebreide interne werkmethoden ontwikkeld om hun kwaliteit te bewaken en te verbeteren. Sommige instellingen maken gebruik van formele managementaanpakken uit de private sector, zoals ‘*Total Quality Management*’, benchmarking en ISO standaarden. Vaker spreken instellingen in hun kwaliteitsbenadering in het algemeen van ‘kwaliteitsmanagement’ en, meer recentelijk, van ‘kwaliteitscultuur’. In wezen hebben alle managementbenaderingen drie elementen met elkaar gemeen: een empirisch onderbouwde benadering van kwaliteit (‘*evidence-based*’), het betrekken van interne en externe belanghebbenden bij het proces, en het benaderen van kwaliteitszorg als een continue activiteit. Interessant is dat er bijna geen voorbeelden zijn van het gebruik van managementaanpakken die tot doel hebben de administratieve last van interne kwaliteitsprocessen in het hoger onderwijs te verminderen.

De literatuur over het effect van interne processen geeft geen uitsluitsel. De meeste studies melden een positief effect op de professionalisering van de interne processen en procedures, op de interne samenwerking tussen het onderwijzend personeel, op het bewustzijn van de doelen en visie van de opleiding, en op de samenhang van het curriculum. De daadwerkelijke verbetering van leerresultaten als gevolg van deze processen is echter moeilijker aan te tonen.

Indicatoren en processen in het Nederlandse hoger onderwijs

Uit twintig interviews met kwaliteitszorgmedewerkers van Nederlandse hoger onderwijsinstellingen blijkt dat kwaliteitsindicatoren uitgebreid gebruikt worden en dat interne kwaliteitscontrole systematisch wordt aanpakt. Veel instellingen hebben afspraken gemaakt over de kernindicatoren op opleidingsniveau die jaarlijks worden gemonitord. De resultaten van de nationale studentenenquête en studiesucces-indicatoren worden het vaakst gebruikt. Sommige instellingen stellen gegevens via een dashboard ter beschikking, maar laten de opleidingen zelf beslissen op welke indicatoren zij zich willen richten bij hun reflectie en verbetering. De meeste instellingen beschikken over grondige processen om de kwaliteit op opleidingsniveau regelmatig te monitoren. Veelgebruikte bouwstenen zijn onder meer jaarlijkse opleidingsrapporten, bespreking van verbeterplannen met de opleidingscommissie (inclusief studenten), en evaluatiegesprekken met een afdelingshoofd of decaan. De gegevens en rapporten op opleidingsniveau zijn in principe alleen bedoeld voor intern gebruik en zijn niet publiekelijk beschikbaar.

Inzichten uit internationale case studies

Ook andere landen zoeken naar effectieve manieren om de kwaliteit van opleidingen te beoordelen en zichtbaar te maken. Er lijken vier beleidsalternatieven voor externe opleidingsaccreditatie te zijn om opleidingskwaliteit te monitoren:

- Zelfaccreditaties waarbij instellingen zelf de accreditatie en het toezicht uitvoeren;
- Publiek toegankelijke ‘*scoreboards*’ en rapportkaarten waarin belangrijke kwaliteitsindicatoren of een eenvoudig kwaliteitsoordeel per opleiding wordt gepubliceerd;
- Thematische evaluaties die opleidingen evalueren in een bepaald vakgebied of een focus hebben op een specifiek aspect van het onderwijs;
- Certificering door beroepsorganisaties of andere professionele, private organisaties die (meestal) op vrijwillige basis opleidingen binnen een bepaald vakgebied accrediteren.

Lessen

Informatie over opleidingskwaliteit kan meerdere doelen dienen: publiekelijk waarborgen dat een diploma aan de verwachte normen voldoet, aankomende studenten van informatie over de opleidingen voorzien, een stimulans en data bieden voor kwaliteitsborging en -verbetering, en de opleidingen evalueren in vergelijking met hun (internationale) gelijken. Geen enkel kwaliteitszorginstrument kan alle doelen op een efficiënte en effectieve wijze dienen. Op basis van de analyse van indicatoren en processen in het hoger onderwijs, internationale ervaringen met kwaliteitsborging en bredere inzichten uit de reguleringstheorie worden in dit rapport de onderstaande lessen voor de toekomst getrokken.

- Goede gegevens en indicatoren over de kwaliteit van de opleidingen blijven waardevol, ondanks de inherente beperkingen ervan. Eigen verantwoordelijkheid voor kwaliteit betekent niet dat kwaliteit verborgen blijft, maar dat de zorg voor het vormgeven van relevante indicatoren, het verzamelen van prestatiegegevens en het handelen op basis van gegevens, bij de instelling zelf ligt. Indicatoren moeten effectief worden gezien als een signaal en als input voor reflectie, niet als een absolute meting van kwaliteit. Voor goede en vergelijkbare indicatoren zijn vaak overheidsinitiatieven nodig. Het regelmatig herzien van indicatoren kan nuttig zijn om vertekeningen als gevolg van een tunnelvisie te voorkomen.
- Instellingen kunnen zelf vrij goed de taak op zich nemen om de kwaliteit van hun onderwijsprogramma's te bewaken. Het delegeren van taken van een extern agentschap naar een instelling leidt echter niet noodzakelijkerwijs tot lagere administratieve kosten of meer flexibiliteit; het verschuift alleen de administratieve last en vereist nieuwe ondersteuningsstructuur binnen de instelling.
- Een extern agentschap speelt een belangrijke rol bij het aanmoedigen van instellingen om systematisch kwaliteitscontroles op opleidingsniveau uit te voeren. De rol hoeft echter niet noodzakelijk intensief te zijn om een substantieel effect te hebben, zoals blijkt uit nieuwere reguleringsmodellen. Duidelijk omschreven verwachtingen en het risico op (redelijke) sancties in geval van niet-naleving zijn waarschijnlijk voldoende, en een arbeidsintensieve controlecyclus is wellicht niet nodig.

- Er moet een moeilijke afweging worden gemaakt tussen enerzijds het formuleren van duidelijke verwachtingen van instellingen ten aanzien van hun interne kwaliteitsborgingssysteem en anderzijds het aan de instellingen overlaten van flexibiliteit bij het definiëren van hun eigen processen. Hoewel flexibiliteit in de hedendaagse reguleringsmodellen wordt aangemoedigd, leidt een dergelijke flexibiliteit tot hoge onzekerheidskosten voor risicomijdende instellingen. Dit betekent dat eigen verantwoordelijkheid met los gedefinieerde verwachtingen kan leiden tot even intensieve of zelfs intensievere (en duurdere) kwaliteitsprocessen binnen een instelling. Een passend sanctiemechanisme is daarom nodig om een evenwicht te vinden tussen nalevingsprikkel en een gezonde risicotolerantie. Nederlandse instellingen ervaren een ‘gebrek aan vertrouwen’ vaak als een contraproductief uitgangspunt dat hen dwingt tot een aantal inefficiënte reacties.
- Nieuwere reguleringsmodellen zoeken naar een evenwicht tussen de administratieve kosten van regulering en de effectiviteit ervan voor het bewaken van publieke waarde. Een kostbare controle van alle aanbieders kan onredelijk zijn als de meeste aanbieders zich aan de regels houden en kwaliteitsproblemen slechts zelden voorkomen. De controlekosten zouden moeten worden toegespitst op aanbieders met een hoog risico. Deze aanpak vereist echter een zekere politieke risicotolerantie. Kwaliteitsincidenten kunnen tot mediaschandalen leiden. Gevoeligheid voor dergelijke schandalen kan ertoe leiden dat regelgevers defensief en te procedureel worden, waardoor het doel van de nieuwe aanpak, namelijk een lichte aanpak, teniet wordt gedaan. De afweging moet door het publiek worden aanvaard: een strenge controlecyclus met hoge kosten voor zowel de instelling als het publiek versus een kleine kans dat incidenten zich af en toe voordoen, maar met extra middelen voor kernactiviteiten in plaats van controle.
- Hoger onderwijsinstellingen zijn geen unitaire actoren. Het succes van het interne kwaliteitsbeleid hangt af van een goede communicatie tussen het centrale management, het management op opleidingsniveau en het onderwijzend personeel. Nieuwe flexibele regelgeving kan het eigenaarschap van interne kwaliteitsprocessen vergroten, maar schept ook hogere verwachtingen bij het centrale management om hun kwaliteitsplannen bekend te maken en anderen te overtuigen van de urgentie ervan binnen de organisatie.
- In een regelgevingsstelsel worden verschillende regelgevingsinstrumenten gecombineerd en het is belangrijk dat met hun wisselwerking rekening wordt gehouden. Er moet duidelijkheid zijn over het doel van elk regelgevend instrument. Informatie op opleidingsniveau om na te gaan of aan de kwaliteitsstandaarden wordt voldaan, heeft andere noodzaak dan informatieverstrekking aan studenten voor de keuze van hun opleiding. Het combineren van traditionele controlemechanismen met nieuwere, flexibele mechanismen zonder het een door het ander te vervangen, maakt de regelgeving ook complexer en meer gefragmenteerd.

1. Inleiding

Het kwaliteitsborgingsysteem in het hoger onderwijs wordt voortdurend tegen het licht gehouden en herzien. In de afgelopen twintig jaar zijn opleidingsaccreditaties het dominante kwaliteitszorginstrument geworden in Europa (Westerheijden et al., 2007). Over het algemeen blijkt het instrument een positief effect te hebben op de professionalisering van kwaliteitsprocessen (Stensaker et al., 2011). Toch wordt het huidige systeem bekritiseerd en staat het onder druk - niet alleen in Nederland maar ook daarbuiten. Het systeem levert veel werkdruk op (Inspectie van het Onderwijs, 2018). Bovendien blijken de voordelen met elke accreditatieronde af te nemen omdat instellingen en opleidingen al geleerd hebben wat er precies van hen verwacht wordt (Stensaker et al., 2011). De vraag of opleidingsaccreditatie past bij de hedendaagse ontwikkelingen zoals toenemende interdisciplinariteit, modulariteit en flexibiliteit in leerroutes is daarnaast ook een punt van discussie (Hopbach, 2014).

Toch is de publieke taak van kwaliteitsborging niet daadwerkelijk veranderd. Er bestaat nog steeds behoefte aan verantwoording en een behoefte om het algemeen belang van het hoger onderwijs te waarborgen. Diploma's moeten voldoen aan de verwachte standaarden en studenten moeten de zekerheid hebben dat hun tijds- en geldinvestering bij een waardevolle opleiding terecht komt. Ideeën over effectieve en efficiënte beleidsbeginselen en -instrumenten zijn echter wel aan het verschuiven. In plaats van traditionele prestatiemetingen en controle is er enerzijds meer aandacht voor goede processen en anderzijds meer aandacht voor meer vertrouwen in de professionaliteit van de (semi-)publieke sector.

Dit leidt tot de hoofdvraag in deze overzichtsstudie: hoe kan de kwaliteit van opleidingen beoordeeld en bewaakt worden op een effectieve en efficiënte wijze? Deze studie zal een antwoord geven middels drie deelvragen:

- 1) Welke indicatoren en databronnen worden gebruikt om de kwaliteit van opleidingen systematisch en regelmatig te monitoren?
- 2) Welke interne processen worden gebruikt binnen hoger onderwijsinstellingen om de kwaliteit van opleidingen te monitoren en waarborgen en wat is het effect daarvan?
- 3) Wat kunnen we leren van internationale ervaringen met alternatieve systemen om de kwaliteit van opleidingen te monitoren?

Dit is een overzichtsstudie. Het belangrijkste doel van deze studie is een inventarisatie en synthese van de huidige kennis en ervaringen in dit vakgebied. Het volgende hoofdstuk geeft achtergrondinformatie over de belangrijkste concepten, argumenten en dilemma's met betrekking tot kwaliteit en kwaliteitszorg. Daarna volgt een beschrijving van de onderzoeksaanpak. De onderzoeksresultaten zijn vervolgens opgedeeld in drie hoofdstukken. Het eerste hoofdstuk behandelt de databronnen en indicatoren die gebruikt worden om opleidingskwaliteit te monitoren. Het middelste hoofdstuk reflecteert op de interne processen en in het laatste hoofdstuk worden op basis van ervaringen in andere landen inzichten gedeeld hoe opleidingskwaliteit op alternatieve manieren beoordeeld kan worden. Tot slot wordt er gereflecteerd op nieuwe ontwikkelingen in het openbaar bestuur en regelgeving, en trekken we lessen daaruit voor de toekomst van kwaliteitszorg in het hoger onderwijs.

2. Kwaliteit en kwaliteitszorg

Binnen de hoger onderwijssector zijn voortdurende dilemma's met betrekking tot de vraag wat kwaliteit is, en wat de doelstellingen van kwaliteitszorg zijn. Verschillende visies leiden tot spanningen, maar nodigen ook uit tot een zoektocht naar betere antwoorden op deze vragen. In dit hoofdstuk wordt achtergrondinformatie gegeven over de belangrijkste concepten, argumenten en dilemma's met betrekking tot kwaliteit en kwaliteitszorg.

2.1 Wat is kwaliteit?

Voordat men iets kan zeggen over de indicatoren of databronnen waarmee kwaliteit in het hoger onderwijs gemeten wordt, is het belangrijk om te weten wat kwaliteit is. Wat wordt er verstaan onder onderwijskwaliteit? Welke kwaliteitscriteria in het hoger onderwijs zijn nastrevenswaardig?

Deze vragen zijn niet eenduidig te beantwoorden. Verschillende belanghebbenden hebben hun eigen ideeën over het doel van hoger onderwijs en daarmee ook over de kwaliteit van hoger onderwijs (ESG 2015, p. 7; Gover & Loukkola, 2018, pp. 6-7). Het idee dat opvattingen over kwaliteit tussen diverse belanghebbenden verschillen, wordt ook wel '*the stakeholder approach*' genoemd (Watty, 2006). Relevante belanghebbenden zijn onder andere studenten, docenten, het bestuur van de onderwijsinstelling, beleidsmedewerkers, (potentiële) werkgevers, beleidmakers en de maatschappij in de brede zin. Bovendien verandert het doel van hoger onderwijs mee met maatschappelijke ontwikkelingen (Beerkens, 2015). Concrete voorbeelden van verschillende doelen van het hoger onderwijs zijn cognitieve ontwikkeling van studenten en loopbaanvoorbereiding, maar ook sociale mobiliteit, kritisch burgerschap (globaal of nationaal) of hooggekwalificeerde arbeidskrachten afleveren aan de samenleving. Doordat men verschillende verwachtingen van het hoger onderwijs heeft, heeft men ook verschillende waarneming van de kwaliteit van het onderwijs en bijbehorende kwaliteitszorg. Los van de verschillende groepen belanghebbenden, moeten we er bovendien rekening mee houden dat elke instelling haar eigen onderwijsvoornemens, onderwijsvisie, historie en context kent (Rijksoverheid, 2018, p. 2). De keuzes voor (en evaluatie van) kwaliteitsbeleid moeten ook binnen deze context begrepen worden.

Om kwaliteit te definiëren wordt in de academische literatuur vaak verwezen naar Harvey en Green (1993) die de verschillende manieren van denken over kwaliteit samenvatten in vijf categorieën. Het gaat daarbij om kwaliteit in termen van uitzondering, perfectie of consistentie, geschiktheid voor het beoogde doel, prijs-kwaliteitverhouding, en kwaliteit als transformatieve kracht (Harvey & Green, 1993). Hoe ziet een opleiding van hoge kwaliteit in deze vijf verschillende denkbelden eruit, en waar moet een kwaliteitszorgsysteem zich dus op richten? Met kwaliteit in termen van uitzondering kan een opleiding laten zien dat die opleiding excellent presteert of dat een opleiding zichzelf onderscheidt van soortgelijke opleidingen. Vanuit dit denkbeld moet een optimaal kwaliteitszorgsysteem opleidingen prikkelen om het excellentieniveau te ambiëren en bereiken. Met kwaliteit als perfectie of consistentie wordt gestreefd naar betrouwbaarheid. Er mogen geen 'defecten' zijn: elk diploma moet aan de standardeisen voldoen. Bij de kwaliteitszorg staat dus preventie centraal in plaats van evaluatie of verbetering. Bij geschiktheid voor het beoogde doel wordt kwaliteit bekeken in termen van de

achterliggende doelstelling van de opleiding en instelling, waarbij de kwaliteit in een instelling niet per se vergelijkbaar is met kwaliteit in een andere instelling met een andere studentenpopulatie, missie of onderwijsvisie. Prijs-kwaliteitverhouding berust op het principe van verantwoording naar zowel belastingbetalers als consumenten. Efficiëntie en effectiviteit worden nagestreefd. De laatste categorie van kwaliteit, kwaliteit als transformatieve kracht, kijkt naar de veranderingen die studenten ondergaan door het onderwijs. Als de opleiding van hoge kwaliteit is, vindt er een continu proces van ontwikkeling en *empowerment* plaats.

De conceptualisatie van kwaliteit heeft een direct effect op de vormgeving en geschiktheid van kwaliteitszorgprocessen. Een conceptualisatie met een focus op consistentie en een nadruk op het vermijden van ‘defecten’ vraagt om een kwaliteitszorginstrument dat test dat geen enkele student een diploma ontvangt zonder te voldoen aan alle benodigde kwalificaties. Interne processen zouden in dat geval strikt en rigoureuus moeten zijn om te waarborgen dat er geen fouten worden gemaakt. Een focus op excellentie vereist een kwaliteitszorginstrument dat verschillen tussen opleidingen en instituten inzichtelijk maakt op een vergelijkbare schaal, zodat men informatie heeft over waar kwaliteit beter is. Daarentegen moet bij een conceptualisatie van kwaliteit als geschiktheid voor het beoogde doel erkennen dat instellingen niet op dezelfde schaal beoordeeld en vergeleken kunnen worden, omdat hun doelstellingen en onderwijsvisies substantieel verschillen. Er is geen enkel kwaliteitszorgmechanisme dat al deze verwachtingen tegelijkertijd kan waarmaken.

In de praktijk spelen de *Europese Standaarden en Richtlijnen voor Kwaliteitszorg in de Europese Ruimte voor Hoger Onderwijs* (ESG) een belangrijke rol in het definiëren van kwaliteit. Aan deze standaarden en richtlijnen moet immers worden voldaan om een internationaal erkend diploma af te kunnen geven. Onderwijsinstellingen kunnen door het halen van deze standaarden laten zien dat een opleiding van voldoende kwaliteit is. De ESG houdt daarbij zo veel mogelijk rekening met de verschillende doelstellingen van onderwijs en de verschillende behoeftes van groepen belanghebbenden (2015). De doelstellingen zijn onder andere studenten voorbereiden op actief burgerschap, studenten klaarstomen voor hun toekomstige carrière en het ondersteunen van hun persoonlijke ontwikkeling. Andere doelstellingen zijn het creëren van een brede en vergevorderde kennisbasis en het stimuleren van onderzoek en innovatie. Het begrip kwaliteit wordt niet expliciet gedefinieerd, maar het wordt beschreven als “de uitkomst van interactie tussen docenten, studenten en de institutionele leeromgeving” (ESG, 2015, p. 7). Deze conceptualisatie en bijbehorende standaarden nemen diverse aspecten en verwachtingen van kwaliteit mee. In recente discussies over kwaliteit en kwaliteitszorg ligt de focus hoofdzakelijk op standaarden en competenties. Middels kwaliteitszorg wordt het hoger onderwijs geacht te evalueren of studenten een bepaald competentieniveau behalen en of hun leeromgeving het behalen van deze competenties ondersteunt. Deze benadering is ook geïntegreerd in de Dublin descriptoren die binnen het Bologna raamwerk de eindtermen voor opleidingen aan universiteiten en hogescholen in Europa beschrijven.

2.2 Wat is kwaliteitszorg?

Onder kwaliteitszorg verstaat men in de literatuur “systematische, gestructureerde en doorlopende aandacht voor kwaliteit in termen van kwaliteitsborging en kwaliteitsverbetering” (Vroeijenstijn, 1995, p. xviii). Dit komt overeen met de definitie vanuit de ESG, die kwaliteitszorg definieert als alle activiteiten die vallen binnen de doorlopende kwaliteitsverbeteringscyclus, waarbij expliciet ook kwaliteitsborging als onderdeel benoemd wordt (ESG, 2015, p. 7).

In de kwaliteitszorgprocessen kan men een onderscheid maken tussen interne kwaliteitszorg en externe kwaliteitszorg. Bij interne kwaliteitszorg gaat het om het beleid voor kwaliteitszorg dat de instellingen zelf ontwikkelen en uitvoeren. Externe kwaliteitszorg is gericht op het afleggen van verantwoording aan een extern orgaan, zoals een kwaliteitszorgorganisatie of een andere organisatie die beoordeelt of een opleiding of instelling aan de overeengekomen standaarden voldoet.

Het feit dat kwaliteitszorg uit zowel interne als externe componenten bestaat en die twee systemen elkaar sterk beïnvloeden, maakt de kwaliteitszorg complex. Uit het voorgaande is immers al gebleken dat verschillende belanghebbenden verschillende verwachtingen hebben bij kwaliteit. Wederom zijn er verschillende denkbeelden over de doelstelling van een kwaliteitszorgsysteem. Er zijn diverse onderzoeken die de doelstellingen in kaart hebben gebracht (Brennan & Shah, 2000; Harvey & Newton, 2004; Schwartz & Wersterheijden, 2004). Het scala aan doelstellingen wordt vaak teruggebracht tot twee belangrijke doelstellingen: verantwoording en verbetering. Deze twee doelstellingen vormen een klassieke dichotomie (Beerkens, 2015, p. 235). Ook de ESG stelt dat deze twee doelstellingen, namelijk ‘*accountability*’ en ‘*enhancement*’, ten grondslag liggen aan alle kwaliteitszorgprocessen. Gezamenlijk moeten deze twee aspecten vertrouwen in de kwaliteit van onderwijs creëren (ESG, 2015, p. 7). Meer recent is er een derde, gelijkwaardige doelstellingen van kwaliteitszorg bijgekomen: transparantie.

Als kwaliteitszorg gericht is op verantwoording wordt voornamelijk gecontroleerd of opleidingen aan de afgesproken standaarden binnen het hoger onderwijs voldoen. De kwaliteit wordt gemonitord met als doel om tijdig aan de bel te kunnen trekken als de kwaliteit van een opleiding ondermaats is of dreigt te worden. Door minimumstandaarden te waarborgen, kan men erop vertrouwen dat het diploma van een opleiding waarde heeft. Ook kan er politieke legitimiteit worden verdiend. Loukkola, Peterbauer en Gover vatten het als volgt samen (2020, p. 1): intern moeten onderwijsinstellingen hun prestaties monitoren om besluitvorming en strategieën voor kwaliteitsverbetering te onderbouwen, terwijl externe belanghebbenden prestaties monitoren om de doelstelling prijs-kwaliteitverhouding te controleren.

Kwaliteitszorg die gericht is op verbetering heeft als doel om de kwaliteit van de opleiding verder te optimaliseren. Het is gebaseerd op de aanname dat kwaliteit continue aandacht en inspanning vraagt, en dat de kwaliteitszorgmechanismen er zijn om ontwikkelingen te monitoren, ondersteunen en te stimuleren. Niet alleen interne kwaliteitszorgprocessen, maar ook externe processen (bijvoorbeeld via externe panels of experts) kunnen zelfreflectie of dialoog over kwaliteitsverbetering stimuleren (Harvey, 2002).

In aanvulling op deze twee aspecten, kan men stellen dat voor het creëren van vertrouwen een aparte, derde doelstelling van kwaliteitszorg nodig is: transparantie. Harvey en Green (1993) noemden eerder het belang van consumentenbescherming, waarbij studenten en ‘indirecte’ consumenten zoals werkgevers de kwaliteit van het onderwijs bepalen op basis van de mate waarin hun eisen en wensen worden vervuld. Om een opleiding te kunnen beoordelen of om een goede geïnformeerde keuze voor een opleiding te kunnen maken, is transparantie nodig. Transparantie overstijgt de doelstelling van verantwoording. Het behalen van de verantwoordingsdoelstelling leidt bijvoorbeeld nog niet automatisch tot de mogelijkheid van benchmarking. Voor echte consumentenbescherming en transparantie moet vergelijkbare data over de sterke en zwakke punten van een opleiding inzichtelijk gemaakt worden. Dit betekent dat een kwaliteitszorgsysteem kwaliteitsverschillen tussen opleidingen en instellingen inzichtelijk zou moeten maken en daarmee waardevolle informatie voor studiekeizers en andere belanghebbenden ter beschikking kan stellen. Deze transparantie kan vervolgens ook weer leiden tot kwaliteitsverbeteringen, bijvoorbeeld doordat men kan leren van ‘*best practices*’ (Middlehurst & Woodhouse, 1995). Als deze derde doelstelling van kwaliteitszorg ook vervuld wordt, kan uiteindelijk het algemene vertrouwen in de kwaliteit een opleiding nog verder toenemen.

2.3 Een publiek belang in kwaliteitszorg

Vandaag de dag wordt kwaliteitszorg voornamelijk gezien als een verantwoordelijkheid van de instellingen zelf. Desondanks verscheen kwaliteitszorg krachtig op de agenda van het hoger onderwijs als een publiek belang waarbij overheidsorganen de kwaliteitszorgprocessen monitorden en stuurden. Recentelijk beïnvloedt het kwaliteitszorgbeleid nog in grote mate hoe internationale kwaliteitszorgprocessen eruit moeten komen te zien. Waarom is interventie vanuit de overheid nodig? En welk doel dient de overheid in de kwaliteitszorg vanuit conceptueel oogpunt? Vanuit een reguleringsperspectief zijn er drie belangrijke functies waardoor interventie vanuit de overheid noodzakelijk kan zijn voor de kwaliteit van het hoger onderwijs.

Consumentenbescherming. Het hoger onderwijs is een speciale soort dienst. In economische termen is het zowel een zeldzame aanschaf, als een ervaringsgoed. Dat laatste betekent dat de kwaliteit van deze dienst niet beoordeeld kan worden voordat het ervaren wordt. Omdat de keuze voor het volgen van een opleiding een eenmalige beslissing is met hoge directe en indirecte kosten, heeft een misplaatst vertrouwen hoge consequenties voor de persoon in kwestie. Vanuit het perspectief van consumentenbescherming is daarom een goed instrumentarium nodig om de kwaliteit van opleidingen te waarborgen. Een opleiding moet voldoen aan legitieme en redelijke verwachtingen over wat een graad in het hoger onderwijs zou moeten bieden.

Faciliteren van een weloverwogen studiekeuze. Veel landen hebben marktmechanismen in de coördinatie van het hoger onderwijs gebracht, vaak met de redenering dat marktwerking kwaliteitsverbeteringen in het onderwijs stimuleert. Een onderliggende aanname van marktwerking is dat er valide en betrouwbare informatie beschikbaar is, zodat een weloverwogen keuze tussen verschillende aanbieders gemaakt kan worden en daardoor hogere kwaliteit beloond en gestimuleerd kan worden. Het is duidelijk dat dergelijke informatie over opleidingen in het hoger onderwijs niet gemakkelijk verkrijgbaar is. Ervaringen uit het verleden hebben laten zien dat

overheidsinvesteringen in dataverzameling en datamonitoring essentieel zijn geweest om enigszins valide data over onderwijskwaliteit te bieden (Dill & Beerkens, 2013).

Efficiëntie en effectiviteit van publieke middelen. Het hoger onderwijs beslaat een omvangrijk gedeelte van de publieke middelen, via directe financiering en gesubsidieerde stelsels van studieleningen, maar ook alternatieve kosten (opportuiniteitskosten) van studenten. Er is dus een publiek belang in het waarborgen dat publieke middelen ingezet worden met maximale effectiviteit, en dat onderwijs het algemeen belang dient.

3. Methodologie

3.1 Literatuurstudie

Er is een overweldigende hoeveelheid literatuur over de kwaliteit van het hoger onderwijs beschikbaar. De onderstaande tabel (tabel 1) laat zien hoeveel hits bepaalde zoektermen geven. Daarbij is uitsluitend gekeken naar wetenschappelijke literatuur en zijn inzichten uit rapporten van onderwijs- en overheidsorganisaties ('grey literature'), artikelen geproduceerd door commerciële organisaties zoals onderzoeksbureaus, of populair wetenschappelijke blogposts niet eens meegenomen.

Tabel 1: Zoekresultaten literatuurstudie

Database	Zoektermen	Waarin zoeken	Tijdsinterval	Aantal hits
Google Scholar	Quality + higher education	Collegiaal getoetste artikelen	1990 - heden	1.830.000
Google Scholar	Quality assurance + higher education	Collegiaal getoetste artikelen	1990 - heden	23.200
Web of Science	Indicator + quality + higher education	Categorie: educational research	1990 - heden	966
Web of Science	Quality management + higher education	Categorie: educational research + education scientific disciplines + management	1990 - heden	4.915
Web of Science	Quality + higher education + case study	Categorie: educational research + education scientific disciplines + Social sciences interdisciplinary	1990 - heden	2.086

Niet alle resultaten zijn relevant voor dit onderzoek. Voor het conceptueel kader hebben wij ons beperkt tot conceptuele artikelen die toonaangevend zijn in de literatuur en vaak geciteerd worden in andere onderzoeken. Om tot een goede selectie van literatuur te komen voor de beantwoording van de onderzoeksvragen is eerst gecontroleerd of de literatuur werkelijk betrekking heeft op kwaliteit van hoger onderwijs op opleidingsniveau, en zelfs dan is niet alle literatuur direct bruikbaar voor dit onderzoek. Veel literatuur is theoretisch of normatief van aard. Zo zijn er bijvoorbeeld veelvoud aan artikelen te vinden die een raamwerk ontwikkelen voor indicatoren die de kwaliteit van onderwijs zouden kunnen meten, maar is er buiten deze voorstellen om weinig aandacht voor welke sets van indicatoren daadwerkelijk gebruikt worden door hoger onderwijsinstellingen en welke impact dit heeft. Het aantal empirische studies dat analyseert wat daadwerkelijk in het hoger onderwijs plaatsvindt, en welke effecten dit heeft, is relatief beperkt. Het is echter de doelstelling van deze overzichtsstudie om tot een synthese te komen van deze literatuur.

