

# Evaluatie gebruik QRiH

29 maart 2019, versie 1.0



## Inhoudsopgave

1. Inleiding.....	3
2. Methode .....	3
3. Resultaten .....	4
3.1 Narratief.....	4
3.2 Indicatoren.....	4
3.3 Lijsten .....	6
3.4 Format.....	7
3.5 Domeinprofielen .....	7
3.6 Algemeen beeld en bekendheid QRiH .....	8
3.7 Overige opmerkingen en suggesties.....	8
4. Conclusies en discussie .....	10
4.1 Algemeen beeld en bekendheid QRiH .....	10
4.2 Narratief.....	10
4.3 Indicatoren.....	10
4.4 Lijsten .....	11
4.5 Format.....	11
4.6 Domeinprofielen .....	12
4.7 Overige opmerkingen en suggesties.....	12
Bijlage 1: lijst van onderzoekseenheden die een bijdrage hebben geleverd aan de evaluatie .....	13

## 1. Inleiding

In het rapport *Duurzame Geesteswetenschappen* (2009), werd geconstateerd dat de geesteswetenschappen, als het gaat om de waardering van hun onderzoek, te zeer zijn overgeleverd aan modellen die ontleend zijn aan de exacte en levenswetenschappen. Het rapport kwam met de aanbeveling om eigen standaarden voor de geesteswetenschappen te ontwikkelen. In 2011 kwam de KNAW met het rapport *Kwaliteitsindicatoren voor onderzoek in de Geesteswetenschappen*, in 2012 gevolgd door *Kwaliteit en relevantie in de geesteswetenschappen. Naar een adequaat systeem voor de beoordeling van wetenschappelijk onderzoek*. Uit deze adviezen van de KNAW bleek dat er nog een aantal zaken waren die verdere uitwerking behoeften alvorens te kunnen spreken van een volledig uitgebalanceerd systeem. Om die reden besloot het landelijke overleg van decanen van de geesteswetenschappelijke faculteiten (DLG), met instemming van het Regieorgaan Duurzame Geesteswetenschappen, eind 2012 in te stemmen met het voorstel om te komen tot een operationeel model voor Kwaliteitsindicatoren voor onderzoek in de Geesteswetenschappen. Daarbij hebben zich vervolgens ook het landelijk Disciplineoverleg Wijsbegeerte (DWB), het Disciplineoverlegorgaan Godgeleerdheid (DGO) en, in een later stadium, de KNAW aangesloten. Het project Kwaliteitsindicatoren in de Geesteswetenschappen (KIG) ontwikkelde, onder leiding van een wetenschappelijke adviesraad<sup>1</sup>, in de periode 2013-2017 het instrument Quality and Relevance in the Humanities (QRiH). De drie discipline overleggen en KNAW besloten tot oprichting van het Samenwerkingsverband QRiH, welke een overkoepelend Landelijk Autorisatiepanel benoemden, belast met de inhoudelijke zeggenschap over het instrument. Het constituerende beraad van het Landelijk Autorisatie Panel (LAP) vond plaats op 5 september 2017. Zij hebben het instrument QRiH gevalideerd. In het najaar 2017 is de tweetalige website [www.qrih.nl](http://www.qrih.nl) gelanceerd, wat inhield dat het instrument gebruikt kon worden voor de onderzoeksevaluaties in 2018.

QRiH kan worden gezien als een verdere uitwerking/verbijzondering van het SEP, met een vast format en handleiding. De kern van QRiH is dat niet alleen de evaluatie van 'societal impact' maar de hele zelfevaluatie is opgezet als een narratief: een narratief waarin de verschillende aspecten van kwaliteit en relevantie in onderlinge samenhang worden behandeld, met als uitgangspunt de missie en ambities van de onderzoekseenheid, en een geëigende set van kwantitatieve en kwalitatieve indicatoren die de robuuste data ter ondersteuning van het betoog leveren.

In 2018 zijn veel geesteswetenschappelijke onderzoekseenheden geëvalueerd. Om te onderzoeken of en hoe QRiH gebruikt is bij de evaluaties heeft het LAP besloten een evaluatie te doen naar het gebruik van QRiH. In dit document wordt eerst de methode (hoofdstuk 2) en vervolgens de resultaten (hoofdstuk 3) en de conclusies (hoofdstuk 4) van de evaluatie beschreven.

## 2. Methode

In de LAP-vergadering van 7 juni 2018 is een concept evaluatieformulier besproken en becommentarieerd. Vanaf de zomer 2018 tot januari 2019 is dit evaluatieformulier opgestuurd naar directeuren en beleidsmedewerkers van geesteswetenschappelijke onderzoekseenheden die recent hun zelfevaluatie geschreven hadden. Er zijn in totaal 21 reacties ontvangen. Sommige evaluatieformulieren zijn minimaal ingevuld, anderen hebben een uitgebreide reactie gegeven. Om een inhoudelijke verdieping te kunnen geven aan de reacties in de evaluatieformulieren is er daarnaast een zevental interviews gehouden. Er zijn vooral interviews gehouden op locaties waar vanuit de vragenlijsten nog geen goed beeld gekregen was. Daarnaast zijn er 16 zelfevaluatierapporten verzameld om te bestuderen of en hoe QRiH gebruikt is. In twaalf van de 16 zelfevaluatierapporten is QRiH gebruikt. In de bijlage is een lijst opgenomen van onderzoekseenheden die een ingevulde vragenlijst opgestuurd hebben en waar een interview gehouden is.

---

<sup>1</sup> Bestaande uit prof. dr Remieg Aerts (RU, UvA), em. prof. dr Geert Booij (Leiden), prof. dr Martha Frederiks (UU), prof. dr Mary Kemperink (RUG), prof. dr Anthonie Meijers, prof. dr Frank van Vree, voorzitter (UvA, NIOD).

