



Vetenskapsrådet

KARTLÄGGNING AV OLIKA NATIONELLA SYSTEM FÖR UTVÄRDERING AV FORSKNINGENS KVALITET

– förstudie inför regeringsuppdraget U2013/1700/F

KARTLÄGGNING AV OLIKA NATIONELLA SYSTEM FÖR UTVÄRDERING AV FORSKNINGENS KVALITET

– förstudie inför regeringsuppdraget U2013/1700/F

Maud Quist, Sara Monaco, Sten Söderberg, John Tumpane

KARTLÄGGNING AV OLIKA NATIONELLA SYSTEM FÖR UTVÄRDERING AV FORSKNINGENS KVALITET
– FÖRSTUDIE INFÖR REGERINGSUPPDRAGET U2013/1700/F

VETENSKAPSRÅDET
Box 1035
101 38 Stockholm

© Vetenskapsrådet
ISBN 978-91-7307-226-7
Diarienummer: 353-2012-6940

INNEHÅLL

SAMMANFATTNING.....	3
Research Excellence Framework (REF2014) i Storbritannien.....	3
Nationella kvalitetsgranskningssystem i andra länder.....	4
Kvalitetsgranskningssystem vid svenska lärosäten	5
SUMMARY	6
Research Excellence Framework (REF2014) in the United Kingdom	6
National quality assessment systems in other countries.....	8
Quality assessment systems at Swedish HEIs.....	8
1. INLEDNING	9
1.1 Bakgrund.....	9
1.2 Arbetet med förstudien.....	9
2. REF2014 I STORBRITANNIEN	11
2.1 Syften och process för att utveckla systemet	11
2.2 Vad som utvärderas	12
2.3 Hur bedömningsprocessen går till	14
2.4 Utvärdering av impact i REF	16
2.5 Expertpanelerna.....	20
2.6 Tekniska lösningar för inrapportering	21
2.7 Hur utvärderingen omsätts till resurser.....	22
2.8 Kostnader för framtagning, implementering och drift av systemet	23
2.9 Konsekvenser av RAE/REF.....	24
3. NATIONELLA KVALITETSGRANSKNINGSSYSTEM AV FORSKNING I ANDRA LÄNDER ..	26
3.1 Dessa länder har vi studerat.....	26
3.2 Syften och process för att utveckla systemet	26
3.3 Vilken instans ansvarar för att genomföra utvärderingarna?	27
3.4 Peer review eller indikatorer?	27
3.5 Vad som utvärderas.....	28
3.6 Hur bedömningsprocessen går till	28
3.7 Hur utvärderingen omsätts till resurser.....	30
4. KVALITETSGRANSKNING AV FORSKNING VID SVENSKA LÄROSÄTEN.....	32
4.1 Syften och process för att utveckla systemet	32
4.2 Vad som utvärderas	32
4.3 Hur bedömningsprocessen går till	33
4.4 Expertpanelerna.....	36
4.5 Tekniska lösningar för inrapportering	36
4.6 Hur utvärderingen omsätts till resurser.....	36
4.7 Kostnader för framtagning, implementering och drift av systemet.....	36
4.8 Konsekvenser	37
5. SLUTSATSER	38
6. FRÅGOR SOM BEHÖVER UTREDAS VIDARE INFÖR FÖRSLAG TILL ETT SVENSKT SYSTEM	39
7. REFERENSER.....	41
Kapitel 1 Inledning.....	41
Kapitel 2 REF2014 i Storbritannien	41
Kapitel 3 Nationella kvalitetsgranskningssystem i andra länder	42
Kapitel 4 Kvalitetsgranskning av forskning på svenska lärosäten.....	44

Kapitel 5 Slutsatser	45
Kapitel 6 Frågor som behöver utredas vidare inför förslag till ett svenskt system.....	45
8. INTERVJUER VID STUDIEBESÖK I STORBRITANNIEN 5–15 FEBRUARI 2013	46
APPENDIX 1 INTERVJUGUIDE	49
APPENDIX 2 UNITS OF ASSESSMENT I REF2014 I STORBRITANNIEN.....	54

SAMMANFATTNING

I denna rapport redovisas resultatet av en kartläggning av olika nationella system för utvärdering av forskningens kvalitet. Studien har genomförts med anledning av propositionen *Forskning och Innovation 2012/13:30* och som en förstudie inför regeringsuppdraget till Vetenskapsrådet att ”*utreda och lämna förslag till en modell för resursfördelning till universitet och högskolor som innefattar kollegial bedömning av forskningens kvalitet och relevans*” (U2013/1700/F).

Förstudien omfattar i huvudsak en fördjupad analys av det brittiska utvärderingssystemet, vilket i olika versioner har varit i funktion under 27 år och därmed är det system som det finns längst erfarenheter från. Vidare redovisas en mer översiktlig kartläggning av åtta andra länders system samt en analys av utvärderingarna vid de svenska lärosäten som har låtit genomföra egna kvalitetsgranskningar av sin forskning. Avsikten med förstudien är att ta tillvara erfarenheter och lärdomar – såväl positiva som negativa – från olika befintliga system för utvärdering av forskning inför arbetet med att utforma Vetenskapsrådets kommande förslag till ett svenskt system.

Research Excellence Framework (REF2014) i Storbritannien

Storbritannien har haft system med nationella utvärderingar av forskningens kvalitet sedan 1986 återkommande med fyra till sex års mellanrum. Myndigheten HEFCE (Higher Education Funding Council for England) ansvarar för dessa utvärderingar. Systemet har successivt utvecklats i en nära dialog med sektorn. Det nuvarande systemet REF2014 (Research Excellence Framework) har tre syften:

1. att vara kvalitetsdrivande genom att identifiera och premiera den mest excellenta forskningen samt att premiera de lärosäten som engagerar sig så att forskningsresultaten kommer till användning i samhället
2. att visa att skattemedel som går till forskning används på ett ansvarsfullt sätt, samt
3. att ta fram ett underlag för fördelning av basanslaget för forskning till universitet och högskolor.

REF syftar *inte* till att utvärdera *all* forskning. Lärosätena väljer själva vilken forskning de skickar in för bedömning. Den andel forskare som omfattas per lärosäte varierar mellan ca 25 och 95 procent. Inför urvalet gör lärosätena ett omfattande arbete. De gör interna utvärderingar, eftersom det endast är den forskning som bedöms vara internationellt excellent och världsledande (de två högsta betygen) som genererar medel, och tar fram fallstudier som visar på forskningens påverkan utanför akademien (*impact*). Systemet ”tvingar” lärosätena att själva ta fram strategier och göra prioriteringar. En utgångspunkt för REF är att samma system ska användas för alla ämnesområden, med vissa anpassningar beroende på disciplin. Inga jämförelser görs mellan ämnesområden. I REF används färre men bredare disciplinära utvärderingsenheter, totalt 36 st., för bedömning jämfört med i tidigare system. Det var 67 st. utvärderingsenheter i föregångaren Research Assessment Exercise (RAE2008).

Bedömningen görs av expertpaneler, totalt fyra huvudpaneler och 36 disciplinära underpaneler, med i huvudsak nationella vetenskapliga experter och användareexperter. Tre aspekter bedöms och utgör tillsammans underlag för lärosätets kvalitetsprofil. Aspekterna är:

1. *output*, forskningens utfall, viktas till 65 procent
2. *impact*, forskningens påverkan på samhället, viktas till 20 procent och
3. *environment*, forskningsmiljön, viktas till 15 procent.

Underpanelerna gör en sammanvägd bedömning där en fyrgradig betygsskala plus kategorin oklassificerad används. De lämnar inga rekommendationer utan endast en kortfattad skriftlig återkoppling som inte är offentlig. Forskningens *output* bedöms genom att panelen läser/tar del av fyra publikationer,

eller andra produkter av forskning, per utvald forskare. Kriterierna för bedömning av forskningens *output* är *rigour* (stringens), *originality* (nytänkande) och *significance* (betydelse). Inom vissa underpaneler tar man även in bibliometriska uppgifter som underlag för bedömning. Forskningens *impact* bedöms dels utifrån fallstudier där samhällelig påverkan av excellent forskning ska föras i bevis och dels genom en beskrivning av hur man arbetar för att åstadkomma *impact* (*impact template*). Kriterierna för bedömning av *impact* är *reach* (räckvidd) och *significance* (betydelse). Bedömning av *environment* sker utifrån kvantitativa data om t.ex. personal, antal doktorsexamina och externa inkomster. Dessutom ska forskningsmiljön och forskningsstrategier beskrivas kvalitativt. Kriterierna för bedömning av *environment* är *vitality* (vitalitet) och *sustainability* (hållbarhet).

I Storbritannien utgör basanslaget till forskning ca 30 procent av de totala forskningsintäkterna till universitet och högskolor (att jämföra med ca 40 procent¹ i Sverige). I England fördelas merparten (ca 2/3) av basanslaget för forskning på basis av resultatet av RAE. Tre parametrar används för att räkna fram resurstilldelningen: kvalitetsprofil, volym och en kostnadsfaktor för ämnesområdet. Kvalitetsprofilen uttrycks som andelen av den forskning inom en utvärderingsenhet vid lärosätet som har sänts in för bedömning och som har bedömts ligga vid respektive betyg (på skalan 1* till 4* och oklassificerad). Volymen avser antal forskare (uttryckt som heltidsekvivalenter) som ingick i aspekten *output*. Kostnadsfaktorn viktas till fördel för mer resursintensiva forskningsämnen, exempelvis har medicin faktorn 1,6 att jämföra med humaniora och samhällsvetenskap som har faktorn 1,0. Den exakta algoritmen som kommer att användas för resursfördelning 2014, baserat på REF-resultatet, är dock inte känd fram till dess att resultatet presenterats.

Lärosätena avgör själva hur de fördelar det basanslag för forskning som de erhåller baserat på REF-resultatet. Det finns alltså ingen begränsning avseende ämnesområde. De kan välja att premiera de bästa miljöerna, stötta svagare områden, eller satsa på strategiska områden. HEFCE fördelar utöver basanslaget till forskning särskilda anslag till lärosätena för utbildning, forskarutbildning och innovation.