Om de onderzoeksvragen te beantwoorden is gekozen voor de volgende aanpak. Allereerst zijn artikelen van de afgelopen 10 jaar uit relevante tijdschriften binnen het vakgebied, namelijk *Quality in Higher Education* en *Higher Education*, gescand. Alle artikelen die aan de hand van *empirisch onderzoek* inzicht bieden in kwaliteit en verantwoording *op het opleidingsniveau* zijn geselecteerd. Zowel institutionele case studies als meta-analyses en systematische reviews voldoen aan de selectiecriteria. Op basis van deze artikelen is verder gewerkt vanuit een ‘*snowballing approach*’ om aanvullende, relevante wetenschappelijke literatuur te vinden. Ook zijn onderzoeksrapporten van belangrijke Nederlandse en Europese onderwijsorganisaties zoals de Onderwijsraad, de European University Association (EUA), en kwaliteitszorgorganisaties van andere landen geraadpleegd. Er is daarbij gericht gezocht op meta-analyses en literatuuroverzichten over het gebruik van indicatoren en managementprocessen in kwaliteitszorg.

3.2 Interviews

Om inzicht te krijgen in wat er speelt op hogescholen en universiteiten in Nederland zijn kwaliteitszorgmedewerkers van in totaal twintig verschillende instellingen geïnterviewd. De interviews zijn bijna allemaal digitaal afgenomen (gebruikmakend van videobellen) en besloegen 45-60 minuten. Het aantal respondenten per instelling wisselde tussen de 1 en 3. De instellingen en respondenten zijn geselecteerd op basis van aanbevelingen van koepelorganisaties en middels de eerdergenoemde ‘*snowballing approach*’ (via-via).

Tijdens het interview zijn respondenten gevraagd naar (a) de wijze waarop indicatoren en databronnen gebruikt worden, (b) hoe indicatoren en databronnen geselecteerd worden, en (c) voor welke doelstelling deze gegevens gebruikt worden. Daarnaast is gevraagd naar de vormgeving van interne processen, en de wisselwerking tussen interne en externe componenten van de kwaliteitszorg. Tot slot werden vragen gesteld over knelpunten bij het gebruik van kwaliteitszorginstrumenten en de toekomstbestendigheid van de kwaliteitszorg.

De interviews waren semigestructureerd. Dat wil zeggen dat de vragen vooraf door de onderzoekers zijn vastgesteld, maar de volgorde waarin de vragen worden gesteld was flexibel en respondenten kregen de ruimte om vrij en open deze vragen te beantwoorden. Een voordeel van deze opzet ten opzichte van een gesloten interview of survey is dat respondenten zelf aangeven en prioriteren wat zij belangrijk vinden. Ook is er ruimte voor nuance of toelichting bij elk antwoord. Een mogelijk nadeel is dat antwoorden niet perfect vergelijkbaar zijn. Als een respondent een bepaald element van de kwaliteitszorg niet benoemt, is dat niet direct een bewijs dat dit element niet aanwezig is bij de instelling waar de respondent werkzaam is. Om voor dit nadeel te compenseren is aan iedere respondent gevraagd om een digitaal overzicht van *key performance indicators* (KPI's), managementrapportages of jaarrapportages na te sturen. Dit draagt bij aan objectivering van antwoorden en een betere vergelijkbaarheid tussen de instellingen. Sommige respondenten is na afloop van het interview bovendien om overige aanvullende informatie gevraagd. In andere gevallen konden antwoorden waar nodig aangevuld worden op basis van openbare informatie, zoals de officiële websites van de onderwijsinstellingen of beoordelingsrapporten van de Nederlands-Vlaamse Accreditatieorganisatie (NVAO).

3.3 Case studies

Dit rapport presenteert twee internationale case studies om de ervaringen van andere landen in meer diepte te kunnen bestuderen. Voorafgaand aan de case studies worden vier alternatieve manieren geïdentificeerd waarop landen met de kwaliteit op het opleidingsniveau omgaan, namelijk middels zelfaccreditatie, kwantitatieve *scoreboards*, thematische evaluaties, en via private accreditatie of certificering door beroepsorganisaties. De eerste twee aanpakken kunnen gezien worden als een alternatief voor opleidingsaccreditatie. Beide instrumenten zijn ontwikkeld voor een vergelijkbaar doeleinde en beoordelen opleidingen op een systematische en regelmatige manier. Om deze reden worden de twee eerste aanpakken uitgewerkt in een case study. De andere twee aanpakken zijn ook relevante alternatieven maar zijn niet gericht op een systematische beoordeling van individuele opleidingen. Thematische evaluaties zijn immers *ad hoc snapshots* van opleidingen binnen een bepaald vakgebied of van een specifiek aspect van onderwijs. Certificaten zijn steeds populairder aan het worden en passen goed bij de huidige cultuur van regelgeving, maar het is geen publiek instrument dat opleidingen verplicht beoordeelt.

Noorwegen is geselecteerd als een casus voor een systeem met een zelfaccreditatie. Het Noorse hoger onderwijssysteem is vergelijkbaar met Nederland wat betreft de binaire structuur (hoger beroepsonderwijs en wetenschappelijk onderwijs), het aantal instellingen en het gebrek aan sterke concurrentie tussen de instellingen. Noorwegen wordt vaak gezien als een succesvol voorbeeld van de instellingsaudit, en is daarom een informatieve casus die laat zien hoe de kwaliteit op het opleidingsniveau benaderd wordt in een kwaliteitszorgsysteem waar goede kwaliteitsprocessen op een instellingsniveau centraal staan. Engeland dient als de tweede casus. In tegenstelling tot Noorwegen heeft Engeland juist een ander type hoger onderwijssysteem, maar de ontwikkeling van het kwaliteitszorgsysteem in Engeland is zeer informatief en leerzaam. Engeland was een voorloper in de kwaliteitszorg in Europa, maar door de hoge administratieve lasten die kwaliteitsbeoordeling op een opleidingsniveau met zich mee bracht, veranderde Engeland het systeem. Engeland kent twee beleidsinstrumenten: een *'light-touch' Quality and Standards Review* als basis, en het aanvullende *Teaching and Excellence Framework (TEF)* dat de kwaliteit van de opleidingen in een simpele goud-zilver-brons categorisatie zichtbaar probeert te maken.

Informatie over de casussen werd verzameld op basis van documentatie op de websites van de betreffende kwaliteitszorgorganisaties, wetenschappelijke literatuur en evaluatiestudies. Voor aanvullende informatie over Noorwegen is ook een expertinterview afgenomen.

4. Databronnen en indicatoren

Er is geen gebrek aan voorstellen voor indicatoren om onderwijskwaliteit te meten. In een recent overzicht zijn meer dan 600 verschillende indicatoren op een rij gezet (SQELT project, 2020). In dit hoofdstuk wordt een poging gedaan om de grote hoeveelheid aan literatuur over het gebruik van indicatoren voor kwaliteitssturing, en de lessen die daaruit getrokken zijn, tot een geheel te brengen. Daarvoor moet allereerst afgebakend worden wat een indicator is. Daarna moet de analyse-eenheid bepaald worden en tot slot wordt gespecificeerd wat precies geëvalueerd wordt.

Op basis van bestuurskundige literatuur definieert dit rapport een indicator als een kwantitatieve of kwalitatieve meting die probeert de belangrijkste aspecten van organisatorische kwaliteit in kaart te brengen (bijv. Van Dooren et al., 2015). Een belangrijke eigenschap van een indicator is dat het niet zo zeer beschrijvende informatie geeft om verschillen uit te lichten, maar dat het een signaal kan afgeven over betere of slechtere kwaliteit. Indicatoren vullen niet noodzakelijk een absoluut kwaliteitsoordeel, omdat indicatoren op basis van context geïnterpreteerd moeten worden. Desondanks is er een vermeend verband tussen de richting van een indicator en een kwaliteitsoordeel.

In dit rapport ligt de focus op het niveau van individuele opleidingen. Dit is niet evident. Er zijn veel voorbeelden en analyses beschikbaar over prestatie-indicatoren op het instellingsniveau. Ook bij beleidsinstrumenten zoals prestatiecontracten of prestatiegebonden financieringsmechanismen ligt de nadruk op indicatoren die kwaliteit op het instellingsniveau meten. Toch zijn indicatoren die gebruikt worden op het instellingsniveau niet noodzakelijk passend of haalbaar op het opleidingsniveau. Het is aangetoond dat overkoepelende instrumenten of rankings niet gegeneraliseerd kunnen worden op individuele opleidingen, omdat er een grote kans is dat er veel variatie is tussen verschillende opleidingen binnen een instelling (Cave et al. 1996). Aan de andere kant is er ook kritiek op een sterke focus op het opleidingsniveau, omdat hedendaagse initiatieven in het onderwijs de nadruk leggen op ‘personalisatie’ van opleidingen, flexibilisering binnen een opleiding en modulair onderwijs (Li & Piltz, 2016). Om deze reden kan een opleiding haar duidelijke scheidingslijnen verliezen.

Dit rapport richt zich op de onderwijskwaliteit van een opleiding in brede zin, zoals eerder besproken. De keuze van de indicatoren die in dit hoofdstuk aan bod komen reflecteert de diverse waarden en doelstellingen van het onderwijs. Op deze heterogeniteit wordt ook gereflecteerd. Voordat echter gekeken wordt naar de verschillende soorten indicatoren die het vaakst gebruikt worden om kwaliteit in kaart te brengen, met hun voor- en nadelen, besteden we aandacht aan de achterliggende databronnen: namelijk wie bepaalde gegevens en informatie verzamelt of bezit.

4.1 Verschillende soorten databronnen

Informatie over de kwaliteit van opleidingen komt voort uit verschillende soorten databronnen. Daarbij is een belangrijk onderscheid te maken tussen de informatie die wordt verzameld door instellingen zelf, en informatie die wordt verzameld door andere actoren. Onder die zogeheten andere actoren vallen zowel overheidsorganisaties

als partijen buiten de overheid. Het ligt voor de hand dat de overheid gegevens wil hebben over geaccrediteerde opleidingen, omdat dit opleidingen zijn die in aanmerking komen voor bekostiging uit publieke middelen. Hierover moet publieke verantwoording worden afgelegd. De overheid heeft er daarnaast belang bij om zelf zicht te houden op de kwaliteit van deze opleidingen om te monitoren in hoeverre nationale beleidsdoelen behaald worden. Deze gegevens kunnen direct door de overheid verzameld worden of via een onafhankelijke organisatie die daarvoor opgericht is. Ook private actoren hebben er soms (commercieel) belang bij om gegevens te verzamelen die de kwaliteit van opleidingen in kaart kunnen brengen. Denk daarbij bijvoorbeeld aan gegevens over de aansluiting tussen opleidingen en de arbeidsmarkt.

Er zijn grote verschillen in het soort data dat verzameld wordt door verschillende actoren. Hier volgen de vier meest belangrijke databronnen.

Informatiebronnen van de instellingen

Instellingen zelf hebben de meeste data over hun opleidingen. Uit de leer- en informatiemanagementsystemen verzamelen instelling gegevens over bijvoorbeeld studiesucces, uitval, studieduur, studentenaantallen, internationale uitwisselingen, honoursonderwijs, contacturen, etc. Instellingen verzamelen ook data via enquêtes, zoals vakevaluaties en programmaevaluaties, en soms ook via een alumni- of medewerkersenquête. De voordelen van deze databron zijn dat die dichtbij de opleiding staat, diep op de kernprocessen kan ingaan, en inhoudelijke informatie over opleidingen kan bevatten. De betrouwbaarheid van de data verschilt echter per indicator en is ook afhankelijk van het niveau van de informatiemanagementsystemen binnen een instelling.

Als gegevens door de instelling zelf verzameld worden, is er vaak sprake van versnippering van de data. Ook zijn lang niet alle data die door onderwijsinstellingen zelf verzameld worden publiek toegankelijk. Het is zelfs mogelijk dat instellingen naast hun eigen interne rapportages een aparte externe rapportage maken, vooral als de resultaten uit rapportages worden gekoppeld aan financiering (Kis, 2005, p. 18). Dit maakt het moeilijker om opleidingen van verschillende instellingen op basis van hun eigen gegevens te vergelijken. De mogelijkheid tot benchmarking valt weg.

Er zijn echter wel voorbeelden van afspraken tussen instellingen om hun eigen verzamelde data te delen, maar het kan soms ook juist een voordeel zijn dat de resultaten uitsluitend voor interne monitorings- en kwaliteitsverbeteringsprocessen gebruikt worden. Volgens de literatuur levert interne dataverzameling een belangrijke bijdrage aan interne zelfreflectie en is het zelfs essentieel voor duurzame verbeteringen (Kis, 2005, p. 18). Allereerst hebben instellingen daarmee de vrijheid om zelf te bepalen wat en hoe ze willen meten, waardoor gegevens beter kunnen aansluiten op de specifieke en actuele context van de instelling. Ook is het gevaar van datamanipulatie geminimaliseerd en is de bereidheid om nauwkeurige en eerlijke data over kernprocessen te verzamelen hoger.

Nationale enquêtes door externe partijen

Middels zelfrapportage kunnen diverse belanghebbenden, zoals studenten, alumni, het beroepenveld of betrokken medewerkers hun ervaring met de opleiding delen. Vaak wordt via vragenlijsten de mate van tevredenheid over verschillende aspecten van het onderwijs in kaart gebracht.

In de afgelopen twintig jaar zijn in veel landen nationale studenten- en alumni-enquêtes opgezet. Op Europees niveau komen alumni-enquêtes het meest voor. Internationaal blijkt dat nationale enquêtes vaker worden gebruikt bij universiteiten dan zogeheten niet-universiteiten (hogescholen). Opvallend is tot slot dat enquêtes vaak worden gebruikt voor evaluaties, maar zelden in verband worden gebracht met accreditatie (ENQA, 2003, p. 30). De enquêtes zijn meestal door de overheid gefinancierd, maar niet altijd. Ze worden uitgevoerd door verschillende type actoren: publieke kwaliteitsagentschappen, semipublieke stichtingen of private organisaties.

De uitkomsten van enquêtes circuleren vaak in keuzegidsen en kunnen daarmee door het publiek worden gebruikt om opleidingen met elkaar te vergelijken. Bij het vergelijken van opleidingen kunnen echter verschillende betrouwbaarheidsproblemen optreden. Zo is het bij nationale enquêtes twijfelachtig in hoeverre opleidingen uit verschillende disciplines met elkaar vergeleken kunnen worden. De verwachtingen en vereisten kunnen bij een kunstopleiding of bij een technische opleiding bijvoorbeeld in grote mate verschillen. Bij internationale enquêtes waarin opleidingen uit verschillende landen met elkaar vergeleken worden komen aanvullende problemen kijken. Volgens de Vereniging van Universiteiten (VSNU) zijn door belangrijke systeemverschillen (bijvoorbeeld de financieringsstructuur van het hoger onderwijs, maar ook door verschillen in aanmeldingseisen of selectievoorwaarden van een opleiding) internationale vergelijkingen altijd lastig te interpreteren (2012, p. 65). Het gevaar is dat het publiek eenvoudig misleid kan worden door zulke data. Tot slot spelen er nog andere methodologische kwesties bij tevredenheidsenquêtes die later aan bod komen als dit rapport nader ingaat op het gebruik van tevredenheidsindicatoren.

Gecentraliseerde statistische informatie

Gecentraliseerde statistische informatie wordt verzameld of verwerkt door statistische bureaus, ministeries van onderwijs of uitvoerende organisaties (bijv. de Dienst Uitvoering Onderwijs (DUO) in Nederland). De meest gebruikte kwantitatieve data in Europa zijn statistieken over studenten- en docentenaantallen, zoals cijfers over de instroom, uitstroom en student-stafratio's (ENQA, 2003, p. 29).

Deze data worden meestal geaggregeerd op instellingsniveau, omdat de gegevens per opleiding te klein kunnen zijn om generieke uitspraken te kunnen doen. Het nadeel van het gebruik van een instellings- of sectorindeling is dat de resultaten vervolgens niet meer te interpreteren zijn op opleidingsniveau. Een ander nadeel is dat de (arbeids)kosten om deze data te verzamelen erg hoog kunnen liggen (Kis, 2005, p. 18). Desondanks besluiten overheden vaak deze data te verzamelen om vergelijkingen te maken binnen het hoger onderwijs of voor benchmarkingsdoeleinden.

Bij kwantitatieve data is de vergelijkbaarheid inderdaad relatief hoog, maar het is niet altijd duidelijk hoe de gegevens samenhangen met de kwaliteit van opleidingen. Er wordt vaak gewaarschuwd voor het door elkaar

halen van prestatie-indicatoren met andere statistieken en managementdata die weinig met kwaliteit van opleidingen te maken hebben (ENQA, 2003, p. 19). Het aantal aanmeldingen zegt bijvoorbeeld niet noodzakelijk iets over de kwaliteit. Dergelijke indicatoren kunnen geen absoluut kwaliteitsoordeel vellen, en deze statistische gegevens krijgen pas betekenis na een proces van contextualisering.

Kwalitatieve beoordelingen door nationale (of internationale) organisaties

Data komt niet alleen voort uit kwantitatieve metingen. Bij kwalitatieve data zijn de afwegingen tegenovergesteld aan die bij kwantitatieve data. Het is op basis van kwalitatieve gegevens weliswaar moeilijker om gegevens tussen opleidingen of door de tijd heen met elkaar te vergelijken, maar er is meer ruimte om verklaringen voor resultaten te vinden en gegevens te contextualiseren.

Een belangrijke informatiebron over de kwaliteit van een opleiding zijn ‘*judgement calls*’, die uitkomsten van complexe evaluatie processen reflecteren (Ewell, 1999). Voorbeelden van kwalitatieve databronnen bij dergelijke processen zijn accreditatierapporten, evaluaties, en thematische overzichten. Deze rapporten zijn het vaakst geproduceerd door landelijke kwaliteitsorganisaties (bijv. NVAO). De rapporten zijn meestal openbaar, zoals ook gevraagd in de ESG. Een andere vorm van een kwalitatieve beoordeling betreft internationale accreditatiesystemen, zoals bijvoorbeeld EQUIS voor *business schools* of ABET voor *engineering* opleidingen (Stensaker & Harvey, 2006). Ook deze oordelen zijn gebaseerd op een grondig evaluatieproces.

Een sterk punt van dergelijke rapporten is dat ze een breed beeld van onderwijskwaliteit schetsen, verschillende databronnen met elkaar combineren, en de mogelijkheid bieden om gegevens in context te plaatsen. Bovendien wordt vaak het oordeel van een externe ‘*peer*’ meegenomen. Aan de andere kant is het kostbaar om dergelijke informatie te verzamelen op het opleidingsniveau, en daardoor is deze informatie niet altijd beschikbaar. Ook is er bewijs dat deze informatiebronnen niet vaak gebruikt worden door externe belanghebbenden, vanwege hun uitgebreidheid en alomvattendheid, en is het zoals eerder benoemd geen gemakkelijk hulpmiddel om tussen opleidingen te vergelijken (Schwartz & Westerheijden, 2004). Er zijn pogingen gedaan om de toegevoegde waarde van rapporten te verhogen, bijvoorbeeld door accreditatierapporten in hun eindoordeel niet alleen het onderscheid tussen voldoende en onvoldoende kwaliteit te laten maken, maar ook te laten differentiëren tussen goede en excellente kwaliteit op verschillende dimensies (Aijing & Westerheijden, 2018). Dergelijke pogingen hebben niet altijd het gewenste resultaat gehad.

Andere initiatieven voor dataverzameling

Initiatieven om ‘value added’ data te verzamelen. Als men simpelweg kijkt naar de eindresultaten van studenten via indicatoren zoals slagingspercentages of kansen op de arbeidsmarkt, kan een vertekend beeld van opleidingskwaliteit ontstaan. Een opleiding kan slechts deels deze uitkomsten verklaren. Een ander, substantieel deel wordt veroorzaakt door karakteristieken van de studentenpopulatie die zich heeft aangemeld voor de opleiding, waaronder hun cognitieve vaardigheden en sociaal kapitaal (Beerkens, 2021, p. 8). Hierdoor werd ‘value added’ een belangrijk concept bij dataverzameling. Door te kijken naar *value added* wordt een poging

gedaan om alleen te kijken naar wat een opleiding heeft toegevoegd aan de ontwikkeling van een bepaalde student door gegevens te controleren voor de beginpositie van studenten (Beerkens, 2021, p. 9).

Australië is een voorloper geweest in deze ontwikkeling, maar in 2006 kondigde de Organisatie voor Economische Samenwerking en Ontwikkeling (OECD) aan dat ook zij in een grootschalig project data wilden verzamelen over hoeveel studenten nu feitelijk in hun (bachelor)opleiding leren, en hoe dit in internationaal verband vergelijkbaar is (OECD, 2015). Het project is bekend onder de naam *Assessment of Learning Outcomes in Higher Education* (AHELO). Hoewel het project een ideaalbeeld schetst van hoe leeruitkomsten transparant, op basis van kwantitatieve gegevens, en zonder vertekening door achtergrondkenmerken van studenten gemeten kunnen worden, is om diverse politieke en technische redenen het initiatief niet uitgegroeid tot een regulier beoordelingsinstrument (Beerkens, 2021, p. 9).

Rankings. Hoger onderwijsinstellingen concurreren met elkaar in een wereldwijde markt. Hierdoor is een behoefte ontstaan om prestaties van hoger onderwijsinstellingen te meten en tegen elkaar af te zetten. Rankings zijn uitgegroeid tot een belangrijk instrument, ondanks kritieken op de methodologische keuzes die achter de rankings schuilgaan (Dill & Soo, 2005, p. 496; Hazelkorn, 2009, p. 4). De belangrijkste kritiek op rankings is dat zij niet op een valide en betrouwbare wijze kwaliteit meten. Oordelen zijn vaak op basis van ruwe methodes of op basis van reputatie tot stand gekomen (Dill & Soo, 2005; Hazelkorn, 2009). Politici gebruiken de rankings echter veelvuldig om sterke punten onder de aandacht te brengen, de instellingen zelf gebruiken rankings om doelstellingen te formuleren, en academici zetten rankings vaak in om hun eigen professionele reputatie en status te versterken (Hazelkorn, 2009, p. 1). Sterker nog, volgens Hazelkorn kunnen zelfs instellingen die niet in de ranking staan door de ‘obsessie met rankings’ beïnvloed worden, bijvoorbeeld uit angst om genegeerd, uitgesloten of gepasseerd te worden (2009, p. 19).

Rankings kunnen worden geproduceerd door publieke organisaties, beroepsorganisaties en commerciële organisaties, waaronder kranten en tijdschriften (Dill & Soo, 2005, pp. 496-498). De meeste rankings rapporteren prestaties op instellingsniveau in plaats van opleidingsniveau. Ook ligt er vaak een focus op onderzoeksactiviteiten en wordt niet (expliciet) naar onderwijskwaliteit gekeken. Dit maakt veel rankings ongeschikt om de kwaliteit van onderwijs op opleidingsniveau te beoordelen.

U-Multirank was een initiatief om een meer valide en betrouwbare ranking te ontwikkelen en daardoor de dysfunctionele effecten van de commerciële rankings te verminderen. Deze opleidingsranking komt tot stand op basis van verschillende databronnen. Om de administratieve lasten van dataverzameling te verlichten wordt zo veel mogelijk gebruik gemaakt van informatie die publiek toegankelijk is. In Nederland gaat het dan om data van DUO, Studiekeuze123, VSNU en de Vereniging Hogescholen. Desondanks blijft aanvullende zelfrapportage van de onderwijsinstellingen en studenten nodig. Vanaf het begin richtten de belangrijkste kritieken op U-Multirank zich op deze soort databronnen. Data op basis van zelfrapportage zou gevoelig zijn voor manipulatie (Rauvargers, 2011, p. 17). Toch probeert U-Multirank op deze manier feiten over de opleiding in eenzelfde format te verzamelen, zodat er geen data in de ranking ontbreekt. Bovendien voorkomt het gebruik van eenzelfde format dat definities onderling van elkaar verschillen. Dit zou de vergelijkbaarheid ten goede komen, volgens U-Multirank. Ook worden tevredenheidsenquêtes gebruikt om de ervaring van studenten mee te nemen. Indien

mogelijk gebruikt U-Multirank nationale tevredenheidsenquêtes (zoals de Nederlandse Nationale Studenten Enquête (NSE)). Verder worden gegevens aangevuld via een eigen enquête (U-Multirank, 2021c). Critici waarschuwen daarbij dat de vergelijkbaarheid van de tevredenheidsenquêtes in het geding kan komen door verschillen in beoordelingscultuur per land (Rauvargers, 2011, p. 17).

Toch zijn er pluspunten van U-Multirank. Anders dan andere grote rankings richt U-Multirank zich nadrukkelijker op meerdere aspecten van het hoger onderwijs en niet hoofdzakelijk op onderzoek. De aspecten naast onderzoek zijn doceren en leren, kennisoverdracht, internationale oriëntatie en regionale betrokkenheid (U-Multirank, 2021d). Ook wil U-Multirank geen gesimplificeerde ranglijsten presenteren, maar worden de scores op indicatoren in vijf brede prestatiegroepen (zeer goed tot zwak) geplaatst. Tot slot kunnen gebruikers zelf kiezen naar welke selectie van indicatoren zij kijken in hun vergelijkingen (U-Multirank, 2021e).

‘Big data’. Door digitalisering is de hoeveelheid beschikbare data sterk toegenomen. Dit geldt ook voor hoger onderwijsinstellingen. De data die voor instellingen beschikbaar zijn omvatten niet alleen administratieve gegevens over cijfers en studiesucces. Door digitalisering kan bijvoorbeeld ook worden opgeslagen hoe vaak studenten inloggen op online leeromgevingen, hoelang zij fysiek aanwezig zijn op campus, en welke onderwerpen zij op *social media* interessant vinden (Beerkens, 2021, p. 11).

Ten opzichte van traditionele data lijkt *‘big data’* dynamisch, minder statisch, en minder arbeidsintensief en kostbaar (Beerkens, 2021, p. 18). In de toekomst kunnen mogelijk modellen ontwikkeld worden die resultaten van accreditatieprocessen kunnen voorspellen (Prades et al., 2015). Er wordt momenteel veel geëxperimenteerd met *big data*, voornamelijk voor de verbetering van leerprocessen, maar er zijn geen concrete voorbeelden van hoe *big data* ingezet kan worden voor de zichtbaarheid van de kwaliteit van opleiding.

4.2 Verschillende soorten indicatoren

Er is geen consensus over welke indicatoren de kwaliteit van een opleiding op een betrouwbare en valide manier kunnen duiden. In deze paragraaf geven we weer welke indicatoren in de praktijk systematisch en regelmatig gebruikt worden om de kwaliteit van onderwijs op opleidingsniveau in kaart te brengen. Omdat hogescholen en universiteiten niet uitsluitend onderwijsinstellingen, maar ook onderzoeksinstituten zijn, zijn er ook diverse sets van indicatoren te vinden die ingaan op de kwaliteit van onderzoek. Ook richt een groot aantal sets zich op de kwaliteit van onderwijs op het instellingsniveau in plaats van het opleidingsniveau. Dit onderzoek richt zich uitsluitend op indicatoren die nadrukkelijk gaan over de kwaliteit van onderwijs bij opleidingen. Het rapport begint met een overzicht van classificaties van de verschillende indicatoren, en daarna wordt gekeken naar de indicatoren die in de praktijk het vaakst voorkomen of als meest relevant gezien worden.

4.2.1 Classificaties uit de literatuur

Er zijn een honderdtal losse indicatoren te onderscheiden om onderwijskwaliteit te meten. Een recent project stelt zelfs een lijst van meer dan 600 indicatoren voor (SQELT project, 2020). Daarmee is het aantal en het soort indicatoren om onderwijskwaliteit te meten veel groter en uiteenlopend dan indicatoren om onderzoekskwaliteit

te meten (Romainville, 1999, p. 418). Om het overzicht te bewaren, wordt gekeken naar de manier waarop deze verschillende indicatoren voor de kwaliteit van opleidingen geassocieerd worden. Het resultaat is een synthese van de literatuur.

Er zijn in de academische literatuur en zogeheten ‘grey literature’ veel verschillende classificaties te vinden van indicatoren die gebruikt worden om de kwaliteit van onderwijs op opleidingsniveau te beoordelen. De meeste auteurs en organisaties classificeren op basis van welk kwaliteitsaspect door een indicator gemeten wordt. Juist omdat onderwijskwaliteit geen eenduidig begrip is, kunnen losse indicatoren zeer uiteenlopende dimensies van kwaliteit meten. Er is in de literatuur echter geen overeenstemming over het aantal dimensies dat onderscheiden kunnen worden en waar de grens tussen deze dimensies ligt. Bovendien is de kwaliteit van een opleiding nauwelijks direct meetbaar. Daarom zijn de indicatoren meestal zogenoemde ‘proxy-indicatoren’ die (hopelijk) op een indirecte manier een signaal geven over betere of slechtere kwaliteit zonder deze direct te kunnen meten.

Harvey en Green maken een belangrijk onderscheid tussen enerzijds het uitkomstaspect van kwaliteit en anderzijds het procesaspect van kwaliteit (1993). Bij het eerste aspect, het uitkomstaspect, kan kwaliteit in absolute zin gemeten worden, bijvoorbeeld aan de hand van concrete standaarden, terwijl bij het procesaspect kwaliteit beoordeeld wordt ten opzichte van de processen die leiden tot die gewenste uitkomsten, ongeacht of een absolute drempel gehaald wordt (Harvey & Green, 1993).

Veel organisaties breiden deze classificatie verder uit tot meerdere categorieën. De *Academic Cooperation Association* (ACA), *The European Association for Quality Assurance in Higher Education* (ENQA) en de *Deutscher Akademischer Austauschdienst* (DAAD) onderscheiden in een rapport aan het Europees Parlement inputindicatoren, procesgerelateerde indicatoren en outputindicatoren (Wächter et al., 2015, p. 57). Bij de inputindicatoren wordt gekeken naar de hulpmiddelen die ten grondslag liggen aan de processen en resultaten van het onderwijs, zoals de selectie van studenten. Onder procesgerelateerde indicatoren vallen de leeromgeving en studenttevredenheid over het leerproces. Outputindicatoren kijkt naar de resultaten van het onderwijs, zoals slagingspercentages, het salaris van afgestudeerden of arbeidsplaatsen voor afgestudeerden.