### 3. Resultaten

In dit hoofdstuk worden de resultaten van de evaluatie beschreven. Allereerst wordt ingegaan op een aantal specifieke onderdelen van QRiH. Afgesloten wordt met een algemeen beeld en overige opmerkingen.

#### 3.1 Narratief

Het narratief lijkt één van de onderdelen van QRiH waar mensen het meest tevreden over zijn. 15 respondenten hebben in de vragenlijst aangegeven de idee van een narratief behulpzaam te vinden. Twee respondenten geven aan dat dit niet behulpzaam was, de ene respondent omdat ze zelf al tot de conclusie gekomen waren dat een narratief een goede vorm zou zijn en de andere respondent omdat ze het SEP hebben gevolgd en niet QRiH. Ook de meeste geïnterviewden hebben een narratief in hun zelfevaluatie(s) gebruikt. Een aantal van hen geven aan dat het narratief (eventueel in combinatie met de indicatoren) goed geholpen heeft bij het vinden van de identiteit van de eenheid (wie zijn we?, wie willen we zijn?, hoe verschillen we van andere eenheden?).

De respondenten van de vragenlijst geven een aantal verbeterpunten. Er zijn verschillende opmerkingen over het eerste deel van het format: dit deel van het format geeft een belemmering om het narratief op een vloeiende manier te beschrijven en de eerste drie inleidende elementen zijn niet altijd goed van elkaar te scheiden. Dit geven verschillende respondenten aan. Het tweede punt is dat er te weinig woorden in het format gegeven worden om het narratief goed te beschrijven. Dit wordt ook door verschillende respondenten aangegeven. De respondenten zijn zich ervan bewust dat dit ook samenhangt met de restricties van het SEP.

Niet iedereen heeft het narratief op dezelfde manier gebruikt. Eén geïnterviewde geeft aan dat het in het begin onduidelijk was hoe het narratief gebruikt moest worden. Later begreep deze persoon dat de bedoeling van QRiH is om de hele zelfevaluatie als narratief te schrijven. Iemand anders geeft aan het narratief met name gebruikt te hebben voor de beschrijving van de maatschappelijke impact en weer iemand anders voor het beschrijven van case studies. Ook uit de bestudeerde zelfevaluaties waarin het QRiH-format gevolgd is, wordt duidelijk dat het schrijven van bijvoorbeeld de inleiding (met daarin een korte beschrijving, profielschets en ambitie van de onderzoekseenheid) in narratieve vorm niet voor iedereen duidelijk was danwel niet iedereen even goed gelukt is. Een aantal respondenten stellen voor om voorbeelden van narratieven op te nemen. Ook in de interviews wordt aangegeven dat er behoefte is aan voorbeelden van zelfevaluaties.

Eén respondent had een meer fundamentele kanttekening bij het narratief: "Bij zelfevaluaties is de neiging groot om sterke punten te benadrukken en zwakke punten niet al te negatief te verwoorden. Het narratief kan die tendensen in potentie nog eens versterken, wat het moeilijk maakt voor de commissie om heel helder te zien hoe sterk of hoe zwak die punten zijn. Tegelijk denk ik wel dat voor ons vakgebied in elk geval het narratief een uitkomst is omdat je beter kunt uitleggen wat je als groep belangrijk vindt. Een ronde waarin ervaringen verzameld worden kan geen kwaad, denk ik, om zo tot een scherper omschrijving te komen van hoe je een narratief aanpakt".

#### 3.2 Indicatoren

Veertien respondenten van de vragenlijst geven aan de set indicatoren behulpzaam te vinden, twee respondenten geven aan het niet behulpzaam te vinden en één respondent weet het niet. Als toelichting geven verschillende respondenten die de indicatoren behulpzaam vonden aan, dat QRiH vele mogelijke soorten indicatoren noemt en dat het goed is dat er vrijheid is om indicatoren te kiezen. Eén iemand geeft aan dat niet iedereen de vrijheid voelde om de indicatoren voor maatschappelijke impact niet te gebruiken (zij dachten dat het verplicht was om zowel indicatoren voor wetenschappelijke kwaliteit en maatschappelijke impact te gebruiken). Twee personen geven aan dat het aan te bevelen is c.q. dat het nuttig was om bij de start van iedere zelfevaluatie na te denken over de indicatoren ("waar wil je op afgerekend worden" of "dit is wat wij willen"). Een respondent die de

indicatoren behulpzaam vond, geeft aan dat de link tussen indicatoren en het narratief niet altijd duidelijk gelegd werd in de zelfstudies, waarschijnlijk omdat de indicatoren voor het narratief als te benauwend ervaren werden. Eén persoon gaf aan dat het belangrijk is de indicatoren regelmatig te actualiseren, maar dat er wel opgelet moet worden voor mogelijk (onbedoelde) perverse prikkels.

Uit de reacties van de respondenten blijkt dat het onderscheid tussen 'geautoriseerde' en 'beredeneerde' indicator niet voor iedereen duidelijk is. Eén respondent stelt voor dit onderscheid te schrappen. Daar tegenover staat dat in één van de interviews aangegeven is, dat het onderscheid juist goed is, omdat je zo "langs je eigen liniaal gelegd wordt". Eén iemand geeft aan dat er meer maatschappelijke indicatoren alleen maar 'beredeneerd' zijn en niet 'geautoriseerd'. Twee mensen geven aan behoefte te hebben aan een set geautoriseerde maatschappelijke indicatoren.

