



UvA-DARE (Digital Academic Repository)

Handleiding

Evaluatie van geesteswetenschappelijk onderzoek volgens het SEP

Prins, A.; Spaapen, J.; van Vree, F.; van Leeuwen, T.; Duindam, D.

Publication date

2017

Document Version

Final published version

License

Other

[Link to publication](#)

Citation for published version (APA):

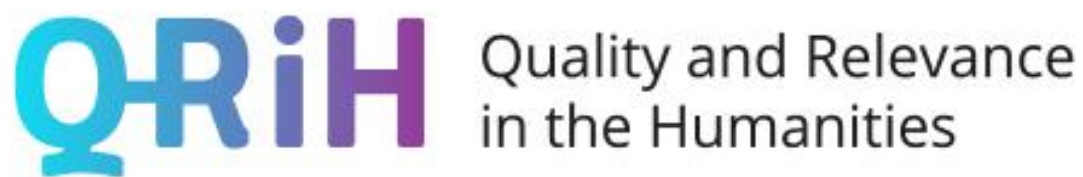
Prins, A., Spaapen, J., van Vree, F., van Leeuwen, T., & Duindam, D. (2017). *Handleiding: Evaluatie van geesteswetenschappelijk onderzoek volgens het SEP*. QRiH.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

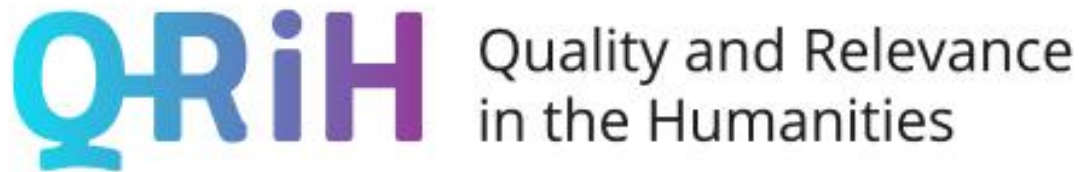
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



Handleiding

Evaluatie van geesteswetenschappelijk onderzoek volgens het SEP

21 september 2017



Handleiding

Evaluatie van geesteswetenschappelijk onderzoek volgens het SEP

Inhoudsopgave

0. Ter inleiding
1. Evaluatie volgens het SEP: uitgangspunten en samenhang
2. Aggregatieniveaus
3. Narratieven als basiselement van het zelfevaluatie-rapport
4. Het instrumentarium: QRiH
5. Aspecten van kwaliteit en relevantie
6. SWOT-analyse
7. Casestudies
8. Tot slot: randvoorwaarden en aanbevelingen

Ter inleiding

De geesteswetenschappen zijn, als het gaat om robuustheid en methodische strengheid, in veel opzichten niet anders dan andere wetenschappen. Tegelijk is er sprake van een eigen, herkenbare wetenschapscultuur met dito praktijken, waarvan de oriëntatie op een breder publiek dan alleen collega-wetenschappers een van de belangrijkste kenmerken is. Gevolg hiervan is dat onderzoekers in de geesteswetenschappen slechts in beperkte mate uit de voeten kunnen met de systemen en instrumenten die in andere wetenschapsdomeinen veel gebruikt worden om onderzoekskwaliteit en – relevantie uit te drukken. *Web of Science* en *Scopus* spelen in grote delen van de geesteswetenschappen – maar ook elders – geen enkele rol, en dat geldt ook voor meetinstrumenten als de *H-index* of citatiescores. Veel geesteswetenschappers herkennen zich net zo min in de eenzijdige oriëntatie op publicaties in Engelstalige tijdschriften. Boeken, tentoonstellingen en documentaires, ook in andere talen dan het Engels, bestemd voor vakgenoten én een breder publiek kunnen in hun domein *de facto* een betere maatstaf voor kwaliteit vormen: dat is waarin grensverleggend onderzoek vaak zijn weerslag vindt. In de ogen van veel onderzoekers ligt daar zelfs een opdracht: wanneer zij zich als onderzoekers zouden beperken tot het domein van Engelstalige *peer reviewed journals* zouden zij de banden doorsnijden met de levende cultuur – en daarmee met hun voedingsbodem.

Dat alles neemt niet weg dat de kwaliteit van onderzoeksresultaten, in termen van wetenschappelijke en maatschappelijke relevantie of impact, ook in de geesteswetenschappen getoetst kan worden. Voorwaarde daarvoor is dat het beoordelingskader en de daartoe behorende instrumenten leiden tot een oordeel dat in de geesteswetenschappen zelf als zinvol en waardevol wordt beschouwd. Het moet, kortom, gaan om een beoordeling die recht doet aan de waarden van de onderzoekers in de verschillende geesteswetenschappelijke domeinen¹ en die het mogelijk maakt op kwaliteit en relevantie te sturen, om te beginnen door de onderzoeksgroepen zelf. Dat beogen we te bereiken door in elk geesteswetenschappelijk domein panels in te stellen, die de kenmerken van kwaliteit en relevantie kunnen vaststellen. Dit ijken en herijken van kwaliteit en relevantie gebeurt dus van onderop.