Scheerens categoriseert indicatoren als volgt: input, proces, output/outcome en impact/long-term outcome (2004, pp. 118-120). Daarnaast maakt Scheerens een aparte categorie voor contextindicatoren voor metingen van kwaliteit op het systeemniveau. Deze contextindicatoren omvatten informatie over de demografie van de maatschappij en karakteristieken van het nationale onderwijsstelsel. Die kunnen in dit rapport buiten beschouwing gelaten worden, omdat deze indicatoren geen informatie geven over de kwaliteit van onderwijs op opleidingsniveau, maar eerder een hulpmiddel zijn om internationale vergelijkingen te maken. De overige categorieën van indicatoren zijn beter toepasbaar en meetbaar op opleidingsniveau. De inputindicatoren zeggen wederom iets over de hulpmiddelen die als investering voor het onderwijs gezien kunnen worden. Procesindicatoren richten zich op de kwaliteit van de leeromgeving en de organisatie van het onderwijs. Tot slot zijn er dus twee verschillende deelaspecten die de resultaten van het onderwijs meten. Output/outcome indicatoren hebben betrekking op directe resultaten die voortvloeien uit het onderwijs en zijn volgens Scheerens statistisch te meten, zoals cijfers over de doorstroom, afstudeerduur of drop-out. Impact/long-term indicatoren zijn vaak complexer en kijken naar de effecten van het onderwijs buiten de onderwijssector om. Naast

kwantitatieve indicatoren kent deze categorie ook kwalitatieve indicatoren die lange termijneffecten van het onderwijs meenemen. Een groot deel van deze indicatoren is echter wederom uitsluitend geschikt voor een beoordeling van de kwaliteit van onderwijs op een systeemniveau en niet voor de kwaliteit van onderwijs op opleidingsniveau, bijvoorbeeld als het gaat om de impact op jeugdwerkloosheid of emancipatie.

In de literatuur is geen eenduidigheid over de begrippen ‘resultaten’ en ‘uitkomsten’, of het Engelse ‘*output*’ en ‘*outcome*’. Regelmatig worden de begrippen door elkaar gebruikt of is de afbakening tussen de begrippen onduidelijk. Ook kan er tussen auteurs met hetzelfde label iets anders bedoeld worden, of gebruiken ze juist verschillende labels voor hetzelfde. In een overzichtsstudie van Chalmers wordt onderscheid gemaakt op basis van het soort data dat gebruikt wordt enerzijds en welk aspect van kwaliteit gemeten wordt anderzijds (2008). Dit leidt tot een relatief eenvoudige classificatie. Volgens haar worden kwantitatieve indicatoren gebruikt om ‘input’ en ‘output’ aspecten van de kwaliteit van het onderwijs te meten, terwijl kwalitatieve indicatoren ‘*process*’ (proces) en ‘*outcome*’ (uitkomst) kunnen meten (Chalmers, 2008, p. 7).

Ook bij Chalmers weerspiegelen de inputindicatoren de (menselijke, financiële of materiële) hulpmiddelen die gebruikt worden om een opleiding te ondersteunen. Belangrijke voorbeelden zijn het aantal studenten en docenten, de samenstelling van de studentenpopulatie (demografische kenmerken, zoals herkomst of geslacht) en de financiële uitgaven per student (Chalmers, 2008, pp. 10-11). Outputindicatoren kwantificeren de resultaten van het onderwijs. Belangrijke voorbeelden zijn cijfers over de doorstroom, het slagingspercentage en de baankans (Chalmers, 2008, pp. 11-12).

Uitkomstindicatoren meten net als outputindicatoren de effecten van het onderwijs, maar het gaat daarbij nadrukkelijk om effecten die complex zijn en niet direct ‘telbaar’ zijn. Als gevolg hiervan omvat deze categorie kwalitatieve indicatoren. Voorbeelden van uitkomstindicatoren die inderdaad vaak kwalitatief gemeten zullen worden zijn indicatoren die de vaardigheden van afgestudeerden, de motivatie van afgestudeerden om vakkennis te blijven actualiseren, of hun betrokkenheid bij de maatschappij in kaart brengen (Chalmers, 2008, p. 13). Chalmers gaat er echter aan voorbij dat het soms ook mogelijk is om complexe effecten of abstracte waarden uit te drukken in een kwantitatieve score. Een belangrijk voorbeeld van zo’n effect dat niet telbaar is, maar in de praktijk wel regelmatig kwantitatief wordt uitgedrukt is de mate van tevredenheid onder diverse belanghebbenden, zoals studenten en werkgevers. Deze vallen bij Chalmers ook onder ‘kwalitatieve’ uitkomstindicatoren. Toch wordt het abstracte begrip ‘tevredenheid’ om praktische redenen vaak numeriek uitgedrukt (bijvoorbeeld op een Likertschaal van 1-5).

Tot slot geven procesindicatoren informatie over hoe het onderwijs is vormgegeven via het officiële beleid en de gebruiken binnen een opleiding, en helpen daarmee bij het interpreteren van output- en uitkomstindicatoren (Chalmers, 2008, pp. 13-14). Belangrijke voorbeelden zijn de onderwijsmissie, onderwijsvisie en onderwijsdoelen, maar ook het beleid omtrent toetsing en feedback (Chalmers, 2008, p. 14). Aan deze indicatoren zullen inderdaad vaak kwalitatieve databronnen en metingen ten grondslag liggen.

De *European University Association* (EUA) gooit het over een andere boeg. Volgens hen kunnen de diverse aspecten van kwaliteit die men wil meten, worden samengevat in drie thema’s. Zij maken daarbij een onderscheid

tussen de kwaliteit van leren, kwaliteit van doceren, en kwaliteit van onderwijs in bredere zin (Loukkola et al., 2020, p. 14). De kwaliteit van leren omvat indicatoren om de resultaten van het leerproces te meten en zijn daarmee vergelijkbaar met het eerdergenoemde uitkomstaspect van Harvey en Green (1993). Het procesaspect van Harvey en Green wordt opgedeeld in twee verschillende thema's. De kwaliteit van doceren zegt iets over de leeromgeving. De kwaliteit van onderwijs in brede zin is een overkoepelende categorie van indicatoren die ervanuit gaat dat de leerervaring van studenten beïnvloed wordt door veel andere factoren buiten doceren om (Loukkola et al., 2020, p. 16).

Tabel 2: Classificaties van indicatoren - synthese op basis van literatuurstudie

Aspecten van kwaliteit	Concrete, telbare aspecten van kwaliteit (vaak kwantitatieve databronnen en metingen)		Complexe, abstracte aspecten van kwaliteit (vaak kwalitatieve databronnen en metingen)	
Harvey & Green (1993)	Procesaspecten			
		Uitkomstaspecten		
EUA (2020)	‘Kwaliteit van doceren’; ‘Kwaliteit van onderwijs in bredere zin’			
		‘Kwaliteit van leren’		
ACA, ENQA & DAAD (2015)	Input	Output		Proces
Scheerens (2004)	Input	Output/outcome (effecten binnen onderwijssector)		Proces
		Impact/long-term outcome (effecten buiten onderwijssector)		
Chalmers (2008)	Input	Output	Uitkomst	Proces
Belangrijkste indicatoren	Samenstelling van studentenpopulatie (herkomst of geslacht), het niveau van inkomende studenten, student-stafratio, de verhouding academisch en niet-academisch personeel, docentkwalificatie, financiële uitgaven per student en het aantal ondersteunende faciliteiten en diensten.	Doorstroomcijfers, uitval/switch, slagingspercentages, baankansen, tijd tot het vinden van een eerste baan en de hoogte van het salaris van alumni.	Vaardigheden van afgestudeerden, tevredenheid van studenten, alumni, werkgevers en andere belanghebbenden, de motivatie van afgestudeerden om vakkennis te blijven actualiseren en hun betrokkenheid bij de maatschappij.	Werkvormen van het onderwijs, de leeromgeving, tevredenheid met het leerproces, diversiteit, onderwijsmissie, onderwijsvisie en onderwijsdoelen, en het beleid omtrent toetsing en feedback.

In de bovenstaande tabel (tabel 2) wordt een overzicht gegeven van de verschillende classificaties. Harvey en Green (1993) en de EUA (2020) classificeren hun indicatoren thematisch. Voor de verschillende dimensies en conceptualisaties van onderwijskwaliteit (zie ook paragraaf 2.1) onderscheiden zij aparte categorieën van indicatoren. De classificaties van ACA, ENQA en DAAD (2015) en Scheerens (2004) presenteren een classificatie van indicatoren op basis van chronologie. Elke fase van het onderwijs heeft eigen indicatoren. Chalmers (2008) doet een poging om de classificatie te verbinden aan het soort databron en de meting die ten grondslag ligt aan de indicator (kwantitatief versus kwalitatief). Hoewel deze scheiding niet zo scherp gemaakt kan worden, omdat in de praktijk ook abstracte aspecten van onderwijs gekwantificeerd worden, houdt tabel 2 wel vast aan een indeling waarbij concrete, telbare aspecten van kwaliteit (vaak kwantitatieve databronnen en metingen) links gesorteerd staan en complexe, abstracte aspecten van kwaliteit (vaak kwalitatieve databronnen en metingen) rechts gesorteerd staan. Daarmee presenteert de tabel een overzicht van hoe indicatoren geclassificeerd kunnen worden op basis van verschillende dimensies van onderwijskwaliteit, de verschillende fases van onderwijs of op basis van databronnen en metingen – met bijbehorende voorbeelden van de belangrijkste indicatoren.

Eén ‘categorie’ van informatie over opleidingskwaliteit valt buiten dit classificatiesysteem, namelijk summatieve kwalitatieve beoordelingen. Om tot een oordeel te komen, worden hierbij meestal alle eerdergenoemde soorten indicatoren meegenomen. Vanuit het oogpunt van validiteit kunnen losstaande prestatie-indicatoren meer beschrijvende meetinstrumenten zoals nationale kwalificatieraamwerken of peer-reviews niet vervangen (Loukkola et al., 2020, p. 24). De ene soort indicator is niet per definitie beter of slechter dan de andere. Voor een volledig beeld zouden deze verschillende soorten indicatoren dus in combinatie met elkaar gebruikt moeten. Op die manier kunnen de kwaliteitszorgprocessen transparantie en vertrouwen in het onderwijs te vergroten.

4.2.2 Meest voorkomende indicatoren in de praktijk

Uit Europees onderzoek komt naar voren dat de indicatoren die het meest voorkomen in het hoger onderwijs de zogeheten inputindicatoren zijn, zoals student-stafratio's. De keuze valt op dit type indicatoren, omdat deze gemakkelijker te kwantificeren is dan indicatoren over de uitkomsten en effecten van onderwijs (Loukkola et al., 2020, p. 21). Haalbaarheid is dus doorslaggevend. Toch worden vrijwel nooit uitsluitend inputindicatoren gebruikt. Hieronder volgt daarom per type indicatoren de meest voorkomende indicatoren.

Inputindicatoren

Inputindicatoren waren een belangrijke informatiebron in het begin van de ‘kwaliteits-revolutie’, omdat er weinig andere indicatoren beschikbaar waren (Beerkens, 2021). Sommige van deze indicatoren worden nog steeds vaak gebruikt. Men moet echter in gedachten houden dat inputindicatoren voornamelijk iets zeggen over de voorwaarden waaronder onderwijs heeft plaatsgevonden en niet noodzakelijk iets over de kwaliteit van het onderwijs zelf (Loukkola et al., 2020, p. 17).

Als indicatoren gekoppeld worden aan financieringsmechanismen zijn studentaantallen de belangrijkste indicator, omdat dit direct aan kosten kan worden gekoppeld (Loukkola et al., 2020, p. 16). Het niveau van de inkomende studenten is ook een veel voorkomende inputindicator in hoger onderwijs systemen waar toelating tot een opleiding competitief is waardoor het gemiddelde niveau van de studenten heel verschillend kan zijn. Daarnaast zijn faciliteiten zoals computerzalen, studieplekken of bibliotheekvoorzieningen ook inputindicatoren die vaak voorkomen, maar die zijn meestal gemeten op een instellingsniveau. Tot slot is student-stafratio een indicator die ook vaak gebruikt wordt. Dit is gezien als indicatie van een intensieve, kleinschalige leerervaring die bijdraagt bij een goed leerproces. In de praktijk is er geen bewijs dat er een verband is tussen een lagere student-stafratio en een intensieve of kleinschalige leerervaring (OECD, 2019). Lage student-stafratio's zijn vaak karakteristiek in een onderzoeks-intensieve omgeving waar staf veel tijd besteed op onderzoekstaken en waar onderwijs vaak juist minder aandacht krijgt.

Outputindicatoren

Outputindicatoren omvatten meerdere belangrijke indicatoren. De meest belangrijke indicatoren zijn gerelateerd aan studiesucces en succes op de arbeidsmarkt. Om studiesucces te meten worden vaak cijfers over uitval (of tegenovergesteld: doorstroom) en de studieduur gebruikt. Deze gegevens zijn belangrijk vanuit efficiëntieoverwegingen. In veel (Europese) landen is de zorg geweest dat studenten te lang blijven studeren. Soms worden deze gegevens ook gezien als een indicatie van effectieve ondersteuning en begeleiding, waardoor studenten in staat zijn om hun studie nominaal af te kunnen ronden.

Succes op de arbeidsmarkt wordt meestal gemeten in de vorm van gemiddelde tijd die het kost tot afgestudeerden een eerste baan vinden. Op basis van goede baankansen is echter niet direct af te leiden dat de kwaliteit van de opleiding goed is. Allereerst kan de aansluiting tussen de opleiding en het verrichte werk niet altijd goed meegenomen worden. Het is in de praktijk niet altijd mogelijk om te kijken naar het niveau van de baan of het startsalaris. Ook is onduidelijk in hoeverre de opgedane kennis en vaardigheden terugkomen in de werkzaamheden. Tot slot kunnen er veel externe factoren een rol spelen zoals de economische situatie, waardoor de cijfers geen reflectie hoeven te geven van de effectiviteit of de kwaliteit van onderwijs (Loukkola et al., 2020, p. 21).

Het bovenstaande toont aan dat outputindicatoren, net als inputindicatoren, niet automatisch de kwaliteit van opleidingen weergeven. Studenten kunnen bijvoorbeeld nominaal afstuderen, omdat er goede ondersteuning vanuit personeel en faciliteiten zijn, maar het kan ook zijn dat de diploma-vereisten laag zijn (Loukkola et al., 2020, p. 21). Bovendien is er wederom sprake van een sterke relatie tussen socio-economisch achtergrond van de studenten en hun studiesucces. Het gevolg is dat opleidingen en instellingen met een diverse studentenpopulatie (bijv. eerste generatie studenten) op basis van deze indicatoren slecht uit de bus kunnen komen, terwijl ze juist een belangrijke bijdrage leveren aan sociale mobiliteit en misschien zelfs extra investeren in studiebegeleiding. Sommige landen hebben daardoor geëxperimenteerd met '*value added*' metingen als alternatief, waar outputindicatoren rekening houden met de diversiteit van de studentenpopulatie (zie ook paragraaf 4.1 over initiatieven om '*value added*' data te verzamelen).

Tevredenheidsindicatoren

Als het gaat om het meten van abstracte uitkomsten van het onderwijs, zoals de motivatie van afgestudeerden om vakkennis te blijven actualiseren of hun betrokkenheid bij de maatschappij, zijn er geen sets van ‘standaardindicatoren’ die daarvoor veelvuldig gebruikt worden. Men is afhankelijk van indirecte meetinstrumenten en uit praktische overwegingen wordt daarom de ervaring van betrokkenen als *proxy*-indicator gebruikt. Tevredenheidsindicatoren vormen daarmee de belangrijkste groep uitkomstindicatoren. Verschillende belanghebbenden, zoals studenten, alumni, medewerkers, potentiële werkgevers of stagebegeleiders kunnen aangeven hoe tevreden zij zijn met de kwaliteit van het onderwijs. Hoewel het mogelijk is om tevredenheid via focusgroepen te meten, is dat relatief duur en arbeidsintensief. Meestal wordt gebruik gemaakt van enquêtes.

Enerzijds kan men stellen dat het belangrijk is dat diegenen die met het onderwijs te maken hebben, dit onderwijs kunnen beoordelen. Anderzijds kan men het principiële bezwaar opwerpen dat men niet te veel belang aan tevredenheid moet hechten, omdat dit consumentisme in het hoger onderwijs kan bevorderen, waarin “de mate van tevredenheid van de student als consument [centraal staat en] waarbij het leren als een tastbaar product wordt gezien” (Chalmers, 2008, p. 12).

In de wetenschappelijke literatuur is met name veel discussie over de link tussen zelfrapportage van studenten enerzijds en de mate van kwaliteit anderzijds. Aan de ene kant wijst empirisch onderzoek uit dat zelfrapportage van studenten over hun leeruitkomsten een betere weergave is van de ‘*value added*’ van het onderwijs dan de behaalde cijfers, omdat die laatste meer zeggen over achtergronden van studenten dan over hun ontwikkeling (Caspersen & Smeby, 2018). Aan de andere kant wordt kritisch gekeken naar het gebruik van deze instrumenten voor kwaliteitsverbetering. Volgens Breetvelt (2021) dragen studentevaluaties niet bij aan de doelstelling van onderwijskwaliteitsverbetering, omdat studentevaluaties niet altijd betrouwbaar zijn en de validiteit gering is (pp. 14-15). Studenten worden in hun oordeel beïnvloed door aspecten waarvan zij onterecht menen dat die effectief zijn voor hun leerproces en leeruitkomsten (zoals gelikte presentaties met veel visualisaties). Hierdoor kunnen studentevaluaties een vertekend beeld geven (Breetvelt, 2021, pp. 13-14; Deslauriers et al., 2019). Indien de evaluaties gebruikt worden voor de doelstelling van onderwijsverbetering, kunnen de evaluaties zelfs een perverse prikkel geven voor docenten om zich aan te passen aan de wensen van de student, en dit kan uiteindelijk leiden tot niveauverlaging van het onderwijs (Breetvelt, 2021, p. 15; Stroebe, 2020). Ook andere auteurs beargumenteren dat de focus op uitkomsten negatieve effecten kan hebben op de ontwikkeling van onderwijskwaliteit (zie o.a. Cort, 2010; Allais, 2014). Er moet dus niet te veel op basis van deze gegevens gestuurd worden.

Tot slot zijn er ook enkele praktische bezwaren met betrekking tot de bruikbaarheid van tevredenheidsindicatoren. Allereerst is het niet altijd haalbaar om een hoge respons te krijgen, terwijl dit wel een noodzakelijke voorwaarde is voor betrouwbaarheid. Daarnaast is het soms lastig om conclusies te trekken over systematische verschillen tussen opleidingen op basis van tevredenheidsindicatoren. Doordat tevredenheid vaak op een schaal van 1-5 wordt gemeten, is er meestal niet veel (significante) variatie tussen de opleidingen. Eventuele verschillen in tevredenheid die wel aan het licht komen, kunnen ten grondslag liggen aan verschillen in de studentenpopulaties per opleiding, waardoor voorzichtigheid bij de interpretatie geboden is.

Procesindicatoren

Alle kwalitatieve meetinstrumenten brengen uitdagingen mee voor de formele toetsing van betrouwbaarheid en validiteit. Volgens sommige auteurs is het bovendien veel lastiger om betrouwbare en valide procesindicatoren te vinden, dan om geschikte uitkomstindicatoren te vinden. Uitkomsten kunnen in theorie immers nog jaren later gemeten en beoordeeld worden, ook als het gaat om de kwaliteit van onderwijs (Elton, 2004, p. 123). Desalniettemin wordt er door sommige auteurs op gewezen dat procesindicatoren “de meest praktische, nuttige en geschikte metingen zijn voor kwaliteit van leren en doceren in hoger onderwijsinstellingen” (Chalmers, 2008, p. 14). Ze helpen bovendien met context geven aan de andere indicatoren. Ondanks de uitdagingen worden procesindicatoren daarom toch veelvuldig gebruikt, met name bij institutionele audits (Chalmers, 2008, p. 14).

Omdat onderwijsprocessen complex zijn, is er echter in beperkte mate sprake van standaardisering van indicatoren. Men kan wederom terugvallen op tevredenheidsindicatoren om een beeld te krijgen van de ervaringen met het leerproces (bijv. de mate waarin onderzoek terugkomt in het onderwijs of interactieve technologieën worden toegepast). Ook ervaringen met toetsing en feedback op de mate van diversiteit kunnen op deze manier in kaart gebracht worden. *Student engagement* indicatoren proberen de mate van actieve betrokkenheid van studenten bij opleidingen te meten (bijv. de mate van interactie met docenten en andere studenten, deelname aan verrijkende onderwijservaringen, etc.), wat een belangrijke voorspeller is gebleken voor een leerproces van hoge kwaliteit (Coates, 2010). Gekwantificeerde metingen van de onderwijsmissie, onderwijsvisie en onderwijsdoelen komen echter veel minder voor. Dergelijke beoordelingen komen vaak tot stand op basis van dialogen met betrokkenen en belanghebbenden. De beoordelingen zijn dan meestal terug te vinden in kwalitatieve rapporten zoals jaarverslagen van opleidingen en minder vaak in dashboards met kernindicatoren.

4.3 Het gebruik en selectie van indicatoren: betrouwbaarheid, validiteit en haalbaarheid

Indicatoren kunnen voor verschillende doelstellingen worden gebruikt. Zo kunnen indicatoren gebruikt worden ten behoeve van transparantie en verantwoording, maar indicatoren kunnen ook gebruikt worden ten behoeve van kwaliteitsverbetering. Een ander belangrijk verschil bestaat tussen indicatoren die uitsluitend voor intern gebruik bedoeld zijn en indicatoren die extern toegankelijk (moeten) zijn. De achterliggende doelstelling bepaalt de geschiktheid van een indicator. De reden hiervoor is dat deze doelstellingen met elkaar op gespannen voet kunnen staan. Idealiter zijn indicatoren betrouwbaar, valide en haalbaar, maar deze waarden zijn in de praktijk *trade-offs*. Op basis van de doelstelling en context zal daarom een juiste balans tussen deze waarden moeten worden gevonden. Hieronder illustreren we dergelijke dilemma's.

Dilemma: vergelijkbaarheid versus maatwerk

Het is duur om gegevens van hoge kwaliteit te verzamelen die goed vergelijkbaar zijn. Deze data wordt daarom vaak bekostigd, afgedwongen of gecoördineerd door de overheid (Dill & Beerkens, 2012). Ook betreft het vaak

relatief korte lijsten van indicatoren, zodat de gegevens hanteerbaar en toegankelijk blijven voor extern publiek. De eenvoud van de indicatoren wordt zelfs belangrijker bevonden dan eventuele methodologische tekortkomingen (Dill & Soo, 2005).

Anderzijds kunnen genuanceerde, gedetailleerde en specifieke gegevens belangrijk zijn om intern te monitoren en problemen te ontdekken (Beerkens, 2021). Indicatoren die worden gebruikt ten behoeve van verbetering van interne processen zijn daarom vaak uitgebreider en complexer. Het betreft dan vaker indicatoren op maatwerk en daarom het is moeilijk om een overzicht te presenteren van het soort indicatoren dat hier systematisch en regelmatig voor gebruikt wordt.

Dit voorbeeld toont eerst dat betrouwbaarheid van indicatoren belangrijk is om vergelijkbaarheid van data te garanderen of om een debat te kunnen voeren over de resultaten op basis van “neutrale statistieken” (Sizer et al., 1992, pp. 135-136). Dit komt echter ten nadele van zowel de haalbaarheid (kosten om gegevens van hoge kwaliteit te verzamelen) als de validiteit (meet een indicator nog wat het beoogt te meten als metingen centraal zijn vastgelegd en worden vereenvoudigd). Daarna wordt aangegeven dat interne kwaliteitszorg die gericht is op verbetering meer behoefte heeft aan gedetailleerde, complexe metingen, maar dat omwille van vergelijkbaarheid en transparantie een externe partij meer waarde hecht aan eenvoud en bruikbaarheid.

Dilemma: haalbaarheid versus validiteit

Er zijn enkele richtlijnen bij het gebruik van indicatoren (Cave et al., 1996). Het belangrijkste is om helder te zijn over wat een indicator kan en zou moeten meten (‘het *proxy*-probleem’) en daarna te analyseren of de keuze voor een indicator de achterliggende doelstelling reflecteert (‘*fitness-for-purpose*’). Tot slot moet bij het interpreteren van de indicatoren het belang van contextualiseren niet onderschat worden.

Dit betekent echter niet dat dat deze afweging altijd rationeel genomen kan worden. Hoewel vanuit rationele besluitvorming de aanname wordt gedaan dat indicatoren ontwikkeld worden op basis van wat men wil meten, blijkt het proces in de praktijk vaak omgekeerd te zijn: de beschikbaarheid van bepaalde data en informatie lokt het gebruik daarvan als indicator uit (Beerkens, 2021; zie verder ook Salemans & Budding (2021) die dit beeld bevestigen op basis van onderzoek van tien hogescholen in Nederland). Idealiter worden verschillende meetinstrumenten en methoden gebruikt om een representatief beeld van onderwijskwaliteit te krijgen, maar haalbaarheid staat hierbij vaak in de weg. In de praktijk zijn middelen beperkt en worden de databronnen en indicatoren gebruikt die simpelweg voor handen liggen – ongeacht of die aansluiten bij de doelstellingen of zij voldoen aan de gewenste standaarden met betrekking tot betrouwbaarheid en validiteit.

Al in 1990 waarschuwde De Weert in een overzichtsstudie naar kwaliteit en verantwoording in het hoger onderwijs dat doelstellingen die gemakkelijk meetbaar zijn een te belangrijke rol kunnen gaan spelen. Hij geeft aan dat dit een onbedoelde en onwenselijke impact op kwaliteitsprocessen heeft, omdat dit kan leiden tot invalide metingen (pp. 64-67). Ook in meer recente bronnen wordt gewaarschuwd voor de risico’s van haalbaarheid. Er wordt geschetst dat er een tekort is aan geschikte indicatoren en dat dat leidt tot het gebruik van vervangende

indicatoren die vaak zeer versimpeld zijn en die de context niet meenemen (Loukkola et al., 2020, p. 1). Het verhogen van de validiteit en betrouwbaarheid zou leiden tot een verhoogde lastendruk bij onderwijsinstellingen, waardoor dit praktisch onhaalbaar is. De *European University Association* waarschuwt ook nadrukkelijk voor het ‘recyclen’ van databronnen en indicatoren voor een ander doeleinde dan waar zij oorspronkelijk voor dienden (Loukkola et al., 2020, p. 25). Dit kan ertoe leiden dat hetzelfde soort indicatoren in de hoger onderwijssector wordt gebruikt ondanks dat er een verschillend begrip van kwaliteit is, en onderwijsinstellingen eigenlijk iets anders proberen te meten.

Dilemma: stabiliteit versus afwisseling van indicatoren

Het is niet mogelijk om kwaliteitszorg zo in te richten dat het aan alle doelstellingen tegelijkertijd voldoet. Toch zal er altijd een behoefte aan informatie blijven bestaan. De ontwikkeling van nieuwe indicatoren en de invoering van nieuwe procedures voor gegevensverzameling vergen veel middelen. Toch is het de overweging waard om regelmatig de selectie van indicatoren te herzien om de zwaktes van bepaalde indicatoren te ondervangen. Bovendien kan daarmee ingespeeld worden op nieuwe ontwikkelingen in het hoger onderwijs. Ook de *European University Association*, bijvoorbeeld, bepleit een periodieke herziening van indicatoren om de validiteit te verhogen (Loukkola et al., 2020, p. 25), ondanks dat dit de vergelijkbaarheid van evaluaties door de jaren heen kan benadelen - en het daarom niet altijd aansluit bij de doelstellingen omtrent de transparantie en helderheid over de kwaliteit van onderwijs.

Salemans en Budding (2021) tonen in hun onderzoek echter aan dat dit in de praktijk een uitdaging is. Zij concluderen dat, ondanks dat alle ondervraagde Nederlandse hogescholen aangaven ‘publieke waarde’ te willen creëren, nog steeds de prestatie-indicatoren gebruikt worden die bewaard zijn gebleven uit de tijd van de prestatieafspraken. Deze indicatoren zijn voornamelijk efficiëntiegericht, komen overeen met het ‘oude’ *New Public Management*-gedachtegoed, en kunnen daardoor een obstakel vormen voor nieuwe ontwikkelingen richting maatschappelijke waardecreatie. Salemans en Budding laten ook zien dat zelfs bij de ontwikkeling van nieuwe indicatoren, bijvoorbeeld op het gebied van arbeidsperspectieven, gegevens worden gebruikt die gemakkelijk beschikbaar zijn, namelijk data uit de HBO-monitor. Dit is een indicatie van hoe groot de uitdaging voor instellingen is om indicatoren te herzien, en passende indicatoren en dataverzamelmethode te ontwikkelen voor nieuwe, abstracte doelstellingen (Salemans & Budding, 2021).

Hoewel de verwachtingen van kwaliteit dynamisch zijn (en zullen blijven), lijkt de kwaliteitsmonitoring en bijbehorende processen zich niet even snel mee te ontwikkelen. Het potentiële gevaar van een mismatch tussen informatie en processen enerzijds en verwachtingen anderzijds, is dat potentieel disfunctionele effecten optreden bij verkeerd gebruik van informatie. Men moet daarom toch ernaar streven om regelmatig de selectie van indicatoren blijven herzien, en daarbij oppassen voor het recyclen van databronnen die oorspronkelijk voor een ander doeleinde verzameld zijn.

4.4 Indicatoren in Nederland

Om een beeld te schetsen van de indicatoren die in Nederland gebruikt worden, zijn interviews met kwaliteitszorgmedewerkers van 11 verschillende hogescholen en 9 universiteiten in Nederland afgenomen. De onderstaande tabel (tabel 3) geeft een overzicht van welke indicatoren Nederlandse hogescholen en universiteiten gebruiken in hun kwaliteitszorg, maar ook hoe zij deze indicatoren gebruiken. Het algemene beeld is dat zowel hogescholen als universiteiten veel doen om de kwaliteit van opleidingen te monitoren. Uit de interviews bleek dat veel prestatie-indicatoren zijn gebaseerd op ESG standaarden en/of bewaard zijn gebleven uit de fase waarin hoger onderwijsinstellingen prestatieafspraken maakten. In lijn met de verwachtingen uit de literatuur worden veelal inputindicatoren gebruikt. Deze zijn immers relatief eenvoudig beschikbaar. Als *proxy* voor proces- of uitkomstindicatoren spelen tevredenheidsindicatoren een grote rol. Het valt op dat beleidsambities of onderwijsvisies zelden kwantificeerbaar worden gemaakt en zelden worden opgenomen als prestatie-indicator in een dashboard of vergelijkbaar overzicht.