Uit de bestudeerde zelfevaluaties blijkt dat bijna alle onderzoekseenheden ervoor gekozen hebben om indicatoren op te nemen in alle zes de SEP-categorieën. In één zelfevaluatie wordt aangegeven niet voor alle zes de categorieën indicatoren gekozen te hebben, omdat zij geen betrouwbare data konden leveren voor alle categorieën. Verder zien we bij veel zelfevaluaties die QRiH gebruikt hebben dat ook aangegeven wordt waarom er voor bepaalde indicatoren gekozen is, maar dat dit zeker niet bij alle zelfevaluaties gebeurt waar QRiH gebruikt is. Ook viel op dat velen gekozen hebben voor kwantitatieve maatschappelijke indicatoren, maar dat een aantal (ook) gekozen hebben voor meer kwalitatieve indicatoren, zoals case studies. Tot slot komt in vijf zelfevaluaties het onderscheid tussen geautoriseerde en beredeneerde indicatoren terug.

Verschillende respondenten vragen om voorbeelden van indicatoren van anderen op de website, omdat dit mensen op ideeën kan brengen om sommige dingen systematisch te verzamelen of waarop een eenheid medewerkers kan stimuleren. Sommigen zien de indicatoren op de website vooral als aanzet om indicatoren voor het eigen domein te ontwikkelen.

In de vragenlijsten en bij de interviews wordt regelmatig ingegaan op de praktische problemen bij het verzamelen van gegevens voor indicatoren. Een aantal mensen geeft aan dat een aantal gegevens (op het gebied van maatschappelijke impact) binnen de onderzoekseenheid nog onvoldoende of niet structureel bijgehouden wordt. Dit komt ook in een aantal zelfevaluaties terug. Er is veel vraag naar praktische tips, instructies of simpele tools om gegevens te verzamelen, beheren en analyseren ter onderbouwing van indicatoren van maatschappelijke impact. Als voorbeeld wordt LexisNexis genoemd voor het doorzoeken van kranten. Er worden verschillende problemen met PURE (en andere Current Research Information Systemen) genoemd. Eén probleem is dat twee publicaties in één journal op twee verschillende manieren in PURE ingevoerd worden, bijvoorbeeld één als wetenschappelijk publicatie en één als vakpublicatie. Dit vergt interne afspraken hierover of controle hierop. Een ander probleem is dat PURE vooral geschikt is om output en activiteiten van wetenschappers te registreren en geen ruimte biedt om indicatoren in de categorieën "gebruik" en "erkenning" op te nemen. Tot slot passen hybride publicaties niet in PURE.

Eén iemand geeft aan dat het aantal indicatoren beperkter mag. Anderen geven juist suggesties om nieuwe indicatoren om op te nemen:

- een indicator die aangeeft hoe vaak wetenschappers gevraagd worden als copromotor of tweede promotor of in de commissie voor de beoordeling van een dissertatie bij een andere universiteit.
- een indicator over onderwijs, bijvoorbeeld het geven van onderwijs bij een andere faculteit/instelling.

De respondenten die de set indicatoren niet behulpzaam vonden geven de volgende dingen aan. Eén persoon geeft aan dat de set afwijkt van de internationale standaarden (wel/niet peer reviewed) wat verwarrend zou kunnen werken voor een visitatiecommissie. De andere persoon is van mening dat er te veel verschillende typen indicatoren zijn: het verschil tussen beredeneerde en geautoriseerde indicatoren zou geschrapt moeten worden en de verschillende labels zoals "products" en "use" zouden niet geschikt zijn om activiteiten van de geesteswetenschappen te beschrijven.

### 3.3 Lijsten

Vanuit de vragenlijsten blijkt dat vier mensen de lijsten met tijdschriften en uitgeverijen behulpzaam vonden, elf mensen vonden de lijsten niet behulpzaam en twee mensen weten het niet. Alle geïnterviewden geven aan dat de lijsten niet gebruikt zijn. Er blijkt een praktisch probleem: het is voor veel mensen niet duidelijk hoe je in technische zin de lijsten vergelijkt met de informatie uit PURE of een ander Current Research Information System (CRIS). Dat wordt nu vaak met de hand gedaan.

Van groter belang is de discussie over de inhoud van de lijsten. Hieronder de belangrijkste punten:

- Eenheden zoals faculteiten beslaan vaak meerdere domeinen, terwijl de lijsten zijn ingedeeld per domein. Dit werkt verwarrend.
- Er is in de lijsten geen onderscheid tussen peer-reviewed en niet peer-reviewed tijdschriften, waardoor mensen alsnog de neiging hebben op deze tweedeling terug te vallen. Aangegeven wordt dat deze indeling internationaal erkend is en breed geaccepteerd en volgens sommigen begrijpelijker voor de internationale evaluatiecommissie. De meerwaarde van de lijsten (ten opzichte van wel/niet peer-reviewed) is niet voor iedereen duidelijk.
- Niet iedereen is het eens over de inhoud van de lijsten. Aangegeven wordt dat men het binnen sommige vakgebieden vaak onderling oneens is over wat de betere tijdschriften of uitgeverijen zijn. Anderen geven aan dat instituten zelf wel weten welke lijsten waardevol zijn. Sommigen geven aan dat er te weinig rekening is gehouden met de subdisciplines en/of dat niet alle relevante tijdschriften in de lijst staan. In één van de interviews werd genoemd dat de evaluatiecommissie commentaar had op de QRiH-lijst. In het nog vertrouwelijke evaluatierapport komt te staan: "Publications in this group make clear that there is an urgent need to update the list or to elaborate a more diversified QRiH-list. Indeed, it must be said that journals such as xxx are internationally recognized as top journals in the field, and that publications in these journals are the result of serious peer review, but that they are simply absent in the QRiH-list".
- Eén iemand stelt dat voor de vakgebieden waar hij ervaring mee heeft onderling veel onenigheid is over wat de betere tijdschriften of uitgeverijen zijn en dat er goede redenen voor die onenigheid zijn. Hij is van mening dat het beter is om evaluatiecommissies dat oordeel te laten vellen, met dien verstande dat ook commissieleden hierover de waarheid niet in pacht hebben. Commissieleden hebben echter veel ervaring wat wel wat vergoedt volgens deze respondent, zeker als ze hun oordelen weten te motiveren.
- Een geïnterviewde geeft aan dat het onmogelijk is om een goede lijst te maken, omdat het vakgebied zo breed en regionaal verschillend is. Deze persoon ziet meer in een beperkte selectie van (top-)uitgeverijen en geeft aan dat het goed zou zijn om bij het aanpassen van de lijsten de mensen van de bibliotheken betrokken zouden moeten worden.
- Eén van de geïnterviewden geeft aan dat het probleem met de lijsten is dat deze als limitatief gezien worden, terwijl dit niet zo bedoeld is.
- Voor het domein Theologie en Religiestudies is er in 2015 door DGO een lijst opgesteld met daarin een A-B-C-waardering. Voor sommigen is het niet duidelijk hoe de QRiH-lijst (opgesteld door NOSTER) zich verhoudt tot deze lijst. Is de lijst hetzelfde of anders? Daarnaast was er een praktisch punt, daar ze in hun eigen CRIS per journal al de A-B-C-waardering van DGO hadden overgenomen en de QRiH-lijst relatief laat beschikbaar kwam.
- Vanuit verschillende personen vanuit de literatuurwetenschappen wordt aangegeven dat de lijsten niet gebruikt zijn en dat zij gekeken hebben naar publicaties in tijdschriften met een hoge impact factor.
- Er was onduidelijkheid over het proces. De aanpassingsmogelijkheden voor de lijsten waren volgens sommigen beperkter dan ze verwacht hadden. Het zou duidelijker gemaakt moeten worden dat de inventarisatie via de landelijke onderzoekscholen bepalend is en vast staat.
- Eén respondent geeft aan dat de lijsten het instrument nodeloos gecompliceerd en ondoorzichtig maken en zijn ook niet nodig. Het risico hiervan is dat de meerwaarde van de andere onderdelen van de QRiH verdwijnt, omdat mensen door de bomen het bos niet meer zien.
- Tot slot is ook aangegeven dat bij het proces van de totstandkoming van de lijsten ook een zekere politiek speelde waardoor er minder vertrouwen is in de lijsten. Verschillende mensen

geven aan dat de landelijke onderzoekscholen niet het juiste gremium is voor het aanleveren van de lijsten, omdat landelijke onderzoekscholen hun eigen politiek kennen of niet de volle breedte van het veld vertegenwoordigen.

### 3.4 Format

Tien mensen geven aan dat het format behulpzaam is geweest, zes mensen geven aan dat het format niet behulpzaam was en één persoon weet het niet. Verschillende mensen geven aan dat het format heel behulpzaam was bij een landelijke visitatie, om alle zelfstudies beter vergelijkbaar te maken. Dit had ook als voordeel dat er geen consensus bereikt hoefde te worden over het format, wat tijd en discussie scheelde.

De mensen die het format behulpzaam vonden hebben wel een aantal kanttekeningen bij het format. Er bleek veel overlap tussen bepaalde onderdelen van het format onder verschillende punten, waardoor het een puzzel was op alles op de juiste plek te krijgen. Eén persoon geeft aan dat het format niet altijd rekening hield met verschillen in aggregatieniveaus. De betreffende evaluatie had een “instituu(deel)” en een “onderzoeksgroep-deel” en het format was in die zin minder flexibel.

De mensen die het format niet behulpzaam vonden geven ook aan moeite te hebben met het format, vooral met het eerste deel. Zoals in eerdere paragrafen aangegeven gaf het format een belemmering om het narratief op een vloeiende manier te beschrijven en waren er te weinig woorden beschikbaar in het format om het narratief goed te beschrijven. Ook is er onduidelijkheid waarom er halverwege conclusies getrokken dienen te worden volgens het format, terwijl sommigen dit aan het eind pas willen doen. Daarnaast is er het voorstel om de bestuurlijke context vóór de resultaten te beschrijven, omdat de bestuurlijke context belangrijk is voor de behaalde resultaten. Tot slot wordt voorgesteld onderdeel 6 (vorige evaluatie, SWOT en toekomstplannen) te plaatsen ná de onderdelen integrity/diversity/phd programmes, omdat deze integraal onderdeel zijn van toekomstplannen en de SWOT.

Er is één iemand die aangeeft dat de onderzoekseenheid meerdere domeinen dekt en dat daarom het QRiH-format niet behulpzaam is.

### 3.5 Domeinprofielen

Zes mensen geven aan de domeinprofielen behulpzaam te vinden, vijf mensen geven aan het niet behulpzaam te vinden en zes mensen weten het niet. De toelichting bij deze vraag is niet veelvuldig ingevuld. Eén persoon geeft aan dat het goed zou zijn gemengde domeinprofielen te maken, omdat de meeste geesteswetenschappelijke faculteiten meerdere domeinen omvatten. Eén persoon geeft aan dat het behulpzaam is als aanzet voor het ontwikkelen van een eigen profiel. Eén persoon geeft aan dat de bedoeling van de domeinprofielen (zachte benchmark) duidelijker gepresenteerd kan worden, omdat dit in eerste instantie niet duidelijk was en ziet het als goede optie. Deze persoon vindt dit een goede optie, omdat het lastig is een instituut te vinden dat vergelijkbaar is, omdat de bibliografische meetmethoden niet goed werken en omdat de diversiteit groot is. Eén persoon vraagt of het mogelijk is om een apart domeinprofiel te ontwikkelen voor “Heritage and Memory Studies”.