¹ In deze handleiding worden ter aanduiding van verschillende onderzoeksgebieden twee termen gebezigd: ‘domeinen’ en ‘subdomeinen’. Met de term *domein* wordt bedoeld op de geesteswetenschappen als geheel alsmede de daarin te onderscheiden vakgebieden of disciplines; specialismen en meer specifieke vakgebieden worden aangeduid als *subdomeinen*. Deze twee min of meer neutrale termen verdienen de voorkeur boven die van ‘disciplines’ of ‘vakgebieden’, omdat die terminologie in de praktijk veel verwarring oproept en bovendien de suggestie in zich draagt van duidelijk afgegrensde terreinen van academische activiteiten, een suggestie die in de praktijk gelogenstraft wordt door het multidisciplinaire karakter van veel onderzoek. Alleen in de lijsten van tijdschriften wordt gesproken van ‘multidisciplinaire tijdschriften’, omdat die term niet met deze connotaties beladen is.

De meest recente versie van het *Standard Evaluation Protocol* (SEP), dat is opgesteld door VSNU, KNAW en NWO als kader voor de evaluatie van wetenschappelijk onderzoek in de periode 2015 - 2021, biedt volop ruimte om het aangereikte evaluatie-instrumentarium geschikt te maken voor uiteenlopende domeinen, niet alleen in de geesteswetenschappen, maar ook in andere domeinen waar de hierboven geschetste eenzijdige, kwantitatieve oriëntatie op publicaties in toptijdschriften ter discussie is gesteld. Het SEP geeft de geesteswetenschappen dus de kans zich in de volle breedte te profileren en te presenteren, niet alleen waar het gaat om de wetenschappelijke kwaliteit, maar ook om hun maatschappelijke betekenis.

Deze handleiding is gemaakt om onderzoekseenheden op verschillende aggregatieniveaus in de geesteswetenschappen in staat te stellen zelfevaluatierapporten te schrijven die voldoen aan het SEP en tegelijk recht doen aan de eigen normen van kwaliteit, zowel waar het gaat om de waardering van onderzoeksresultaten in de wetenschap als in de samenleving.

1. Evaluatie gebaseerd op het SEP: uitgangspunten en samenhang

Het SEP 2015-2021 wil een gebalanceerde beoordeling mogelijk maken van de wetenschappelijke kwaliteit én de maatschappelijke relevantie van geleverd werk. Aan de basis van de beoordeling staat het zelfevaluatierapport, dat op pagina 13 van het SEP en ook in *Appendix D Format self-assessment report* besproken wordt. De door SEP voorgeschreven Tabel D1 (*output indicators*) op pagina 25 vormt de grondslag van de feitelijke onderbouwing van het zelfevaluatierapport. De indicatoren in deze tabel zijn leidend voor het zelfevaluatierapport en de informatie die ervoor wordt verzameld is de basis voor de evaluatie door de beoordelingscommissie.

		Quality domains	
		Research quality	Relevance to society
Assessment dimensions	Demonstrable products	1. Research products for peers	4. Research products for societal target groups
	Demonstrable use of products	2. Use of research products by peers	5. Use of research products by societal target groups
	Demonstrable marks of recognition	3. Marks of recognition from peers	6. Marks of recognition by societal target groups

Tabel 1 Soorten outputindicatoren volgens het SEP. Verkorte versie van tabel D1: format self-assessment report, blz. 25

Op het eerste gezicht lijkt zo'n zelfevaluatie aan de hand van outputindicatoren eenvoudig. Maar dat is het niet. Ten eerste is het verleidelijk (én begrijpelijk) deze tabel te vullen met zo veel mogelijk data. Daaruit komt echter niet vanzelfsprekend een helder beeld van de gerealiseerde ambities te voorschijn, in het bijzonder waar het de samenhang tussen de wetenschappelijke en maatschappelijke doelen van het onderzoek betreft.

Ten tweede zijn de genoemde indicatoren vooralsnog lege categorieën, die weliswaar veldafhankelijk kunnen worden gebruikt, maar die op zichzelf nog niets zeggen over de kwaliteit en de relevantie van de resultaten die daaronder worden geschaard. Anders gezegd: het SEP neemt weliswaar duidelijk afstand van de overwegend kwantitatieve, sterk op de exacte wetenschappen georiënteerde en betrekkelijk eenzijdige wijze van kwaliteitsmeting, maar dat laat onverlet dat de criteria voor kwaliteit en relevantie die recht doen aan de eigenheid van de geesteswetenschappen en de afzonderlijke domeinen wel moeten worden geoperationaliseerd.

Precies daarin voorziet deze handleiding door:

1. het bieden van een leidraad bij het ordenen en profileren van de onderzoeksresultaten, waarin de nadruk komt te liggen op 'narratieven' waarin de verschillende aspecten van kwaliteit en relevantie in samenhang kunnen worden behandeld;
2. het bieden van een set van indicatoren van kwaliteit en relevantie die in de verschillende onderzoeksdomeinen van de geesteswetenschappen kunnen worden ingezet bij het schrijven van een zelfevaluatierapport. Deze indicatoren kunnen zowel kwantitatief als kwalitatief zijn.