Det är svårt att uppskatta kostnaden för att genomföra REF. Utöver de rent monetära kostnaderna nämns även kostnader för utebliven forskning och psykosociala arbetsmiljökostnader på lärosätena.

Lärosätena lägger ned mycket tid och resurser på att justera dimensionering av och innehållet i de tre aspekterna (*output*, *impact* och *environment*) i förhållande till varandra, och därmed försöka optimera utfallet vad gäller pengar och/eller prestige. Trots att flera omgångar av nationell kvalitetsutvärdering genomförts sedan 1986 innebär det en ryckighet i lärosätenas verksamhet med en kraftsamling ett par år inför varje ny RAE/REF.

I Storbritannien anser de som har intervjuats inom ramen för denna förstudie, att systemet uppnår alla sina tre syften (kvalitetsdrivande, ansvarsutkrävande och underlag för resursfördelning), men att den kvalitetsdrivande effekten planat ut med tiden. Systemet ger inga drastiska omfördelningseffekter mellan lärosäten men förstärker den befintliga koncentrationen av forskningsresurser till fyra av totalt 165 lärosäten. Systemet förmår dock även att identifiera små excellenta miljöer som därmed kan premieras. I Storbritannien har systemet lett till en ökad klyfta mellan undervisande och forskande personal. Systemet har också lett till en ökad konkurrens om kollegiala experter (*peers*) som behövs på flera håll – till REF, till lärosätenas egna förutvärderingar, till forskningsrådets ansökningsprocesser och utvärderingar samt till granskning av publikationer. Publikationsvolymen har ökat totalt, vilket leder till köbildning på tidskrifter och förlag, särskilt inför deadline för att skicka in material till RAE/REF.

Nationella kvalitetsgranskningssystem i andra länder

Utöver kartläggningen av Storbritannien omfattar denna förstudie en mer översiktlig kartläggning av åtta andra länder: Norge, Danmark, Finland, Australien, Nya Zeeland, Italien, Flandern (norra Belgien) och Nederländerna. Prestationsbaserade resurstilldelningssystem har införts i flera länder och blivit allt vanligare på senare år. Det finns också länder som genomför nationella kvalitetsgranskningar av forskningskvaliteten vid universitet och högskolor utan att de är (direkt) kopplade till resurstilldelning. Vid jämförelse av olika länders system bör man beakta kontexten, t.ex. systemets syfte och di-

¹ I detta är inte de s.k. ALF-medlen, som betalas ut för landstingens medverkan i läkarutbildning och klinisk forskning, om ca 6 procent inräknade.

mensionering i förhållande till landets övriga instrument för forskningsfinansiering, vilka torde påverka både utfallet och även designen av modellen.

Det är vanligt att myndigheter som inte är forskningsråd ansvarar för utvärderingsuppdraget. Utvärdering på lärosätetsnivå på basis av ämnesområden är den vanligaste modellen men indelningen av ämnen skiljer sig åt mellan länderna. Indikatorer används som underlag till en kollegial bedömning eller som ett självständigt parallellt spår. Vad gäller val av indikatorer inkluderas i de flesta system mått som inte enbart har med vetenskaplig produktion och forskningens kvalitet att göra, såsom forskarutbildning, internationalisering, m.m. Utvärderingsmetoder och indikatorer avsedda att uppskatta forskningens nytta och samverkan med samhället/näringslivet är inte lika utvecklade och förekommer i betydligt mindre utsträckning än de som är avsedda att utvärdera forskningens vetenskapliga kvalitet.

Kvalitetsgranskningssystem vid svenska lärosäten

Sju svenska lärosäten har hittills genomfört egna kvalitetsgranskningar av sin forskning, och fyra lärosäten har pågående utvärderingar. Ingen av dem har haft som främsta syfte att omfördela resurser, däremot kan det sägas vara ett indirekt syfte. Ett genomgående övergripande syfte för samtliga lärosätets utvärderingar har varit att ligga till grund för ett strategiskt utvecklingsarbete. Genom att identifiera styrkor och svagheter i forskningen har lärosätena önskat få underlag för att förbättra den vetenskapliga kvaliteten, exempelvis genom att stödja forskargrupper med potential eller avveckla mindre framgångsrik verksamhet.

Samtliga svenska lärosäten som genomfört utvärderingar av sin forskning vittnar om att det varit lärorika men arbetsamma processer. En av de stora utmaningarna har varit *hur* man definierar en utvärderingsenhet och *vem* som gör det. Det har varit problematiskt att de utvärderade enheterna har skilt sig åt vad gäller exempelvis storlek och förutsättningar. När de utvärderade enheterna inte sammanfallit med lärosätets organisatoriska enheter har det försvårat implementeringen av eventuella åtgärder till följd av utvärderingens resultat. Vidare har det gjort jämförelser inom lärosätet i princip omöjliga.

Lärosätena understryker vikten av att från början klargöra både syfte och hur man hanterar resultaten. Det har inte varit ovanligt att de utvärderade enheterna haft orimliga förväntningar på vad man kunde förvänta sig efter avslutad utvärdering i termer av extraresurser och liknande, vilket i sin tur lett till viss frustration och besvikelse.

SUMMARY

This report presents the findings of a survey of various national systems for evaluating the quality of research. The study was carried out as a result of the Swedish Government Bill 2012/13:30 (*Research and Innovation*) and as a preliminary study in advance of the directive to the Swedish Research Council to “investigate and present proposals for a model for the allocation of funding to universities and university colleges that includes peer review of the quality and relevance of research” (U2013/1700/F) Dnr. 111-2013-2496. The preliminary study consists mainly of an in-depth analysis of the British assessment system, which has been in use in various versions for 27 years and is thus the system with the longest experience from which to learn. Furthermore, the study reports from a more general survey of eight other countries’ systems and a review of the evaluations at Swedish institutions of higher education that have arranged to have the quality of their research assessed. The purpose of the preliminary study is to gather experiences and lessons learned – both positive and negative – from various existing systems for assessing research in advance of framing the Swedish Research Council’s coming proposal for a Swedish system.

Research Excellence Framework (REF2014) in the United Kingdom

The United Kingdom (UK) has had a national system for evaluating the quality of research since 1986, recurring at intervals of four to six years. HEFCE (Higher Education Funding Council for England), is the public body responsible for these assessments. The system has successively been refined in close dialogue with the sector. The current system, REF2014 (Research Excellence Framework), has three purposes:

1. to foster quality by identifying and rewarding the most excellent research and to reward higher-education institutions that endeavour to ensure that research findings will be put to use in society
2. accountability, to show that money from the public purse spent on research is being used in a responsible manner, and
3. to provide a basis for allocating block grants for research at universities and university colleges.

REF does *not* aim to evaluate *all* research. Higher-education institutions (HEIs) themselves select which research they submit for assessment. The proportion of researchers covered per HEI varies from ca. 25 to 95 per cent. HEIs devote considerable effort to making their selection. They perform internal evaluations, as it is only research that is judged to be internationally excellent or world-leading (the two highest grades) that generates funding, and they highlight case studies that demonstrate the *impact* of the research outside the academy. The system ‘compels’ HEIs themselves to devise strategies and establish priorities. One point of departure for REF is that the same system is to be used for all subject areas, with certain adaptations, depending on the discipline. No comparison is made between subject areas. REF uses fewer but broader disciplinary units of assessment, a total of 36, compared with earlier systems. REF’s predecessor, Research Assessment Exercise (RAE2008), comprised 67 units of assessment.

Assessment is made by expert panels, a total of four main panels and 36 disciplinary sub-panels, with largely domestic research experts and user experts. Three aspects are judged and, taken together, constitute the basis for the quality profile of the HEI. These aspects are:

1. *output*, outcome of research, weighted at 65 per cent
2. *impact*, effect of research on society, weighted at 20 per cent and
3. *environment*, research environment, weighted at 15 per cent.

The sub-panels make an overall assessment, using a four-point grading scale plus the category unclassified. They make no recommendations but merely provide short written feedback, which is not made public. The research *output* is evaluated by the panel reading/reviewing four publications, or other products of research, per selected researcher. The criteria for assessing research *output* are *rigour*, *originality* and *significance*. Some sub-panels also consider bibliometric data as a basis for evaluation. Research *impact* is judged, on the one hand, on the basis of case studies where the effect of excellent research is to be established and, on the other hand, by a description of how the HEI works to achieve this societal *impact*, that is, its *impact template*. The criteria for assessing *impact* are *reach* and *significance*. Evaluation of *environment* is based on quantitative data on research staff, number of doctoral degrees and external income, for example. Moreover, the research environment and research strategies are to be described qualitatively. The criteria for assessing the *environment* are *vitality* and *sustainability*.

In the UK the block grants for research constitute ca. 30 per cent of total research income at HEIs (compared with ca. 40 per cent² in Sweden). In the UK the majority (ca. 2/3) of research funding allocated via HEFCE is based on the results of RAE. Three parameters are used to calculate resource allocation: quality profile, volume and a cost factor for the subject area. The quality profile is expressed as the respective proportions of the research included in an evaluation unit submitted by the HEI that has been judged to warrant each grade (on the scale 1* to 4* and unclassified). Volume refers to the number of researchers (expressed as full-time equivalents) that were included in the *output* aspect. The cost factor is weighted in favour of more resource-intensive research subjects; for example, medicine has a factor of 1.6, to be compared with 1.0 for the humanities and social sciences. The exact algorithm to be used for allocating resources in 2014, based on REF results, will remain unknown until the results have been presented.

HEIs determine themselves how to distribute internally the research funding that they receive based on the REF results. There are thus no constraints regarding subject areas. They can choose to reward the best environments, support weaker areas, or invest in strategic areas. On top of basic allocations for research, the HEFCE grants special funding to HEIs for education, postgraduate education and innovation.

It is difficult to estimate the cost of carrying out REF. Besides purely monetary costs, mention is made of costs for disruptions to research and psychosocial working environment costs at HEIs.

HEIs devote large amounts of time and resources to adjusting the dimensioning of and contents of the three aspects (*output*, *impact* and *environment*) in relation to each other, thereby trying to optimise outcomes in terms of funding and/or prestige. Even though several rounds of national quality assessments have been carried out since 1986, HEI operations are perturbed, with an intensifying of activity for a couple of years in advance of each new RAE/REF.