Tabel 3 laat zien dat vrijwel alle onderwijsinstellingen ten minste één indicator hebben over uitval, rendement en studenttevredenheid. Soms worden gegevens verder uitgesplitst. 1 op de 3 instellingen monitort bijvoorbeeld expliciet hoe het aantal ingeschreven studenten in verhouding staat tot het aantal studenten dat een vergelijkbare opleiding aan een andere onderwijsinstelling volgt (marktaandeel) en ongeveer de helft van de instellingen splitst gegevens over uitval uit per leerjaar.

Bij hogescholen is het veel gebruikelijker om de ervaringen en tevredenheid van staf of medewerkers mee te nemen als kwaliteitsindicator dan bij universiteiten. Hoewel hogescholen een meer beroepsgerichte oriëntatie hebben, lijken er geen grote verschillen te zijn in het relatieve aantal hogescholen en universiteiten dat de tevredenheid van potentiële werkgevers monitort als indicator van kwaliteit. Tot slot valt op dat in beide sectoren opvallend weinig student-stafratio wordt gemonitord.

Tabel 3: Veel gebruikte indicatoren in Nederland

Indicatoren	Percentage hogescholen	Percentage universiteiten
Aantal studenten/toename aantal studenten	72	100
Marktaandeel opleiding	36	33
Uitval (of tegenovergesteld: doorstroom)	100	100
Uitval eerste jaar	45	44
(Opleidings)rendement/studiesucces	82	100
Rendement BSA/propedeuse	45	78
Mobiliteitsgegevens	54	78
Studenttevredenheid	100	100
Tevredenheid alumni	73	78
Tevredenheid van staf/medewerkers	64	11
Tevredenheid werkgevers/beroepenveld	27	22
Indicatoren voor onderwijsvisie/kernwaarden/beleid	27	33
Studenten in honoursonderwijs	18	11

Student-stafratio	18	11
<i>Duidelijke doelstellingen of targets in dashboard</i>	54	33
<i>Vergelijking over tijd/trends</i>	64	89
<i>Vergelijking of benchmarking met andere opleidingen/instellingen</i>	45	44

Er zijn ook wat verschillen waar te nemen in het gebruik van de indicatoren. Hogescholen hebben vaker expliciete en absolute streefwaardes (targets) bij hun prestatie-indicatoren geformuleerd (54% tegenover 33% van de universiteiten). Universiteiten lijken hun scores echter relatief vaker af te zetten tegen de scores die zij de afgelopen jaren behaald hebben (89% tegen 64%). Iets minder dan de helft van de hogescholen en universiteiten zet cijfers af tegen andere opleidingen en/of andere instellingen.

Als men gedetailleerder kijkt naar de indicatoren die gebruikt worden, blijkt er tussen universiteiten onderling meer sprake van homogeniteit te zijn dan tussen de hogescholen. Dit komt overeen met de verwachting, omdat er tussen hogescholen ook meer verschillen zitten met betrekking tot grootte, mate van centralisatie, aard van de opleiding, etc.

De resultaten die in het rapport worden gepresenteerd zijn een momentopname en ze laten de ontwikkeling van kwaliteitszorg niet zien (wat hebben de instellingen voorheen gedaan, en waar werken zij op dit moment naartoe). Tot slot bevat de tabel met veel gebruikte indicatoren (tabel 3) geen voorbeelden van indicatoren die docentkwalificaties meten, omdat deze door sommige instellingen expliciet niet als indicatoren voor onderwijskwaliteit worden beschouwd. Dergelijke gegevens werden in dat geval samen met *human resources* en financiële data in andere overzichten gemonitord waar de respondenten soms geen inzicht in of toegang tot hadden.

Het gebruik van indicatoren in hogescholen

Vrijwel alle hogescholen hebben kernindicatoren (KPI's) om het rendement of studiesucces in kaart te brengen. Eén hogeschool maakt echter een omslag van rendementssucces naar studentsucces en monitort daarom niet de uitval bij de opleiding, maar de '*survival rate*' van studenten binnen het gehele HO systeem. De helft van de hogescholen monitort apart de rendementscijfers van het eerste jaar (BSA/propedeuse). Cursusrendement of slagingspercentages van toetsen worden soms wel gemonitord, maar zelden in een dashboard opgenomen om de kwaliteit van de opleiding als geheel te monitoren (respectievelijk 9% en 27%).

Meer dan de helft van hogescholen monitort de mobiliteit van studenten. Daarbij wordt vervolgens altijd gekeken naar het aantal studenten dat uit het buitenland komt studeren, maar soms (27%) wordt ook geregistreerd hoeveel studenten van de opleiding aan een buitenlandse instelling studeren.

Alle hogescholen verwerken scores van de tevredenheidsenquête van NSE als een of meerdere KPI's. Een kwart van de hogescholen kijkt daarnaast nog naar vergelijkbare landelijke metingen zoals die van de Elsevier Keuzegids. Ruim 70% van de hogescholen monitort de tevredenheid van alumni en hoe zij denken over de

aansluiting op de arbeidsmarkt met data uit de *HBO Monitor*. Ruim een kwart monitort daarbovenop de tevredenheid van stage- of afstudeerbegeleiders.

Ongeveer een derde van de hogescholen (36,4%) verzamelt gegevens over studentbinding of de ervaring van beginnende studenten en maakt daar een KPI van. 27,3% monitort het aantal contacturen van eerstejaarsstudenten. 18,2% monitort het aantal studenten dat deelneemt aan honoursonderwijs.

Alle onderzochte hogescholen reflecteren op basis van kwalitatieve gegevens op de mate waarin de onderwijsvisie, het onderwijsbeleid of kernwaarden van de instelling terugkomen in de opleiding. Verder wordt vaak genoemd dat men bij deze kwalitatieve evaluaties kijkt naar de uitvoering van het huidig curriculum, de input van docenten, toetsing, onderwijsinnovatie, het niveau van eindwerken of kwaliteitszorg in het algemeen. Soms wordt ook gekeken naar internationalisering, ervaringen vanuit het werkveld, opleidingsspecifieke voorzieningen, opleidingsvarianten en flexibilisering, de onderwijsorganisatie of de toegevoegde waarde aan de maatschappij (onderwijsmissie). Een enkele hogeschool kijkt ook naar de begeleiding aan studenten en de samenstelling van studenten die instromen.

Het gebruik van indicatoren in universiteiten

Bij de universiteiten valt op dat bepaalde indicatoren ofwel heel vaak, ofwel heel weinig worden gebruikt. Dit duidt op een hoge mate van homogeniteit tussen universiteiten onderling. Alle universiteiten houden gegevens bij over het aantal studenten dat zich inschrijft voor de opleiding. Bijna 80% kijkt naar het land van herkomst van de studenten en meer dan de helft naar de vooropleiding (bijvoorbeeld VWO diploma of HBO diploma). Alle universiteiten houden gegevens bij over de uitval of doorstroom van studenten, hoewel het (zoals eerder aangegeven) kan verschillen of deze cijfers worden uitgesplitst per opleidingsjaar. Alle universiteiten hebben een KPI voor het rendement van de opleiding. Bijna 80% houdt ook bij hoeveel studenten hun BSA of propedeuse behalen in hun eerste jaar. Een gelijk aantal universiteiten kijkt naar de totale studieduur van studenten.

Alle universiteiten maken gebruik van landelijke enquêtes over de studenttevredenheid. Zonder uitzondering wordt de NSE gebruikt, twee derde van de universiteiten de Nationale Alumni Enquête (NAE), zo'n 20% de *International Student Barometer* (ISB) en iets minder dan de helft maakt nog gebruik van overige landelijke databronnen zoals de Keuzegids waarin gegevens uit de NSE, NAE en het 1cijferHO-bestand samenkomen. Alle universiteiten nemen zelf enquêtes bij studenten af waarin onder andere de tevredenheid over de opleiding wordt bevraagd. Opvallend is dat slechts een derde van de universiteiten de resultaten van deze enquêtes meenemen in hun indicatorenoverzicht.

Slechts een enkele universiteit neemt de tevredenheid van eigen staf of tevredenheid vanuit het werkveld mee als KPI. Op één universiteit na kijkt geen enkele universiteit naar de verhouding tussen het aantal onderwijzend personeel en studenten (student-stafratio). Wat betreft het monitoren van excellente studenten, kunnen universiteiten verschillende indicatoren kiezen. Hoeveel studenten honoursonderwijs volgen wordt slechts door één universiteit bijgehouden als indicator, maar een derde van de universiteiten rapporteert wel het percentage studenten dat (summa) cum laude afstudeert.

Alle universiteiten hebben naast hun kwantitatieve indicatoren ook kwalitatieve analyses over de kwaliteit van opleidingen. Meestal betreft het conclusies over het behalen van, of de voortgang op, ambities uit de onderwijsvisie van een instelling of faculteit of het betreft een reflectie op de jaarafspraken die een opleiding gemaakt heeft. Regelmatig wordt gekeken naar de internationale inbedding of de positionering van een opleiding. Ook de voorbereiding of aansluiting op de arbeidsmarkt komt vaak terug in dergelijke analyses. Slechts een enkele keer worden de volgende onderwerpen genoemd: samenwerking in het docententeam, de organisatie van rekruteringsactiviteiten, de verbondenheid tussen docenten en het onderzoeksveld, de samenstelling of het functioneren van de medezeggenschap, sociale veiligheid, de mate van ondersteuning bij een opleiding, de informatievoorziening voor studenten of de (maatschappelijke) impact van een opleiding.

Er zijn dus geen grote verschillen tussen de universiteiten en hogescholen. Het meest opvallende verschil is dat 64% van de hogescholen tevredenheid van de eigen staf als een kwaliteitsindicator van onderwijs gebruikt – tegenover slechts één universiteit. Op detailniveau zijn ook een paar verschillen. Een kleine minderheid van de universiteiten neemt KPI's in een dashboard op die het rendement per vak of toets laten zien. Bij de hogescholen wordt iets vaker nauwkeurig bijgehouden wat het slagingspercentage bij de eerste toetskans is, oftewel het toetsrendement (27%). Universiteiten houden daarentegen vaker dan hogescholen gedetailleerde gegevens bij over de instroom van hun studenten, zoals hun geslacht of vooropleiding (respectievelijk 44% en 67% van de universiteiten monitort dit).

5. Interne processen

Sinds het einde van de vorige eeuw wordt meer aandacht besteed aan de kwaliteitszorg in het hoger onderwijs. Als reactie daarop hebben veel hoger onderwijsinstellingen kwaliteitsmanagementsystemen geïmplementeerd om de kwaliteitszorg vorm te geven (Pratasavitskaya & Stensaker, 2010, p. 37). Op basis van een systematische literatuurstudie komen Manatos, Sarrico en Rosa tot de volgende algemeen geaccepteerde definitie van kwaliteitsmanagement: “een filosofie of managementaanpak die bestaat uit een set van principes die elkaar wederzijds versterken en elk onderbouwd zijn door een set van gebruiken en technieken” (2017, p. 160). Pratasavitskaya en Stensaker vullen aan dat deze gebruiken en technieken verschillende doelstellingen kunnen hebben. Ze kunnen er rechtstreeks op gericht zijn om de kwaliteit van onderwijs te behouden en te verbeteren, maar het doel kan ook zijn om verantwoording af te leggen, en transparantie en efficiëntie te verhogen (2010, p. 39).

In dit hoofdstuk wordt gekeken naar een aantal specifieke managementaanpakken, bijbehorende uitdagingen en mogelijke impact. Daarna wordt kort stilgestaan bij meer algemene vormen van kwaliteitsmanagement in hoger onderwijs en het begrip ‘kwaliteitscultuur’. Tot slot wordt gereflecteerd op de helderheid, transparantie en efficiëntie van interne processen.

5.1 Algemeen beeld van kwaliteitsmanagementprocessen: eenheid in verscheidenheid

Er is veel literatuur over verschillende mogelijkheden, uitdagingen en problemen met betrekking tot kwaliteitsmanagement in hoger onderwijs. Helaas is relatief weinig empirisch onderzoek gedaan naar specifieke managementaanpakken binnen hoger onderwijsinstellingen en meetbare impact van deze managementaanpakken. Hoewel het aandeel empirisch onderzoek toeneemt, zijn de meest geciteerde artikelen nog steeds theoretische en normatieve artikelen (Manatos et al., 2007). Hierdoor is het lastig om in kaart te brengen welke managementpraktijken in hoger onderwijsinstellingen plaatsvinden, en kunnen bevindingen over de impact moeilijker gegeneraliseerd worden.

Er zijn diverse mogelijke verklaringen waarom er weinig systematische en vergelijkende studies over dit onderwerp zijn (Manatos et al., 2007; Pratasavitskaya & Stensaker 2010). Allereerst betreft het een multidisciplinair onderzoeksgebied waardoor wetenschappers met verschillende perspectieven, theorieën en vocabulaire het debat betreden. Hoewel auteurs zich soms met dezelfde thematiek bezighouden, kunnen zij hiervoor toch verschillende concepten en theorieën gebruiken. Daarnaast zijn er relatief veel *single case studies* en kleinschalige onderzoeken in het veld waardoor een overkoepelend theoretisch raamwerk moeilijker tot stand komt.

Desondanks is in 2010 een survey afgenomen onder 222 hoger onderwijsinstellingen uit 36 Europese landen om de bestaande kwaliteitszorgprocessen in kaart te brengen (Loukkola & Zhang, 2010, p. 9). Het doel van het onderzoeksproject was om een overzicht te bieden van de institutionele processen en structuren die een interne

kwaliteitscultuur stimuleren (Loukkola & Zhang, 2010, p. 9). De resultaten laten zien dat er veel verschillende combinaties van processen en structuren bestaan (Loukkola & Zhang, 2010, p. 35). Bijna 65% van de hoger onderwijsinstellingen die zijn onderzocht gebruikte een systeem van kwaliteitszorg dat naar eigen zeggen ‘instellingsspecifiek’ is. Ruim een kwart zag de kwaliteitszorgprocessen voor het onderwijs als maatwerk en gebruikte daarom geen enkel pasklaar model.

Hoewel de labels en terminologie per managementaanpak kunnen verschillen, blijkt kwaliteitsmanagement in het hoger onderwijs in de praktijk echter toch vrij homogeen te zijn. Pratasavitskaya en Stensaker (2010) analyseerden 18 empirische artikelen over managementpraktijken en zien veel overeenkomsten, ongeacht of er sprake is van een bekende managementaanpak uit de private sector of een meer algemeen label van ‘kwaliteitsmanagement’. Ook Manatos, Sarrico en Rosa (2017) concluderen op basis van een systematische literatuurstudie van 127 artikelen dat ondanks een verschillende focus in de artikelen, de praktijk relatief homogeen is. Er is dus minder verscheidenheid in de hoger onderwijssector dan de labels doen vermoeden.

Binnen de Europese context dragen de ESG bij aan de homogenisering van interne managementprocessen. In de herziene versie van 2015 zijn 10 standaarden met bijbehorende richtlijnen ten aanzien van de interne processen geformuleerd. Zo is onder andere vastgelegd dat instellingen externe belanghebbenden moeten betrekken bij het beleid van kwaliteitszorg, dat instellingen relevante informatie moeten verzamelen en analyseren over opleidingen, en dat zij hun opleidingen regelmatig moeten monitoren en beoordelen om bij te dragen aan de continue verbetering van de opleiding (ESG, 2015).

5.2 Specifieke managementaanpakken

Een algemeen patroon is dat hoe langer een instelling bezig is met kwaliteitszorg, des te geavanceerder en meer ontwikkeld de processen en structuren zijn (Loukkola & Zhang, 2010, p. 10). Hoewel minder dan 10% van de 222 onderzochte hoger onderwijsinstellingen aangeeft een pasklaar managementmodel te gebruiken (Loukkola & Zhang, 2010, p. 28), kunnen maatschappelijke ontwikkelingen zoals toenemende competitie tussen onderwijsinstellingen, toenemende vraag naar verantwoording, en bezuinigingen ertoe leiden dat instellingen overstappen naar een pasklaar model (Brookes & Becket, 2007). Hierom mag de rol van specifieke managementaanpakken in de hoger onderwijssector niet onderschat worden.

In dit rapport wordt gekeken naar enkele specifieke managementbenaderingen, namelijk *Total Quality Management*, de *Balanced Scorecard*, en *Lean Sigma Six* benaderingen. Daarna wordt stilgestaan bij niet nader gespecificeerde aanpakken van kwaliteitszorg (en wat deze alsnog met elkaar gemeen hebben), en wordt kort gereflecteerd op kwaliteitscultuur als tegenhanger van ‘traditioneel’ kwaliteitsmanagement.

Total Quality Management

Sinds de jaren 1980-1990 werd kwaliteit in de hoger onderwijssector steeds meer gezien als iets dat verbeterd en gemanaged moest worden. In deze tijd ontstond het idee dat populaire kwaliteitsmodellen uit de private sector

ook in het hoger onderwijs toepasbaar zouden zijn (Pratasavitskaya & Stensaker, 2010, p. 38). Het eerste en meest populaire kwaliteitsmanagementmodel was *Total Quality Management* (TQM).

Wat is TQM? Harvey en Green tonen aan dat TQM bij diverse organisaties en personen verschillende betekenissen kan hebben en dat er een veelvoud aan literatuur beschikbaar is met praktische handvatten voor implementatie (1993). De filosofie achter TQM wordt echter simpelweg samengevat als “competitiekracht en winst naar behoren maximaliseren” (Harvey & Green, 1993, p. 18). De klant bepaalt wat kwaliteit is en men streeft ernaar om dit doel zo efficiënt mogelijk te behalen. TQM overstijgt de principes ‘*quality control*’ (detectie van problemen) of ‘*quality assurance*’ (acties om kwaliteit te behouden), omdat bij TQM vanuit efficiëntieoverwegingen gestreefd wordt naar preventie van problemen, en omdat TQM zoekt naar continue verbetering van het systeem (kwaliteit als bewegend doel).

Naast TQM zijn ook ISO 9001:2000 en EFQM populaire benaderingen die vanuit de private sector zijn toegepast in het hoger onderwijs (Pratasavitskaya & Stensaker, 2010, p. 39). ISO 9000 series kunnen als een springplank voor de implementatie van TQM worden gezien. De ISO certificering richt zich slechts op één aspect van kwaliteit, namelijk consistentie in de levering van een product of service, en richt zich niet op de (consument gedreven) aard van kwaliteit (Zhu, 1999, p. 294). Het EFQM Excellence Model wordt regelmatig door managers en academici als *proxy* voor de implementatie van TQM gebruikt (Gómez et al., 2015). Toch richt dit rapport zich op het bredere, overkoepelende begrip TQM, temeer omdat in de praktijk de ISO standaarden voor productieprocessen als te algemeen worden beschouwd om toe te passen in het hoger onderwijs, en het probleem met excellentiemodellen als EFQM is dat zij alleen succesvol toegepast kunnen worden voor excellente instellingen (Csizmadia, 2006, p. 61).

Impact. Van TQM modellen is bekend dat ze kunnen bijdragen aan positieve veranderingen op verschillende gebieden binnen de hoger onderwijsinstelling (Pratasavitskaya & Stensaker, 2010, p. 38). Door empirisch onderzoek is bekend dat TQM kan leiden tot betere prestaties van studenten, betere serviceverlening en een afname van de gemaakte kosten (Kanji & Tambi, 1999). Aan de andere kant kan de implementatie van TQM mislukken als er sprake is van te weinig toewijding of ervaring. Diverse onderzoeken vinden bewijs dat de implementatie van TQM faalt als er weerstand tot verandering is, een cultuuromslag ontbreekt (vaak door leiderschapsproblemen), als vooruitgang niet gemeten kan worden of als er te weinig waarde wordt gehecht aan de resultaten (Kearney, 1991; Seymour, 1991). Als de managementaanpak niet goed afgestemd wordt op de huidige cultuur, kunnen er zelfs negatieve effecten voor de kwaliteit van onderwijs optreden (Skelton, 2012; Sulaiman et al., 2013). De potentiële relevantie voor de hoger onderwijssector, de manier van implementatie, en de onderwijskundige en sociale effecten werden in de praktijk vaak ter discussie gesteld (Stensaker, 2007; Pratasavitskaya & Stensaker, 2010, p. 39; Manatos et al., 2017, p. 159). Volgens Nadeau nam de populariteit van TQM vooral af toen men merkte dat deze aanpak moeilijk te koppelen was aan de strategische ambities van de instellingen (2017, p. 592).

Voorbeelden. In de jaren 1990 werden TQM en ISO 9000 toegepast op de University of Wolverhampton vanuit het geloof dat hoger onderwijsinstellingen moesten leren van de *best practices* uit het bedrijfsleven om eigen ambities te bereiken (Doherty, 1993, p. 337). Hoewel de universiteit stuitte op problemen met betrekking tot de

communicatie en cultuurverandering, was een positief effect dat de samenwerking tussen academici en administratieve medewerkers verbeterde. De studenten zouden van deze verbeteringen profiteren (Doherty, 1993, pp. 335-336). In de literatuur zijn echter ook voorbeelden dat TQM niet alleen efficiëntieverbeteringen voor de staf oplevert, maar ook directe positieve effecten hebben op het leren van studenten (Williams, 1993, p. 232).

Balanced scorecard

De *Balanced Scorecard* (BSC) is een aanpak die is ontwikkeld door Kaplan en Norton (1996, 2001, 2004). De BSC wordt veelvuldig gebruikt om op effectieve wijze de implementatie van strategisch plannen te monitoren en te communiceren (naar externen). Aanvankelijk richtte de BSC zich op private bedrijven met een winstoogmerk, maar later werd de BSC bij veel diverse non-profit organisaties geïmplementeerd, waaronder (publieke) hoger onderwijsinstellingen (Sayed, 2013, p. 203).

Wat is de BSC? De BSC is “een geïntegreerde en resultaatgerichte set van prestatie-indicatoren (KPI’s) [...] die de huidige prestaties en de aanjagers van toekomstige prestaties kunnen monitoren” (Beard, 2009, p. 275). De aanname die ten grondslag ligt aan de BSC is dat doelen meetbaar moeten zijn, omdat “je niet kunt managen wat je niet kunt meten” (Kettunen, 2008, p. 328). Het management moet meetinstrumenten (meestal KPI’s) opzetten om te analyseren in hoeverre doelen behaald worden, maar ook interne processen en structuren ontwikkelen om die doelen te behalen. Het uitgangspunt van de BSC is bovendien dat strategische doelen op basis van het strategische plan van de instelling worden opgesteld (Kettunen, 2008, pp. 328-329).

Volgens Kaplan en Norton zouden de prestatie-indicatoren vier organisatorische perspectieven moeten reflecteren (Beard, 2009, p. 276): het financiële perspectief (“*how should we appear to our stakeholders?*”, bijvoorbeeld indicatoren over bekostiging en maatregelen voor kostenefficiëntie), het consumentenperspectief (“*how should we appear to our customers?*”, bijvoorbeeld student- of werkgeverstevredenheid), het interne bedrijfsprocessenperspectief (“*what processes must we excel at?*”, bijvoorbeeld gemiddelde studieduur, *drop-out rate* of het percentage buitenlandse studenten) en het leer- en groeiperspectief (“*how can we sustain our ability to change and improve?*”, bijvoorbeeld het percentage docenten met doctoraat of basiskwalificatie onderwijs/basiskwalificatie didactische bekwaamheid) (voorbeelden uit Kettunen, 2008, pp. 329-330). Volgens Kettunen selecteren hoger onderwijsinstellingen meestal zo’n twintig indicatoren voor hun BSC (2008, p. 330). De meeste gegevens voor deze indicatoren zijn beschikbaar via nationale databases (Kettunen, 2008, p. 330).

Impact. Volgens sommigen kan de BSC relatief eenvoudig dienen als basis voor zogeheten ‘*plan-do-check-act*’ managementpraktijken (PDCA) binnen hoger onderwijsinstellingen (Hladchenko, 2015). Het voordeel van de BSC is dat informatie van verschillende perspectieven verzameld kan worden, maar tegelijkertijd moet een informatieoverbelasting worden tegengegaan doordat uitsluitend indicatoren worden geselecteerd die relevant zijn om het strategisch plan te behalen (Kettunen, 2008, p. 329). Op deze manier draagt de BSC er ook aan bij dat hoger onderwijsinstellingen hun visie en beleidsplannen verhelderen en vertalen in meetbare, operationele doelstellingen. Bovendien biedt het een kans voor alle belanghebbenden om te evalueren waarom de instellingen bestaat, wat belangrijk is voor de instelling, en waar de instelling naartoe wil gaan (Beard, 2009, p. 278). Anderen

laten echter zien dat de BSC aanpak niet altijd goed aansluit op de hoger onderwijssector. Volgens Sayed (2013) voldeed de BSC aanpak niet aan de verwachtingen en werd deze losgelaten bij sommige hoger onderwijsinstellingen, omdat de aanpak niet altijd goed begrepen werd, de juiste middelen niet vrijgemaakt konden worden en/of de mate van succes die ermee geboekt werd onduidelijk was.

Voorbeelden. Uit interviews met senior managers van vier universiteiten in het Verenigd Koninkrijk die de BSC gebruikten, komt naar voren dat de BSC op alle universiteiten een positief effect had op verdere integratie van strategieën en beleidsplannen, de communicatie over prestatie management binnen de instelling, en de ontwikkeling van leiderschap (Taylor & Baines, 2012, p. 123). Deze bevinding vonden de auteurs opvallend omdat de vier universiteiten zeer verschillend waren wat betreft hun organisatiecultuur en -tradities (Taylor & Baines, 2012, p. 123). Ook andere case studies rapporteren naar aanleiding van de implementatie van BSC verbeteringen op het gebied van communicatie, verwachtingen en samenwerking tussen docenten en het management (bijvoorbeeld Yu et al., 2009). Door het beperkte aanbod van empirische studies is het lastig om te concluderen welke mogelijk (directe) effecten de BSC op de kwaliteit van onderwijs heeft. Het gebrek aan empirische studies met deze specifieke focus is mogelijk te verklaren omdat de BSC van origine uit de private sector komt. Uit een systematische literatuurstudie in Latijns-Amerika blijkt dat hoger onderwijsinstellingen de neiging hebben om vast te houden aan de originele structuur van het BSC model zoals Kaplan en Norton dat ontwierpen voor de private sector, waarbij het financiële perspectief het eerste strategische level is (Hinojosa, 2021, p. 138). Dit perspectief sluit het minste aan op onderwijskwaliteit. Er zijn recentelijk echter ook enkele voorbeelden van onderwijsinstellingen die geen indicatoren op dit financiële perspectief ontwikkelden. De nadruk bij deze instellingen bleek over het algemeen te liggen op het consumentenperspectief (Hinojosa, 2021, p. 138). Deze ontwikkeling laat zien hoe belangrijk interne en externe belanghebbenden voor hoger onderwijsinstellingen zijn, en wellicht draagt deze focus bij aan meer empirische studies in de toekomst over de mogelijke effecten van de BSC op bijvoorbeeld studenten- en werkgeverstevredenheid.

Lean, Six Sigma en Lean Six Sigma

Sinds 2000 is de belangstelling voor *Lean*, *Six Sigma* en *Lean Six Sigma* gegroeid (Nadeau, 2017, p. 592). Ook in het hoger onderwijs zijn deze managementbenaderingen in opkomst (Albliwi et al., 2014).

Wat is Lean, Six Sigma en Lean Six Sigma? *Lean* en *Six Sigma* zijn twee onafhankelijke managementbenaderingen met een oorsprong in Japan. Bij *Lean* is het doel om consumenten een hogere toegevoegde waarde te bieden en bij *Six Sigma* is het doel om kostenefficiënt te produceren. Beide benaderingen hebben nadelen, maar de combinatie van *Lean* en *Six Sigma*, genaamd *Lean Six Sigma* (LSS), wordt gezien als een effectieve aanpak die elkaars nadelen kunnen ondervangen. Bovendien is deze benadering zeer geschikt voor dienstverlenende sector (Sunder, 2013). Hoewel onderwijs in algemene zin gezien kan worden als een dienst, worden deze managementbenaderingen door diverse onderzoeken ook als veelbelovend gezien om specifieke doelstellingen binnen de hoger onderwijssector te behalen. Zo zouden *Lean Six Sigma* benaderingen bijvoorbeeld gebruikt kunnen worden om *drop-out rates* te verlagen (Gupta et al., 2020) of om het curriculumstelsel te verbeteren zodat deze beter aansluit op de verwachtingen van de omgeving (Shanshan et al., 2021).

Impact. Op basis van een literatuurstudie komt Nadeau op een lijst van uitdagingen die deze aanpakken met zich mee kunnen brengen (2017, p. 596). Bekende uitdagingen die men uit andere sectoren kent, maar men ook in het hoger onderwijs terugvindt, zijn: een gebrek aan toewijding of training bij betrokkenen (het management, medewerkers of studenten), algemene weerstand bij verandering of slecht management bij conflicten binnen de organisatie, een gebrek aan heldere communicatie, misverstanden over concepten en instrumenten, een gebrek aan middelen dat wordt toegewezen, onvoldoende planning en coördinatie, en een slechte koppeling met strategische ambities en interventies binnen de instelling (Nadeau, 2017, p. 596). Uitdagingen die specifiek voor de hoger onderwijssector gelden, komen voort uit: de complexiteit van het werkveld en de processen, een gebrek aan gedocumenteerde ervaringen met deze aanpak, moeilijkheden met het definiëren wie de consument is en wat deze wil (de vraag wat de toegevoegde waarde voor de consument is), en een negatieve beleving bij faculteiten en academici die voelen dat hun (academische) vrijheid wordt beperkt (Nadeau, 2017, p. 596).

De vraag wie de consument is in het hoger onderwijs laat zich inderdaad niet gemakkelijk beantwoorden, maar is wel cruciaal voor een goede implementatie van *Lean* managementbenaderingen (Radnor & Bucci, 2011). Sommige auteurs stellen zonder twijfel dat de student de consument is, terwijl anderen ook wijzen naar hun ouders, het werkveld, of de maatschappij in het geheel (Nadeau, 2017, p. 597). Hoewel veel universiteiten proberen de implementatie van een *Lean* aanpak te versimpelen door alleen de student als consument te beschouwen (El-Sayed et al., 2005), wordt in de literatuur ook gewaarschuwd voor de gevaren voor het behandelen van studenten als consumenten, omdat dit onderwijsstandaarden kan verlagen en een negatief effect kan hebben op de relatie tussen studenten en docenten (Eagle & Brennan, 2007). Studenten zouden daarom misschien beter op een gelijk niveau als andere belanghebbenden kunnen worden beschouwd (zie Bay & Daniel, 2001; Clayson & Haley, 2005), maar dit zal de implementatie van *Lean* management bemoeilijken. Een studie in Noorwegen laat immers zien hoe de impact en effectiviteit van kwaliteitszorg verschillend wordt beoordeeld door studenten, academische staf en kwaliteitsmanagers binnen universiteiten (Stensaker et al., 2011). Daaruit bleek dat over het algemeen leden van het institutioneel bestuursbureau en administratieve medewerkers het meest positief over de effecten van kwaliteitszorg waren. Studenten zagen daarentegen het minste positieve resultaten. Interessant is bovendien dat het institutioneel bestuursbureau en de administratie ook positief waren over de effecten van kwaliteitszorg op het onderwijs en het doceren, terwijl de studenten en academische staf – degenen die het dichtste bij het onderwijs staan – deze effecten niet waarnamen (Stensaker et al., 2011, p. 473). Het is onmogelijk om deze verschillende perspectieven te valideren met objectieve data over de effecten van kwaliteitszorg, juist omdat kwaliteit niet eenduidig of neutraal te definiëren is. Deze complexiteit leidt echter tot dilemma's in de implementatie van deze managementbenadering en het monitoren van de effecten daarvan.