In een aantal zelfevaluaties wordt gerefereerd aan de domeinprofielen. Veelal gebeurt dit aan het begin van de zelfevaluatie, bijvoorbeeld bij de beschrijving van het profiel van de onderzoekseenheid. De domeinprofielen worden gebruikt om de verschillen en overeenkomsten aan te geven. De domeinprofielen lijken echter niet als zachte benchmark gebruikt te zijn; in deze evaluaties staat vaak een apart stuk over benchmarks.

Omdat de toelichting niet veelvuldig ingevuld is, is bij de interviews verder gevraagd naar de ervaringen met de domeinprofielen. Eén persoon geeft aan dat er in het voorwoord van de zelfevaluatie verwezen is naar het domeinprofiel. Omdat ze zich in grote lijnen herkenden in het domeinprofiel is er verder weinig aandacht aan besteed. Eén persoon geeft aan dat de faculteit zich

heel erg herkent in het domeinprofiel en dat er daarom geen noodzaak was om in de evaluatie aandacht te besteden aan het domeinprofiel. Eén persoon geeft aan dat ze het niet gebruikt hebben bij hun zelfevaluaties, omdat de eenheden domeinoverstijgend zijn. Na uitleg over hoe het gebruikt had kunnen worden, geeft de persoon aan meer meerwaarde te zien in een benchmark met een internationaal instituut. Tot slot geeft iemand aan het domeinprofiel niet gebruikt te hebben, omdat ze dit onderdeel niet goed op hun netvlies hadden staan.

### 3.6 Algemeen beeld en bekendheid QRiH

Over het algemeen is er bij de benaderde personen een zeer grote bereidheid ervaren om mee te doen aan de evaluatie. Slechts van enkele onderzoekseenheden is (ook na een eenmalige herhaalde oproep) geen reactie ontvangen, danwel een negatieve reactie ontvangen. Daarnaast is de bekendheid van QRiH onder de bevroegde directeuren en beleidsmedewerkers groot. Van de benaderde personen hebben slechts twee personen aangegeven niet bekend te zijn met QRiH. De ene persoon is een nieuwe beleidsmedewerker, waardoor het niet verwonderlijk is dat deze persoon nog niet bekend is met QRiH. De andere persoon is een directeur die verderop in het evaluatieformulier aangeeft zijdelings van QRiH gebruik te hebben gemaakt, dus blijkbaar toch bekend is met QRiH.

In de reacties lijkt er een verschillend beeld te zijn van QRiH. De meeste personen staan heel positief tegenover QRiH in z'n algemeenheid. Velen geven aan dat ze blij zijn dat dit initiatief voor de geesteswetenschappen ontwikkeld is en waarderen de tijd en moeite die erin gestoken is om het te ontwikkelen. Daarnaast zijn er ook meer neutrale reacties. Zij geven aan dat SEP voor hun prima werkt en dat er daarom geen gebruik is gemaakt van QRiH. Ook zijn er mensen die aangeven dat QRiH onderscheid maakt in domeinen en dat het daardoor niet of minder bruikbaar is voor onderzoekseenheden die multidisciplinair van karakter zijn. Er zijn ook een aantal respondenten die negatief staan tegenover QRiH. Met name vanuit de filosofie zijn negatieve reacties ontvangen. Belangrijkste argumenten zijn dat het SEP goed werkt en dat filosofie een goedlopende onderzoeksschool heeft waar discussies over kwaliteit gevoerd worden, waardoor er geen nieuw orgaan (het LAP) nodig is.

Op de vraag of QRiH gebruikt is bij de zelfevaluatie geven 17 mensen aan QRiH (in meer of mindere mate) gebruikt te hebben. Ook de meeste van de zeven geïnterviewden geven aan QRiH gebruikt te hebben. Vier respondenten geven aan QRiH niet gebruikt te hebben. Daarvan geven twee respondenten geen reden hiervoor op en twee respondenten geven aan dat QRiH nog niet ver genoeg ontwikkeld was ten tijde van het schrijven van de zelfevaluatie. Deze laatste twee respondenten lijken echter niet zo positief over QRiH waardoor het de vraag is of ze QRiH in een toekomstige evaluatie wel zullen gebruiken.

Op de vraag of QRiH behulpzaam is geweest bij de zelfevaluatie antwoorden 5 mensen 'zeer behulpzaam', 12 mensen 'enigszins behulpzaam' en één respondent heeft het antwoord 'niet behulpzaam' gekozen. Deze laatste heeft aangegeven het SEP te hebben gevolgd bij de zelfevaluatie en niet QRiH. Bij de toelichting op de vraag of QRiH behulpzaam is geweest bij de zelfevaluatie geven mensen aan met name het format, het narratief en de indicatoren behulpzaam te vinden. Overigens is er ook commentaar op deze punten, zoals in bovenstaande paragrafen beschreven is.

### 3.7 Overige opmerkingen en suggesties

Veel mensen hebben geen goed beeld van de organisatie rondom QRiH of suggesties voor verbetering hiervoor, anders dan meer/betere communicatie, omdat ze na de lancering van QRiH niets meer vernomen hebben. Eén persoon geeft aan dat de autoriteit van QRiH nog niet voldoende gesanctioneerd is door de faculteiten en besturen en dat hier te winnen valt. Eén persoon geeft aan dat er weinig vertrouwen is (geweest), omdat QRiH vanaf het begin een onduidelijk proces heeft gekend, met name voor de lijsten: elke onderzoeksschool heeft de lijsten op een andere manier aangeleverd.