In the UK, those who have been interviewed within the framework of this pilot study consider that the system achieves all three of its goals (fostering quality, accountability and establishing a basis for resource allocation), but that the quality-fostering effect has levelled off with time. The system does not lead to any drastic redistribution effects among HEIs but reinforces the existing concentration of research resources in four of a total of 165 HEIs. However, the system is able to identify small excellent environments that can thereby be rewarded. In the UK the system has led to a growing gap between teachers and researchers. The system has also brought with it greater competition for peers, who are needed in many contexts for REF, for HEIs' own pre-evaluations, for research councils' application processes and assessments, and for reviewing publications. Publication volume has increased overall, which leads to queue-forming at journals and publishing companies, especially ahead of the deadline for submitting material to RAE/REF.

² This does not include so-called ALF funding, which covers the participation of county councils in medical education and clinical research, at ca. 6 per cent.

National quality assessment systems in other countries

Besides the survey of the United Kingdom, this preliminary study comprises a more general overview of eight other countries: Norway, Denmark, Finland, Australia, New Zealand, Italy, Flanders (northern Belgium) and the Netherlands. Performance-based resource allocation systems have been introduced in several countries and have become more and more common in recent years. There are also countries that carry out national quality assessments of research quality at HEIs without linking them (directly) to resource allocation. When various countries' systems are compared, the context should be borne in mind – for example, the purpose and dimensioning of the system in relation to the country's other instruments for research funding – which may affect both the outcome and the design of the model.

It is common that public bodies that are not research councils are responsible for the evaluation process. Evaluation at the HEI level on the basis of subject areas is the most common model, but the division into subjects differs across countries. Indicators are used as a basis for peer review or as an independent parallel track. Regarding the choice of indicators, most systems include measures that have to do with more than academic production and quality of research, such as postgraduate education, internationalisation, etc. Evaluation methods and indicators designed to estimate the utility of research and its interaction with the wider society and business community are not as fully developed and occur to considerably lesser extent than those devised to assess the scientific quality of research.

Quality assessment systems at Swedish HEIs

Seven Swedish HEIs have thus far carried out their own quality assessments of their research, and four HEIs currently have such assessments underway. None of them has had the redistribution of resources as the primary purpose, although this can be said to be an indirect purpose. A consistent overall purpose for all these HEIs' assessments has been to provide a basis for strategic development. By identifying strengths and weaknesses in research, these HEIs have strived to establish a platform for enhancing academic quality, for example by supporting research teams with potential or by phasing out less successful operations.

All Swedish HEIs that have carried out assessments of their research testify that it is an informative but arduous process. One of the major challenges has been *how* to define a unit of assessment and *who* is to do so. It has been problematic that the evaluated units have differed in terms of size and circumstances, for example. When the evaluated units have not coincided with the HEI's organisational units, this has complicated the implementation of any measures prescribed as a result of the evaluation. Furthermore, this has rendered comparisons within the HEI virtually impossible.

HEIs emphasise the importance of clarifying from the outset both the purpose of the assessment and how findings are to be dealt with. It has not been uncommon for evaluated units to have unrealistic expectations about what might ensue once the assessment has been completed, in terms of extra resources and the like, which in turn has led to a certain amount of frustration and disappointment.

1. INLEDNING

1.1 Bakgrund

I Vetenskapsrådets remissvar (dnr 112-2012-6) på utredningen *Prestationsbaserad resurstilldelning för universitet och högskolor* (U2011/7356/UH) ställde Vetenskapsrådet sig positivt till att en del av de direkta anslagen till lärosätena baseras på prestation avseende vetenskaplig kvalitet och forskningens samverkan och påverkan på samhället. Till skillnad från utredningens förslag ansåg Vetenskapsrådet att bedömning av prestation borde ske i en process med sakkunniggranskning. På uppdrag från utbildningsdepartementet skickades en skiss in i maj 2012 (dnr 111-2012-6066) på hur ett uppdrag till Vetenskapsrådet skulle kunna utformas för att utreda ett förslag till *Utformning av ett system för peer review utvärdering som underlag för prestationsbaserad resurstilldelning för universitet och högskolor*. Vetenskapsrådet utgick ifrån de synpunkter som framförts i tidigare remissvar och i skrivelsen *Insatser för att stärka Sverige som forskningsnation* (dnr 311-2011-1306).

I propositionen *Forskning och Innovation 2012/13:30* som kom i oktober 2012 föreslog regeringen (punkt 6.2.4) följande: ”Vetenskapsrådet bör få i uppdrag att, i samråd med Forskningsrådet för arbetsliv och socialvetenskap, Forskningsrådet för miljö, areella näringar och samhällsbyggande samt Verket för innovationssystem, utreda och lämna förslag till ett resursfördelningssystem som gör det möjligt att fördela anslag till universitet och högskolor innefattande kollegial bedömning av universitets och högskolors kvalitet och prestation i forskning, inkluderande såväl vetenskapliga bedömningar som bedömning av dess relevans och nytta för samhället”. För att förbereda det förutsedda regeringsuppdraget (som kom att beslutas 14 mars 2013³) startades i november 2012 föreliggande förstudie.

1.2 Arbetet med förstudien

Förstudien innehåller en kartläggning av det brittiska kvalitetsgranskningssystemet REF2014 (Research Excellence Framework) och till viss del dess föregångare RAE (Research Assessment Exercise). Förstudien ger också en översikt av system för utvärdering av forskningens kvalitet i ett antal andra länder samt av de svenska lärosätenas egenutvärderade kvalitetsgranskningar av sin forskning. Kartläggningen gör inte anspråk på att vara vare sig fullständig, uttömmande eller heltäckande.

Förstudien har genomförts under perioden november 2012 – mars 2013 genom dokumentstudier och studiebesök i England och Nordirland 5–15 februari 2013. Vid studiebesöken intervjuades sammanlagt drygt 50 personer⁴ på Department of Business, Innovation and Skills (BIS); Higher Education Funding Council for England (HEFCE); fyra forskningsråd och paraplyorganisationen Research Councils UK (RCUK); sju universitet och den övergripande organisationen Universities UK samt Quality Assurance Agency (QAA). Kriterierna för urval av lärosäten att besöka var att de skulle ha viss ämnesbredd (de skulle ha deltagit i fler än 20 av utvärderingsenheterna i RAE2008), de skulle ha deltagit i pilotförsök avseende *impact* och de skulle representera såväl ”vinnare” som ”förlorare” i RAE2008. De skulle dessutom ligga geografiskt så till att det gick att genomföra studiebesök under begränsad tid. På lärosätena träffade vi rektorer, pro-rektorer, centrala REFansvariga och REF-handläggare, deltagare i pilotstudien om forskningens *impact* och ledamöter i expertpanelerna. I VINNOVA:s regi har ett arbete påbörjats med att göra telefonintervjuer med användarrepresentanter i panelerna. Vi har deltagit i flera externa möten och seminarier i Sverige där kvalitetsgranskning av

³ Innebörden i regeringsuppdraget är densamma som i propositionen men något annorlunda formulerat: *Regeringen uppdrar åt Vetenskapsrådet att i samråd med Forskningsrådet för arbetsliv och socialvetenskap, Forskningsrådet för miljö, areella näringar och samhällsbyggande samt Verket för innovationssystem utreda och lämna förslag till en modell för resursfördelning till universitet och högskolor, som innefattar kollegial bedömning och som gör det möjligt att fördela resurser på ett sådant sätt att det premierar kvalitet och prestation i forskning. Ett förslag till en modell ska innefatta såväl vetenskapliga bedömningar som bedömning av forskningens relevans och nytta för samhället.* U2013/1700/F dnr 111-2013-2496.

⁴ Se lista över intervjuade i kapitel 8.

forskning har diskuterats. Samtal med Universitetskanslersämbetet om utvärdering av forskarutbildning har inletts.

Arbetet har genomförts av en grupp på Vetenskapsrådet vars kärna utgjorts av Sara Monaco, Sten Söderberg, John Tumpane och Maud Quist (projektledare). I den utvidgade gruppen har även Carl Jacobsson, Jenny Nordquist och Jan Bolin ingått. Ragnhild Romanus har varit knuten till gruppen i egenskap av kommunikatör. Generaldirektör Mille Millnert har varit gruppens uppdragsgivare. Samråd har skett underhand med VINNOVA via Rolf Nilsson, Maria Landgren och Peter Eriksson samt med Formas via Erik Roos och Peter Schilling. Vilka personer som deltog vid de olika besöken i Storbritannien framgår av schemat för intervjuer i kapitel 8. FAS har informerats om arbetet via Tommy Dahlén. Presentation för Vetenskapsrådets ledningsråd har skett vid två tillfällen. Myndigheternas analyschefsgrupp har informerats om förstudien vid två tillfällen. Rapporten har presenterats vid ett seminarium på Vetenskapsrådet 8 april 2013 dit även personal från FAS, Formas och VINNOVA var inbjudna.

2. REF2014 I STORBRITANNIEN

Följande avsnitt baseras på information från www.ref.ac.uk⁵ samt intervjuer⁶ vid studiebesöket i Storbritannien i februari 2013.

2.1 Syften och process för att utveckla systemet

Storbritannien har haft system för nationella utvärderingar av forskningens kvalitet sedan 1986 återkommande med fyra till sex års mellanrum. Vi har i huvudsak satt oss in i det nuvarande systemet REF (Research Excellence Framework) som kommer att genomföras 2014 och som ersätter RAE (Research Assessment Exercise) och dess föregångare Research Selectivity Exercise.

Syftet med REF2014 är trefaldigt:

1. Att vara kvalitetsdrivande genom att identifiera och premiera den mest excellenta forskningen samt att premiera de lärosäten som engagerar sig så att forskningsresultaten kommer till användning i samhället. REF syftar *inte* till att utvärdera *all* forskning.
2. Att gentemot skattebetalarna, politikerna och finansdepartementet visa att de pengar som avsätts till forskning används på ett ansvarsfullt sätt (*accountability*). Att visa på forskningsresultat och att det är en god investering att fortsätta satsa offentliga medel på forskning var det ursprungliga syftet när man införde ett nationellt system för kvalitetsutvärdering av forskning 1986, i ett skede då många nya universitet tillkom.
3. Att få fram ett robust underlag för selektiv fördelning av merparten av basanslaget till universitet och högskolor (resursfördelning baserat på kvalitet infördes runt 1990).