De kern van iedere zelfevaluatie – ongeacht of dit betrekking heeft op een onderzoeksgroep, een programma, een onderzoeksinstituut of een faculteit – wordt gevormd door wat in deze handleiding, aansluitend bij de terminologie in het Angelsaksisch taalgebied, wordt aangeduid als een 'narratieven': overkoepelende, beargumenteerde en onderbouwde beschrijvingen van de onderzoekseenheid en haar missie en werkterrein.

Deze benadering gaat dus een stap verder dan het SEP formeel voorschrijft. In het SEP worden narratieven alleen genoemd in relatie tot de evaluatie van de maatschappelijke relevantie. Door de idee van narratieven breder in te zetten, wordt benadrukt dat het in de geesteswetenschappen juist gaat om de samenhang tussen wetenschappelijke en maatschappelijke doelen belangrijk.

Narratieven zoals bedoeld in deze handleiding laten zien welke ambities een onderzoekseenheid in den brede heeft, hoe deze zich wil profileren in termen van productievormen, publiek, taal en valorisatie, en in hoeverre de eenheid daarin is geslaagd in termen van kwaliteit en relevantie. Deze narratieven dienen te worden ondersteund door robuuste data, geordend volgens de SEP indicator-categorieën, en, indien gewenst, door casestudies. Als geheel biedt dit materiaal de basis voor de *Strengths, Weaknesses, Opportunities en Threats* (SWOT)-analyse die conform het SEP deel uitmaakt van het zelfevaluatierapport en kan fungeren als vertrekpunt voor een beleid van sturen op kwaliteit en relevantie. Met de keuze om zelfevaluaties in de geesteswetenschappen primair de vorm te geven van narratieven, wordt invulling gegeven aan de ruimte die het SEP biedt.

Narratieven zijn beargumenteerde, explicatieve verhalen. Deze vergen een onderbouwing met robuuste data, die dienen als evidentie – data die betrekking hebben op de verschillende categorieën van indicatoren die het SEP protocol benoemt. Narratieven raken daarmee aan elk van de zes dimensies van kwaliteit en relevantie van onderzoek in de overzichtstabel hierboven: de kwaliteit van het onderzoek (kolom 1) en de maatschappelijke relevantie daarvan (kolom 2), elk gezien vanuit drie invalshoeken: de feitelijke resultaten of *output*, het gebruik daarvan en de erkenning daarvan (rij 1, 2 en 3). Narratieven kunnen ook mede worden onderbouwd door casestudies, die op hun beurt uiteraard worden gestut door robuuste data.

De indicatoren voor de bepaling van kwaliteit en relevantie, zoals voorgesteld in de zes cellen van SEP tabel D1 – met lijstjes van voorbeelden die, zoals het SEP stelt, kunnen worden uitgebreid – zijn deels veldafhankelijk. In deze handleiding wordt aangegeven hoe uiteenlopende indicatoren kunnen worden ingezet voor de beoordeling van het onderzoek in het domein van de geesteswetenschappen. Op de website *Quality and Relevance in the Humanities* (www.qrih.nl) worden de indicatoren concreet uitgewerkt (zie paragrafen 4 en 5 van deze handleiding).

Zelfevaluatierapporten in de geesteswetenschappen bestaan aldus uit de volgende onderdelen:

1. Inleiding met korte beschrijving, profielschets en ambitie
2. Keuze van relevante indicatoren
3. Wetenschappelijke en maatschappelijke resultaten
4. Eigen conclusies van de zelfevaluatie

Deze eerste vier onderdelen hebben het karakter van een samenhangend narratief: hierin worden karakter, profiel en ambitie van te beoordelen eenheid beschreven, tegen welke achtergrond vervolgens wordt ingegaan op de wetenschappelijke en maatschappelijke kwaliteit en relevantie van de onderzoeksresultaten. Daarmee vormen deze eerste vier paragrafen het hart van het zelfevaluatie-rapport.

De opbouw van de rest van het zelfevaluatie-rapport volgt het SEP-format, met de kanttekening dat in ieder geval ook de onderdelen 6 en 9 hoofdzakelijk de vorm van een narratief krijgen.

5. (Bestuurlijke) context
6. Uitkomsten van vorige evaluaties, SWOT analyse en toekomstplannen
7. PhD programma
8. Diversiteit
9. Integriteit, ethiek, onderzoeksdata management
10. Robuuste data, waaronder
 - Case studies
 - tabellen en appendices

De onderdelen 1 tot 9 van het zelfevaluatie-rapport omvatten in totaal 15 pagina's; dit is dus exclusief casestudies, tabellen en andere robuuste data.²

Op de QRiH-website is de structuur van dit rapport uitgewerkt in een *handzaam format*, inclusief een indicatie voor de omvang van de verschillende onderdelen en een output tabel. Ook vindt men hier de contouren van een zelfevaluatie-rapport en voorbeelden van casestudies.