Voorbeelden. Soms worden *Lean Six Sigma* benaderingen weliswaar toegepast in het hoger onderwijs, maar wordt het alleen geïmplementeerd op financiële en administratieve afdelingen binnen de instelling, op personeelszaken, IT-gebieden of bibliotheekservices, waardoor de impact van deze benaderingen op het verbeteren van onderwijskwaliteit moeilijk in kaart te brengen zijn (Antony et al., 2018). Toch laat een case study dat *Lean Six Sigma* “veelbelovende resultaten” bij een universiteit in Iran zien als het gaat om het verhogen van de onderwijskwaliteit, het verbeteren van studenttevredenheid, het verkorten van wachttijden bij studieadviseurs, en het verhogen van het aantal aanmeldingen (Haerizadeh & Sunder, 2019). Ook een case study in Saudi-Arabië

laat zien dat *Lean Six Sigma* gebruikt kan worden ter verbetering van zogeheten ‘*assurance of learning processes*’ (Adeinat et al., 2022). Voorwaarden zijn echter wel dat onderlinge verantwoordelijkheden duidelijk moeten worden vastgesteld, werknemers goed kunnen samenwerken middels gemeenschappelijke vocabulaire, en dat kwantificeerbare doelen worden gezet (Adeinat et al., 2022).

Niet nader gespecificeerde managementaanpakken van kwaliteitszorg

Zoals eerder gezegd, zijn er volgens Loukkola en Zhang (2010) veel hoger onderwijsinstellingen die niet een van de bovenstaande pasklare managementsystemen gebruiken. Deze instellingen hanteren een systeem van kwaliteitszorg dat instellingsspecifiek wordt genoemd (2010, p. 28). Toch blijkt uit een literatuurstudie dat er in de praktijk weinig diversiteit in hun methoden en werkwijzen geobserveerd kan worden. De benadering van kwaliteitszorg lijkt ondanks verschillen in terminologie vrij homogeen te zijn wat betreft de aanpak en doelstellingen (Pratasavitskaya & Stensaker, 2010, p. 45).

Veelgenoemde managementpraktijken om onderwijsactiviteiten te ondersteunen zijn het ontwikkelen van beleidsvisies, de ontwikkeling van leiderschapsvaardigheden onder academici, het opzetten van een kwaliteitsinformatiesysteem, en het aanmoedigen van een dialoog met belanghebbenden binnen en buiten de instelling (Pratasavitskaya & Stensaker, 2010, pp. 45-46). Daarbij is een grote behoefte aan verdere ontwikkeling van strategische documenten en beleidsvisies binnen een instelling, verbetering van de organisatiecultuur, ontwikkeling van leiderschapsvaardigheden onder academici, beter gebruik van resource management en een ondersteunend kwaliteitsinformatiesysteem, en een stimulans voor effectieve communicatie tussen belanghebbenden binnen en buiten de instelling (Pratasavitskaya & Stensaker, 2010, p. 46). Manatos, Sarrico en Rosa concluderen dat van alle verschillende principes van kwaliteitsmanagement (2017, p. 161), drie principes het vaakst centraal staan, namelijk: consumentenfocus, het betrekken van belanghebbenden, en het streven naar continue verbetering van organisatorische prestaties (Manatos et al., 2017 pp. 167-168). Als het vervolgens gaat om criteria om kwaliteitsmanagement te verbeteren (het laatste punt van Manatos, Sarrico en Rosa), vatten Pratasavitskaya en Stensaker dit samen als leiderschap, samenwerking en resource management (2010, p. 46).

Kwaliteitszorg bestaat uit verschillende aspecten en principes, en kan bovendien plaatsvinden op verschillende niveaus (zoals opleidingsniveau, departementsniveau, instellingsniveau of systeemniveau). Dit zijn als het ware bouwstenen om kwaliteitszorgprocessen vorm te geven. De literatuur richt zich steeds meer op integratieve vormen van kwaliteitsmanagement, waarbij een holistische aanpak met zo veel mogelijk bouwstenen wordt gekozen om de kwaliteit van het hoger onderwijs te verbeteren (Rosa & Amaral, 2007). Toch is er volgens Manatos, Sarrico en Rosa in de praktijk nog geen sprake van een ‘totale integratie’ van kwaliteitsmanagement (2017, p. 171). Veelal is de aanpak slechts ‘gedeeltelijk’ of ‘beperkt’ holistisch, waarmee bedoeld wordt dat er weliswaar meerdere principes en niveaus van kwaliteitsmanagement worden geïntegreerd, maar dat ze nog niet allemaal met elkaar in verbinding worden gebracht (Manatos et al., 2017, p. 170). In vergelijking tot andere sectoren wordt kwaliteitsmanagement nog relatief vaak behandeld als een losstaand veld dat apart binnen hoger onderwijsinstellingen opereert (Manatos et al., 2017, p. 171). Dit kan verdere integratie in de weg staan.

Kwaliteitscultuur als tegenhanger van kwaliteitsmanagement

Uit het voorgaande is gebleken dat in de afgelopen decennia veel hoger onderwijsinstellingen wereldwijd interne kwaliteitsmanagementsystemen opgetuigd met formele regels en routines om het onderwijsbeleid te ondersteunen (Brennan & Shah, 2000). Als tegenhanger is het concept kwaliteitscultuur geïntroduceerd. Bij kwaliteitscultuur wordt verder gekeken dan de managementpraktijken en -systemen en worden de aspecten van kwaliteitszorg die buiten de regels, procedures en managementstrategieën om plaatsvinden bestudeert (Harvey & Stensaker, 2008; Elken & Stensaker, 2018). Deze culturele benadering voor kwaliteitszorg heeft aan populariteit gewonnen, omdat het als een aantrekkelijk alternatief werd gezien voor de bureaucratisch, gecentraliseerde universiteit met rigide management routines (Elken & Stensaker, 2018, p. 192). In Europa wordt het concept kwaliteitscultuur door onder andere de *European University Association* gebruikt, en zijn er evaluatieraamwerken ontwikkeld waarin het monitoren en stimuleren van kwaliteitscultuur centraal staat (Rosa et al., 2011).

Wat is 'kwaliteitscultuur'? Uit een overzichtsstudie naar kwaliteitscultuur blijkt dat de werkingsmechanismen van een kwaliteitscultuur bestaan uit toewijding van de medewerkers, gedeeld eigenaarschap, empowerment en kennis (Bendermacher et al., 2017, pp. 48-50). Managers hebben een aanjagende functie, omdat zij de toewijding van middelen kunnen beïnvloeden (Bendermacher et al., 2017, p. 39). Daarnaast is goede communicatie een noodzakelijke voorwaarde voor een goede kwaliteitscultuur omdat dit bijdraagt aan het verspreiden van kwaliteitsbeleid, het evalueren van resultaten, en het achterhalen van waarden en overtuigingen van medewerkers (Bendermacher et al., 2017, p. 39). Als aan deze voorwaarden voldaan wordt, kan een kwaliteitscultuur positief effect hebben op de medewerkers- en studenttevredenheid, bijdragen aan continue verbetering van het leerproces en verdere ontwikkeling van docenten en studenten stimuleren (Bendermacher et al., 2017, pp. 50-51).

5.3 Interne processen in Nederland

Alle publieke hogescholen en universiteiten worden eens in de zes jaar extern gevisiteerd door de NVAO om in aanmerking te komen voor bekostiging uit publieke middelen. Daarnaast geven alle hoger onderwijsinstellingen die geïnterviewd zijn aan dat zij op nog frequentere basis de kwaliteit van het onderwijs monitoren. Het wisselt per instelling, maar rapportages worden meestal uitgebracht op kwartaal-, semester- of jaarbasis.

Precies tussen twee externe visitaties van de NVAO, voeren instellingen vaak een aparte monitoringsronde door, een zogeheten 'mid-term review of audit'. Middels een dergelijke tussenrapportage kan expliciet gekeken worden in hoeverre aanbevelingen uit de vorige externe visitatie zijn verwerkt, waar de instelling nu staat en waar de instelling in de komende drie jaar naartoe wil. Alle bevraagde hogescholen geven aan een dergelijke mid-term te organiseren, maar niet alle universiteiten noemen dit als onderdeel van het kwaliteitszorgproces.

Tabel 4 geeft een totaaloverzicht van welke kwaliteitszorgprocessen de hogescholen en universiteiten in de interviews naar voren hebben gebracht.

Tabel 4: Interne processen in Nederland

Processen	Percentage hogescholen	Percentage universiteiten
Externe visitaties (NVAO)	100	100
Periodieke review en rapportages (kwartaal/semester/jaar)	100	100
Mid-term reviews/audits	100	67
Interne (tevredenheids)enquetes studenten	100	100
Medezeggenschap (studentraden, opleidingscommissies, etc.)	100	100
Focus- of panelgroep, klankbordgroep	45	44
Externe peers/critical friends/adviesorganen	36	11
Exitgesprekken/vragenlijst na uitval	27	0
Beroepenveldcommissie	27	22
Workshops/benchlearning/platform	18	11
Werkbelevingsonderzoek onder medewerkers	73	33
Intern vergelijkingsonderzoek (benchmarking)	73	89
Thema-onderzoeken	45	44

Om een indicatie te krijgen van wat er speelt, nemen alle hogescholen zelf (tevredenheids)enquêtes af bij studenten. Zij zijn dus niet geheel afhankelijk van landelijke databronnen zoals NSE. Opvallend is echter dat deze enquêtes van de instelling zelf niet altijd expliciet als KPI in het dashboard terug zijn te zien.

Alle instellingen betrekken standaard medezeggenschap bij de kwaliteitszorg. De manier waarop verschilt per hogeschool, maar meestal gaat het om opleidingscommissies of studentraden. Ongeveer de helft van de instellingen betreft daarnaast ook een focus- of panelgroep of stelt een klankbordgroep in. Iets minder dan de helft van de instellingen voert met regelmaat eigen thema-onderzoeken uit, bijvoorbeeld op het gebied van diversiteit en inclusie of, recentelijk, op het gebied van de impact van COVID-19.

Een opvallend verschil is dat hogescholen veel vaker een werkbelevingsonderzoek onder medewerkers als onderdeel van het kwaliteitszorgproces zien (73%) dan universiteiten (33%). Ook wordt door hen vaker systematisch een beroep gedaan op externe *peers* en adviesorganen (36% om 11%). Drie van de ondervraagde hogescholen verzamelen informatie van studenten die bij de opleiding uitvallen (middels een gesprek of een vragenlijst) om daaruit lessen te trekken voor de kwaliteitsverbetering. Dit werd bij de universiteiten niet genoemd.

5.4 Het gebruik en vormgeving van processen: helderheid, transparantie en efficiëntie

Hoger onderwijsinstellingen kiezen opvallend vaak dezelfde bouwstenen bij het vormgeven van hun eigen interne processen. Zelfs aan verschillende concepten en benaderingen liggen veelal dezelfde kwaliteitsprincipes ten grondslag. Het gevolg is dat instellingen te maken hebben met soortgelijke voorwaarden voor een succesvolle implementatie van hun aanpak, zoals goede communicatie en toewijding van betrokkenen. Ook de bijbehorende

uitdagingen vertonen veel overeenkomsten, zoals een gebrek aan ervaring met de desbetreffende managementbenadering, of een slechte aansluiting op de strategische ambities van de hoger onderwijsinstelling. Als instellingen een druk van buitenaf voelen om voor een bepaalde managementaanpak te kiezen, lijkt de kans op succes kleiner, omdat de aanpak dan waarschijnlijk minder aansluit op de huidige werkcontext en cultuur. Dit kan nadelige effecten hebben op de helderheid, transparantie en efficiëntie. Bij het vormgeven van hun eigen interne processen, moeten instellingen daarom het volgende in overweging nemen:

Een duidelijke rolverdeling binnen de kwaliteitszorg is essentieel om de helderheid van processen te waarborgen

Brink observeert dat er een tegenstelling is ontstaan tussen wat gevraagd wordt van kwaliteitszorg door externe belanghebbenden, zoals politici en de media, enerzijds en de kwaliteitszorg die door professionals geleverd wordt anderzijds (2010). In het publieke debat wordt gevraagd om een ontwikkeling van kwaliteitszorg die door professionals nog wordt gezien als naïef (Brink, 2010, p. 140). Dit komt de helderheid van de processen niet ten goede. Om dit te verbeteren moet volgens Sursock allereerst de verdeling van rollen en verantwoordelijkheden binnen de kwaliteitszorg helder zijn (2011, pp. 50-51).

De Onderwijsraad concludeert dat ook in Nederland de rollen onvoldoende helder zijn en daarnaast dat er te weinig onderling respect en begrip is (2015, p. 35). De Onderwijsraad constateert dat de zeggenschap weg is gedreven van de onderwijspraktijk, maar ook dat “het alternatief voor sturing vanuit de overheid en de instelling te veel gezocht [wordt] in individuele vrijheid, terwijl een kwaliteitscultuur op opleidingsniveau, gedragen door de gemeenschap van docenten en studenten, er nu juist voor kan zorgen dat professionele vrijheid niet uitmondt in vrijblijvendheid” (2015, p. 35). Hoewel de verantwoordelijkheid voor onderwijskwaliteit op dit moment gedeeld wordt, dient deze primair op opleidingsniveau te liggen, aldus de Onderwijsraad. Als opleidingen deze rol oppakken, kan dit de helderheid van processen ten goede komen.

Een gebrek aan transparantie kan op de lange termijn een ondermijnend effect hebben voor interne kwaliteitsprocessen

Loukkola en Zhang (2010, p. 26) concluderen dat alle data die over onderwijskwaliteit verzameld worden, niet automatisch publiek gedeeld worden. Het blijkt daarbij dat resultaten van externe evaluaties vaker openbaar gemaakt worden dan resultaten van interne evaluaties. Een mogelijke verklaring ligt in de aard van interne evaluaties. Interne evaluaties hebben als doel om tekortkomingen aan het licht te brengen en verbeteringen te vinden waardoor ze wellicht liever als intern werkdocument behouden worden (Loukkola & Zhang, 2010, p. 27). Ook wordt informatie over docentevaluaties in meer dan de helft van de gevallen vertrouwelijk gehouden en alleen voor het management op institutioneel of faculteitsniveau beschikbaar gemaakt. Slechts bij 14 van de 222 instellingen wordt dergelijke informatie publiek beschikbaar gemaakt (Loukkola & Zhang, 2010, p. 27).

Voor een hoge mate van transparantie is het echter belangrijk dat belanghebbenden zien wat de resultaten zijn én hoe deze resultaten gebruikt worden. Bij studentevaluaties gaat het bijvoorbeeld niet alleen om het zien van de resultaten, maar is het ook belangrijk dat studenten zicht hebben op hoe de resultaten gebruikt worden om

verbeteringen door te voeren. Het afleggen van verantwoording via externe processen kan deze behoefte naar transparantie niet geheel ondervangen. Lange academische *peer review* rapporten lijken bijvoorbeeld niet tegemoet te komen aan de vraag om meer transparantie (Westerheijden et al., 2008). Een potentieel risico is dat te veel tijd en energie wordt gestoken in dergelijk papierwerk om externe verantwoording af te leggen of accreditatie op orde te brengen (Chu & Westerheijden, 2018, p. 267), terwijl deze niet bijdraagt aan de werkelijke ervaring van transparantie van diverse belanghebbenden. Op termijn kan dit de kwaliteitszorg ernstig verzwakken. Volgens Loukkola en Zhang kan een gebrek aan transparantie over resultaten en beperkt gebruik van deze resultaten immers leiden tot een vicieuze cirkel. Studenten zullen bijvoorbeeld minder vaak evaluaties invullen, als ze niet zien hoe resultaten gebruikt worden, en onderwijsinstellingen hebben dan geen bruikbare, betrouwbare data meer tot hun beschikking voor hun eigen kwaliteitsprocessen of om verantwoording over af te leggen (2010, p. 38). Bij het vormgeven van interne processen moet daarom gestreefd blijven worden naar voldoende transparantie.

Focus niet alleen op verbetering in verantwoording en efficiëntie, maar monitor ook de impact van processen op de kwaliteit zelf

De introductie van kwaliteitsmanagementaanpakken uit de bedrijfs- en industriesector is voornamelijk extern gedreven. Dergelijke aanpakken zijn geïmplementeerd om tegemoet te komen aan de toenemende vraag naar verantwoording en efficiëntie in de hoger onderwijssector (Taylor & Baines, 2012). Het is dan ook niet verrassend dat deze aanpakken voordelig kunnen zijn om verantwoording naar externen te verbeteren en de efficiëntie te verhogen. Een literatuurstudie van Brookes en Becket bevestigt dat hoger onderwijsinstellingen verschillende managementpraktijken vanuit de private sector hebben geleend om hun interne kwaliteitszorg te professionaliseren, en dat deze benaderingen inderdaad kunnen bijdragen aan het verbeteren van systematische aandacht voor kwaliteit en verantwoording (2007). Desondanks is er een zorg dat modellen als TQM weinig bijdragen aan de verbetering van doceren en leren (Brookes & Becket, 2007). Omdat het toepassen van specifieke managementbenaderingen als *Lean*, *Six Sigma* of *Lean Six Sigma* aanpakken in het hoger onderwijs een nog relatief nieuwe ontwikkeling is, is er in de literatuur nog weinig overtuigend en dwingend bewijs over de impact van deze vormen van kwaliteitsmanagement (Nadeau, 2017, p. 593). Zowel de toepassing als de evaluatie van dergelijke aanpakken hebben zich tot op het heden het vaakst gericht op financiële, administratieve, HR- en IT-afdelingen binnen onderwijsinstellingen.

Onderzoeken die wel nadrukkelijk impact op de kwaliteit van onderwijs bestuderen, richten zich veelal op het ontwerp, de relevantie en uitkomsten van nationale kwaliteitsborgingsregelingen en -raamwerken in plaats van op managementpraktijken of interne processen (Pratasavitskaya & Stensaker, 2010, pp. 37-38). Uit deze onderzoeken blijkt onder andere dat in landen waar externe kwaliteitszorg en accreditatie gebaseerd is op instellingsniveau, vrijwel alle instellingen de nationale kaders en richtlijnen volgen (Loukkola & Zhang, 2010, p. 29). In termen van efficiëntie zou accreditatie op instellingsniveau dus voordelen kunnen bieden. Hiervoor zijn minder experts nodig en liggen de kosten in termen van tijd en geld lager (Kis, 2005, p. 17). Daartegenover staat dat er minder ruimte is voor betrokkenheid op ‘*grass-roots level*’ van individuele medewerkers (Kis, 2005, p. 17). Ook staat deze efficiëntiewaarde op gespannen voet met de aanwijzing dat kwaliteitsprocessen die tussen

instellingen en binnen instellingen verschillen bevorderlijk kunnen zijn voor de mate van creativiteit en innovatie in het onderwijs (Sursock et al., 2010, p. 88). Dit zou kunnen leiden tot de conclusie dat er is geen ideale, *one-size fits-all* aanpak voor het hoger onderwijs is (Berlin Communiqué, 2003, geciteerd in Manatos et al., 2017), en dat dit ook niet nagestreefd zou moeten worden middels een nationaal raamwerk. In elk geval is bij het vormgeven van processen belangrijk dat een balans gevonden wordt tussen de behoefte aan efficiëntie en diverse aspecten van kwaliteit, waaronder het verbeteren van leren en doceren. Binnen de huidige gestelde wettelijke kaders en richtlijnen is ruimte voor interpretatie en eigen invulling die door instellingen verder benut kan worden. Bovendien kunnen instellingen op eigen initiatief activiteiten ondernemen die de gestelde vereisten, van bijvoorbeeld de ESG, overstijgen.

6. Internationale case studies

Door ervaringen in het buitenland te bestuderen kan inzicht vergaard worden in de verschillende manieren waarop de kwaliteit van de opleidingen in kaart gebracht en beoordeeld kan worden. Opleidingsaccreditatie is een van de bekendste vormen van externe kwaliteitszorg op het opleidingsniveau. In sommige landen is er echter geen proces om de kwaliteit op het niveau van opleidingen publiekelijk aan te tonen. Zoals hierboven besproken vertrouwen veel landen op interne processen: hoger onderwijsinstellingen moeten een goed proces hebben om de kwaliteit op opleidingsniveau te bewaken via voortdurende kwaliteitscontrole en -verbetering, maar van hen wordt niet per se verwacht dat ze bewijsmateriaal voor de kwaliteit van hun opleidingen leveren. Op basis van internationale ervaringen hebben we vier alternatieve hoofdmodellen geïdentificeerd voor het monitoren van kwaliteit die elkaar niet noodzakelijkerwijs uitsluiten:

- Zelfaccreditatie van opleidingen
- Publiek toegankelijke *scoreboards* en rapportkaarten
- Regelmatige thematische evaluaties
- Professionele en andere particuliere certificering

In dit hoofdstuk presenteren we kort de vier typen kwaliteitsinstrumenten en geven we een samenvatting van het bewijsmateriaal over hun effecten – indien deze beschikbaar is. In de tweede helft van het hoofdstuk gaan we dieper in op twee case studies, namelijk het zelfaccreditatiesysteem in Noorwegen en een openbaar, op informatie gebaseerd monitoringssysteem in Engeland.

6.1 Vier alternatieven voor het beoordelen van opleidingskwaliteit

Zelfaccreditatie van opleidingen

Hoewel opleidingsaccreditatie in Europa een dominante vorm van kwaliteitszorg leek te worden (Schwartz & Westerheijden, 2004), zijn er veel systemen zonder opleidingsaccreditatie of voorbeelden te vinden waar opleidingsaccreditatie de laatste jaren is afgeschaft. Een alternatief voor externe opleidingsaccreditatie is een zelfaccreditatiesysteem waarbij instellingen de bevoegdheid hebben om hun eigen opleidingen te accrediteren. Zelfaccreditatie kan verschillende vormen aannemen. In sommige gevallen zijn de criteria voor het proces openbaar gedefinieerd (bijv. in Noorwegen en Australië), in andere gevallen hebben instellingen meer vrijheid bij het opzetten van hun processen (bijv. in Taiwan). Vaak worden zelfaccrediterende instellingen aan een evaluatie of audit onderworpen om ervoor te zorgen dat hun interne processen solide genoeg zijn om de zelfaccreditatiefunctie te vervullen.

Het lijkt erop dat het proces van externe en interne accreditatie geen grote verschillen kent. Het grootste verschil ligt in de taakverdeling, maar niet in het gebruik van informatie of in de striktheid van de processen. Het lijkt erop dat een zelfaccreditatieproces de neiging heeft om de opleidingsaccreditatienormen van de externe

kwaliteitszorginstelling na te bootsen (Chen & Su, 2014) en als gevolg daarvan niet veel afwijkt van externe opleidingsaccreditatie. Om de taken uit te kunnen voeren zou zelfaccreditatie de interne bureaucratie kunnen vergroten, maar het zou daarentegen wel iets meer flexibiliteit kunnen opleveren dan een extern georganiseerde accreditatie (Elken et al., 2014).

Publiek toegankelijke scoreboards en rapportkaarten

In verschillende landen zijn er openbaar toegankelijke informatieportalen over de kwaliteit van opleidingen, vaak vormgegeven op een vergelijkbare manier. De studiegidsen zijn meestal geïnitieerd door de overheid, zoals *Student Barometer* in Noorwegen, *Studiekeuze123* in Nederland, of *Discover Uni* in het Verenigd Koninkrijk. Er zijn echter uitzonderingen. De Duitse *CHE-ranking* van universiteiten is bijvoorbeeld geïnitieerd door een stichting. Het overgrote deel van dergelijke portaalsites maakt gebruik van gegevens uit nationale enquêtes, zoals een nationale studentenenquête of een alumni-enquête. Soms worden in dergelijke overzichten ook aanvullende parameters opgenomen, zoals studierendement of student-stafratio's.

De motieven voor dergelijke publieke initiatieven zijn veelvoudig. In de eerste plaats zijn zij er om studenten informatie te bieden voor een studiekeuze, maar tegelijkertijd ook om het publiek te informeren over de kwaliteit, en om de kwaliteitsverbetering in de instellingen op gang te brengen. Het effect van dergelijke informatie-instrumenten hangt af van de aard ervan. Potentieel disfunctionele effecten van vereenvoudigde particuliere rankings hebben veel aandacht gekregen (zie Hazelkorn, 2009). Aan de andere kant zullen genuanceerdere platforms waarschijnlijk een positiever effect hebben (Dill & Soo, 2010). Meestal wordt de onderliggende informatie opgenomen in interne kwaliteitsmanagementprocessen van de instelling, en het publiekelijk beschikbaar maken van deze gegevens in dergelijke gidsen en ranglijsten draagt waarschijnlijk bij aan hun effect.

Thematische evaluaties

Verschillende landen hebben thematische evaluaties ingevoerd om informatie te verzamelen over de kwaliteit van hun opleidingen. De thematische evaluaties worden meestal uitgevoerd in disciplinaire blokken, maar kunnen zich ook richten op een specifieke kwestie in de hoger onderwijssector. In Noorwegen bijvoorbeeld verzocht het *Ministerie van Onderwijs en Onderzoek* in 2014 om een gezamenlijke evaluatie van onderwijs en onderzoek op diverse studiegebieden. Het doel van zo'n evaluatie is om de huidige stand van zaken op dit gebied in kaart te brengen, eventuele belemmeringen te identificeren en zo bij te dragen aan de verwezenlijking van onderwijs van hoge kwaliteit.

Finland is een land waar regelmatig vakspecifieke evaluaties worden uitgevoerd en universiteiten verplicht zijn aan dergelijke evaluaties deel te nemen. De evaluaties worden uitgevoerd door het FINEEC (*Finnish Educational Evaluation Center*), een semionafhankelijk agentschap dat zijn taken uitvoert op grond van een regeringsdecreet. Het doel van de thematische evaluatie is als volgt: het schetsen van een algemeen beeld van de sterke punten en verbeterpunten (in de opleidingen die worden onderzocht en binnen het vakgebied in het algemeen) wat betreft de aangeleerde vaardigheden en het belang daarvan voor het werkende leven en voortdurend leren. De evaluaties kunnen informatie bieden over in hoeverre de instelling in staat is om onderwijs te bieden dat past bij

veranderende eisen en past bij toekomstige werkomgevingen. De evaluaties leggen dus sterk de nadruk op leerresultaten en op het verband met de behoeften van de arbeidsmarkt. Enkele recente evaluaties betreffen bijvoorbeeld het hoger onderwijs in de rechten en het universitair medisch onderwijs. Een thematische evaluatie is een omvangrijke exercitie die een tot drie jaar in beslag neemt, en die kwantitatieve indicatoren, een zelfbeoordeling door de instellingen, en feedback van verschillende belanghebbenden omvat.

Private accreditatie en certificering door beroepsorganisaties

Er is een veelvoud aan accreditatiebureaus en beroepsorganisaties die opleidingen op een specifiek vakgebied accrediteren. Vaak zijn de criteria voor accreditatie in samenwerking met belanghebbenden uit de relevante vakgebieden ontwikkeld. Als gevolg van de internationalisering is er een toename van internationale accreditatieregelingen die in het vakgebied goed ingeburgerd en bekend zijn, zoals bijvoorbeeld in *business and management* (EQUIS), of *engineering*-opleidingen (ABET). Dit zijn vrijwillige certificaten die opleidingen kunnen aanvragen als ze hun kwaliteit en standaarden willen aantonen. Er zijn niet per se grote verschillen tussen dergelijke private accreditaties en hun publieke tegenhangers en ze kunnen een waardevolle aanvulling zijn op een publiek kwaliteitszorgsysteem (Stensaker & Harvey, 2006; De Paor, 2016). Bovendien kunnen vakspecifieke criteria relevanter zijn voor opleidingen dan generieke normen die door een verplichte opleidingsaccreditering aan alle programma's worden opgelegd (Salto, 2021). Meestal publiceren private accreditoren hun evaluatierapporten niet, maar alleen hun uiteindelijke besluit, waardoor hun rol in het vergroten van transparantie enigszins beperkt is.

Er zijn ook nationale accreditatieregelingen die een lange traditie hebben. In het Verenigd Koninkrijk bijvoorbeeld, moeten alle medische opleidingen worden geaccrediteerd door de *General Medical Council*, een beroepsorganisatie die criteria vaststelt en controle op de standaarden uitvoert. Er kunnen ook meerdere, met elkaar concurrerende accreditoren in een vakgebied zijn. In de Verenigde Staten (VS) bijvoorbeeld, accrediteren drie verschillende beroepsverenigingen lerarenopleidingen – NCATE, TEAC, en AACTE – die alle drie geen publieke organisaties zijn, maar door de overheid wel als legitieme accreditatie-instanties worden erkend. Verder zijn de VS ook een voorbeeld waar instellingen worden geaccrediteerd door particuliere accreditatie-instanties die zijn voortgekomen uit het initiatief van instellingen en opleidingen zelf. Niettemin spelen zij een belangrijke rol in het publieke kwaliteitsborgingssysteem en hebben zij een wettelijke betekenis voor het aanvragen van overheidsgelden en – subsidies, en met het oog daarop worden zij door de overheid getoetst en erkend.

Naast accreditatieregelingen per vakgebied zijn er ook certificaten voor een specifiek kwaliteitsaspect. Het *Certificate for Quality in Internationalisation* (CeQuInt), bijvoorbeeld, biedt een keurmerk aan opleidingen die blijf geven van excellentie in de internationale oriëntatie van hun opleiding (Beerkens, 2022).