Inför varje RAE/REF har ett utvecklingsarbete gjorts, vilket innebär att systemet successivt förfinats. Processen präglas av ett omfattande konsultationsarbete med sektorn vid flera tillfällen. Förankring sker genom t.ex. öppna webbaserade konsultationer, workshops riktade till specifika grupper och användning av rådgivande expertgrupper.

Under perioden 2004–2009 togs initiativ till att förändra RAE i riktning mot ett enklare system som i högre grad baserades på kvantitativa mått och indikatorer för att minska det administrativt tungrodda peer review-systemet. Ett pilotförsök med bibliometri-indikatorer⁷ genomfördes därför under 2008. Slutsatsen från pilotförsöket var att citeringsdata inte är tillräckligt robust för att användas systematiskt som en fristående indikator i REF, men att det kan användas som underlag för kollegial bedömning inom vissa ämnesområden. En utgångspunkt har varit att samma utvärderingssystem ska användas för hela sektorn oavsett ämnesområde. System baserade på indikatorerna externa inkomster och citeringar anses slå alltför olika mellan ämnen. Förslaget att i större utsträckning utgå från kvantitativa mått och indikatorer väckte protester från akademien, särskilt inom humaniora och samhällsvetenskap. På grund av pilotförsökets resultat blev det en stor uppslutning kring att fortsätta använda peer review som verktyg. Sektorn upplever det också som att man med peer review behåller kontroll över hur forskningsmedlen prioriteras och att det skapar en önskvärd distans till den politiska makten. En del av dem vi

⁵ Se särskilt

REF01.2010. Units of assessment and recruitment of expert panels.

REF02.2011. Assessment framework and guidance on submissions.

REF01.2012. Panel criteria and working methods.

⁶ Se kapitel 8. Semistrukturerade intervjuer genomfördes utifrån de frågor som redovisas i Appendix 1.

⁷ <http://www.ref.ac.uk/background/bibliometrics/>; HEFCE (2009/39) *Report on the pilot exercise to develop bibliometric indicators*; Technopolis (2009) *First report on 'lessons learned' from the bibliometrics pilot exercise: data collection*; Technopolis (2009) *Second report on 'lessons learned' from the bibliometrics pilot exercise: outcomes*.

intervjuade på vissa lärosäten sa dock att inom medicin och naturvetenskap borde det kunna fungera lika bra med enbart indikatorer och kvantitativa data, utan peer review. Men utgångspunkten för HEFCE är som sagt att alla discipliner ska granskas med samma metod. Vissa paneler använder dock bibliometri som ett av flera underlag inom REF.

Krav från regeringen, näringslivet och allmänheten på att forskningsresultatens vidare samhällseffekter ska kunna påvisas ledde till att *impact* inkluderades i utvärderingssystemet. Ett pilotförsök genomfördes för att prova ut lämplig metod för att mäta och bedöma forskningens *impact* 2010⁸. Slutsatsen var att det är möjligt att införa *impact* som en aspekt i systemet för alla ämnesområden, genom att använda fallstudier där man beskriver forskningsresultatens effekter i ett vidare samhällsperspektiv. För en grundligare genomgång av införandet av *impact* som en aspekt i REF, se avsnitt 2.4.

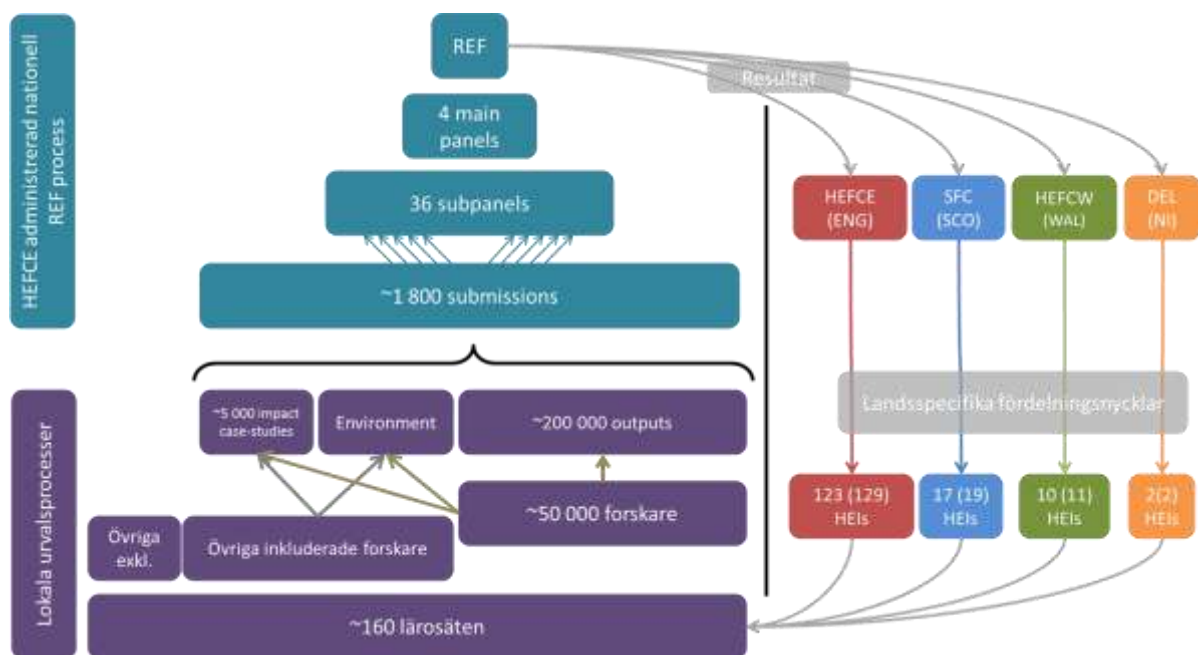
2.2 Vad som utvärderas

Utvärderingen i REF är uppdelad i totalt 36 disciplinära utvärderingsenheter (*units of assessments*) (se Appendix 2) som är kopplade till lika många underpaneler av experter som gör bedömningen. Underpanelerna är i sin tur fördelade på fyra huvudpaneler. Detta är en kraftig minskning jämfört med RAE2008, som bestod av 67 underpaneler fördelade på 15 huvudpaneler. Minskningen anses vara en effektivisering av processen. Avgränsning av en utvärderingsenhet avspeglar inte nödvändigtvis universitetens egna organisatoriska och disciplinära indelningar, vilket kan innebära vissa svårigheter att använda befintlig statistik. Ambitionen finns hos HEFCE att samordna efterfrågade statistiska uppgifter med forskningsråden och statistikmyndigheten Higher Education Statistics Agency (HESA).

Det samlade material som ett lärosäte skickar in avseende en disciplinär utvärderingsenhet kallas för en *submission*. Tillsammans kommer de drygt 160 brittiska lärosätena att skicka in ca 1800 *submissions* till REF 2014. Uppgifter om alla typer av forskning, oavsett finansieringskälla, kan skickas in. Det är dessa *submissions*, en för varje utvärderingsenhet vid respektive lärosäte, som utvärderas och får en kvalitetsprofil i REF. Varje *submission* omfattar minst en individ och de största omfattar mellan 100 och 200 individer. En del lärosäten menar att det vore önskvärt att ha en princip om minimalt antal forskare för att en *submission* ska komma ifråga, något man använder sig av i Wales. Uppskattningsvis 85 procent av Storbritanniens totala kår av forskare, som definieras genom kriterier avseende anställningens art, omfattas av REF.

En *submission* innehåller underlag för att bedöma tre aspekter – *output* (forskningens utfall), *impact* (forskningens påverkan på samhället) och *environment* (forskningsmiljön), se vidare avsnitt 2.3. I figur 1 presenteras en översikt av processen för REF2014.

8 Research Excellence Framework. (2010). *Impact pilot exercise. Findings of the expert panels.*



Figur 1. Skiss över processen för REF2014. Siffrorna inom parantes i rutorna nere till höger anger hur många lärosäten (HEIs) som deltog i RAE2008 i respektive land (England, Skottland, Wales och Nordirland), siffrorna utanför parantes anger hur många som fick medel via RAE2008.

Enligt lärosätenas preliminära uppskattning kommer de sammantaget alltså att skicka in ca 1800 *submissions*. Dessa innefattar ca 200 000 s.k. *research outputs* producerade av ca 50 000 individuella forskare, ca 5 000 fallstudier som visar på forskningens *impact* och ca 1800 beskrivningar av forskningsmiljöer. Det är antalet heltidsekvivalenter avseende de ca 50 000 forskarna knutna till *output* som sedan används för beräkning av volymmåttet som ingår i fördelningsnyckeln (se avsnitt 2.7). Av dessa ca 200 000 *outputs* är ungefär 2 000 på ett annat språk än engelska och ca 20 000 är monografier. Merparten av alla *outputs* skickas in digitalt men HEFCE räknar med att ca 50 000 kommer att deponeras fysiskt i deras lager. Dessa 50 000 *outputs* är på ett eller annat sätt omöjliga att ladda upp elektroniskt och innefattar monografier, konstverk och annat material (även artiklar) som måste skickas i den ursprungliga fysiska formen eller via USB om filen är för stor för ansökningsystemet. Vid nästa REF år 2020 kommer HEFCE att kräva att publikationerna är öppet tillgängliga elektroniskt.