Met deze opzet van zelfevaluaties in de geesteswetenschappen, waarbij wetenschappelijke en maatschappelijke aspecten worden gecombineerd, wordt een specifieke invulling gegeven van het SEP. De reden om hiervoor te kiezen, is om een grotere samenhang te creëren en de dynamiek binnen de onderzoeksprogramma's te laten zien.

² Voor een uitgebreide omschrijving van de opbouw van zelfevaluatie-rapporten, zie het SEP, pagina 23 en 24.

2. Aggregatieniveaus

Zelfevaluaties, opgezet volgens de in deze handleiding geschetste methodiek, kunnen in beginsel op verschillende aggregatieniveaus – faculteiten, instituten, programma's en groepen – worden uitgevoerd. Het is evenwel belangrijk te beseffen dat het karakter van de narratieven en het gebruik van indicatoren in belangrijke mate afhankelijk zijn van het niveau waarop evaluaties worden uitgevoerd.

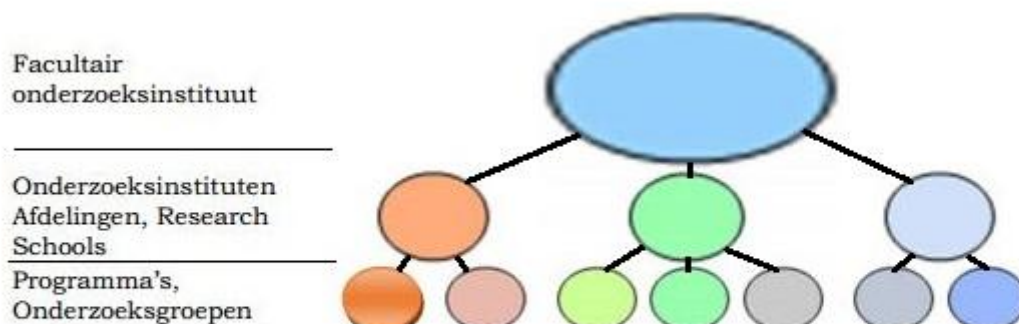
In de handleiding van het SEP (pagina 10) wordt gesteld dat het aan het verantwoordelijk bestuur is te beslissen op welk aggregatieniveau de evaluatie plaatsvindt: 'the assessment will concern a research group, a research institute, a research cluster or the research carried out within a faculty'. Vereiste daarbij is dat de eenheden voldoende samenhang en volume moeten hebben.

De eis van samenhang is vanzelfsprekend: evalueren op het niveau van een disciplinair homogene onderzoeksgroep of –instituut creëert de mogelijkheid een scherp en eenduidig narratief op te stellen, met robuuste data gebaseerd op veldafhankelijke indicatoren. Of deze eenheden in de geesteswetenschappen, met haar betrekkelijk grote disciplinaire verscheidenheid, steeds de door het SEP genoemde omvang van tenminste 10 fte aan vaste onderzoeksformatie zullen hebben, is de vraag; om een hoogwaardige en toegesneden zelfevaluatie mogelijk te maken, zal deze randvoorwaarde, met het oog op de noodzakelijke homogeniteit, met soepelheid gehanteerd moeten worden.

Tegelijk is het goed mogelijk de hier beschreven methodiek te gebruiken voor evaluaties op een hoger aggregatieniveau, waarbij resultaten van onderzoek uit meerdere vakgebieden met verschillende publicatieculturen worden geëvalueerd. Dat laatste vraagt om een gedifferentieerder narratief dan in het geval van een homogene onderzoekseenheid, omdat alleen zo recht kan worden gedaan aan de diversiteit van de onderscheiden vakgebieden.

Die eis van differentiatie geldt te meer wanneer sprake is van zelfevaluaties van zeer grote en gevarieerde eenheden – zoals faculteitsbrede onderzoeksinstituten. Om te voorkomen dat dergelijke zelfevaluaties zich beperken tot het gebruik van een set van enkele algemene, gedeelde indicatoren voor kwaliteit en relevantie – met voorbijgaan aan het principe van diversiteit – verdient het aanbeveling zelfevaluatie-rapporten op een dergelijk hoog aggregatieniveau op te zetten aan de hand van een 'overkoepelend narratief'. Zo'n overkoepelend narratief zal noodzakelijkerwijze een synthetiserend en dus abstracter karakter krijgen,

waarbij de zelfevaluaties van de onderliggende, meer homogene eenheden fungeren als de robuuste data vormen.



Figuur 1 Gelaagdheid zelfevaluaties

3. Narratieven als basiselementen van het zelfevaluatie rapport

Het SEP biedt een overkoepelend model waarbinnen – afhankelijk van de keuzes van onderzoekers en instellingen, maar ook van de maatschappelijke context – diverse legitieme onderzoeksprofielen mogelijk zijn. Dit profiel wordt ‘neergezet’ in een narratieve vorm, waarmee ruimte wordt gecreëerd voor een grote variatie.