6.2 Casus 1: Zelfaccreditatie in Noorwegen

De basis van het Noorse kwaliteitsborgingssysteem

Op grond van de Wet op het hoger onderwijs van 2005 berust de verantwoordelijkheid voor systematische kwaliteitsborging in Noorwegen bij de instelling. Hoger onderwijsinstellingen zijn wettelijk verplicht te beschikken over “een adequaat intern kwaliteitszorgsysteem om de kwaliteit van onderwijs te waarborgen en verder te ontwikkelen” (hoofdstuk 1-6). De verantwoordelijkheid voor het toezicht op de processen berust bij de Noorse kwaliteitsborgingsorganisatie NOKUT, die normen voor kwaliteitsborging vaststelt en richtsnoeren ontwikkelt. De NOKUT is een onafhankelijk uitvoerend agentschap dat de taken uitvoert die in wet- en regelgeving zijn vastgelegd, maar de organisatie functioneert onder het ministerie van onderwijs, en accreditatiebesluiten moeten door de desbetreffende minister worden goedgekeurd.

Noorwegen heeft een binair stelsel voor hoger onderwijs en de specifieke kwaliteitsborgingsregeling die een instelling moet volgen, hangt af van het type instelling. Zowel universiteiten als hogescholen zijn onderworpen aan een institutionele audit, d.w.z. een regelmatige beoordeling van hun kwaliteitsbeleid. In principe zijn universiteiten zelfaccrediterende instellingen, wat betekent dat zij zelf het mandaat hebben om nieuwe opleidingen te accrediteren en dat zij geen opleidingsaccreditatie hoeven aan te vragen bij de NOKUT. Hogescholen kunnen geaccrediteerd zijn in een specifiek domein en binnen dit domein kunnen zij zelf nieuwe opleidingen accrediteren. Bij niet-geaccrediteerde instellingen of in domeinen die buiten de accreditatie vallen, kan alleen de NOKUT een nieuwe opleiding accrediteren. Een instellingsaccreditatieprocedure wordt alleen uitgevoerd als een instelling haar status wil opwaarderen, maar instellingen kunnen in principe hun accreditatie verliezen als zij niet voldoen aan de wettelijke kwaliteitseisen en onvoldoende maatregelen nemen om de problemen te verhelpen.

Institutionele audit

Een sleutelement in het Noorse kwaliteitsborgingssysteem is de institutionele audit. De NOKUT voert regelmatig, om de zes tot acht jaar, verplichte institutionele audits uit. Deze audits onderzoeken de eigen kwaliteitsborgingsprocessen van de instellingen en zorgen ervoor dat de instellingen hun verantwoordelijkheden voor de kwaliteitscontrole van de instelling serieus nemen. De auditcyclus omvat het aanleveren van documentatie door een panel van deskundigen, bezoeken ter plaatse, en vervolgens suggesties voor aanpassingen en een besluit.

De NOKUT-verordening ‘*Toezicht op de Onderwijskwaliteit*’ definieert de normen en doelstellingen voor een intern kwaliteitsborgingssysteem, die vervolgens worden geëvalueerd middels de institutionele audit:

- De kwaliteitsborgingspraktijken moeten worden vastgelegd in een strategie en alle gebieden bestrijken die van belang zijn voor de kwaliteit van de leerresultaten van de studenten;
- De kwaliteitsborgingspraktijken moeten op alle managementniveaus van de instellingen worden toegepast;

- Door middel van hun kwaliteitsborgingspraktijken moeten de instellingen een kwaliteitscultuur onder personeel en studenten bevorderen;
- De instellingen moeten er systematisch voor zorgen dat alle opleidingen voldoen aan de eisen die in de nationale wet- en regelgeving zijn vastgelegd, met daarbij inbegrepen eventuele aanvullende regelgeving die door de instellingen zelf is vastgelegd;
- De instellingen moeten systematisch informatie uit relevante bronnen verzamelen om de kwaliteit van alle opleidingen te beoordelen;
- De via kwaliteitsborging verworven kennis moet worden gebruikt om de kwaliteit van de opleidingen te verbeteren en gevallen van ontoereikende kwaliteit aan het licht te brengen. Tekortkomingen in de kwaliteit moeten binnen een redelijke termijn worden verholpen;
- De resultaten van kwaliteitsborgingspraktijken maken deel uit van de kennisbasis die wordt gebruikt bij de beoordeling en ontwikkeling van het totale portfolio van studieprogramma's van de instelling.

Instellingen zijn dus verplicht om er systematisch voor te zorgen dat alle opleidingen aan de formele eisen voldoen, en moeten informatie verzamelen uit relevante bronnen om de kwaliteit van elke opleiding te beoordelen. Instellingen hebben de autonomie om hun eigen kwaliteitsborgingsprocessen vorm te geven, dus er is geen voorgeschreven formule hoe zij een evaluatie moeten uitvoeren. Ze zijn ook vrij in het bepalen van het bewijsmateriaal dat ze verzamelen en de indicatoren die ze monitoren om aan de eisen te voldoen. De NOKUT biedt echter wel richtlijnen voor wat als belangrijk wordt beschouwd en wat als *best practices* wordt gezien (NOKUT, 2016).

In de tweede ronde van de kwaliteitsaudits omvatte het proces twee locatiebezoeken, die als een sterk punt in het systeem worden gezien (NOKUT, 2016). Het eerste bezoek richt zich meestal op het kwaliteitssysteem in het algemeen, en het tweede op de implementatie ervan in verschillende opleidingen en op de verschillen binnen een instelling. Voor de ronde 2018-2024 zijn enkele herzieningen doorgevoerd in het proces. Met name het aspect van leeruitkomsten werd versterkt.

Zelfaccreditatie en kwaliteitsborging op opleidingsniveau

Noorwegen kent een systeem van accreditatie op opleidingsniveau: alle nieuwe opleidingen moeten formeel worden geaccrediteerd en de vereisten voor opleidingsaccreditatie zijn vastgelegd in de '*Academic Supervision Regulation*' van de NOKUT. Opleidingen worden in principe door de NOKUT geaccrediteerd maar de Wet op het hoger onderwijs staat universiteiten en geaccrediteerde hogescholen echter toe om het accreditatieproces zelf uit te voeren, en mandateert het college van bestuur voor het uiteindelijke accreditatiebesluit. Het accreditatieproces voor opleidingen verloopt volgens dezelfde procedure, ongeacht of het door de instelling of door de NOKUT wordt uitgevoerd.

Het zelfaccreditatieproces bestaat uit vier stappen:

- Documenteren van de basis van kwaliteit van de opleiding en hoe deze voldoet aan de eisen voor accreditatie;

- Beoordeling, waarbij een duidelijk antwoord wordt gegeven op de vraag of en hoe aan de eisen wordt voldaan, en wat er zou kunnen en/of moet worden veranderd aan de opleiding om voldoende kwaliteit te bereiken;
- Aanpassingen, wijzen op veranderingen die moeten worden gemaakt en stappen die beschrijven hoe dergelijke aanpassingen moeten worden gemaakt;
- Beslissing door het bestuur van de instelling.

Een opleidingsaccreditatie moet worden uitgevoerd door een gekwalificeerde commissie van deskundigen die de opleiding evalueert, en deze moet een formele, schriftelijke beoordeling geven.

Hoewel opleidingsaccreditatie een belangrijk element is in het Noorse kwaliteitsborgingssysteem, is het alleen relevant voor nieuwe opleidingen. De kwaliteit van lopende opleidingen moet worden gewaarborgd via een continu kwaliteitsborgingsproces. Ondanks de autonomie van de instellingen om hun eigen kwaliteitsprocessen te ontwerpen, lijken de instellingen tamelijk homogeen te zijn in hun kwaliteitsprocessen. Niettemin bestaan er enkele variaties en meestal zijn er ook verschillen tussen faculteiten binnen dezelfde instelling waar te nemen.

Aan de *University of Oslo* worden bijvoorbeeld alle opleidingen om de zes jaar intern geëvalueerd. Bij de evaluatie ligt de nadruk op de relatie tussen de leerresultaten, de academische inhoud, de onderwijsmethoden, en de beoordelingsregelingen van de opleiding. De evaluatie is gebaseerd op een zelfstudie en een beoordeling door een extern evaluatiepanel. In de evaluatieverslagen worden diverse gegevens bijeengebracht, zoals gegevens over leren en beoordeling en resultaten van studenten en enquêtes op cursus- en opleidingsniveau, een aantal werkgever en enquêtes, en de nationale studenten en enquête. Naast de zesjarige evaluatiecyclus worden alle opleidingen jaarlijks geëvalueerd. Deze evaluatie is toegespitst op programmabeschrijvingen en leerresultaten, maar omvat ook evaluatieresultaten en andere gegevens waaruit zou kunnen blijken dat veranderingen nodig zijn. Voor sommige opleidingen worden ook regelmatig werkgevers en enquêtes gehouden om feedback te krijgen over de leercompetenties. Hoe geformaliseerd de verschillende indicatoren zijn, lijkt binnen de organisatie te variëren. Sommige faculteiten hebben een set prestatie-indicatoren geselecteerd die regelmatig worden gemonitord, maar in het institutionele auditrapport uit 2016 werd erop gewezen dat kwaliteitsindicatoren niet in het algemeen managementsysteem zijn opgenomen.

Een ander voorbeeld van een nieuwere en kleinere universiteit, de *University of Adger*, benadrukt het belang van relevante kwaliteitsindicatoren en informatie uit verschillende bronnen. Minimaal om de zes jaar wordt elke opleiding geëvalueerd door een externe commissie bestaande uit internationale, externe en studentleden. De evaluatie is gebaseerd op documentatie die door de opleiding wordt verstrekt. Daarnaast schrijven alle opleidingen een jaarverslag met daarin relevante studiegegevens, docentevaluaties en feedback van de opleidingscommissie. Dit verslag wordt geëvalueerd door het afdelingshoofd voor een follow-up.

Hoewel het kwaliteitszorgsysteem van de NOKUT gericht is op kwaliteitsprocessen op instellingsniveau, vormt de kwaliteit op opleidingsniveau nog steeds de hoeksteen van de kwaliteitsfilosofie van de NOKUT. Van instellingen wordt verwacht dat zij de kwaliteit van elk van hun opleidingen nauwlettend in de gaten houden en de NOKUT biedt instellingen richtlijnen voor het denken over en beoordelen van de kwaliteit van hun

opleidingen. De NOKUT definieert ook de benadering van kwaliteit in het hoger onderwijs: dat studenten leerresultaten behalen op een hoog academisch niveau, dat zij hun studie in een nominale tijd afronden en dat hun leerresultaten relevant zijn voor hun professionele ontwikkeling en een leven lang leren (NOKUT, 2016). Bovendien erkent de aanpak dat eindresultaten sterk afhankelijk zijn van het eigen vermogen, de motivatie en de inspanning van studenten en dat daarom niet alleen naar eindresultaten kan worden gekeken, maar dat ook activiteiten om het leren van studenten te vergemakkelijken en hun motivatie te ondersteunen van belang zijn. In de nieuwste ronde kwaliteitsaudits is de nadruk op het leren van studenten verder versterkt, samen met de arbeidsmarktrelevantie en systematisch kwaliteitswerk in het hele stelsel.

Impact

Wat weten we over de doeltreffendheid van een zelfregulerend kwaliteitszorgstelsel? Er zijn verschillende effecten waargenomen. Het kwaliteitsinstrument lijkt enig positief effect te hebben. Uit een evaluatie uit 2008 blijkt dat respondenten van instellingen gematigd positief zijn over de resultaten, maar zij ervaren het systeem wel als meer gericht op controle dan op verbetering (Stensaker et al., 2011). De grootste effecten werden waargenomen op het ontwikkelen van nieuwe procedures en kwaliteitsborgingspraktijken, met betrekking tot het voeren van wetenschappelijke discussies over kwaliteit, maar ook op reputatie. In mindere mate werden effecten op de kwaliteit zelf waargenomen. Het is interessant om op te merken dat de resultaten niet significant verschillen tussen de verschillende elementen in de complexe Noorse systemen, zoals institutionele audits of opleidingsaccreditaties, wat leidt tot de conclusie dat de specifieke kenmerken van het instrument relatief weinig effect hebben, en dat het grootste effect voorkomt uit de interne processen die ze op gang brengen, zoals een kritische reflectie en collegiale discussies over een opleiding.

Een ander effect betreft de interne bureaucratie en procedures. Elken et al. (2020) melden dat de instellingsaudit en zelfaccreditatie aanleiding waren voor het creëren van een ‘lokale NOKUT’ binnen instellingen, met andere woorden een eenheid die de activiteiten van een externe kwaliteitszorginstantie weerspiegelt en zich bezighoudt met onafhankelijk toezicht op de kwaliteit van opleidingen. Dit wordt gezien als een waarborg dat er bij de officiële controle door de NOKUT geen problemen naar voren zullen komen. De toename van de professionele bureaucratie binnen de instellingen als gevolg van het interne kwaliteitszorgsysteem lijkt een gevolg te zijn van de hoge verwachtingen ten aanzien van kwaliteitszorg. Op faculteitsniveau zijn verschillende nieuwe functies gecreëerd om de kwaliteit bevorderende taken uit te voeren (Elken et al., 2020).

Een verwant punt van kritiek betreft het feit dat zelfaccreditatie een aanzienlijke hoeveelheid expertise en capaciteit vereist en dat kleinere instellingen wellicht niet zijn uitgerust om een dergelijke taak uit te voeren en een dergelijke capaciteit op te bouwen (Van den Broek et al., 2019). In dergelijke gevallen kan het efficiënter zijn om de NOKUT nieuwe opleidingen te laten accrediteren en een zelfaccreditatiestatus niet als strategisch doel te nemen.

Zelfs in zo’n geavanceerd systeem lijken er problemen te zijn met het integreren van kwaliteitsmanagement in strategisch management en leiderschap van de instelling. Omdat de taken op het gebied van kwaliteitszorg nog vaak geïsoleerd zijn van institutionele en strategische taken, worden ze soms als bureaucratisch en procedureel

ervaren (Elken et al., 2020). Er zijn pogingen geweest om kwaliteitsinformatie beter in de managementsystemen te integreren, zoals ook in institutionele auditverslagen is aanbevolen, vooral om meervoudige plannings- en rapportagekringen te vermijden.

Hoewel universiteiten autonoom kunnen beslissen welke indicatoren zij in hun kwaliteitsmonitoring gebruiken, lijken hun keuzes tamelijk homogeen te zijn. Twee initiatieven moeten afzonderlijk worden vermeld die de homogeniteit kunnen beïnvloeden: de nationale studenten- en docentenenquête die door de NOKUT worden uitgevoerd, en de kernprestatie-indicatoren op systeemniveau die door het ministerie worden uitgevoerd.

De *Studiebarometer* is een nationale studentenenquête die wordt geïnitieerd door het ministerie van Onderwijs en jaarlijks wordt uitgevoerd door de NOKUT. Er wordt gevraagd naar verschillende aspecten van kwaliteit op opleidingsniveau. De vragen betreffen de tevredenheid van studenten over de kwaliteit van de lessen, de beoordeling en de inhoud van het curriculum, maar ook hun eigen motivatie en vele andere aspecten. De resultaten worden op een portaalsite gepubliceerd en maken het mogelijk om opleidingen te vergelijken en ontwikkelingen in de tijd te observeren. Het doel van het informatie-instrument is tweeledig: om de inspanningen voor kwaliteitsverbeteringen in het hoger onderwijs te ondersteunen en om studiekeizers nuttige informatie te verstrekken voor het kiezen van een opleiding en instelling om te studeren.

De docentenenquête wordt sinds 2017 uitgevoerd. Deze enquête peilt naar onderwijsactiviteiten, prioriteiten en beschikbare middelen, en hier reflecteren docenten ook op de kwaliteit van opleidingen. De resultaten van deze enquête worden ook opgenomen in de *Studiebarometer*. Instellingen zijn niet verplicht deze informatie te gebruiken in hun kwaliteitsmonitoring, maar het behoort wel tot de aanbevolen bronnen en lijkt breed gebruikt te worden.

Naast kwaliteitsborging op opleidingsniveau kent Noorwegen ook een systeem van sectorbrede prestatieafspraken. Een element in het model zijn de kernprestatie-indicatoren, die de gebieden aangeven die het ministerie wil controleren en verbeteren. De indicatoren op institutioneel niveau omvatten bijvoorbeeld nominale afstudeercijfers, tevredenheid van studenten over de opleidingen, contacturen, maar ook publicaties, beurzen, het aandeel vrouwelijke professoren, enzovoorts. Daarnaast zijn er bewuste pogingen gedaan om het aantal universiteiten via fusies te verminderen om de internationale concurrentiepositie van de universiteiten te versterken. In de prestatiecontracten tussen het ministerie en de hoger onderwijsinstelling zijn ook verschillende streefcijfers verwerkt, die indirect van invloed zijn op het belang van indicatoren op opleidingsniveau.

6.3 Casus 2: Publiek toegankelijke *scoreboards* en rapportkaarten in Engeland

Het kwaliteitssysteem in Engeland

Het kwaliteitssysteem in Engeland is waarschijnlijk een van de meest dynamische kwaliteitssystemen in Europa. Het Verenigd Koninkrijk was een voorloper bij het opzetten van een systematisch extern kwaliteitssysteem en heeft het systeem voortdurend bijgesteld op basis van opkomende knelpunten en veranderende ideeën over effectieve kwaliteitsregulering. In de afgelopen 30 jaar heeft Engeland met verschillende kwaliteitsborgingsinstrumenten geëxperimenteerd. Engeland heeft een complex hoger onderwijssysteem met een groot aantal diverse instellingen, dat zich kenmerkt door een hoge mate van autonomie voor de instellingen en een hoge mate van internationalisering in het systeem.

Het huidige kwaliteitssysteem is vanaf 2018 van kracht, samengevat in het beleidsdocument ‘*Securing student success: Regulatory framework for higher education in England*’ en uitgewerkt in de *Higher Education and Research Act*. De belangrijkste motivatie voor de recente beleidswijziging is vermindering van regeldruk en administratieve lasten voor instellingen. Als gevolg hiervan heeft het beleid twee belangrijke bouwstenen voor het kwaliteitssysteem. Ten eerste wordt het idee van ‘*risk-based regulation*’ (risico gebaseerde regulering) ingevoerd, wat inhoudt dat niet alle aanbieders in gelijke mate moeten worden gecontroleerd, maar dat aanbieders met een hoog risico moeten worden gecontroleerd en dat de rapportageverplichtingen voor aanbieders met een laag risico kunnen worden beperkt. De nadruk van de controle ligt dus op het verzekeren van de basis, d.w.z. dat alle aanbieders van onderwijs voldoen aan de verwachte normen voor diploma’s en goede kwaliteit van het onderwijs, maar er wordt niet toegezien op continue verbetering. Ten tweede ziet het beleid concurrentie als de drijvende kracht voor kwaliteitsverbetering: “*the dynamic of providers responding to informed student choice [is] the best mechanism for driving quality and improvement*”, en het beleid is opgezet om op sectorniveau te reguleren, om op die manier de ‘*informed choice*’ mogelijk te maken.

Beoordeling van kwaliteit en standaarden

Een belangrijk element in het kwaliteitsborgingssysteem is de kwaliteitscode voor het hoger onderwijs (‘*Quality Code for Higher Education*’). De Code omvat twee belangrijke dimensies: ervoor zorgen dat het niveau van een diploma overeenstemt met de eisen van het nationale kwalificatiekader en zoals erkend in de sector; en waarborgen dat het onderwijs van hoge kwaliteit is, een goede en ondersteunende leeromgeving biedt en de studenten bij de kwaliteitsevaluatie betreft. De Code is een *principle-based* regelgevend instrument wat betekent dat de verwachtingen kwalitatief worden gedefinieerd zonder streefcijfers of benchmarks, en dat de instellingen de flexibiliteit wordt gelaten om de criteria voor hun context te interpreteren, en de bijbehorende bewijzen te leveren.

Alle instellingen worden geacht zich aan de kwaliteitscode te houden en de *Quality Assurance Agency* (QAA) heeft de opdracht om dit te controleren. De QAA is een onafhankelijk agentschap en heeft zijn mandaat van het *Office for Students* (OfS). Het proces van ‘*Quality and Standard Review*’ (QSR - Beoordeling van Kwaliteit en

Standaarden) is opgezet om na te gaan of de aanbieders inderdaad voldoen aan de kernpraktijken van de Code, en het volgt drie belangrijke beginselen:

- Gericht op de belangen van studenten en de resultaten die voor studenten van belang zijn, in plaats van op het onderliggende beleid of op de onderliggende procedures die door instellingen zijn ingevoerd;
- Flexibel, in de zin dat zij ruimte biedt aan een breed scala van verschillende soorten instellingen en voorzieningen en elke instelling beoordeelt op de verdiensten van zijn eigen voorziening;
- Op risico gebaseerd en doelgericht, in de zin dat de aanpak wordt toegesneden op risico's van niet-naleving bij verschillende instellingen.

Het is belangrijk om te benadrukken dat deze benadering expliciet waarschuwt voor een focus op interne processen en procedures, die waarschijnlijk tot onnodige bureaucratie, administratieve lasten en uniformiteit zal leiden. Ten tweede is het van belang op te merken dat de *risk-based* en gerichte aanpak niet alle instellingen via de QSR controleert. Alle instellingen worden gecontroleerd op hun risicoprofiel, gebaseerd op leidende indicatoren (bijv. uitvalpercentages, studenttevredenheidsscores, enz.), noemenswaardige gebeurtenissen, en andere gegevens zoals klachten. Instellingen met een hoog risicoprofiel worden aan de QSR onderworpen. Daarnaast zal een klein deel van de instellingen (ongeveer 5%) willekeurig voor beoordeling worden geselecteerd, met als doel om de sector ertoe aan te zetten om systematisch betrokken te blijven bij kwaliteitsactiviteiten.

Het QSR-proces omvat de volgende stappen: een beoordeling door deskundigen gebaseerd op bewijsmateriaal en kwaliteitscriteria, toetsing aan de kernpraktijken van de kwaliteitscode, visitatie, het betrekken van studenten, en publicatie van het verslag. De administratieve belasting is ook een prominent aandachtspunt in de QSR-procedures. Er zijn duidelijke overeenkomsten wat betreft de gegevens die van de instellingen mag worden verlangd. Bovendien is er een duidelijke beperking van de lengte van het document dat de instellingen in de eerste fase aan de commissie voorleggen – een beschrijving van hun profiel en een samenvatting van de wijze waarop zij aan de criteria van de Code voldoen – en dat is 20 bladzijden A4.

Kader voor onderwijsexcellentie en studieresultaten

Naast het beperken van risico's en het waarborgen dat aan normen wordt voldaan, steunt de beleidsaanpak in Engeland op concurrentie als mechanisme voor innovatie en verbetering. Om de concurrentie te bevorderen, heeft het *Office for Students* in 2018 het *Teaching Excellence and Student Outcome Framework* (TEF) ingevoerd, dat de onderwijskwaliteit van instellingen beoordeelt en hen een gouden, zilveren of bronzen kwaliteitsmerk toekent. Het doel van het instrument was om:

- Studenten beter te informeren bij hun keuzes over wat en waar te studeren;
- De waardering voor het onderwijs te verhogen;
- Uitstekend onderwijs te erkennen en te belonen;
- Beter tegemoet te komen aan de behoeften van werkgevers, het bedrijfsleven, de industrie en de vrije beroepen (Vivian et al., 2016).

Het TEF is een instrument om onderwijskwaliteit op institutioneel niveau te meten. Naar verwachting zal het TEF zich verder ontwikkelen tot een meting op opleidingsniveau. Het is gebaseerd op kwantitatieve indicatoren, maar de indicatoren worden aangepast aan een benchmark die differentieert naar het type instelling. Verder dienen instellingen een document van maximaal 15 pagina's in om hun onderwijskwaliteit te presenteren en wordt de uiteindelijke toekenning van het kwaliteitscijfer gedaan door een onafhankelijke commissie, dus niet puur op basis van kwantificeerbare indicatoren.

Het TEF maakt onderscheid tussen drie aspecten van onderwijskwaliteit, namelijk kwaliteit van doceren, de leeromgeving en de leerresultaten. Elk aspect wordt vervolgens opgesplitst in meerdere criteria (zie Office for Students, 2018). Naast contextuele data en de schriftelijke aanmeldingsstukken gebruikt het TEF de volgende indicatoren om opleidingen te beoordelen (Office for Students, 2021, p. 15):

- Doceren bij vakken (*National Student Survey data*);
- Toetsing en feedback (*National Student Survey data*);
- Inspraak studenten (*National Student Survey data*);
- Academische ondersteuning (*National Student Survey data*);
- Hulpmiddelen bij het leren (*National Student Survey data*);
- Doorstroom van de opleiding (*Higher Education Statistics Agency/Individualised Learner Record data*);
- Toegang tot een hooggekwalificeerde baan of vervolgopleiding (*Destination of Leavers from Higher Education data*);
- Behoud van werkgelegenheid of verdere studie (*Longitudinal Education Outcomes data*);
- Meer dan mediaan inkomen of verdere studie (*Longitudinal Education Outcomes data*).

In 2017-2018 en 2018-2019 voerde TEF een pilot uit om onderwijs ook op het opleidingsniveau te beoordelen (Office for Students, 2021, p. 4). De pilot is gebaseerd op eerdere TEF beoordelingen op instellingsniveau en heeft als doelstelling om ook verschillen in ervaringen en uitkomsten tussen diverse opleidingen binnen een instelling aan het licht te kunnen brengen (Office for Students, 2021). Daarbij was de verwachting dat het inzichtelijk maken van deze gegevens ook zou leiden tot verbeteringen in het onderwijs en excellentie zou stimuleren (Office for Students, 2021, p. 6).

Op basis van de tweede ronde van de pilot blijft het TEF zich verder ontwikkelen. Er wordt enerzijds gesuggereerd om meer begeleiding te geven bij het interpreteren van de data (Office for Students, 2021, p. 7). Anderzijds zou het interpreteren van data waarschijnlijk gemakkelijker gaan als het TEF meer interne impact zou kunnen hebben (Office for Students, 2021, p. 100). Een andere suggestie is om sommige contextuele data achterwege te laten en te vervangen door gegevens over regionale werkgelegenheidsfactoren (Office for Students, 2021, p. 7-8).

Impact

Het TEF heeft zowel positieve als negatieve feedback ontvangen. Een onafhankelijke evaluatie in 2019 concludeerde voorzichtig dat het instrument potentie heeft om de kwaliteit in instellingen te verbeteren, maar deed een aantal suggesties over enkele cruciale aspecten. De aanbevelingen waren onder andere: het veranderen van de goud-zilver-brons labeling, het veranderen van de naam, meer duidelijkheid over wat er precies gemeten wordt met de indicatoren, en het tonen van gevoeligheid voor disciplinaire en regionale verschillen. Vivian et al. (2019) laten zien dat een grote meerderheid van docenten denkt dat het TEF een ‘positieve’ of ‘neutrale’ impact had en dat een kleine minderheid de impact van het TEF negatief vond voor hun instelling of de sector als geheel. Het vaakst werd een positieve impact gerapporteerd over toegenomen aandacht voor studentresultaten, onderwijskwaliteit en de leeromgeving, en arbeidsmarktvoorbereiding. Het merendeel van de medewerkers van onderwijsinstellingen meldde dat het TEF heeft bijgedragen aan meer aandacht voor kwantitatieve indicatoren bij het toezicht op bijvoorbeeld National Student Survey-scores, doorstroompercentages en werkgelegenheidsgegevens.

Kritischere studies plaatsen vraagtekens bij de conceptuele basis van het instrument, d.w.z. bij de vraag of de vastgestelde indicatoren inderdaad iets zeggen over onderwijsexcellentie (Cui et al., 2021). Bovendien bleek uit een andere medewerkersenquête dat het TEF heeft bijgedragen aan centralisatie en bureaucratisering van bepaalde onderwijsprocessen, zoals standaardisatie van onderwijsplannen en studiegidsen of vakwijzers, koppeling van vakken aan arbeidsmarktvaardigheden, en tot slot wordt de nadruk op de arbeidsmarktresultaten van een opleiding gezien als het belangrijkste effect (Cui et al., 2021). Volgens de critici versterkte het TEF met name de “*‘relentless focus’ on monitoring and improving the student experience, student satisfaction and student outcomes*”, hoewel het moeilijk te zeggen is of deze trends specifiek met TEF verband hielden of met het algemene gevoel van bureaucratisering in universiteiten en instrumenten zoals de National Student Survey (Cui et al., 2021, p. 1764).

Het TEF is overigens niet het enige middel dat voor studenten beschikbaar is. Ook het *Discover Uni* platform biedt de mogelijkheid om opleidingen gemakkelijk te vergelijken op basis van verschillende datapunten. De bijdrage van het TEF is echter het kwaliteitsoordeel dat wordt gepresenteerd met het kwaliteitslabel goud-zilver-brons.

7. Nieuwe ontwikkelingen in het openbaar bestuur en de regelgeving

Het kwaliteitsborgingsbeleid in het hoger onderwijs heeft altijd de bredere trends in het openbaar bestuur en de regelgeving gevolgd. De eerste generatie kwaliteitsborgingshervormingen werd beïnvloed door de hervormingen van de overheidssector in de jaren negentig. Instellingen kregen meer autonomie in de besluitvorming, maar daardoor was er ook meer controle nodig op hun activiteiten (Van Vught, 1994). De ideeën van het *New Public Management* benadrukten het belang van prestatiemetingen en prestatie management. Bovendien ontstond door de massificatie van het hoger onderwijs de behoefte aan meer aandacht voor kwaliteitsborging. In korte tijd ontwikkelde de kwaliteitszorg in het hoger onderwijs zich tot een eigen vakgebied. Op dit moment is de dynamiek van de sector weer aan het veranderen. Bovendien zijn er tendensen in de publieke management-, bestuurs- en reguleringstheorie aan het verschuiven op een manier waarbij de geschiktheid van oude beleidsinstrumenten ter discussie worden gesteld.

Kwaliteitsborging in het hoger onderwijs is een vorm van regulering. Regelgeving kent drie belangrijke elementen: de standaarden vaststellen waaraan moet worden voldaan, informatie verzamelen om toezicht te houden op de naleving van de vastgestelde standaarden, en positieve veranderingen teweegbrengen door passende sancties, prikkels en facilitering. Alle drie de elementen kunnen een uitdaging vormen. Een succesvol kwaliteitsborgingsstelsel voor het hoger onderwijs moet kwaliteitsnormen vaststellen, informatie bij de instelling verzamelen om de naleving van de normen te controleren, en ten slotte sancties, beloningen of richtlijnen vaststellen om de standaarden te bereiken.