Urval på lärosätetsnivå

REF i sig utvärderar endast ett urval av brittisk forskning. Andelen forskare som inkluderas i REF varierar mellan lärosäten. I *submissions* från stora etablerade forskningsuniversitet som t.ex. Cambridge omfattas ca 95 procent av forskarna, medan andelen kan vara så låg som ca 25 procent vid mindre, inte så forskningsintensiva lärosäten. Urvalsprocessen på lärosätena för att ta fram underlaget till REF omfattar dock betydligt mer forskning och flera forskare än de som kommer att utvärderas inom REF. Lärosätena är fria att välja om de vill rapportera in till en disciplinär utvärderingsenhet överhuvudtaget. De väljer också själva vilka forskare som ska ingå i underlaget. Detta sker enligt en hierarkisk struktur på lärosätena med beslut av en rektorskommitté eller motsvarande om vilka som får delta. Eftersom det endast är den forskning som bedöms vara internationellt excellent och världsledande (betyg 3* och 4*) som genererar medel, har lärosätena själva en omfattande urvalsprocedur för att optimera utfallet. På samtliga lärosäten som vi besökte förekommer interna utvärderingsprocesser som består av en intern eller extern peer review, eller både och. Det är ganska omfattande processer som börjar långt innan materialet ska vara inskickat till HEFCE.

2.3 Hur bedömningsprocessen går till

På HEFCE sköts REF i praktiken av ett litet team om åtta personer (plus IT-personal). I samband med att lärosätena skickar in sina bidrag rekryteras mer personal till HEFCE centralt och ett 30-tal personer rekryteras från lärosätena till ett sekretariat. Dessutom anlitas ca 1000 experter till panelerna för att göra granskningen.

REF fokuserar på uppnådda resultat. Varje *submission* bedöms ur tre aspekter som viktas till en sammantagen kvalitetsprofil för lärosätet:

1. Kvaliteten på forskningens *output* (utfall), viktas till 65 procent
2. Forskningens *impact* (påverkan) utanför akademien, viktas till 20 procent
3. *Environment* (forskningsmiljö), viktas till 15 procent

Kvalitetsprofilen visar hur stor proportionen av varje *submission* som graderats enligt bedömningen:

- 4* *World leading* (världsledande)
- 3* *Internationally excellent* (internationellt excellent)
- 2* *Internationally recognised* (internationellt erkänd)
- 1* *Nationally recognised* (nationellt erkänd)
- U *Unclassified* (oklassificerad)

Endast betygen 3* och 4* genererar basanslag till forskningen. För att forskningens *impact* ska komma ifråga för bedömning ska den vara baserad på forskning av minst 2* standard. Även om bedömningskriterierna och betygsskalan är generiska, kan inte och ska inte REF jämföra eller gradera bedömningen mellan olika ämnesområden.

Som underlag för bedömning av aspekten *output* ska fyra produkter av forskningen från perioden 2008–2013 per utvald forskare skickas in. Det kan vara t.ex. vetenskapliga artiklar, konferensbidrag, monografier, patent, konstnärliga produktioner etc. Varje *output* ska läsas och bedömas av minst två experter. De forskare vars *output* skickas in ska vara anställda på det aktuella lärosätet på minst 20 procent och i huvudsak bedriva forskning, alternativt undervisning och forskning. Undantag görs för s.k. kategori C personal, t.ex. forskare som är anställda av ett sjukhus för att forska på ett lärosäte. Om giltiga skäl finns – t.ex. föräldraledighet, sjukdom, att forskaren är karriärmässigt ung m.m. – kan färre än fyra *output* vara tillräckligt. Omfattande monografier kan räknas som två *outputs*, eftersom de är tidskrävande att producera. Vissa underpaneler använder bibliometriska underlag som komplement för bedömning av forskningens vetenskapliga kvalitet. Förlaget Elsevier har fått i uppdrag av HEFCE att ta fram det bibliometriska underlaget. Om forskare flyttar mellan lärosäten följer forskarens produkter (som alltså kan tillgodoräknas aspekten *output*) med individen. Upp till sex forskare (tio inom underpanelen för fysik) kan var och en tillgodoräkna sig en samförfattad/samproducerad *output* såvida de inte ingår i samma lärosätes *submission*; varje forskare ska då kunna visa på sitt bidrag. Experterna betygsätter visserligen individer, men resultaten av granskningen publiceras per *submission*, aldrig per individ. HEFCE publicerar dock i efterhand namnen på alla som ingått i en *submission*.

Kriterierna för bedömning av *output* är:

- a) *rigour* (stringens)
- b) *originality* (nytänkande) och
- c) *significance* (betydelse).

Som underlag för bedömning av aspekten *impact* ska minst två fallstudier per *submission* skickas in. När antalet heltidsekvivalenta forskare överstiger 15 får ytterligare en fallstudie skickas in, och ytterli-

gare en för var tionde heltidsekvivalent utöver detta. Vid några av studiebesöken uppmärksammades vi på att detta kan vara en eventuell tröskelmekanism som även påverkar aspekten *output*, så tillvida att mängden fallstudier kan komma att styra antalet forskare och *outputs* som skickas in till REF. Fallstudierna måste dock inte avse samma forskare som har ingått i aspekten *output*. Fallstudierna ska basera sig på forskningsresultat av minst 2* kvalitet som publicerats under perioden 1993–2013. 75 procent av bedömningen baseras på beskriven och bevisad påverkan i samhället utanför akademien. 25 procent av bedömningen görs på basis av de ansträngningar man gjort för att åstadkomma *impact* redovisad i en *impact template*. Aspekten *impact* följer institutionen där forskningen bedrevs och följer alltså inte med den enskilde forskaren vid en ev. flytt mellan lärosäten. *Impact* används som en aspekt för första gången i REF2014 och många upplever en osäkerhet kring hur bedömningen i realiteten ska gå till. För en grundligare genomgång av *impact* i REF se avsnitt 2.4.

Kriterierna för bedömning av *impact* är:

- a) *reach* (räckvidd) och
- b) *significance* (betydelse).

Som underlag för aspekten *environment* skickas kvantitativa uppgifter för perioden 1 augusti 2008 till 31 juli 2013 in. HEFCE:s REF-team förser institutionerna med vissa data från statistikmyndigheten HESA. Uppgifterna avser:

- a) antal doktorsexamina
- b) inkomster från externa forskningsmedel
- c) inkomster till forskningen ”*in-kind*”.

Utöver dessa data skickas också en beskrivning enligt mall av hela miljön som utgör en *submission*, avseende bestämd tidsperiod, med följande innehåll:

- a) översikt
- b) forskningsstrategi
- c) personal
- d) inkomster, infrastruktur och utrustning samt
- e) samarbete och bidrag till disciplinen.

Kriterierna för bedömning av *environment* är:

- a) *vitality* (vitalitet) och
- b) *sustainability* (hållbarhet).

Environment förefaller vara den minst kontroversiella av bedömningsaspekterna. Det är tillåtet att inkludera forskare som inte publicerat 3* och 4* *outputs*. Det kan t.ex. vara personer som handlett många doktorander fram till disputation. Däremot kan inte personal som enbart undervisar inkluderas. Det finns tydliga instruktioner för vad som krävs för att forskare ska vara kvalificerade för att räknas med och kunna inkluderas. Personaldelen utgör 25 procent av aspekten *environment*. Hela forskningsmiljön ska avspeglas, vilket ibland leder till konstgjorda narrativ ifall universitetets egen indelning inte motsvarar HEFCE:s indelning i utvärderingsenheter.

Utvärdering av forskarutbildning ingår i viss mån som en del av bedömningen av aspekten *environment*. Dels i det kvantitativa underlaget genom uppgifter om antal doktorsexamina, dels som narrativ i mallen för forskningsmiljön, där förutsättningar och genomförande av forskarutbildningen ska beskrivas i termer av t.ex. ”*effective and sustainable doctoral research training, evidence of a strong and integrated research student culture och evidence of CASE awards and application of technology generated by research students*”. Panelerna har i flera fall lagt till ytterligare bedömningsgrunder för forskarutbildningen. Doktorsavhandlingar tas inte in som underlag för bedömning. Doktoranders prestationer kan dock indirekt bli föremål för bedömning i de fall de är medförfattare på publikationer som ingår i en *submission*.

2.4 Utvärdering av impact i REF

Definition och avgränsning av impact

Impact definieras i REF mycket brett som ”an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia”. Impact ska inte inkludera inomvetenskapliga effekter; sådana effekter räknas som output eller i förekommande fall som del av *environment*. Inte heller inkluderas effekter på utbildning m.m. inom lärosätet i fråga.

Vid bedömning av *impact* finns inga begränsningar avseende hur stort geografiskt område som påverkas. En *impact* som omfattar t.ex. Hampton Court eller Nordirland (båda exemplen är faktiska, från piloten) kan räknas lika högt som en *impact* som omfattar hela världen eller mänskligheten⁹.

Den *impact* som ska redovisas och bedömas ska redan vara inträffad och bevisad under en viss specificerad tidsperiod (1 januari 2008–31 juli 2013), alltså ingår inte potential för framtida *impact*.¹⁰ Där emot kan den underliggande forskningen ligga så långt tillbaka i tiden som 1 januari 1993.

Initiativ att inkludera bedömning av impact

Som tidigare nämnts framfördes under perioden 2004–2009 en rad förslag – från bl.a. det engelska finansdepartementet och några större lärosäten – att förändra det tidigare RAE i framför allt två viktiga avseenden. För det första att avskaffa eller förenkla RAE:s peer review-princip till förmån för ett system som i högre grad baserades på kvantitativa mått och indikatorer, vilket senare förkastades, för det andra att i högre grad inkludera mått på forskningens *impact* (och inte enbart *output* som i RAE 2008)¹¹. HEFCE uppdrog till RAND Europe att genomföra en studie vars syfte var att ge en översikt över “*approaches to evaluating the impact of research*”.¹² En av de rekommendationer som gavs i RAND-rapporten var att använda en kombination av narrativ/fallstudier och indikatorer för bedömning av *impact*. Detta förslag gavs med hänvisning till det dåvarande planerade australiensiska utvärderingssystemet Research Quality Framework (RQF), vilket dock aldrig genomfördes p.g.a. ett regeringsskifte.

Den metod som HEFCE kom att föreslå för redovisning av *impact* var *case studies* (fallstudier), kombinerat med ett s.k. *impact statement* vars syfte skulle vara att “*give a rounded picture of the full range of impacts secured by the unit [of assessment] and illustrated in the case studies*” (*impact statement* har nu ersatts av en mer standardiserad *impact template*, se nedan). Indikatorer/mått på *impact* sägs kunna användas i form av underlag till fallstudier – men alltså inte utgöra själva redovisningen av *impact*.