Zoals in de eerste paragraaf uiteen is gezet, spelen narratieven een cruciale rol in de vormgeving van de aldaar genoemde onderdelen van een zelfevaluatie rapport. Het instituut, de groep of het programma geeft in een narratief aan wat de kern van het onderzoek is, hoe men zich positioneert en welke strategie wordt gevolgd om de ambities te realiseren en de resultaten van het onderzoek te delen met de wetenschappelijke wereld en de samenleving, en wat het succes daarvan is.

De narratieve delen van de zelfevaluatie bieden volop gelegenheid om verder te komen dan opsommingen van de gerealiseerde uitkomsten van projecten en programma's, zoals publicaties en andere uitkomsten van onderzoek; dergelijke kwantitatieve overzichten komen verderop in het zelfevaluatie rapport aan de orde (deel 9). In de narratieve delen gaat om de onderzoeksrichting en ambitie en om de ‘processen’ die daarmee verbonden zijn en daaraan voorafgaan: zo vergt de ambitie om bij te dragen aan specialistische kennis door artikelen in internationale tijdschriften andere inspanningen en ook een andere tijdsboog dan het ontwikkelen en onderhouden van databases, het maken van tentoonstellingen en catalogi, of

het schrijven van monografieën die zowel gericht zijn op vakgenoten als toegankelijk zijn voor niet-academisch belangstellenden (zogenaamde hybride publicaties). Ook kan in het narratief worden ingegaan op de impact van vroeger onderzoek over een langere termijn.

Narratieven bieden dus ruimte voor een beschrijving van de samenhang van ambitie en de diverse inspanningen die in de loop van het onderzoek gedaan zijn bij het streven naar wetenschappelijke kwaliteit en maatschappelijke relevantie, en het proces waarin deze samenhang tot stand komt.

Narratieven vormen daarmee de basiselementen van zelfevaluatierapporten in de geesteswetenschappen.

4. Het instrumentarium: QRiH

QRiH is de naam van het instrumentarium op de website: *Quality and Relevance in the Humanities* (www.qrih.nl). De website levert de instrumenten om zelfevaluatierapporten in de geesteswetenschappen op te stellen, volgens de methodiek die in de voorgaande paragrafen is beschreven. De website bevat profielen van de onderscheiden publicatieculturen, definities van indicatoren en hun toepassingsmogelijkheden, voorbeelden van toepassingen, lijsten van door panels geautoriseerde indicatoren, voorbeelden van narratieven en casestudies, naast achtergrondinformatie over het SEP en het ontwikkelde QRiH-instrumentarium.

De vraag waarvoor iedere onderzoekseenheid zich geplaatst zal zien, is welke categorieën van indicatoren zij zelf het meest geschikt acht om haar werk te representeren en de aspecten van kwaliteit en relevantie daarvan in kaart te brengen. De lijsten met categorieën op de website QRiH zijn dus, evenmin als de voorbeelden die in het SEP zelf worden genoemd, bedoeld als ‘afvinklijstjes’. Integendeel, niet de lijst van indicatoren in de cellen van het SEP dienen richting te geven aan de zelfevaluatie, maar het narratief.

Welke categorieën van indicatoren voor een specifiek (sub)domein vooral relevant zijn, wordt in sterke mate bepaald door de heersende onderzoeks- en publicatiecultuur van het betreffende onderzoeksgebied of specialisme. Daarbij tekenen zich niet alleen grote verschillen af ten opzichte van de publicatieculturen in bijvoorbeeld het bèta en medische domein, maar ook tussen onderzoeksgebieden en domeinen in de geesteswetenschappen zelf. Deze verschillen hebben met name betrekking op de communicatiekanalen, in termen van de aard, de vorm en de taal van de publicatie en het beoogde publiek en het feitelijk gebruik van deze producten.

Dit laat zich illustreren aan de hand van een paar voorbeelden:

- in een groot deel van de geesteswetenschappen zijn monografieën en *edited volumes* de belangrijkste publicatiekanalen, ook als het gaat om de impact op het eigen vakgebied;
- in een groot deel van de geesteswetenschappen wordt ‘hybride output’, bijvoorbeeld wetenschappelijke publicaties, maar ook tentoonstellingen of rapporten, die zich tegelijkertijd richten op onderzoekers, opdrachtgevers en een breder publiek, als een zeer belangrijk communicatiekanaal beschouwd, juist vanwege de nagestreefde verbinding tussen de wetenschap en de ‘levende cultuur’;
- publicaties in andere talen dan het Engels kunnen in verschillende vakgebieden van grote waarde zijn, zowel waar het gaat om kwaliteit als impact.

Door de beschrijvingen van verschillende categorieën indicatoren, hun toepasbaarheid en technische aspecten, biedt het instrumentarium QRiH de opstellers van zelfevaluatierapporten concrete handvatten om deze vorm te geven op een wijze die niet alleen voldoet aan deze handleiding, maar ook recht doet aan de geïnternaliseerde waarden van kwaliteit en relevantie in de onderscheiden vakgebieden.