In één van de interviews bleek dat de evaluatiecommissie QRiH af en toe lastig te gebruiken vond, waardoor de commissie teruggreep op het SEP. De suggestie werd gedaan om vanuit QRiH een aparte tekst of handleiding te maken voor de evaluatiecommissie. In de meeste andere interviews werd dit idee aan de geïnterviewden voorgelegd en werd dit idee gesteund. Een enkeling gaf aan hier geen behoefte aan te hebben, omdat ze de naam QRiH niet in het zelfevaluatie-rapport gebruikt hadden.

Een persoon gaf aan dat een positief punt van QRiH is, dat er een brede consultatie gedaan is toen het instrument opgezet werd. QRiH zou volgens deze persoon verbeterd kunnen worden door ook te focussen op de waardering voor de lange termijn, bijvoorbeeld door te vragen naar de nationale meerwaarde of naar wat de onderzoekseenheid uniek maakt.

Eén persoon geeft aan dat het goed zou zijn om richtlijnen te geven over de (strategische) keuzes die gemaakt worden bij het evaluatieproces. Er hangt vaak erg veel van visitaties af en dat vergroot de druk op commissies en ook op de onderzoeksleiders. Het liefst zou deze persoon het visitatiesysteem als geheel meer als advies dan als beoordeling zien, zodat visitaties minder als bedreigend gezien worden. Twee andere personen geven aan graag meer een gesprek te willen met de visitatiecommissie over het narratief/de zelfevaluatie, willen graag een meer collegiale visitatie en verwijzen daarbij naar Vlaanderen waar dit gebruikelijk zou zijn.

Eén persoon geeft aan behoefte te hebben aan expliciete voorbeelden voor de narratieven over maatschappelijke relevantie. Meerdere personen geven aan dat voorbeelden van zelfevaluaties nuttig zouden zijn.

Eén persoon geeft als suggestie mee om in een zelfevaluatie een bijlage op te nemen waarin de tien 'key publications' per 'principal investigator' opgenomen worden. Dit geeft de evaluatiecommissie snel een goed overzicht van het onderzoek.

Eén persoon noemt dat benchmarks een behulpzaam onderdeel van QRiH is gebleken en één iemand noemt dat case studies goed zijn om "bijzondere aspecten neer te zetten en het instituut 'smoel' te geven".

Eén persoon geeft aan dat SEP te dicht bij de 'afrekencultuur' ligt en minder zou moeten focussen op het aantal publicaties.

Met betrekking tot de website wordt aangegeven dat op de Engelse versie geen correct Engels staat.

## 4. Conclusies en discussie

### 4.1 Algemeen beeld en bekendheid QRiH

Er is een grote bekendheid van QRiH onder de bevroagde directeuren en beleidsmedewerkers. In hoeverre QRiH bekend is bij het bredere veld, is op basis van deze evaluatie niet te bepalen. Uit persoonlijke contacten vernemen we dat in het bredere veld QRiH een stuk minder bekend is. Een grote groep personen staat heel positief tegenover QRiH in z'n algemeenheid. Velen spreken hun waardering uit voor het initiatief. Uit persoonlijke contacten maken we op dat dit met name komt doordat veel geesteswetenschappers het gevoel hebben dat in evaluatiesystemen te veel nadruk ligt op indicatoren en gebruikelijke praktijken uit de medische en natuurwetenschappen. Dat er waardering is voor het instrument blijkt ook uit het feit dat het merendeel van de respondenten en geïnterviewden (22 van de 27) aangeeft QRiH (in meer of mindere mate) gebruikt te hebben bij de zelfevaluatie en ook uit dat het merendeel aangeeft QRiH (enigszins) behulpzaam te vinden.

Ook zijn er meer neutrale en negatieve(re) geluiden over QRiH. De neutrale geluiden komen met name vanuit mensen die aangegeven prima uit de voeten te kunnen met het SEP. De negatieve geluiden komen met name vanuit de filosofie. Vanuit de filosofie lijken de grootste bezwaren te komen doordat ze geen behoefte hebben aan een Landelijk Autorisatiepanel, omdat ze zelf goed georganiseerd zijn.

Meerdere mensen geven aan dat QRiH onderscheid maakt in domeinen en daardoor niet bruikbaar is voor onderzoekseenheden die multidisciplinair van karakter zijn. Blijkbaar is de domeinindeling - die bijvoorbeeld goed zichtbaar is bij de tijdschriftenlijsten en de domeinprofielen - verwarrend. QRiH is in eerste aanleg bedoeld voor de geesteswetenschappen, maar de uitgangspunten zijn evengoed toepasbaar binnen andere domeinen en eenheden met een multidisciplinair karakter. Blijkbaar is dit momenteel nog niet duidelijk genoeg.

### 4.2 Narratief

De kern van QRiH is dat de gehele zelfevaluatie opgezet wordt als een narratief. In het narratief worden de verschillende aspecten van kwaliteit en relevantie in onderlinge samenhang behandeld, met als uitgangspunt de missie en ambities van de onderzoekseenheid.

Bijna alle respondenten geven aan een narratief gebruikt te hebben in hun zelfevaluatie(s) en het narratief lijkt één van de onderdelen van QRiH waar het mensen het meest tevreden over zijn. Velen geven aan het narratief erg behulpzaam te vinden. Een aantal mensen geven aan dat het narratief behulpzaam is bij het vinden van de identiteit van de onderzoekseenheid. Uit de reacties in de vragenlijst wordt wel duidelijk dat niet iedereen de idee om de gehele zelfevaluatie als narratief te beschrijven goed begrepen (c.q. opgevolgd) heeft.