Pilotstudie av impact

En pilotstudie av fallstudiemetodiken för redovisning och bedömning av *impact* i REF genomfördes under år 2010. Syftet med piloten var att testa följande fyra frågor: innehåll; bedömningsprocess och –kriterier; viktning: “*the weighting of the impact sub-profile, and how it will influence the overall outcomes*”; och frågor om genomförande: “*identifying the operational implications of assessing impact (including burden) for institutions, research users, panels and the funding bodies*”.¹³

29 lärosäten från hela Storbritannien deltog i piloten, som avsåg fem ämnesområden (utvärderingsenheter): *Clinical Medicine; Physics; Earth Systems and Environmental Sciences; Social Work and Social Policy & Administration*; och *English language & Literature*. Dessa utvärderingsenheter redovisade *impact* enligt den av HEFCE föreslagna modellen, och underlaget (*impact statement* plus fallstudier) bedömdes av fem paneler med ungefär lika stor andel vetenskapliga experter och använ-

⁹ HEFCE (2011, updated 2012) *REF2014 Assessment framework and guidance on submissions*. Annex A.

¹⁰ HEFCE (2011, updated 2012) *REF2014 Assessment framework and guidance on submissions*.

¹¹ HEFCE (2009/38). *Research Excellence Framework: Second consultation on the assessment and funding of research*.

¹² RAND Europe (Grant, J., Brutscher, P.-B., Kirk, S., Butler, L. & Wooding, S.), (December 2009): *Capturing Research Impacts – A review of International Practice*. Prepared for the Higher Education Funding Council for England.

¹³ HEFCE (March 2010): *Impact Pilot Exercise. Guidance to expert panels*.

darexperter. Resultat och erfarenheter från piloten dokumenterades (november 2010).¹⁴ Som goda exempel på fallstudier kan nämnas: ”Life-changing therapeutic intervention in patients with neonatal diabetes” (Clinical Medicine); ”Impact on the consumer electronics industry: The Floating Low-energy Ion Gun” (Physics); ”Conservation of Bumblebees” (Earth Systems and Environmental Sciences); ”The Review of the Public Administration in Northern Ireland” (Social Work and Social policy & Administration); och ”Henry VIII at Hampton Court Palace” (English). Piloten kompletterades med tre särskilda workshops under sommaren 2010, vilka syftade till att ytterligare belysa impact-frågor inom *Arts and Humanities och viss Social Sciences*.¹⁵

Såväl piloten som nämnda workshops anses ha gett ytterligare stöd för ”*the case study approach*” och för ”*the narrative approach*”, men sammanfattningsvis konstaterades att tillvägagångssättet måste kompletteras med utförlig vägledning, att (oönskade) konsekvenser för forskares beteende kan förekomma, och att man behöver vara uppmärksam på detta.

Så redovisar lärosätena impact¹⁶

Den *impact template* som ingår i lärosätenas redovisning ska avse hela utvärderingsenhetens ”*approach to impact during the assessment period*”. Den innehåller information om ”*context*”, ”*the unit’s approach to impact*”, ”*strategy and plans for supporting impact*” och ”*the relationship between the unit’s approach to impact and the submitted case studies*”. Omfattningen av *impact template* (normalt 3–5 sidor) står i proportion till personalantalet.

Huvudelementet är dock fallstudier. Antalet tillåtna fallstudier står i proportion till antalet personal som ingår i en *submission*: två fallstudier för upp till 14.99 heltidsekvivalenter, tre fallstudier för 15–24.99 heltidsekvivalenter osv., upp till sex fallstudier för 45 eller fler heltidsekvivalenter (plus ytterligare en fallstudie per varje ytterligare 10 heltidsekvivalenter). Detta är en förändring från piloten, där man begärde in en fallstudie per 10 heltidsekvivalenter.

- Kriterier för att fallstudier ska kunna tillgodoräknas är: *Impacten* ska ha inträffat under en angiven femårsperiod (i REF2014 åren 2008–2013)
- Forskningsaktiviteten ska ha inträffat under en 20-årsperiod föregående *impacten*, således inte tidigare än 1993. (I piloten användes ett 15-årskriterium, åren 1993–2008).
- Den underliggande forskningen ska vara av hög kvalitet (2* eller högre) – ett kvalitetskriterium som också kommer att bedömas av panelen.
- Det ska göras klart att en specifik angiven forskningsaktivitet vid den enhet som en *submission* omfattar ska ha gjort en ”*distinct and material contribution*” till *impacten*. I riktlinjerna finns en rad preciseringar av vad detta kan innebära. Det finns t.ex. möjlighet att inkludera ”*interim impacts*” om det inte är realistiskt att redovisa sluteffekter, vid långa/tidsmässigt utdragna effektkedjor.
- Det ska finnas evidens för att påvisa sambandet mellan forskningsaktiviteten och *impact* enligt ovan.

¹⁴ En rapport med erfarenheter och synpunkter från panelerna (ordförandena i de fem pilotpanelerna *Research Excellence Framework impact pilot exercise: Findings of the expert panels* (2010); Lärosätenas synpunkter och erfarenheter sammanställdes i en trappport Technopolis (November 2010) *REF Research Impact Pilot Exercise. Lessons-Learned Project: Feedback on Pilot Submissions*, Final report. Fem rapporter redovisades (en för varje utvärderingsenhet i piloten) med 5-6 ”*impact case studies*” för respektive utvärderingsenhet, vilka hade fått goda omdömen av respektive pilotpanel och där syftet var att ”*indicate good practice in terms of the pilot submissions*”. Se vidare <http://www.ref.ac.uk/background/pilot/>

¹⁵ HEFCE (December 2010). Workshops on the impacts of research in the practice-based creative and performing arts, the humanities and social sciences.

¹⁶ Citaten i detta avsnitt är från HEFCE (2011, updated 2012) *REF2014 Assessment framework and guidance on submissions*, om inte annat anges.

Fallstudierna ska redovisas enligt följande mall:

1. *Summary of the impact* (max. 100 ord)
2. *Underpinning research* (max. 500 ord)
3. *References to the research* (max. 6 referenser)
4. *Details of the impact* (max. 750 ord)
5. *Sources to corroborate the impact* (max. 10 referenser)

Jämfört med piloten har i REF2014 ett tydligare och större ansvar lagts på lärosätena att redovisa, strukturera och avgränsa information på ett sådant sätt att panelerna kan göra sina bedömningar utan att behöva söka eller efterfråga ytterligare information – bl.a. genom detaljerade anvisningar kring referenser (antal, utformning etc.).¹⁷

Erfarenheter, tidsåtgång och arbetsinsats på lärosätena

En utvärderingsrapport från Technopolis om feedback från de deltagande lärosätena i impactpiloten redovisar mycket positiva resultat och uppfattningar: ”*In conclusion, the feedback from Pilot Institutions has confirmed the feasibility of the approach tested through the pilot exercise; it is clear that HEIs (Higher Education Institutions) can document non-academic impacts and that in doing so a great majority will derive insight and local benefits. The use of subject-specific Assessment Panels means a generic model should work for all disciplines.... Overall, we recommend HEFCE move forward with the generic model tested through the pilot, retaining the impact case studies as the centre-piece of each submission, for assessment by subject-specific Assessment Panels.*”¹⁸

En slutsats från piloten är att redovisning av *impact* medför en reell och icke försumbar arbetsinsats för lärosätena men att arbetsinsatsen bedömdes som rimlig i förhållande till värdet. Ett genomsnitt på 0,5 persondagar per heltidsekvivalent vid respektive enhet som omfattas av en *submission* redovisas, vilket motsvarar i genomsnitt ca 6 persondagar per fallstudie för de som producerar tio eller färre fallstudier. Vid några av de lärosäten som vi träffade vid studiebesöket i februari 2013 menade man dock att Technopolis-rapportens slutsatser är en klar underskattning av tids- och arbetsinsatsen.

Användarexperternas roll i panelerna¹⁹

Impact submissions bedöms i första hand av de 36 underpanelerna. I underpanelerna ingår förutom vetenskapliga experter (”*academics*”, som är i majoritet) även användarexperter (”*users*”). Se vidare avsnitt 2.5.

Panelerna, som alltså är blandade, bedömer samtliga tre aspekter (*output*, *impact* och *environment*). Andelen användarexperter tycks ligga runt 10–15 procent (dvs. klart lägre än i pilotstudien där dock enbart *impact* bedömdes). De vetenskapliga experterna förväntas bedöma alla aspekter, medan användarexperterna endast förväntas bedöma *impact*. Dock kan de sistnämnda i undantagsfall bidra till bedömningen av de andra aspekterna om de har rätt kompetens.

Utvärderingen av impactpiloten visade att sammansättningen av panelerna fungerade bra och att bedömningarna i princip inte skilde sig åt mellan de vetenskapliga experterna och användarexperterna.²⁰ Att finna lämpliga användarexperter till panelerna har visat sig möjligt inom praktiskt taget alla ämnesområden (utvärderingsenheter) och detta beskrevs också under vårt studiebesök i februari 2013 som den minst kontroversiella delen i REF:s introduktion av *impact*.

¹⁷ Intressant är vidare att i piloten bedömdes ett litet antal fallstudier som oklassificerade, antingen för att informationen var ofullständig eller att *impacten* inte hade inträffat inom den specificerade tiden, eller också att det var frågan om “*case studies that displayed no substantive impact, impact only within academia, only potential impact (rather than impact already achieved), or made vague claims about impact without sufficient supporting evidence*”. (HEFCE *Research Excellence Framework impact pilot exercise: Findings of the expert panels* (2010)). Således har såväl inomvetenskapliga/inomakademiska effekter som potentiella effekter bedömts som ”*unclassified*” av panelerna i piloten.

¹⁸ Technopolis (November 2010). REF Research Impact Pilot Exercise. Lessons-Learned Project: Feedback on Pilot Submissions, Final report.

¹⁹ Citaten i detta avsnitt är från HEFCE (2012) *REF2014 Panel criteria and working methods*, om inte annat anges.