5. Aspecten van kwaliteit en relevantie

Het SEP biedt vakgebieden de gelegenheid om instrumenten te (laten) ontwikkelen en het gebruik van robuuste indicatoren van kwaliteit en relevantie te bevorderen. De hieronder uitgewerkte aspecten van kwaliteit en relevantie (met de bijbehorende uitwerkingen op de website) vormen een samenhangend instrumentarium ten behoeve van het opstellen van zelfevaluatierapporten in het kader van externe onderzoeksevaluaties in de geesteswetenschappen volgens het SEP.

Uitgangspunt is dat in een zelfevaluatierapport alle aspecten van onderzoek die de kwaliteit en relevantie van het onderzoek kunnen aantonen in samenhang beschreven moeten worden. Om die reden hebben de instrumenten voor de beoordeling van kwaliteit niet alleen betrekking op directe resultaten of uitkomsten, maar ook op de processen die daaraan noodzakelijkerwijs voorafgaan of op andere wijze intrinsiek met de uitkomsten verbonden zijn.

Indicatoren voor kwaliteit en relevantie

De gegevens met betrekking tot de verschillende kwaliteitsaspecten dienen gezamenlijk ter onderbouwing van de narratieven. De bijbehorende

indicatoren vormen als het ware een *toolbox* of gereedschapskist waaruit een keuze kan worden gemaakt voor de zelfevaluatie-rapporten. Deze indicatoren zijn systematisch uitgewerkt en gedefinieerd op de QRiH website, met daarbij een uitleg van het mogelijke gebruik, de reikwijdte en, waar noodzakelijk, de technische aspecten. De validiteit en bruikbaarheid van de indicatoren vergen, in het licht van de dynamiek van de publicatie- en onderzoekscultuur, voortdurende toetsing respectievelijk aanpassing.

Uitgangspunt is dat juist die indicatoren worden gebruikt die het beste passen bij het werk dat de onderzoekseenheid doet. Dat neemt niet weg dat voor een aantal indicatoren geldt dat die zo kenmerkend zijn voor wetenschappelijk onderzoek (bijvoorbeeld publicaties in wetenschappelijke tijdschriften), dat die voor alle vakgebieden gelden. Daarnaast zijn er gebiedsspecifieke indicatoren.

Het is aan de onderzoekseenheden of instellingen zelf om te besluiten welke indicatoren zij voor hun zelfevaluatie willen gebruiken. De keuze voor bepaalde indicatoren zou logisch moeten volgen uit de heersende onderzoeksculturen in de onderscheiden vakgebieden. Beschrijvingen van die verschillende onderzoeksculturen zijn eveneens te vinden op de QRiH website.

Geautoriseerde en beredeneerde indicatoren

Omdat verschillende categorieën van indicatoren van kwaliteit en relevantie voor de geesteswetenschappen minder eenduidig en ook minder uitgewerkt zijn dan voor sommige andere wetenschapsgebieden – met als gevolg dat zij vaak worden afgerekend naar maatstaven die niet of nauwelijks op hen van toepassing zijn – worden ter bepaling van de robuustheid van indicatoren twee vormen van toetsing gehanteerd, namelijk ‘geautoriseerd’ of ‘beredeneerd’. Beide categorieën kunnen in elk van de zes cellen in Tabel D1 voorkomen.

Voor wat betreft geautoriseerde indicatoren gaat het om instrumenten die zijn opgesteld door representatieve panels van landelijke onderzoekscholen in de geesteswetenschappen en vervolgens zijn gevalideerd door het Landelijk Autorisatie Panel (LAP), een gezaghebbend en breed samengesteld landelijk panel, zoals het geval is bij lijsten van toptijdschriften en uitgeverijen in een bepaald vakgebied. Waar mogelijk zullen de panels hierin aansluiten bij internationale standaarden, zoals bijvoorbeeld de ERIH Plus lijsten of de

Vlaamse (VABB) of Noorse lijsten (Cristin).³ Geregelde toetsing en bijstelling van de geautoriseerde indicatoren zijn, zoals gesteld, noodzakelijk.

In het geval van *beredeneerde* indicatoren gaat het om indicatoren die een deugdelijke argumentatie vereisen omtrent hun gebruik om de kwaliteit te kunnen toetsen. Onder deze categorieën van indicatoren vallen bijvoorbeeld bijzondere publicaties, publicaties in andere talen dan Engels of Nederlands, tentoonstellingen, samenwerking met maatschappelijke partners enzovoort. Overigens kunnen ook deze beredeneerde indicatoren kwantitatief van aard zijn, zoals uit hun beschrijving op de website zal blijken. De onderzoekseenheid zal zelf de argumentatie voor het gebruik van deze indicatoren moeten opstellen, waarbij in een aantal gevallen een beroep kan worden gedaan op elders ontwikkelde standaarden, zoals de Vlaamse en Noorse lijsten van gekwalificeerde tijdschriften.