Naast de suggestie om duidelijker aan te geven dat het de bedoeling is om de gehele zelfevaluatie als narratief te beschrijven, zijn er een aantal andere suggesties voor verbetering. De respondenten geven aan dat het format op een aantal punten aangepast zou moeten worden. Daarnaast is er behoefte om ervaringen te delen en om voorbeelden van narratieven te zien. Dit is deels praktisch: zodat duidelijk wordt hoe een narratief het beste opgezet kan worden. Deels is deze behoefte meer algemeen, omdat in een zelfevaluatie de neiging groot kan zijn om sterke punten te benadrukken en zwakke punten niet al te negatief te verwoorden, wat in een narratief versterkt kan worden.

### 4.3 Indicatoren

De bedoeling van QRiH is om in een zelfevaluatie bij het narratief een geëigende set van kwantitatieve en kwalitatieve indicatoren te kiezen die de robuuste data leveren ter ondersteuning van

het betoog. QRiH heeft een brede range aan indicatoren voor productie, gebruik en erkenning. Deze indicatoren zijn de uitkomst van een brede consultatie in het geesteswetenschappelijk veld. De indicatoren zijn onderscheiden naar geautoriseerd (aangedragen door de domeinpanels en geautoriseerd door het LAP) en beredeneerd (te omschrijven en te beargumenteren door de opstellers van de zelfevaluatie).

Zoals QRiH bedoelt, wordt in veel (maar lang niet alle) zelfevaluaties waar QRiH gebruikt is, aangegeven waarom er voor bepaalde indicatoren gekozen is. Dit lijkt dus veelal goed te gaan.

In de vragenlijst geven de meeste mensen aan de set indicatoren behulpzaam te vinden. Wel wordt er heel verschillend tegen de indicatoren aangekeken. Sommigen geven aan dat er te veel indicatoren zijn, anderen doen suggesties om extra indicatoren toe te voegen. Sommigen gebruiken de indicatoren van QRiH, anderen zien de indicatoren vooral als aanzet om indicatoren voor het eigen domein te ontwikkelen.

Op basis van de evaluatie wordt duidelijk dat de meningen over het onderscheid naar geautoriseerde en beredeneerde indicatoren uiteen lopen. In vijf zelfevaluaties komt het onderscheid tussen geautoriseerde en beredeneerde indicatoren terug. Dat het onderscheid niet in meer zelfevaluaties gebruikt wordt, komt mogelijk doordat in de vragenlijst aangegeven wordt, dat het onderscheid niet voor iedereen duidelijk is. Ook zijn er reacties om dit onderscheid te schrappen. Daarnaast geven een aantal mensen aan behoefte te hebben aan geautoriseerde maatschappelijke indicatoren.

Regelmatig worden er praktische problemen genoemd bij het verzamelen van gegevens voor indicatoren. Er is behoefte aan praktische tips, instructies of tools om gegevens te verzamelen, met name voor de indicatoren voor maatschappelijke impact. Daarnaast zijn er een aantal praktische problemen met PURE (en soms andere Current Research Information Systemen). Vanuit QRiH kunnen de problemen met de systemen niet opgelost worden. Wel zou QRiH kunnen overwegen meer aandacht te besteden aan praktische tips of instructies voor het verzamelen van maatschappelijke impact.

#### 4.4 Lijsten

Een onderdeel van de indicatoren vormen de lijsten van vooraanstaande publicatiekanalen (uitgeverijen en tijdschriften). Deze lijsten zijn ingevoerd om iets te kunnen zeggen over de kwaliteit van publicaties. De lijsten zijn opgesteld door de domeinpanels.

De lijsten blijken het onderdeel waar de respondenten het minst tevreden over zijn. Naast praktische problemen hoe de lijsten vergeleken moeten worden met de informatie uit een CRIS, zijn er vele inhoudelijke problemen. Een aantal problemen kan relatief eenvoudig opgelost worden (zoals door het laten vervallen van het onderscheid in domeinen), maar velen zijn fundamenteeler van aard. Er is veel discussie over de inhoud van de lijsten. De commentaren lopen hierbij sterk uiteen. Het is de vraag of het mogelijk is een lijst samen te stellen die erkend wordt als kwalitatief robuuste representatie van het gehele veld. Daarnaast is er kritiek op het proces van de totstandkoming van de lijsten, waardoor er minder vertrouwen is in de lijsten.

Gezien het vele commentaar op de lijsten vraagt dit om nadere analyse en discussie met het veld.

#### 4.5 Format

QRiH heeft een format ontwikkeld voor de zelfevaluaties met daarin een hoofdstuk- en paragrafen-opbouw en een aantal woorden. Het format is bedoeld als handreiking en geen verplichting om alles letterlijk op te volgen.

Het format heeft bij een landelijke evaluatie een goede rol gespeeld. Door verschillende mensen is aangegeven dat het format heel behulpzaam was bij deze landelijke evaluatie, om alle zelfstudies beter vergelijkbaar te maken. Dit had ook als voordeel dat er geen consensus bereikt hoefde te worden over het format, wat tijd en discussie scheelde.