²⁰ Research Excellence Framework impact pilot exercise: Findings of the expert panels (2010).

Kriterier och bedömningsskalor för bedömning av impact

Underpanelernas uppgift vad gäller *impact* är att bedöma kriterierna ”*reach and significance*” av *impact* “*on the economy, society and/or culture that were underpinned by excellent research conducted in the submitted unit, as well as the submitted unit’s approach to enabling impact from its research*”. De två kriterierna ska bedömas gemensamt och inte separat.²¹

Alla huvudpaneler har exemplifierat/specificerat olika typer av *impact* som kan vara aktuella.²² För t.ex. område D (Arts and Humanities) exemplifieras 18 olika typer av *impact* (t.ex. “*generating new ways of thinking that influence creative practice*”, “*preserving, conserving and presenting cultural heritage*”, “*providing expert advice to governments, NGOs, charities and the private sector in the UK and internationally, and thereby influencing policy and/or practice*”).

Kommentarer angående impact

I samband med studiebesöket i Storbritannien i februari 2013 framkom en mängd synpunkter och kommentarer på *impact*-delen i REF. Några av de viktigare och mer frekventa synpunkterna var:

- 1) *Impact* har aldrig förekommit förut. Trots noggranna instruktioner och utprovningar från HEFCE:s sida finns betydande osäkerhet på lärosätena på hur *impact* de facto kommer att bedömas av panelerna.
- 2) Uppfattningen finns att det framförallt är finansdepartementet och näringslivet som har tryckt på för att få med *impact* som en substantiell del i bedömningen inför allokering av basanslaget. Det finns ett politiskt tryck från regeringen att forskningsresultaten ska komma samhället till nytta och leda till tillväxt så snart som möjligt. Av den anledningen måste en bra balans av kortsiktighet och långsiktighet uppmuntras, vilket har lett till *impact*agendan. Finansdepartementet har velat att *impact* ska ha en substantiell andel av viktningen, 25–30 procent, ”*to make a difference*”. Eftersom *impact* är ett oprövat kort, är en vanlig uppfattning att viktningen snarare borde vara 15 eller rentav 10 procent. Det var också det råd vi fick på flera håll vad gäller viktning i ett eventuellt svenskt system. Den andel som HEFCE till slut fastställt (20 procent) ses som en kompromiss och många tror att andelen kommer att höjas i nästa REF 2020.
- 3) Trycket på akademien att tydligare redovisa *impact* finns inte bara via HEFCE/REF utan mer generellt. Forskningsråden var först ut med detta i form av krav på s.k. ”*impact statements*” i alla ansökningar, liksom genom uppbyggnad av de särskilda rapporteringssystemen e-val/ResearchFish (se vidare avsnitt 2.6) respektive ROS (Research Outcomes System) som används.²³
- 4) Många vi träffade menar att fallstudier är ett bättre och mer rättvisande sätt att redovisa och bedöma *impact* på än indikatorer. Det är en vanlig uppfattning att akademien har slagit vakt om fallstudier som metod inför hotet att bli bedömd enbart utifrån ”*metrics*”, och att HEFCE har stått på akademins sida i detta. Det finns också uppfattningar att *impact*-piloten har varit riggad för att visa på ett positivt resultat och mindre tidsåtgång än vad som i praktiken blir fallet.
- 5) En kritik som riktas mot fallstudie-metodiken är att den innebär ett mycket omfattande arbete, inte minst för lärosätena, att ta fram och redovisa information som sedan inte kan återanvändas, vilket skulle tala till förmån för ett mer indikatorbaserat system där information insamlas och redovisas mer kontinuerligt. HEFCE tillbakavisar denna kritik och menar att det visst finns planer för återanvändning av fallstudier.
- 6) Det finns en tydlig generationsklyfta i synen på att redovisa och bedöma *impact*. Äldre akademiker bjuder motstånd medan det av många yngre forskare ses som mer naturligt.
- 7) Fler lärosäten har genomfört mycket omfattande interna s.k. ”*mock exercises*” för att kartlägga *impact*. En inte ovanlig uppfattning är att det har gått lättare än väntat att få fram exempel på

21 HEFCE (2011, updated 2012) *REF2014 Assessment framework and guidance on submissions*, Annex A.

22 HEFCE REF01:2012 *Panel criteria and working methods*

23 Förkortningar för forskningsråden: MRC *Medical Research Council*, STFC *Science and Technology Facilities Council*, NERC *Natural Environment Research Council*, AHRC *Arts and Humanities Research Council*, BBSRC *Biotechnology and Biological Sciences Research Council*, EPSRC *Engineering and Physics Science Research Council* och ESRC *Economic and Social Research Council*.

impact, däremot finns frågetecken kring kvaliteten på fallstudierna liksom kring redovisningen av evidens/orsakssamband.

- 8) I kölvattnet på introduktionen av *impact* i REF har det vuxit fram en ny konsultmarknad, ”*impact-konsulter*” vilka erbjuder tjänster som att t.ex. formulera slagkraftiga fallstudier, liksom att kartlägga, spåra och verifiera *impact*.

2.5 Expertpanelerna

Jämfört med i RAE2008 har man i REF2014 färre och bredare paneler som arbetar på ett mer likartat sätt. Antalet experter är dock ungefär detsamma, totalt ca 1000. Den absoluta majoriteten av panelernas medlemmar är brittiska. Den officiella versionen är att Storbritannien är så pass stort att det går att undvika jäv och att de nationella experterna har tillräckligt med internationella perspektiv. På en del lärosäten uttrycktes dock tveksamhet om att det går att undvika jäv, även om man inte alls deltar vid bedömning av det egna lärosätet. En fördel med att använda brittiska experter är att de känner till systemet väl. Huvudargumentet mot att i högre utsträckning anlita internationella experter är att det anses som alltför kostsamt.

De 36 underpanelerna har ca 10–30 ledamöter vardera. De flesta ledamöterna är vetenskapliga experter, men mellan två och sex är s.k. användareexperter. En poäng med att akademiker och användareexperter sitter i samma panel (och inte i separata) uppges bl.a. vara de konstruktiva diskussionerna som förs om vad som utgör kvalitet. Varje underpanel genomför detaljerad granskning av alla *submissions* inom respektive utvärderingsenhet. Underpanelerna arbetar under ledning av fyra huvudpaneler, med ca 15–25 ledamöter i varje. I huvudpanelerna ingår underpanelernas ordförande, dessutom finns där internationella vetenskapliga experter och användareexperter. Representanter för forskningsråden finns med som observatörer i huvudpanelerna.

Ordförandena för huvudpanelerna och underpanelerna har tillsatts genom en öppen ansökningsprocess. De övriga vetenskapliga experterna i panelerna tillsattes hösten 2010 genom en öppen nomineringsprocess, där 1950 instanser inbjöds att nominera. Inga lärosäten fick nominera experter.²⁴ Panelerna kompletteras med ett begränsat antal ytterligare ledamöter under 2013, vid behov, för att utöka panelens bredd och djup. Detta tar man ställning till när man tagit del av lärosätenas *submission intentions* där de gör en uppskattning av vad de avser att skicka in. Det finns farhågor på lärosätena att panelerna är för breda i sin sammansättning och inte tillräckligt specialiserade för att kunna göra en rättvisande bedömning. Rekrytering av användareexperter sker under 2013 efter nominering. Som tidigare nämnts deltar användareexperter endast i bedömningen av *impact*, men om en individ har dubbla kompetenser kan hen bedöma både vetenskaplig kvalitet och *impact*. Panelerna är sammansatta för att även kunna bedöma tvärvetenskap, dessutom kan samråd mellan paneler ske i begränsad utsträckning.

Panelernas uppgift är att ta fram en kvalitetsprofil för varje *submission*. De har inte i uppgift att ge rekommendationer, bara en kort skriftlig återkoppling (ca en halv till en sida) som inte är offentlig. De kan inte begära in mer material utöver det som skickats in, inte heller be om fysiska möten eller intervjuer. Panelerna träffas vid ca 15 tillfällen, bl.a. för att kalibrera sina bedömningar och bli klara över var gränsen mellan de olika betygsstegen ligger (vissa tenderar att övervärdera, andra att undervärdera). Samma generiska kriterier för bedömning gäller för alla paneler. Inom varje huvudpanel operationaliserar man dessa kriterier och det är möjligt att lägga till särskilda kriterier för vissa underpaneler om det är motiverat p.g.a. disciplinens särart.

Vid studiebesöket intervjuade vi flera personer som ingår i panelerna. Vi fick uppfattningen att de tagit på sig uppdraget av solidaritet, intresse och för prestigens skull. Uppdraget följs inte av någon större ekonomisk ersättning. Arbetet innebär att experterna får avsätta upp till ett halvår för att läsa alla *submissions* inklusive ca 200–800 *outputs* (i extrema fall upp till 1000 st.) per expert och typiskt ca

24 HEFCE identifierade 1950 s.k. ”*nominating bodies: academic associations and other bodies with an interest in research and in nominating candidates to be REF panel members*”. Härutöver utvidgade man inbjudan att nominera panelmedlemmar till ”*any other association or organisation with a clear interest in the conduct, quality, funding or wider benefits of publicly funded research – except for mission groups, individual UK HEIs and groups within or subsidiaries of individual UK HEIs*.” HEFCE (2010) *Units of assessment and recruitment of expert panels*.

25–30 fallstudier som visar på *impact*. Det sker inom ramen för deras ordinarie anställning och de blir under den tiden ”befriade” från undervisning och administrativa åtaganden på sina lärosäten.

2.6 Tekniska lösningar för inrapportering

I Storbritannien finns generellt ett stort fokus på att kunna visa för samhället att offentliga forskningsmedel använts på ett ansvarsfullt sätt och att de lett till resultat. Sådana uppgifter används också vid tilldelning av nya medel till forskningen. Detta är en av de underliggande drivkrafterna inom REF2014, men också för forskningsråden, som även de förväntas redogöra för resultat från den forskning de finansierar. Mot bakgrund av detta har ett antal system utvecklats för att inhämta information om forskningsresultat, dvs. elektroniska system där forskare rapporterar in t.ex. publikationer, patent och samarbeten.