Indicatoren en interdisciplinariteit

Interdisciplinair onderzoek vergt speciale aandacht bij de beoordeling. In een review artikel over deze kwestie wordt samenvattend geconstateerd:

Existing output measures alone cannot adequately capture this process. Among the quantitative measures considered, bibliometrics (co-authorships, co-inventors, collaborations, references, citations and co-citations) are the most developed, but leave considerable gaps in understanding of the social dynamics that lead to knowledge integration.⁴

De meeste geesteswetenschappelijke onderzoeksdomeinen vertonen een grote mate van onderlinge – interdisciplinaire – verwevenheid, ook al is er soms sprake van specialisatie in bijvoorbeeld tijdvakken, talen of geografische gebieden. Dat geldt voor cultuurgeschiedenis, filosofie, archeologie of literatuurwetenschap even goed als voor multidisciplinaire domeinen zoals mediastudies, mediëvistiek, *cultural studies*, religiewetenschap en theologie of gender studies. Ook is er in veel domeinen sprake van een inhoudelijke, methodologische en theoretische verwantschap met sociale wetenschappen zoals antropologie en sociologie. De verbanden met dergelijke domeinen worden actief onderhouden, met als gevolg dat de onderzoekscultuur in het

³ <https://www.ecoom.be/en/vabb>; <http://www.cristin.no/>
<https://dbh.nsd.uib.no/publiseringskanaler/erihplus/> Voor het laatst geraadpleegd op 15 juli 2017. Het CRISTiN wordt ook in andere Scandinavische landen gebruikt,

⁴ Caroline S. Wagner, J. David Roessner, Kamau Bobba, Julie Thompson Klein, Kevin W. Boyack, Joann Keyton, Ismael Rafols, Katy Börner, 'Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature', *Journal of Informetrics* Vol. 5/1 (January 2011) pp 14–26.

ene domein naast specifieke eigenschappen ook sterke overeenkomsten toont met die van andere domeinen. De KNAW waarschuwde er in het rapport *Kwaliteit en Relevantie in de Geesteswetenschappen* (2012, blz. 46) al voor dat speciale aandacht nodig is het gebruik van interne (facultaire) ranglijsten bij interdisciplinair onderzoek:

Daarom zou men bij het opstellen van classificaties een aggregatieniveau moeten kiezen dat de faculteiten overstijgt. Voorgesteld wordt om lijsten die nu al her en der circuleren te bekijken en te bezien wat daarin de overeenkomsten zijn en wat voor effecten die lijsten hebben.

Het is ook daarom van belang dat de panels door de onderzoekscholen worden samengesteld omdat die vaak interdisciplinair georiënteerd zijn. Zij zijn zich er ook van bewust dat *peer review* negatieve consequenties kan hebben voor interdisciplinair onderzoek en zullen er daarom zorg voor dragen dat de panels voldoende breed zijn samengesteld.

6. SWOT-analyse

Het vierde, verplichte onderdeel van het zelfevaluatierapport is de *Strengths, Weaknesses, Opportunities en Threats* (SWOT)-analyse. Deze richt zich op het prospectieve (sturings)perspectief, en is dus primair gericht op beleidsaspecten en managementaspecten door middel van een beschrijving van de sterke en zwakke kanten van de te beoordelen eenheid, zowel voor wat betreft interne als externe factoren. Het SEP geeft in algemene termen aan hoe men dit kan doen (pagina 30).

Dit deel van het zelfevaluatierapport krijgt de vorm van een tabel die de vier elementen van een SWOT bevat, met korte toelichting, waarin expliciet wordt ingegaan op de toekomst en het beleid en de middelen die nodig zijn om het onderzoek kwalitatief te versterken. Het spreekt vanzelf dat de SWOT-analyse voortbouwt op de voorgaande paragrafen van het zelfevaluatierapport.

7. Casestudies

Casestudies, zoals genoemd in het SEP, vormen een facultatief onderdeel van het zelfevaluatierapport. Tegelijk kunnen ze een cruciale rol spelen in de opbouw daarvan. Ze kunnen immers worden gebruikt als illustratie⁵ en

⁵ Er zijn overigens ook andere manieren om ondersteunende, illustratieve elementen in te brengen in de eerdere paragrafen van het zelfevaluatierapport, zoals het werken met *kaderstukjes*: korte casusbeschrijvingen waarin de impact van het onderzoek of van een specifiek werk of project op het wetenschappelijke domein of in de samenleving wordt beschreven.

tegelijk als ondersteuning van het betoog waarmee de onderzoekseenheid zich met name in de eerste vier paragrafen van de zelfevaluatie presenteert. Casestudies hebben een narratieve vorm en kunnen betrekking hebben op bijzondere projecten of programma's van de onderzoekseenheid, maar ook op bepaalde aspecten van de onderzoeksactiviteiten, zoals de interactie tussen onderzoek en samenleving of het promotieonderwijs. Door middel van casestudies is de onderzoekseenheid dus in staat specifieke delen van het onderzoek in beeld brengen, met name waar men dit belangrijk acht voor het beeld dat beoordelaars zich kunnen vormen van het werk van de te beoordelen eenheid.