Als we kijken naar de vragenlijst dan zijn de meningen verdeeld over of het format behulpzaam is. Als naar het inhoudelijke commentaar gekeken wordt, dan kan geconcludeerd worden dat door aanpassingen van het format gemakkelijk tegemoet gekomen kan worden aan de commentaren. Belangrijkste commentaar is dat het eerste deel van het format (het narratief en de indicatoren) aangepast dient te worden. Daarnaast zijn er verschillende andere concrete, praktische voorstellen om het format aan te passen. De meeste van deze voorstellen kunnen gemakkelijk overgenomen c.q. overwogen worden. De voorstellen gaan met name de plek en volgorde van de conclusies, de bestuurlijke context en de SWOT en toekomstplannen. Bij het aanpassen van het format zou ook bekeken kunnen worden of het opportuun is een uitleg toe te voegen, zodat duidelijk wordt hoe het format gebruikt kan worden als een evaluatie verschillende aggregatieniveaus heeft (bijvoorbeeld een “instituu(de)el” en een “onderzoeksgroep-deel”).

#### 4.6 Domeinprofielen

Domeinprofielen zijn bedoeld ter ondersteuning bij het schrijven van het narratief van de zelfevaluatie, zowel ter onderbouwing van de overeenkomst met wat gebruikelijk is in het domein, als ter onderbouwing van bijzondere contrasten. De domeinprofielen kunnen gebruikt worden als soort “zachte benchmark”: een nationaal/internationaal referentiekader, voor het beschrijven van de overeenkomsten en contrasten met het eigen profiel.

De meningen over of de domeinprofielen behulpzaam zijn lopen sterk uiteen. Niet iedereen begrijpt de idee van een ‘zachte benchmark’. Het lijkt erop dat mensen die zich herkennen in het profiel er geen/weinig aandacht aan besteden, omdat ze daar geen reden toe zien. En mensen die zich niet herkennen lijken te kiezen voor een andere oplossing. Zij kiezen bijvoorbeeld voor een vergelijking met een internationaal instituut of geven aan dat er beter een gecombineerd profiel opgesteld kan worden dat bij hun onderzoekseenheid aansluit. Een aantal mensen geeft aan dat het profiel van de onderzoekseenheid niet aansluit bij één domein, waardoor ze aangeven de domeinprofielen niet kunnen gebruiken.

Het is de vraag of een betere toelichting bij de domeinprofielen zal resulteren in een frequenter gebruik. In de interviews is geprobeerd om hier meer inzicht te krijgen, maar de reacties waren niet eenduidig. Sommigen zien meer in vergelijking met een internationaal instituut, terwijl anderen een dergelijke zachte benchmark in potentie wel als waardevol beschouwen.

#### 4.7 Overige opmerkingen en suggesties

De belangrijkste punten uit de overige opmerkingen gaan over de website, de communicatie naar het veld en uitleg over QRiH voor de evaluatiecommissies. De Engelse versie van de website zou geredigeerd moeten worden en mogelijk zouden er zelfevaluaties en voorbeelden van narratieven op de website geplaatst kunnen worden. Daarnaast is de communicatie naar het brede veld, na de lancering van QRiH, stilgevallen. Dit zou opgepakt kunnen worden. Tot slot is de suggestie gedaan een aparte tekst of handleiding te schrijven voor de evaluatiecommissies. Deze suggestie is in de interviews door verschillende mensen ondersteund.

## Bijlage 1: lijst van onderzoekseenheden die een bijdrage hebben geleverd aan de evaluatie

Lijst met onderzoekseenheden die een reactie of een ingevuld evaluatieformulier terug gestuurd hebben (in alfabetische volgorde):

1. Internationaal Instituut voor Sociale Geschiedenis
2. Koninklijk Instituut voor Taal-, Land- en Volkenkunde
3. Maastricht University, Faculty of Arts and Social Sciences
4. Meertens Instituut
5. NIOD Instituut
6. Rijksuniversiteit Groningen, Faculteit Godgeleerdheid en Godsdienstwetenschap, Centrum voor Religiewetenschap
7. Rijksuniversiteit Groningen, Groningen Research Institute for Philosophy
8. Rijksuniversiteit Groningen, Research Institute for the Study of Culture
9. Tilburg University, School of Humanities and Digital Science
10. Tilburg University, Tilburg School of Catholic Theology
11. Universiteit Leiden, Faculteit Archeologie
12. Universiteit Leiden, Instituut voor Geschiedenis
13. Universiteit Leiden, Leiden University Centre for Linguistics
14. Universiteit Leiden, Leiden University Centre for the Arts in Society
15. Universiteit Utrecht, Faculteit Geesteswetenschappen
16. Universiteit Utrecht, Faculteit Geesteswetenschappen, Departement Filosofie en Religiewetenschap
17. Universiteit Utrecht, Faculteit Geesteswetenschappen, Instituut voor Cultuurwetenschappelijk Onderzoek
18. Universiteit van Amsterdam, Faculteit der Geesteswetenschappen, Amsterdam School for Cultural Analysis
19. Universiteit van Amsterdam, Faculteit der Geesteswetenschappen, Amsterdam School for Heritage, Memory and Material Culture
20. Universiteit van Amsterdam, Faculteit der Geesteswetenschappen, Amsterdam School for Regional, Transnational and European Studies
21. Universiteit van Amsterdam, Faculteit der Geesteswetenschappen, Amsterdam School of Historical Studies

Lijst van onderzoekseenheden van de geïnterviewden (in alfabetische volgorde):

- Meertens Instituut
- Protestantse Theologische Universiteit
- Radboud Universiteit, Centre for Language Studies (CLS) en Institute for Historical, Literacy and Cultural Studies (HLCS)
- Theologische Universiteit Kampen
- Universiteit Utrecht, Faculteit Geesteswetenschappen, Departement Geschiedenis en Kunstgeschiedenis
- Universiteit Utrecht, Faculteit Geesteswetenschappen, Utrecht Institute of Linguistics OTS
- Vrije Universiteit Amsterdam, Faculteit Religie en Theologie