Een doeltreffend en efficiënt regelgevingsstelsel hangt af van een aantal contextvariabelen, zoals hoe arbeidsintensief het toezicht is, hoe nauwkeurig de verwachte kwaliteitsnormen kunnen worden gedefinieerd, hoe acceptierend of afwijzend de organisaties tegenover de regelgeving staan, hoe homogeen of heterogeen de instellingen zijn, en hoe dynamisch en complex de sector is. Deze factoren bepalen bijvoorbeeld of de standaarden en het toezicht rechtstreeks moeten worden afgestemd op de eindprestaties (bijv. afgestudeerden die bepaalde onderwijsresultaten behalen, of succes op de arbeidsmarkt) of op het proces van het verstrekken van onderwijs (bijv. contacturen, kwalificaties van het personeel). Deze factoren bepalen ook of effectieve standaarden de vorm moeten hebben van een concrete checklist van specifieke normen of dat algemene beginselen moeten worden vastgesteld die de instellingen meer beoordelingsruimte laten. Deze factoren bepalen ook de intensiteit van de monitoringcyclus.

Effectieve regelgeving heeft de afgelopen decennia veel aandacht gekregen en er zijn nieuwe ideeën en trends naar voren gekomen. De traditionele opvatting van regelgeving is gebaseerd op het idee van een hiërarchisch, door de staat geleid systeem: de overheid bepaalt de normen en oefent toezicht en controle uit via haar agentschappen. Het traditionele model van regelgeving is de laatste jaren echter ernstig in twijfel getrokken. De kosten van toezicht zijn vaak onredelijk hoog, het instrument is niet flexibel genoeg om innovatie mogelijk te maken, het kan een symbolische nalevingcultuur creëren in plaats van een transformatieve verandering, de

deskundigheid die nodig is voor effectieve regelgeving kan de zich snel ontwikkelende sectoren of de toenemende complexiteit niet bijhouden, en internationale onderlinge afhankelijkheden verzwakken dergelijke inspanningen op nationaal niveau aanzienlijk.

Daarnaast is duidelijk geworden dat vele andere organisaties – NGO's, beroepsverenigingen, particuliere informatieplatforms – ook regelgevende taken vervullen en soms kostbare overheidsinitiatieven kunnen vervangen. Bovendien blijkt dat in veel sectoren aanbieders bereid en gemotiveerd zijn om zich aan de regels te houden. De schaarse overheidsmiddelen moeten dus niet worden gebruikt voor uitgebreide controle, maar voor facilitering en informatie-uitwisseling en voor stimulering van degenen die zich niet aan de regels houden.

Het hoger onderwijs is niet immuun voor deze trends in de regelgeving. In dit hoofdstuk gaan we op zoek naar inspiratie uit deze nieuwe trends op het gebied van regelgeving en denken we na over de lessen die daaruit kunnen worden getrokken voor de kwaliteitszorg in het hoger onderwijs. Hieronder worden vier nieuwe, onderling samenhangende benaderingen besproken: *management-based* regulering, *risk-based*-regulering, zelfregulering, en certificering en polycentrische regulering. Tot slot bespreken we kort de relatie tussen verantwoordingsplicht en 'vertrouwen' als een alternatief op prestatiemetingen en prestatie management voor (semi-)publieke organisaties en vatten we lessen voor de beoordeling van opleidingskwaliteit samen.

7.1 Nieuwe benaderingen van regelgeving

Management-based regulering (bijv. instellingsaudits)

Kwaliteitsborging is een moeilijke taak. Vaak kan kwaliteit niet rechtstreeks worden gemeten, kunnen de processen om tot kwaliteit te komen dubbelzinnig zijn, en kan er te veel variatie zijn tussen aanbieders om uniforme normen te ontwikkelen. Om in te spelen op zo'n complexe omgeving is de *management-based* benadering ontstaan (Gilad, 2006), die in uiteenlopende sectoren is toegepast, zoals gezondheid en veiligheid op het werk, duurzame bosbouw of voedselveiligheid (Gunningham & Sinclair, 2009; Coglianese & Starobin, 2020) en die conceptueel ook relevant is voor de kwaliteit in het hoger onderwijs. Uitgangspunt voor de benadering is dat kwaliteit en risicobeperking in de eerste plaats de verantwoordelijkheid van de aanbieder zelf is. De rol van externe controle is dus om na te gaan of het management deze rol serieus neemt en robuuste interne processen voor kwaliteitszorg heeft ontwikkeld. Het instrument controleert de kwaliteit dus niet rechtstreeks, maar zet het management ertoe aan 'na te denken' over de acties en handelingen die in hun organisatie nodig zijn om de kwaliteit te waarborgen en de risico's te beperken, en om hun praktijken voortdurend te verbeteren. De directie legt haar plannen voor kwaliteitsborging en risicobeheersing ter evaluatie voor aan een externe regelgevende instantie.

Management-based regulering vindt ook ingang in het hoger onderwijs. De heroriëntering naar de evaluatie van niet de kwaliteit van de opleidingen, maar de interne kwaliteitsborgingspraktijken van de instellingen is een voorbeeld van deze vorm van regulering. Het is daarom relevant om lessen te trekken uit andere sectoren over de sterke en zwakke punten van een dergelijk systeem zoals dat relevant is voor het hoger onderwijs, en na te denken over de uitgangspunten voor succes. De belangrijkste lessen kunnen als volgt worden samengevat:

- Als de kwaliteit van het onderwijs redelijk goed en efficiënt meetbaar is, of als het mogelijk is te definiëren welke processen en activiteiten in alle instellingen tot kwaliteit leiden, kan traditionele regelgeving (bijv. meting van leeruitkomsten of leerprocessen) een effectievere oplossing zijn. Als dit niet het geval is, kan een *management-based* aanpak een oplossing zijn voor de bestaande ambiguïteit. Het is mogelijk dat bepaalde door de overheid vastgestelde basisnormen nog steeds verplicht zijn in de *management-based* aanpak. Zowel conceptueel als pragmatisch moet de regulering vanuit de overheid echter nooit gericht zijn op zowel het meten van de eindkwaliteit als op de managementpraktijken om kwaliteit te bereiken.
- Hoe gedetailleerd of algemeen de managementstandaarden zijn, is een kwestie van afweging. De kracht van deze aanpak is dat de instellingen de ruimte krijgen om hun eigen unieke praktijken te definiëren. Te gedetailleerde standaarden zouden het onderliggende idee van flexibiliteit, openheid voor innovatie en zelfverantwoordelijkheid ondermijnen. Aan de andere kant brengen losjes gedefinieerde normen ernstige onzekerheidskosten met zich mee. Uit andere sectoren is gebleken dat managementplannen vaak worden afgewezen omdat de regelgever geen duidelijk idee heeft over hoe een aanvaardbaar voorstel eruit zou zien (Gilad, 2010). Ook blijkt dat het erg veel tijd kost om de plannen te evalueren en erover te onderhandelen als de normen niet duidelijk zijn (Aalders, 2002). Bovendien brengen onduidelijke verwachtingen kosten met zich mee voor instellingen vanwege hun risicomijdend gedrag. Om koste wat kost een negatieve beslissing te vermijden, zullen meer middelen in interne processen worden geïnvesteerd dan optimaal is. Gebleken is dat instellingen voor hoger onderwijs door onzekerheid en risicomijding geneigd zijn intensievere interne processen te ontwikkelen dan misschien nodig is.
- *Management-based* regelgeving wordt vaak als efficiënter beschouwd omdat er geen arbeidsintensieve controleactiviteiten op microniveau nodig zijn. Vaak betekent dit echter dat de kosten van monitoring daardoor niet worden vermeden, maar worden doorgeschoven van een extern agentschap naar de aanbieder (Gilad, 2020). *Management-based* regulering omvat een rigoureuus toezicht op de kwaliteit binnen de organisatie, ondersteund door goed bewijsmateriaal. Dit punt van zorg is ook aan de orde gesteld in de context van het hoger onderwijs. Zelfverantwoordelijkheid heeft ertoe geleid dat instellingen sterke interne eenheden in het leven roepen om een deel van de toezichthoudende taken van externe agentschappen over te nemen. Dit betekent niet noodzakelijkerwijs dat kosten worden bespaard, en de zelfregulerende aanpak kan voor kleine instellingen zelfs onbetaalbaar zijn.
- Een *management-based* aanpak is gebaseerd op de veronderstelling dat organisaties over inzet, motivatie en capaciteit beschikken om hoge standaarden vast te stellen en deze na te leven. Het veronderstelt ook een zeker vertrouwen om informatie te versterken aan regelgevers, inclusief het blootleggen van hun kwetsbaarheden. Onafhankelijke informatieverzameling door de regelgever, met inbegrip van informatie van belanghebbenden en andere bronnen, is een belangrijk element in de controlecyclus.
- De interne organisatiecultuur en het vertrouwen tussen de verschillende organisatielagen kunnen een ernstige belemmering vormen voor het succes van deze aanpak. Gebleken is dat een gebrek aan communicatie, verschillende subculturen binnen de organisatie, en het falen van het management om de

kwaliteitsagenda deel te laten uitmaken van de kernprocessen in de hele organisatie, een groot obstakel kunnen zijn voor een positief effect (Gilad, 2006). Het paradigma van de ‘kwaliteitscultuur’ in het hoger onderwijs probeert precies dit probleem op te lossen: voortdurende verbetering tot een doel maken dat door de hele organisatie wordt gedeeld.

Al met al is gebleken dat een *management-based* aanpak een aantal positieve effecten heeft en potentie heeft om meer flexibiliteit en innovatievermogen mogelijk te maken, evenals om de kosten van het overheidstoezicht te verlagen (Coglianese & Starobin, 2020). Toch verdwijnen veel van de uitdagingen van de traditionele regulering niet.

Risk-based regulering

Risk-based regulering pakt het probleem van inefficiëntie van regelgeving aan. Traditionele regelgevingsinstrumenten zijn vaak administratief omslachtig, inflexibel en daardoor vaak te duur om het beoogde resultaat te bereiken. Toezicht is meestal erg duur als het uniform en systematisch in de hele sector moet worden uitgevoerd. Als een sector goed presteert en ernstige kwaliteitsproblemen slechts zelden voorkomen, is een dergelijke uitgebreide controle verspilling. Als antwoord op dit probleem wordt in *risk-based* regelgeving beweerd dat de controle in de eerste plaats moet worden toegespitst op aanbieders met een hoog risico, waardoor de kosten voor zowel de regelgever als de instellingen zelf worden gedrukt. Bovendien ontwikkelen regelgevers in toenemende mate een *risk-based* aanpak om keuzes te maken voor verschillende soorten interventieactiviteiten, zoals het verstrekken van informatie en advies, monitoring en het geven van prikkels (Black, 2010). Het instrument is gebaseerd op risicoberekening op basis van verschillende signalen en track-records uit het verleden. Aanbieders met een laag risico zijn vrijgesteld van controle. Toch worden vaak steekproefsgewijze controles, thematische toetsingen of voorlichtingscampagnes uitgevoerd om alle aanbieders betrokken te houden bij het bereiken van de doelstellingen van de regelgeving.

Risk-based regulering is vanaf de jaren 2000 ingevoerd in de kwaliteitszorg van het hoger onderwijs, in landen als het Verenigd Koninkrijk en Australië (King, 2018). Er is zeer weinig bewijs over het succes van *risk-based* regulering in het hoger onderwijs. Op basis van ervaringen uit andere sectoren kunnen de volgende lessen worden afgeleid:

- *Risk-based* regelgeving vereist enige politieke risicotolerantie. Aangezien er geen externe controle van alle opleidingen is, kan het gebeuren dat een slecht presenterende aanbieder er tussendoor glipt. Anderzijds zou het controleren van alle aanbieders leiden tot overregulering en onnodige kosten voor de sector als geheel. Het is dus een bewuste afweging en er moet publiekelijk overeenstemming worden bereikt over een aanvaardbaar risiconiveau.
- De grootste uitdaging van het instrument betreft een adequate risicobeoordeling. *Risk-based* regelgeving vereist bewijs en gegevens om een adequaat risicoprofiel op te stellen, dus de kwestie van goede indicatoren blijft belangrijk. Met name bestaat het gevaar dat het risicokader niet anticipeert op risico's in de toekomst, maar zich te veel concentreert op het verleden en het heden.

- Risicobeoordelingen vereisen een andere vaardigheid dan een reguliere *compliance check*. Het vereist ook een andere denkwijze over het doel van het regelgevingskader. Herhaaldelijk is opgemerkt dat onvoldoende bijscholing van ‘inspecteurs’ een belemmering kan vormen voor een doeltreffende invoering van *risk-based* regelgeving.

Zelfregulering

Bij zelfregulering stellen organisaties zelf gedetailleerde regels op voor hun functioneren om de doelstellingen van de regelgeving te waarborgen. Bij gedwongen zelfregulering geeft de regulerende instantie toestemming voor de regelgeving en controleert zij of de organisaties zich inderdaad aan hun eigen regels houden. Het toezicht kan ook worden overgelaten aan gebruikers van de diensten, de media of andere private partijen. Zelfregulering heeft verschillende voordelen: de sector heeft meer expertise om normen vast te stellen, het is efficiënter in het verzamelen van gegevens, en vaak wordt het beter geaccepteerd en dus nageleefd door de organisaties (Baldwin et al., 2012).

Zelfregulering neemt vaak de vorm aan van een gedragscode. Een code kan worden aangenomen door een enkele organisatie of door de hele sector. Vaak bepaalt de sector de normen waaraan moet worden voldaan en de processen om dit te bereiken. In verschillende gevallen zijn voor de gehele sector geldende normen overeengekomen om meer opdringerige externe regelgeving te vermijden.

Zelfregulering heeft sterke wortels in de hoger onderwijssector. Het kwaliteitszorgsysteem in Nederlandse universiteiten begon door het vaststellen van standaarden door de Nederlandse vereniging van universiteiten, aantoonbaar ook om externe controle te voorkomen (Beerkens, 2015). De normen werden vervolgens gehandhaafd door de overheidsinstantie. Er zijn ook andere gedragscodes die niet extern worden gehandhaafd, zoals de gedragscode voor de omgang met internationale studenten die instellingen voor hoger onderwijs gezamenlijk hebben afgesproken en naleven.

De belangrijkste lessen kunnen als volgt worden samengevat:

- Zelfregulering veronderstelt dat aanbieders in staat zijn en bereid zijn om aan hogere eisen te voldoen, en dat zij door elkaar of door betrokken belanghebbenden ter verantwoording worden geroepen. Als er van buitenaf weinig belangstelling is voor de kwaliteitsstandaarden, zal een zelfreguleringssysteem waarschijnlijk weinig bijdragen;
- Onder de verkeerde omstandigheden kan zelfregulering leiden tot het vooropstellen van het eigenbelang, zoals het niet stellen van eisen die ambitieus genoeg zijn, of het stellen van te hoge eisen waardoor concurrenten in een nadelige positie worden geplaatst;
- Zelfregulering werkt beter als er geen grote sociale gevolgen verbonden zijn aan niet-naleving. Het voldoen aan de basisvereisten voor een graad in het hoger onderwijs moet nog steeds een wettelijke basis hebben en niet aan zelfregulering worden overgelaten. Aan de andere kant kan zelfregulering zeer effectief zijn bij het verspreiden van nieuwe normen en het vergroten van het bewustzijn.

Certificering en polycentrische regelgeving

Nieuwe benaderingen van regelgeving stellen de opvatting ter discussie dat regelgeving alleen iets is tussen de overheid en de aanbieders van het onderwijs. Het concept van polycentrische regulering (Black, 2010) presenteert het idee dat aanbieders tegelijkertijd door verschillende reguleringsinstrumenten en actoren ter verantwoording worden geroepen. Het belangrijkste punt daarbij is dat aanbieders niet alleen verantwoording moeten afleggen op grond van de overheidsnormen, maar ook op grond van de eisen die door verschillende maatschappelijke actoren zijn vastgesteld. Een van de meest prominente vormen van dergelijke private regulering komt in de vorm van certificerings- en etiketteringssystemen, bijvoorbeeld ecologische en duurzaamheidsnormen, normen voor eerlijke handel, en normen voor maatschappelijk verantwoord ondernemen. Certificaten zijn ook prominent aanwezig in het hoger onderwijs, zoals certificaten die worden afgegeven door internationale beroepsorganisaties (bijv. EQUIS voor *business schools*), door algemene kwaliteitsorganisaties (bijv. ISO 9000), of door organisaties met thematische doelstellingen (bijv. het internationaliseringscertificaat van het *European Consortium for Accreditation* (ECA)). Op basis hiervan kunnen de volgende lessen met betrekking tot certificering worden getrokken:

- Certificaten helpen alleen om hoge eisen te handhaven als het publiek op de hoogte is van de certificaten en ze belangrijk vindt voor het maken van keuzes. Aangezien certificaten meestal vrijwillig zijn voor aanbieders, moeten aanbieders ze om intrinsieke of extrinsieke redenen nuttig vinden;
- Vrijwillige certificaten helpen niet om ondermaats presenterende aanbieders op te sporen. Ze kunnen echter wel zeer effectief zijn om nieuwe standaarden in een sector in te voeren en hogere standaarden voor de hele sector te bevorderen, die op lange termijn tot een gangbare praktijk kunnen uitgroeien. Certificaten kunnen een zeer doeltreffend instrument zijn voor bewustmaking van zowel aanbieders als gebruikers;
- In sommige sectoren is gebleken dat er behoefte is aan ‘certificatie van certificeerders’. Concurrerende certificeringssystemen, financiële afhankelijkheid van certificeringsorganisaties ten opzichte van aanbieders die een certificaat aanvragen, en hoge kosten in verband met het toezicht op aanbieders kunnen een situatie creëren waarin certificaten niet op betrouwbare wijze een indicatie geven van hogere standaarden. Dit zal het vertrouwen in certificaten doen afnemen en aanbieders niet aanzetten tot het bereiken van hogere standaarden;
- Certificaten kunnen door aanbieders beter worden aanvaard dan licentiebeheer of accreditatie vanuit de overheid indien de certificerende instantie wordt gezien als een organisatie die beter op de hoogte is van de professionele normen en standaarden van de sector. Er kan ook een positieve leercyclus ontstaan als de certificeringsinstantie met de sector samenwerkt om goede en zinvolle standaarden aan te bieden.

Verantwoordingsplicht en vertrouwen

De relatie tussen verantwoordingsmaatregelen en vertrouwen is veelzijdig. Verantwoordingsmechanismen zoals prestatiemetingen werden ingevoerd om de prestaties te verbeteren, maar ook om vertrouwen bij het publiek te wekken. De golf van hervormingen in de publieke sector die prestatiemeting heeft verspreid, heeft echter

wisselende resultaten opgeleverd (Laupente & Van de Walle, 2020). Men hoopte dat verantwoordingsmechanismen en specifiek prestatiemetingen “*efficient control, democratic legitimacy, and effective performance*” zouden bieden, maar de beloften worden vaak op alle drie punten niet nagekomen (Dubnick & Frederickson, 2011).

De impact van het kwaliteitszorgbeleid in het hoger onderwijs is vaak in twijfel getrokken. Het is duidelijk dat het beleid kwaliteitszorgpraktijken heeft geprofessionaliseerd, interne procedures heeft geformaliseerd, systematische evaluatieactiviteiten heeft ontwikkeld, en aantoonbaar heeft bijgedragen aan de nieuwe ‘kwaliteitscultuur’ (Stensaker et al., 2004; Lui et al., 2015). Toch blijft het effect op de feitelijke onderwijskwaliteit moeilijk te bewijzen. De uitgebreide verantwoordingsinstrumenten hebben desondanks hoge kosten voor instellingen en voor het publiek. Bovendien worden ze bekritiseerd omdat ze tot onnodige bureaucratisering leiden en daardoor professionals demotiveren. Als gevolg daarvan gaan er vaak stemmen op voor meer ‘vertrouwen’ in plaats van controle en metingen.

Publieke verantwoording op zich is echter geen modegril of hervormingstendens. Verantwoordelijkheid voor de besteding van publieke middelen betekent verantwoording afleggen. Er is echter een verschil tussen instrumenten die uitgaan van vertrouwen en instrumenten die uitgaan van wantrouwen. Of op vertrouwen gebaseerde mechanismen gerechtvaardigd en effectief zijn, is geen moreel oordeel, maar gebaseerd op een empirische evaluatie van de motivatie en de missie van instellingen. Instellingen voor hoger onderwijs hebben meerdere doelstellingen die met elkaar in strijd kunnen zijn, wat de complexiteit van hun missie en motivatie vergroot. Evenwichtige stimulansen kunnen nodig zijn om de aandacht voor deze verschillende doelen te waarborgen. Aan de andere kant is vertrouwen – indien gerechtvaardigd – een troef die het mogelijk maakt toezichtskosten te besparen, deze schaarse middelen te investeren in een productiever gebruik, en aantoonbaar ook meer eigenaarschap en doelrealisatie aan te moedigen.

7.2 Lessen voor de beoordeling van opleidingskwaliteit

De nieuwste trends in de regelgeving hebben de oriëntatie van de regelgeving veranderd. In plaats van controle is het primaire doel van regelgeving het definiëren en communiceren van publieke doelen, en het ondersteunen van het eigen vermogen van organisaties om hun prestaties te waarborgen en te verbeteren. Benaderingen zoals *management-based* regulering, *risk-based* regulering of zelfregulering zijn gebaseerd op de constatering dat organisaties gewoonlijk in staat zijn, en zich ertoe verbinden, om aan de verwachte standaarden te voldoen. Een strenge controle is daarom wellicht niet het meest efficiënte gebruik van schaarse middelen. Bovendien delen de benaderingen het idee dat flexibiliteit en discretie nodig zijn om innovatie aan te moedigen en unieke oplossingen te ontwikkelen op basis van individuele behoeften. Traditionele instrumenten voor kwaliteitsborging in het hoger onderwijs hebben met soortgelijke kritiekpunten te kampen en de nieuwe ideeën zullen waarschijnlijk een inspiratiebron zijn voor verdere ontwikkelingen in de sector.

Wat betekenen deze benaderingen voor het meten en aantonen van de kwaliteit van opleidingen in het hoger onderwijs? Op basis van de eerdere analyse van indicatoren en processen in het hoger onderwijs, internationale

ervaringen met kwaliteitsborging, en bredere inzichten uit de reguleringstheorie stellen wij de volgende lessen voor:

- Goede gegevens en indicatoren om de kwaliteit van de opleidingen in kaart te brengen blijven waardevol. Zelfverantwoordelijkheid voor kwaliteit betekent niet dat kwaliteit verborgen blijft, maar dat de last om relevante indicatoren te definiëren, prestatiegegevens te verzamelen en op basis van gegevens te handelen, bij de instelling zelf ligt. In de nieuwe benaderingen hoeven de gegevens niet noodzakelijk vergelijkbaar te zijn tussen de verschillende aanbieders, zodat de aanbieders over de nodige flexibiliteit beschikken. Niettemin blijft een goed toezicht op de kwaliteit belangrijk.
- Nieuwe benaderingen verwachten enige politieke risicotolerantie. Kwaliteitsincidenten hebben de neiging door de media te worden geschandaliseerd. Gevoeligheid voor dergelijke schandalen kan ertoe leiden dat regelgevers defensief worden, overdreven procedureel, en daarmee het '*light-touch*' doel van de nieuwe benaderingen tenietdoen (King, 2018). De afweging moet duidelijk zijn en geaccepteerd worden: overregulering met hoge kosten voor de aanbieders en het publiek versus een kleine kans dat incidenten zich incidenteel kunnen voordoen.
- Niet alleen politieke actoren maar ook instellingen zelf hebben de neiging risicomijdend te zijn. Dit kan vaak betekenen dat eigen verantwoordelijkheid met los gedefinieerde verwachtingen kan leiden tot een even intensief of zelfs intensiever (en duurder) kwaliteitsregime binnen de organisatie. Een optimale sanctionering is daarom belangrijk. Sommige sancties zijn waarschijnlijk nodig om de aanbieders aan hun verantwoordelijkheden te houden, maar als ze prohibitief zijn, kunnen ze een dure cyclus van interne overregulering creëren. De praktijk van 'responsieve regulering' bouwt daarom sancties op in de logica van het ondersteunen en faciliteren van onwetende aanbieders tot het verharden van sancties om opzettelijke niet-naleving aan te pakken en te ontmoedigen.
- Hoger onderwijsinstellingen zijn geen unitaire actoren. Het succes van het interne kwaliteitsbeleid hangt af van de communicatie tussen het centrale management en het management op opleidingsniveau en het onderwijzend personeel. Deze nieuwe flexibele regelgeving kan de eigen verantwoordelijkheid voor kwaliteitsprocessen intern vergroten, maar schept ook hogere verwachtingen ten aanzien van het centrale management om hun kwaliteitsplannen bekend te maken, en te overtuigen van de urgentie ervan.
- In een regelgevingsstelsel worden verschillende regelgevingsinstrumenten gecombineerd en het is belangrijk dat met hun wisselwerking rekening wordt gehouden. Er moet duidelijkheid zijn over het doel van elk regelgevend instrument. Informatie op het opleidingsniveau om te controleren of aan de kwaliteitsnormen wordt voldaan, vervult andere behoeften dan informatieverstrekking aan studenten om hen te helpen met het maken van een studiekeuze. Het laatste kan het eerste ondersteunen, of het tweede kan doelbewust naast het eerste worden opgezet om de concurrentieprikkels te vergroten, maar de twee moeten niet worden samengevoegd.

8. Conclusie

Het beoordelen en bewaken van opleidingskwaliteit in het hoger onderwijs is geen gemakkelijke opgave. Het is niet mogelijk om kwaliteitszorg zo in te richten dat het aan alle doelstellingen en verwachtingen tegelijkertijd voldoet. Toch probeert de sector zo goed mogelijk om te gaan met uitdagingen die bij kwaliteitszorg komen kijken, en blijft men streven naar kwaliteitszorg die gericht is op verbetering én waarmee publieke verantwoording kan worden afgelegd - idealiter via processen die helder, transparant en efficiënt zijn.

In Nederland, maar ook daarbuiten, worden kwaliteitszorgsystemen voortdurend tegen het licht gehouden en herzien. Hierdoor heeft de kwaliteitszorg zich de laatste 25 jaar flink geprofessionaliseerd en ontwikkeld. Een belangrijke trend die daarbij zichtbaar is, is dat onderwijsinstellingen zelf steeds meer verantwoordelijkheid nemen.

Ook in Nederland staat de kwaliteitszorg niet stil. De kwaliteitszorg wordt zeer serieus genomen. Om een oordeel te kunnen vellen over de kwaliteit van opleidingen wordt gebruik gemaakt van verschillende soorten databronnen. Onderwijsinstellingen verzamelen zelf informatie, maar monitoren ook gecentraliseerde statistische informatie of informatie uit landelijke enquêtes.

Uit de interviews met twintig verschillende Nederlandse hoger onderwijsinstellingen bleek echter ook dat de keuze om bepaalde informatie te verzamelen veelal gedreven was op basis van haalbaarheid, beschikbaarheid of conventie. Hierdoor verzamelen vrijwel alle instellingen gemakkelijk kwantificeerbare input-, output- of (student)tevredenheidsindicatoren. Hoewel dergelijke informatie een hoge mate van vergelijkbaarheid heeft, kunnen deze cijfers geen volledig beeld van opleidingskwaliteit schetsen. Voor een goede interpretatie van de gegevens moet daarom naar aanvullende kwalitatieve informatie gezocht worden. Instellingen doen dit via studentraden, opleidingscommissies of via focus- of panelgroepen met daarin studenten of externe belanghebbenden. De verschillende databronnen worden vervolgens met elkaar gecombineerd in accreditatierapporten en jaarrapporten die instellingen zelf opstellen. Echter, juist vanwege de uitgebreidheid en alomvattendheid van deze informatiebronnen worden zij niet vaak geraadpleegd door het publiek. Het gevolg is dat informatie soms strikt genomen wel ‘zichtbaar’ is, maar daarmee nog niet direct ‘inzichtelijk’.

De indicatoren die op dit moment het meest toegankelijk zijn voor het publiek weerspiegelen slechts een beperkt aantal aspecten van onderwijskwaliteit. Het blijft een uitdaging voor instellingen om passende kwalitatieve indicatoren en dataverzamelingsmethoden te ontwikkelen om abstracte doelstellingen en aspecten van kwaliteit te meten. Uit de internationale literatuur blijkt dat meer geavanceerde metingen doorgaans arbeidsintensief zijn, maar ook dat de interpretatie ervan weinig intuïtief is.

Hoewel de verwachtingen van kwaliteit dynamisch zijn (en zullen blijven), is het dus moeilijk om de kwaliteitsmonitoring en bijbehorende processen even snel te laten mee-ontwikkelen. Het potentiële gevaar van een mismatch tussen informatie enerzijds en verwachtingen anderzijds, is echter dat potentieel disfunctionele effecten optreden bij verkeerd gebruik van informatie. Men moet daarom regelmatig de selectie van indicatoren

blijven herzien, en daarbij oppassen voor het recyclen van databronnen die oorspronkelijk voor een ander doeleinde verzameld zijn.

Hoewel men overal worstelt met de vormgeving van kwaliteitsmonitoring en bijbehorende processen, kunnen lessen getrokken worden uit de ervaringen in andere landen. Kwaliteitszorgprocessen bestaan uit verschillende aspecten en principes (zoals consumentenfocus, het betrekken van belanghebbenden en het streven naar continue verbetering) en kunnen bovendien plaatsvinden op verschillende niveaus (zoals opleidingsniveau, departementsniveau, instellingsniveau of systeemniveau). Veel hoger onderwijsinstellingen in Europa gebruiken een systeem van kwaliteitszorg dat naar eigen zeggen ‘instellingsspecifiek’ is. Toch zijn er in de praktijk ook veel overeenkomsten te vinden tussen de processen van hoger onderwijsinstellingen die geen pasklaar managementmodel te gebruiken. Kwaliteitszorg volgt immers trends in publiek management en wordt beïnvloed door bredere maatschappelijke ontwikkelingen.

In veel landen wordt het traditionele reguleringsmodel waarbij de overheid standaarden bepaalt en deze laat controleren via organisaties of agentschappen ter discussie gesteld, omdat monitoringskosten hoog zijn, het systeem inflexibel is, en het mogelijk kan leiden tot een symbolische nalevingscultuur in plaats van transformatieprocessen of kwaliteitsverbeteringen. Alternatieve reguleringsmodellen die zijn ontstaan als reactie hierop zijn management-based regulering, risk-based regulering, zelfregulering en polycentrische regelgeving (waaronder certificering door beroepsorganisaties of andere private actoren). Concreet zijn deze ontwikkelingen terug te zien in het zelfaccrediterende systeem in Noorwegen en het monitoringssysteem dat gebaseerd is op publieke informatie in Engeland.