Casestudies kunnen op diverse aggregatieniveaus worden uitgevoerd: project, programma, eenheid als geheel. Waar het SEP het gebruik van casestudies beperkt tot maatschappelijke impact, wordt er hier vanuit gegaan dat casestudies zowel informatie over de wetenschappelijke als de maatschappelijke doelen en output kunnen bevatten, en bij uitstek geschikt zijn om de samenhang tussen beide aan te geven – een samenhang die juist in de geesteswetenschappen als essentieel wordt beschouwd. Tot hier tekstvak

In het Britse *Research Excellence Framework* (REF)⁶ van 2014 is ervaring opgedaan met zogenaamde *impact case studies*. Die studies richten zich op de maatschappelijke impact en geven een beschrijving van onder andere het project, de deelnemers en hun aandeel, de aard van de impact en de reikwijdte ervan, en de concrete opbrengsten. Voorbeelden zijn te vinden op (<http://impact.ref.ac.uk/CaseStudies/>).

Impact kan vele vormen aannemen. In het REF wordt dit als volgt omschreven:

Impact is an effect on, change or benefit to:

- the activity, attitude, awareness, behaviour, capacity, opportunity, performance, policy, practice, process or understanding
- of an audience, beneficiary, community, constituency, organisation or individuals
- in any geographic location whether locally, regionally, nationally or internationally
- the reduction or prevention of harm, risk, cost or other negative effects.

⁶ Het Research Excellence Framework (REF) 'is the new system for assessing the quality of research in UK higher education institutions. It replaced [in 2014] the Research Assessment Exercise (RAE), last conducted in 2008 [...] The assessment provides accountability for public investment in research and produces evidence of the benefits of this investment'. Zie <http://www.ref.ac.uk/>

- impacts within the higher education sector, including on teaching or students, are included where they extend significantly beyond the submitting institute.

Op de website QRiH worden voorbeelden gegeven.

8. Tot slot: randvoorwaarden, borging en aanbevelingen

Deze handleiding en het daarbij ontwikkelde instrumentarium QRiH stellen onderzoekers in de geesteswetenschappen in staat om de kwaliteit en relevantie van hun onderzoek te beschrijven en te laten beoordelen op een wijze die recht doet aan de waarden en normen in de wetenschapsdomeinen waarin zij werkzaam zijn. Om deze handleiding en het instrumentarium met vrucht te gebruiken, dienen onderzoekseenheden te zorgen voor een deugdelijke registratie van onderzoeksresultaten; zonder deze registratie zal het niet mogelijk zijn om robuuste data te produceren. Opgemerkt kan worden, dat dit instrument breder kan worden toegepast dan alleen voor evaluaties volgens het SEP. De ervaring leert dat deze wijze van rapporteren zich ook uitstekend leent in andere verbanden, bijvoorbeeld van kleinere onderzoeksprojecten en -programma's, onder meer voor NWO of publiek-private organisaties.

Nog belangrijker is, dat dit instrument kan worden ingezet om – al dan niet aanknopend bij de SWOT-analyses – het onderzoeksbeleid op verschillende aggregatieniveaus vorm te geven. Dat zou zich kunnen uitstrekken tot het personeelsbeleid, om, bijvoorbeeld, uiteenlopende doelstellingen in termen van kwaliteit en relevantie op het niveau van onderzoeksgroepen of -teams vorm te geven.

Ten slotte zou het ontwikkelde instrumentarium – in al zijn veelzijdigheid waar het gaat om kwaliteit en relevantie – individuele onderzoekers behulpzaam kunnen zijn zich een beeld te vormen van hun *Umfeld* en daarmee van hun mogelijkheden en loopbaanperspectieven.

Om het instrumentarium van indicatoren voor kwaliteit en relevantie in de geesteswetenschappen voor toekomstige evaluaties te kunnen gebruiken, is het belangrijk dat de landelijke organisaties op het gebied van de geesteswetenschappen, i.c. het disciplineoverleg Letteren en Geesteswetenschappen (DLG), het disciplineoverleg Wijsbegeerte (DWB) en het disciplineoverleg Godgeleerdheid (DGO) en de KNAW, die dit instrument als *Samenwerkingsverband QRiH* hebben ingesteld, alsmede de landelijke onderzoekscholen, zich aan dit stelsel committeren en borg staan voor de instandhouding en verdere ontwikkeling daarvan.

Inmiddels heeft het Samenwerkingsverband besloten tot instelling van een kleine organisatie rond het Landelijk Autorisatiepanel, waarvan de taken en bevoegdheden zijn vastgelegd in een reglement (zie hiervoor de website www.QRiH.nl). Het is aan het LAP te zorgen voor het up-to-date houden en eventueel reviseren van de geautoriseerde indicatoren, zoals de relevante wetenschappelijke tijdschriften en uitgeverijen, maar ook voor de bijstelling en de nadere uitwerking, of het operationeel maken van andere, mogelijk nog niet geautoriseerde indicatoren, zulks in samenwerking met de Domeinpanels en de Werkgroep van deskundigen.