Ook in Nederland kan een systeemverandering een vernieuwingsslag tot stand brengen. Echter moet er rekening mee gehouden worden dat het delegeren van accrediterende taken van een externe organisatie naar de instelling zelf niet noodzakelijkerwijs de werkdruk verlaagt of autonomie voor opleidingen hoeft te vergroten.

Ongeacht de vraag of systeemverandering wordt doorgevoerd zijn belangrijke randvoorwaarden voor effectieve en efficiënte beoordeling en bewaking van onderwijskwaliteit het hebben van zekerheid en vertrouwen. Als processen als onvoorspelbaar worden ervaren, kan dit leiden tot onzekerheid. Als men uit onzekerheid meer werkdruk naar zich toetrekt, wordt onderwijskwaliteit mogelijk niet verbeterd, maar juist ondergesneeuwd. Daarnaast is vertrouwen nodig om gebruik te kunnen maken van vrijheden binnen een systeem. Als er een ‘afrekencultuur’ heerst of wantrouwen als uitgangspunt wordt genomen, wordt het gevoel van eigenaarschap ondermijnt. Er kan risico vermijdend gedrag ontstaan waarbij er juist minder ruimte wordt genomen voor maatwerk, maar men uit angst gaat standaardiseren. Tot slot is het van belang dat men in de hoger onderwijssector van elkaar kan leren. Omdat kwaliteitszorg op verschillende manieren kan worden vormgegeven, is het nuttig om ervaringen met elkaar te delen, te leren van elkaar, en best practices over te nemen. Facultaire of juist internationale platformen waarin men, het liefst in samenwerking met externe belanghebbenden, processen kan bestuderen en erop kan reflecteren, kunnen expertise, kennis en het zelfvertrouwen om te experimenteren binnen de gestelde kaders verder laten groeien.

Literatuurlijst

Adeinat, I., Al Rahahleh, N., & Al Bassam, T. (2022). Lean Six Sigma and Assurance of Learning (AoL) in higher education: a case study. *International Journal of Quality & Reliability Management*, 39(2), 570–587. <https://doi.org/10.1108/IJQRM-01-2021-0017>

Albliwi, S., Antony, J., Halim Lim, S.A., & Von der Wiele, T. (2014). Critical Failure Factors of Lean Six Sigma: a Systematic Literature Review. *International Journal of Quality & Reliability Management*, 31, 1012–1030. <https://doi.org/10.1108/IJQRM-09-2013-0147>

Allais, S. (2014). *Selling Out Education: National qualifications frameworks and the neglect of knowledge*. Rotterdam, Sense.

Antony, J., Ghadge, A., Ashby, S.A., & Cudney, E.A. (2018). Lean Six Sigma journey in a UK higher education institute: a case study. *International Journal of Quality & Reliability Management*, 35(2), 510–526. <https://doi.org/10.1108/IJQRM-01-2017-0005>

Baldwin, R., Cave, M., & Lodge, M. (2012) *Understanding Regulation: Theory, Strategy, and Practice*. Oxford: Oxford University Press.

Bay, D., & Daniel, H. (2001). The Student Is Not the Customer—An Alternative Perspective. *Journal of Marketing for Higher Education*, 11, 1–19. https://doi.org/10.1300/J050v11n01_01

Black, J. (2010). Risk-based Regulation: Choices, Practices and Lessons Being Learnt. In *Risk and Regulatory Policy: Improving the Governance of Risk*, pp. 185–224. Parijs, OECD.

Beard, D.F. (2009). Successful Applications of the Balanced Scorecard in Higher Education. *Journal of Education for Business*, 84(5), 275–282. <https://doi.org/10.3200/JOEB.84.5.275-282>

Beerkens, M. (2015). Quality assurance in the political context: in the midst of different expectations and conflicting goals. *Quality in Higher Education*, 21(3), 231–250.

Beerkens, M. (2021). An evolution of performance data in higher education governance: a path towards a ‘big data’ era?. *Quality in Higher Education*. <https://doi.org/10.1080/13538322.2021.1951451>

Beerkens, M. (2022). Steering internationalization by labels and certificates: The role of the European Internationalisation Certificate in higher education quality. *Journal of Studies in International Higher Education*, forthcoming.

Beerkens, M., & Udam, M. (2017). Stakeholders in higher education quality assurance: Richness in diversity?. *Higher Education Policy*, 30(3), 341–359. <https://doi.org/10.1057/s41307-016-0032-6>

- Bendermacher, G.W.G., Oude Egbrink, M.G.A., Wolfhagen, I.H.A.P., & Dolmans, D.H.J.M. (2017). Unravelling quality culture in higher education. *Higher Education*, 73(1), 39–60. <https://doi.org/10.1007/s10734-015-9979-2>
- Brennan, J., & Shah, T. (2000). *Managing Quality in Higher Education: An international perspective on institutional assessment and change*. Buckingham, SRHE/Open University Press.
- Breetvelt, I. (2021). *Studentevaluatie basis voor docentbeoordeling en onderwijsverbetering?*. Kohnstamm Instituut. <https://kohnstammstituut.nl/wp-content/uploads/2021/07/Breetvelt-2021-Studentevaluatie-basis-voor-docentbeoordeling-en-onderwijsverbetering.pdf>
- Brink, C. (2010). Quality and Standards: Clarity, Comparability and Responsibility. *Quality in Higher Education*, 16(2), 139-152. <https://doi.org/10.1080/13538322.2010.487698>
- Brookes, M., & Becket, N. (2007). Quality Management in Higher Education: A Review of International Issues and Practice. *International Journal of Quality and Standards*, 1, 1-37.
- Cameron, J. (2009). *An Exploration of Key Performance Indicators for Academic Quality*. Paper presented at INQAAHE Conference 2009. University of Canterbury, Christchurch, New Zealand. https://www.inqaahe.org/sites/default/files/pictures/3_Jan%20Cameron_An%20exploration%20of%20key%20performance%20indicators%20for%20academic%20quality.pdf
- Caspersen, J., & Smeby, J.-C. (2018). The relationship among learning outcome measures used in higher education. *Quality in higher education*, 24(2), 117–135. <https://doi.org/10.1080/13538322.2018.1484411>
- Cave, M., Henkel, M., Kogan, M. & Hanney, S. (1996). *The Use of Performance Indicators in Higher Education: The Challenge of the Quality Movement*. London, Jessica Kingsley.
- Chalmers, D. (2008). *Indicators of University Learning and Teaching Quality*. Australian Learning and Teaching Council. https://www.academia.edu/20497802/Indicators_of_university_teaching_and_learning_quality
- Chen, K. H.-J., & Hou, A. Y.-C. (2016). Adopting self-accreditation in response to the diversity of higher education: quality assurance in Taiwan and its impact on institutions. *Asia Pacific Education Review*, 17(1), 1–11.
- Chu, A., & Westerheijden, D.F. (2018). Between quality and control: what can we learn from higher education quality assurance policy in the Netherlands. *Quality in Higher Education*, 24(3), 260-270. <https://doi.org/10.1080/13538322.2018.1559513>
- Clayson, D.E., & Haley, D.A. (2005). Marketing Models in Education: Students as Customers, Products, or Partners. *Marketing Education Review*, 15, 1-10. <https://doi.org/10.1080/10528008.2005.11488884>

Coates, H. (2010). Development of the Australasian survey of student engagement (AUSSE). *Higher Education*, 60(1), 1–17.

Coglianesi, C. and Starobin, S. M. (2020). Management-based regulation. In *Elgar Encyclopedia of Environmental Law*, pp. 292-307. Cheltenham, Edward Elgar Publishing.

Cort, P. (2010). Stating the obvious: the European qualifications framework is not a neutral evidence-based policy tool. *European Educational Research Journal*, 9(3), 304–16.

Csizmadia, T.G. (2006). *Quality Management in Hungarian Higher Education: Organisational Responses to Governmental Policy*. CHEPS/University of Twente. <https://www.utwente.nl/en/bms/cheps/phd-page/cheps-alumni-and-their-theses/thesiscsizmadia.pdf>

Cui, V., French, A., & O’Leary, M. (2021). A missed opportunity? How the UK’s teaching excellence framework fails to capture the voice of university staff. *Studies in Higher Education*, 46(9), 1756-1770. <https://doi.org/10.1080/03075079.2019.1704721>

de Paor, C. (2016). The contribution of professional accreditation to quality assurance in higher education. *Quality in Higher Education*, 22(3), 228-241. <https://doi.org/10.1080/13538322.2016.1263925>

Deslauriers, L., McCarty, L.S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *PNAS*, 116(39), 19251-19257. <https://doi.org/10.1073/pnas.1821936116>

Dill, D.D., & Beerkens, M. (2013). Designing the framework conditions for assuring academic standards. *Higher Education*, 65(3), 341–357. <https://doi.org/10.1007/s10734-012-9548-x>

Dill, D. & Soo, M. (2005). ‘Academic quality, league tables and public policy: a cross-national analysis of university ranking systems’, *Higher Education*, 49(4), 495–537.

Doherty, G.D. (1993). Towards Total Quality Management in higher education: A case study of the University of Wolverhampton. *Higher Education the International Journal of Higher Education and Educational Planning*, 25(3), 321–339. <https://doi.org/10.1007/BF01383857>

Dubnick, M., & Frederickson, H.G. (2011) *Public accountability: Performance measurement, the extended state, and the search for trust*. National Academy of Public Administration. Kettering Foundation.

Eagle, L. & Brennan, R. (2007) Are Students Customers? TQM and Marketing Perspectives. *Quality Assurance in Education*, 15, 44-60. <https://doi.org/10.1108/09684880710723025>

Eaton, J.S. (2015). *An Overview of U.S. Accreditation*. Council for Higher Education Accreditation.

Edwards, F. (2012). The evidence for a risk-based approach to Australian higher education regulation and quality assurance. *Journal of Higher Education Policy and Management*, 34(3), 295-307.

Elken, M., Frølich, N., Maassen, P., & Stensaker, B. (2020). Quality Systems in Higher Education Institutions: Enabling and Constraining Quality Work. In *Quality Work in Higher Education: Organisational and Pedagogical Dimensions*, Elken, M., Maassen, P., Nerland, M., Prøitz, T.S., Stensaker, B., & Vabø, A. (Eds.). Dordrecht: Springer.

Elken, M., & Stensaker, B. (2018). Conceptualising 'quality work' in higher education. *Quality in Higher Education*, 24(3), 189–202. <https://doi.org/10.1080/13538322.2018.1554782>

El-Sayed, M., Zgorzelski, M., Berry, K.J., & Zang, P.H. (2005). *Lean Thinking and Quality Control Strategies for Improving Engineering Educational Processes*. ASEE Annual Conference and Exposition, Conference Proceedings American Society for Engineering Education, Portland.

Elton, L. (2004). Goodhart's Law and Performance Indicators in Higher Education. *Evaluation & Research in Education*, 18(1-2), 120-128. <https://doi.org/10.1080/09500790408668312>

ESG (2015). *Standards and Guidelines for Quality Assurance in the European Higher Education Area*. Brussel.

<https://www.eua.eu/downloads/content/standards%20and%20guidelines%20for%20quality%20assurance%20in%20the%20european%20higher%20education%20area%20esg%202015.pdf>

European University Association (2006). *Quality Culture in European Universities: A Bottom-Up Approach: Report on the Three Rounds of the Quality Culture Project 2002-2006*. <https://eua.eu/downloads/publications/quality%20culture%20in%20european%20universities%20a%20bottom-up%20approach.pdf>

Ewell, P.T. (1999). Linking Performance Measures to Resource Allocation: exploring unmapped terrain. *Quality in Higher Education*, 5(3), 191–209. <https://doi.org/10.1080/1353832990050302>

Gilad, S. (2010). It runs in the family: Meta-regulation and its siblings. *Regulation & Governance*, 4, 485-506. <https://doi.org/10.1111/j.1748-5991.2010.01090.x>

Gómez, J., Martínez-Costa, M., & Martínez-Lorente, A. (2015). EFQM Excellence Model and TQM: an empirical comparison. *Total Quality Management & Business Excellence*, 1(2), 1-16. <https://doi.org/10.1080/14783363.2015.1050167>

Gover, A., & Loukkola, T. (2018). Enhancing quality: From policy to practice. *Enhancing Quality through Innovative Policy & Practice (EQUIP)*. <https://eua.eu/downloads/publications/enhancing%20quality%20from%20policy%20to%20practice%20equip%20publication%20final.pdf>

- Gunn, A. (2018a). The UK Teaching Excellence Framework (TEF): The Development of a New Transparency Tool. In *European Higher Education Area: The Impact of Past and Future Policies*, Curaj A., Deca L., & Pricopie R. (eds). Springer, Cham. https://doi.org/10.1007/978-3-319-77407-7_31
- Gunn, A. (2018b). Metrics and methodologies for measuring teaching quality in higher education: developing the Teaching Excellence Framework (TEF). *Educational Review (Birmingham)*, 70(2), 129–148. <https://doi.org/10.1080/00131911.2017.1410106>
- Gupta, S.K., Antony, J., Lacher, F., & Douglas, J. (2020). Lean Six Sigma for reducing student dropouts in higher education - an exploratory study. *Total Quality Management & Business Excellence*, 31(1-2), 178–193. <https://doi.org/10.1080/14783363.2017.1422710>
- Haerizadeh, M., & Sunder, V.M. (2019). Impacts of Lean Six Sigma on improving a higher education system: a case study. *International Journal of Quality & Reliability Management*, 36(6), 983–998. <https://doi.org/10.1108/IJQRM-07-2018-0198>
- Harvey, L. (2002). The End of Quality?. *Quality in Higher Education*, 8(1), 5-22. <https://doi.org/10.1080/13538320220127416>
- Harvey, L. & Green, D. (1993). Defining Quality. *Assessment & Evaluation in Higher Education*, 18(1), 9-34. <https://doi.org/10.1080/0260293930180102>
- Harvey, L., & Newton, J. (2004). Transforming quality evaluation. *Quality in Higher Education*, 10(2), 149-165. <https://doi.org/10.1080/1353832042000230635>
- Harvey, L., & Stensaker, B. (2008). Quality culture: understandings, boundaries and linkages. *European Journal of Education*, 43(4), 427–42. <https://doi.org/10.1111/j.1465-3435.2008.00367.x>
- Hazelkorn, E. (2009). Rankings and the battle for world-class excellence: Institutional strategies and policy choices. *Higher Education Management and Policy*, 21(1), 47–68. <https://doi.org/10.1787/hemp-v21-art4-en>
- Hildesheim, C., & Sonntag, K. (2020). The Quality Culture Inventory: a comprehensive approach towards measuring quality culture in higher education. *Studies in Higher Education*, 45(4), 892-908. <https://doi.org/10.1080/03075079.2019.1672639>
- Hinojosa V., M.F. (2021). Adaptation of the Balanced Scorecard to Latin American Higher Education Institutions in the Context of Strategic Management: A Systematic Review with Meta-analysis. In *Production Research*, pp. 125–140. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-76310-7_10

Hladchenko, M. (2015). Balanced Scorecard – a strategic management system of the higher education institution. *International Journal of Educational Management*, 29(2), 167–176. <https://doi.org/10.1108/IJEM-11-2013-0164>

Kaplan, R.S., & Norton, D.P. (1996). *Translating Strategy into Action: The Balanced Scorecard*. Boston, MA, Harvard Business School Press.

Kaplan, R.S., & Norton, D.P. (2001). *The Strategy-focused Organization: How Balanced Scorecard Companies Thrive in the New Business Environment*. Boston, MA, Harvard Business School Press.

Kaplan, R.S., & Norton, D.P. (2004). *Strategy Maps: Converting the Intangible Assets into Tangible Outcomes*. Boston, MA, Harvard Business School Press.

Keamey, A.T. (1991). *Total Quality: time to take off the rose-coloured spectacles*. London, Kearney Ltd & IFS.

Kettunen, J. (2008). A conceptual framework to help evaluate the quality of institutional performance. *Quality Assurance in Education*, 16(4), 322-332. <https://doi.org/10.1108/09684880810906472>

King, R. (2018). Risk-based regulation in higher education: why, how, when, and what else? In *Handbook on the Politics of Higher Education*, Brendan Cantwell, Hamish Coates and Roger King (eds). Cheltenham, Edward Elgar.

Kis, V. (2005). Quality Assurance in Tertiary Education: Current Practices in OECD Countries and a Literature Review on Potential Effects. *OECD Thematic Review of Tertiary Education*. <https://www.oecd.org/education/skills-beyond-school/38006910.pdf>

Lapiente, V., van de Walle, S. (2020). The effects of new public management on the quality of public services. *Governance (Oxford)*, 33(3), 461–475. <https://doi.org/10.1111/gove.12502>

Leiber, T. (2019). A general theory of learning and teaching and a related comprehensive set of performance indicators for higher education institutions. *Quality in Higher Education*, 25(1), 76–97. <https://doi.org/10.1080/13538322.2019.1594030>

Li, J., & Pilz, M. (2017). Modularisation in the German VET system: a study of policy implementation. *Journal of Education and Work*, 30(5), 471-485. <https://doi.org/10.1080/13639080.2016.1243233>

Liu, S., Tan, M., & Zharoui, M. (2015). Impact of Quality Assurance on Higher Education Institutions: A Literature Review. *Higher Education Evaluation and Development*, 9(2), 17-34.

Lomas, L. (2001). Does the development of mass education necessarily mean the end of quality?. *Paper presented at The Sixth QHE Seminar: The End of Quality?*, Birmingham, 25-26 May.

- Loukkola, T. Peterbauer, H., & Gover, A. (2020). *Exploring higher education indicators*. European University Association (EUA).
<https://www.eua.eu/downloads/publications/indicators%20report.pdf>
- Loukkola, T., & Zhang, T. (2010). *Examining Quality Culture: Part 1- Quality Assurance Processes in Higher Education Institutions*. European University Association (EUA).
<https://eua.eu/downloads/publications/examining%20quality%20culture%20part%20i%20quality%20assurance%20processes%20in%20higher%20education%20inst.pdf>
- Manatos, M.J., & Sarrico, C.S., & Rosa, M.J. (2017). The integration of quality management in higher education institutions: a systematic literature review. *Total Quality Management*, 28(2), 159-175.
<https://doi.org/10.1080/14783363.2015.1050180>
- Middlehurst, R., & Woodhouse, D. (1995). Coherent Systems for External Quality Assurance. *Quality in Higher Education*, 1(3).
- Nadeau, S. (2017). Lean, Six Sigma and Lean Six Sigma in Higher Education: A Review of Experiences around the World. *American Journal of Industrial and Business Management*, 7(5), 591-603.
<https://doi.org/10.4236/ajibm.2017.75044>
- NOKUT (2016). *NOKUT's evaluation of institutions' quality assurance systems for educational provision (audit): Summary of analysis of the second round of audits (2009–2016)*.
https://www.nokut.no/globalassets/nokut/rapporter/evalueringer/2018/nokuts_evaluation_of_institutions_quality_assurance_systems.pdf
- NOKUT (2016). *Kvalitetsområder for studieprogram: Veiledende document*.
https://www.nokut.no/siteassets/akkrediteringhu/kvalitetsomrader_for_studieprogram_300816.pdf
- OECD (2015). *AHELO Main Study*. <https://www.oecd.org/education/skills-beyond-school/ahelo-main-study.htm>
- OECD (2019). Indicator D2. What is the student-teacher ratio and how big are classes?, In *Education at a Glance 2019*. Parijs, OECD.
- Office for students (2018). *Teaching Excellence and Student Outcomes Framework: Subject-level pilot guide*.
https://www.officeforstudents.org.uk/media/57eb9beb-4e91-497b-860b-2fd2f39ae4ba/ofs2018_44_updated.pdf
- Office for Students (2021). *Teaching Excellence and Student Outcomes Framework (TEF): Findings from the subject-level pilot 2018-19*. <https://www.officeforstudents.org.uk/media/b5257a90-2992-465f-ba5c-d53fb68ca082/tef-findings-from-the-subject-level-pilot-2018-19.pdf>

- Onderwijsraad (2015). *Kwaliteit in het hoger onderwijs: Evenwicht in ruimte, regels en rekenschap*. <https://www.onderwijsraad.nl/binaries/onderwijsraad/documenten/adviezen/2015/08/28/kwaliteit-in-het-hoger-onderwijs/Kwaliteit-in-het-hoger-onderwijs-1.7.pdf>
- Prades, A., Quirós, M., Giné M., & Bernáldez, L. (2015). *Big (and Small) Data meets Quality Assurance*. 10th European Quality Assurance Forum. https://www.aqu.cat/doc/doc_10794542_1.pdf
- Radnor, Z., & Bucci, G. (2011). *Analysis of Lean Implementation in UK Business Schools Universities*. A to Z Business Consultancy.
- Rauhvargers, A. (2011). *Global University Rankings and Their Impact*. European University Association. <https://www.eua.eu/downloads/publications/global%20university%20rankings%20and%20their%20impact.pdf>
- Rijksoverheid (2018). *Investeren in Onderwijskwaliteit: Kwaliteitsafspraken 2019-2024*. Amsterdam. <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2018/04/09/investeren-in-onderwijskwaliteit-kwaliteitsafspraken-2019-2024/investeren-in-onderwijskwaliteit-kwaliteitsafspraken-2019-2024.pdf>
- Romainville, M. (1999). Quality Evaluation of Teaching in Higher Education. *Higher Education in Europe*, 24(3), 414-424. <https://doi.org/10.1080/0379772990240309>
- Rosa, M.J., & Amaral, A. (2007). A self-assessment of higher education institutions from the perspectives of EFQM model. In *Quality assurance in higher education: Trends in regulation, translation and transformation*, Westerheijden, D.F. (Ed.), pp. 181-207. Dordrecht: Springer.
- Rosa, M.J., Cardoso, S., Dias, D., & Amaral, A. (2011). The EUA institutional evaluation programme: an account of institutional best practices. *Quality in Higher Education*, 17(3), 369–86. <https://doi.org/10.1080/13538322.2011.625207>
- Salemans, L., & Budding, G. (2021). Operationalizing public value in higher education: The use of narratives as an alternative for performance indicators. *Journal of Management and Governance*. <https://doi.org/10.1007/s10997-021-09596-4>
- Salto, D.J. (2021). Beyond national regulation in higher education? Revisiting regulation and understanding organisational responses to foreign accreditation of management education programmes. *Quality in Higher Education*, 27(2), 206-221. <https://doi.org/10.1080/13538322.2020.1833420>
- Sayed, N. (2013). Ratify, reject or revise: balanced scorecard and universities. *International Journal of Educational Management*, 27(3), 203–220. <https://doi.org/10.1108/09513541311306440>
- Scheerens, J. (2004). Perspectives on Education Quality, Education Indicators and Benchmarking. *European Educational Research Journal*, 3(1), 115-138. <https://doi.org/10.2304/eeerj.2004.3.1.3>

Schwarz, S., & Westerheijden, D.F. (2004). *Accreditation and Evaluation in the European Higher Education Area* (Eds.). Dordrecht: Kluwer.

Seymour, C. (1991). TQM on campus: What the pioneers are finding. *AAHE Bulletin*, 44, 10-18.

Seymour, C. (1992). *On Q: Causing Quality in Higher Education*. New York, American Council on Education/Macmillan.

Shanshan, S., Wenfei, L., & Lijuan, L. (2021). Applying lean Six Sigma incorporated with big data analysis to curriculum system improvement in higher education institutions. *International Journal of System Assurance Engineering and Management*. <https://doi.org/10.1007/s13198-021-01316-3>

Skelton, A. (2012). Colonised by quality? Teacher identities in a research-led institution. *British Journal of Sociology of Education*, 33(6), 793–811. <https://doi.org/10.1080/01425692.2012.692047>

Sizer, J., Spee, A., & Bormans, R. (1992). The role of performance indicators in higher education. *Higher Education*, 24, 133-155. <https://link.springer.com/content/pdf/10.1007/BF00129438.pdf>

SQELT project (2020). *ERASMUS+ Project Sustainable Quality Enhancement in Higher Education Learning and Teaching*. <https://www.evalag.de/forschung/sqelt/the-project/>

Stensaker, B. (2007). Quality as fashion. In *Quality Assurance in Education*, Westerheijden, D., Stensaker, B., & Rosa, M. (Eds.). Dordrecht: Springer.

Stensaker, B., & Harvey, L. (2006). Old wine in new bottles? A comparison of public and private accreditation schemes in higher education. *Higher Education Policy*, 19(1), 65-85. <https://doi.org/10.1057/palgrave.hep.8300110>

Stensaker, B., Langfeldt, L., Harvey, L., Huisman, J., & Westerheijden, D. (2011). An In-depth Study on the Impact of External Quality Assurance. *Assessment & Evaluation in Higher Education*, 36(4), 465-478. <https://doi.org/10.1080/02602930903432074>

Stroebe, W. (2020). Student evaluations of teaching encourages poor teaching and contributes to grade inflation: A theoretical and empirical analysis. *Basic and Applied Social Psychology*, 1-19. <https://doi.org/10.1080/01973533.2020.1756817>

Sulaiman, N.F., Manochehri, N.-N., & Al-Esmail, R.A. (2013). Level of Total Quality Management adoption in Qatari Educational Institutions: Private and semi-government sector. *Journal of Education for Business*, 88(2), 76–87. <https://doi.org/10.1080/08832323.2011.649311>

Sunder, M.V. (2013). Synergies of Lean Six Sigma. *The IUP Journal of Operations Management*, 12(1), 21-31.

Sursock, A. (2011). *Examining Quality Culture - Part II: Processes and Tools - Participation, Ownership and Bureaucracy*. European University Association (EUA). Brussel.

Sursock, A., Smidt, H., Davies, H., Korhonen, J., Madill, G., & Wilson, L. (2010). *Trends 2010: A decade of change in European Higher Education*. European University Association (EUA).
<https://www.eua.eu/downloads/publications/trends%202010%20a%20decade%20of%20change%20in%20european%20higher%20education.pdf>

Tavares, O., Sin, C., Videira, P., & Amaral, A. (2017). Academics' perceptions of the impact of internal quality assurance on teaching and learning. *Assessment & Evaluation in Higher Education*, 42(8), 1293-1305.
<https://doi.org/10.1080/02602938.2016.1262326>

Taylor, J., & Baines, C. (2012). Performance management in UK universities: implementing the Balanced Scorecard. *Journal of Higher Education Policy and Management*, 34(2), 111–124.
<https://doi.org/10.1080/1360080X.2012.662737>

U-Multirank (2021a). *Basic Facts about U-Multirank*. <https://www.umultirank.org/export/sites/default/press-media/documents/Basic-Facts-about-U-Multirank-2021.pdf>

U-Multirank (2021b). *University Ranking by Subject*. <https://www.umultirank.org/university-rankings/rankings-by-subject/>

U-Multirank (2021c). *U-Multirank's Data Sources and Verification*.
<https://www.umultirank.org/about/methodology/data-sources/>

U-Multirank (2021d). *Catalogue of Indicators*. <https://www.umultirank.org/about/methodology/indicators/>

U-Multirank (2021e). *Our Approach*. <https://www.umultirank.org/about/methodology/our-approach/>

van den Broek, A., Lodwick, J., Termorshuizen, T., de Vries, H., Braam, C., Wiertsema, S., Vlaar, V., & van Casteren W. (2019). *Instellingsaccreditatie in het buitenland: Een overzicht van de invoering van instellingsaccreditatie in het hoger onderwijs in Vlaanderen, Denemarken, Noorwegen, Letland, Estland en Schotland*. ResearchNed. <http://www.researchned.nl/wp-content/uploads/2019/07/Instellingsaccreditatie-in-het-buitenland.pdf>

van Dooren, W., Bouckaert, G., & Halligan, J. (2015). *Performance Management in the Public Sector*. London, Routledge.

van Vught, F. (1994). Autonomy and Accountability in Government/University Relationships. In *Revitalizing Higher Education*, Salmi, J. and Verspoor, A. (Eds.), pp. 322–362. International Association of Universities (IAU) Press.

- Vivian, D., James, A.S., Salamons, D., Hazel, Z., Felton, J., & Whittaker S. (2019) *Evaluation of Provider-Level TEF 2016-17 (Year 2): Measuring the Initial Impact of the TEF on the Higher Education Landscape Research Report*. London, Department for Education. Accessed August 2019.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/914169/TEF_Year_2_Evaluation_Report__6_.pdf
- Vroeijenstijn, A.I. (1995). Improvement and accountability: navigating between Scylla and Charybdis. *Higher Education Policy Series*, 30.
- VSNU (2012). *Prestaties in perspectief: Trendrapportage universiteiten 2000-2020*. Den Haag.
https://www.universiteitenvannederland.nl/files/documenten/Publicaties/Trendrapportage_DEF.pdf
- Wächter, B., Kelo, M., Lam, Q.K.H., Effertz, P., Jost, C., & Kottowski, S. (2015). *University Quality Indicators: A Critical Assessment*. European Parliament's Committee on Culture and Education.
[https://www.europarl.europa.eu/RegData/etudes/STUD/2015/563377/IPOL_STU\(2015\)563377_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2015/563377/IPOL_STU(2015)563377_EN.pdf)
- Watty, K. (2006). Want to Know about Quality in Higher Education? Ask an Academic. *Quality in Higher Education*, 12(3), 291-301. <https://doi.org/10.1080/13538320601051101>
- Westerheijden, D.F., Cremonini, L., Kolster, R., Kottmann, A., Redder, L., Soo, M., Vossensteyn, H., & de Weert, E. (2008). *New Degrees in the Netherlands: Evaluation of the bachelor-master structure and accreditation in Dutch higher education: Final report*. Den Haag, Ministerie van Onderwijs, Cultuur en Wetenschap.
- Williams, G. (1993). Total Quality Management in Higher Education: Panacea or Placebo? *Higher Education*, 25(3), 229–237. <https://doi.org/10.1007/BF01383852>
- Yu, M.L., Hamid, S., Ijab, M.T., & Soo, H.P. (2009). The E-Balanced Scorecard (e-BSC) for Measuring Academic Staff Performance Excellence. *Higher Education*, 57(6), 813–828. <https://doi.org/10.1007/s10734-009-9197-x>
- Zhu, Z. (1999). A comparison of quality programmes: Total quality management and ISO 9000. *Total Quality Management*, 10(2), 291–297. <https://doi.org/10.1080/0954412998018>