År 2009 lanserade det medicinska forskningsrådet (MRC) sitt återrapporteringssystem e-Val som sedan dess har utvecklats till en gemensam tjänst kallad Researchfish²⁵ som numera används av flera forskningsfinansiärer. Ett antal andra forskningsråd i Storbritannien, hade vid samma tid påbörjat ett arbete med att utveckla ett motsvarande system, Research Outcomes system (ROS²⁶). Dessa valde att inte ansluta sig till Researchfish p.g.a. att det system man själv börjat utveckla var öppet tillgängligt, vilket man bedömde skulle bli billigare. Man var inte heller helt i fas med MRC som behövde lansera ett system snabbare än övriga forskningsråd. Parallellt med detta skapade HEFCE ett rapporteringssystem för REF2014. Inför den utvärdering som utförs inom REF2014 levererar universiteten som nämnts en rad olika typer av data. Uppgifterna skickas främst in elektroniskt men man har även utfört ett omfattande arbete med att inrätta ett lager för att ta emot fysiska *outputs* (ca 25 procent av alla *outputs*). Data om aspekten *environment* bygger främst på standardiserad data från HESA, den statistikansvariga myndigheten för högskoleväsendet.

Någon motsvarighet till den svenska publikationsdatabasen SwePub finns inte i Storbritannien. Det finns inte heller någon egen centraliserad citeringsdatabas motsvarande den som upprätthålls av Vetenskapsrådet²⁷. Detta innebär att man saknar en central databas där publikationsdata kan hämtas eller analyseras. Det är inte heller alla brittiska universitet som har egna publikationsdatabaser. HEFCE:s bibliometripilot visar att det vore resurskrävande för lärosätena att införa sådana.

Resultatet av detta är att man i Storbritannien skapat ett antal parallella system som i huvudsak har samma syfte och till stora delar innehåller samma data. En lärdom av studiebesöket i Storbritannien är att detta inte är ett önskvärt scenario för Sverige. Forskarna/lärosätena ska inte behöva rapportera samma uppgifter i flera olika system och vi ska inte behöva ha flera organisationer som underhåller liknande, men separata system.

Förutsättningarna för att skapa ett enhetligt system i Sverige är relativt goda. De statliga forskningsråden (Vetenskapsrådet, FAS och Formas) utvecklar just nu ett gemensamt ansöknings- och återrapporteringssystem, Prisma²⁸, som kommer att vara tillgängligt för fler organisationer att ansluta sig till och använda. Prisma innefattar ett online-CV där forskarna kan lägga in alla sina forskningsresultat som de sedan kan koppla till återrapporteringar för enskilda projekt. Om Sverige ska införa en motsvarighet till det brittiska REF2014 eller annat nationellt kvalitetsgranskningssystem bör därför Prisma, och framför allt det CV som finns som en del av Prisma, utredas som möjligt system för inhämtning av information om forskningsresultat. Principen för det svenska systemet bör vara att data endast ska matas in en gång, men kunna användas för flera ändamål (se även utredningen Forskningsinformation 3.0²⁹). I detta bör också framförallt samspelet mellan Prisma och SwePub beaktas.

25 <https://www.researchfish.com/> (Används bl.a. av MRC, STFC, British heart foundation, Arthritis research UK, och Alzheimers research UK)

26 <http://www.rcuk.ac.uk/research/researchoutcomes/pages/ROS.aspx> (Används för discipliner som finansieras av bl.a. AHRC, BBSRC, EPSRC, ESRC och NERC)

27 Vetenskapsrådets databas är i och för sig internationell och tillgänglig att köpa in från Thomson Reuters för den som så önskar.

28 Prisma väntas lanseras under hösten 2013.

29 Vetenskapsrådet dnr 111-2012-203.

2.7 Hur utvärderingen omsätts till resurser

På nationell nivå

HEFCE ansvarar för genomförandet av utvärderingen av forskningen för alla fyra länder inom Storbritannien men resurstilldelningen sköts av respektive lands egna myndigheter (SFC, HEFCW, DELNI och HEFCE [dock en annan enhet inom HEFCE än den som sköter REF])³⁰ (se Figur 1). Länderna bestämmer själva hur resultaten omsätts till resurser och det är möjligt för ett land att ändra fördelingsnyckel efter själva utvärderingen och mellan utvärderingstillfällena. Detta hände t.ex. inför 2012/13 års fördelning då 2* forskning (baserat på RAE 2008) inte längre skulle räknas in som underlag vid beslut om fördelning av resurser utan endast forskning som fått betygen 3* och 4*. Den exakta algoritmen som kommer att användas för resursfördelning 2014 baserad på REF resultatet är ännu okänd.

England³¹

I England utgör basanslaget ca 30 procent³² av den totala forskningsfinansieringen till lärosätena. 2/3 av basanslaget baseras på resultatet av RAE/REF. Resurserna tilldelas universitetet som helhet och är inte öronmärkta till särskilda ämnen, men varje utvärderingsenhetens bidrag till universitetets aggregerade basanslag är känt. Tre parametrar används för att räkna fram resurstilldelningen: kvalitetsprofilen, volymen och kostnadsfaktorn.

Kvalitetsprofilen tas, som tidigare nämnts, fram från tre underprofiler. För RAE2008 var dessa *output*, *environment* och *peer esteem*, i REF2014 är dessa *output*, *impact* och *environment*. Profilen uttrycks som hur stor andel av forskningen som har bedömts ligga vid varje betyg på skalan 1* till 4*. Volymen avser antalet forskare (heltidsekvivalenter) som ingick i *output*. Kostnadsfaktorn viktas som tidigare nämnts till fördel för mer resursintensiva forskningsämnen (1,6 för natur, teknik och medicin; 1,3 för vissa samhällsvetenskaper och 1,0 för övriga samhällsvetenskaper och humaniora).

Fördelningen av basanslaget 2012/13 baseras på resultatet av RAE2008 och sker i flera steg. I varje steg sker fördelningen till varje enskild enhet relativt hur alla andra enheter har presterat då storleken på potten bestäms utifrån andelen av det totala resultatet. Först bestäms storleken på potten (andelen av den totala potten) för varje ämnesområde (motsvarande 15 huvudpaneler i RAE). Detta görs utifrån volymen multiplicerat med andelen forskning som uppnår nivå 2* eller högre multiplicerat med kostnadsfaktorn. (Åren 2009/2010/2011 fanns det dock ett politiskt beslutat randvillkor att andelen av finansieringen till STEM-ämnen³³ inte kunde minskas).

Varje huvudpanels pott fördelas i sin tur mellan de tillhörande utvärderingsenheterna (underpanelerna) och inom varje utvärderingsenhet till lärosätena. Detta görs utifrån samma formel där kvalitetsprofilen multipliceras med kostnadsfaktorn och volymen. Sedan 2012/13 räknas endast forskning med betygen 3* och 4* in. Skillnaden här (jämfört med steg 1) är att 4* räknas tre gånger mer än 3*. Varje universitets tilldelning av kvalitetsrelaterad finansiering (quality related – QR) räknas genom att aggregera tilldelningen till varje utvärderingsenhet för universitetet.

Hur lärosätenas kvalitetsprofiler räknas fram

De tre aspekterna *output*, *impact* och *environment* vägs aritmetiskt samman till en gemensam kvalitetsprofil "Overall Quality Profile", som anger hur stor procentuell andel av respektive utvärderingsenhetens sammanvägda betyg för bedömningsaspekterna som når upp till respektive betygsnivå.

30 SFC – Scottish funding council, HEFCW – Higher Education Funding Council for Wales, DELNI – Department of Education and Learning Northern Ireland.

31 Guide to Funding – How HEFCE allocates its funds, HEFCE 2010

32 <http://www.hesa.ac.uk>

33 Science, technology, engineering and mathematics.

Tabell 1. Exempel på kvalitetsprofiler.³⁴

Utvärderingsenhet A (t.ex. fysik)	Kategori A personal (hel-tidsekvivalenter) ingående i <i>submission</i>	Procentandel av forskningen som ingår i <i>submission</i> som bedöms uppnå standarden för respektive betyg				
		4*	3*	2*	1*	Oklassificerad
Universitet X	50.45	18	41	25	16	0
Universitet Y	65.2	12	32	45	10	1

Underpanelerna sammanställer dessa övergripande kvalitetsprofiler genom att först sammanställa underprofiler för varje aspekt (*output, impact, environment*). Underprofilerna aggregeras enligt en kumulativ avrundningsprocess.

På lärosätetsnivå

Den interna resurstilldelningen ser olika ut på olika brittiska lärosäten. Vissa lärosäten betonade vikten av en transparent intern tilldelning av de kvalitetsrelaterade (QR)-anslagen, medan andra tvärtom menade att det var ett medvetet val att inte göra detta transparent för att kunna motverka ”*game-playing*” internt. På några av universiteten vi besökte gick QR-intäkterna till den organisatoriska enhet som ”tjänade in” dem, dvs. universiteten tillämpade samma fördelningsprincip som HEFCE även internt. Några behöll en viss procentsats (upp till 25 procent i ett fall) av QR-intäkterna centralt för strategiska eller andra ändamål. Några uppgav också att det var svårt ur ledningssynpunkt att omfördela resurser på annat sätt internt. Andra lärosäten vi besökte angav att det inte fanns någon direkt koppling mellan den interna resurstilldelningen och HEFCE:s tilldelning. Strategiska argument som framfördes för att stödja enheter som inte presterat väl i REF kunde vara att dessa var viktiga för att bygga upp en kritisk massa, eller för att universitetet ville bevara sin disciplinära bredd.

2.8 Kostnader för framtagning, implementering och drift av systemet

Det är svårt att få klarhet i hur stora kostnaderna för systemet är. På departementet BIS sade man att kostnaderna för RAE2008 var sammanlagt £ 40 miljoner för att fördela forskningsresurser om sammanlagt £ 20 miljarder, således 0,2 procent. På HEFCE sa man däremot att kostnaderna för RAE2008 var £ 12 miljoner. HEFCE uppskattade kostnaderna för REF2014 till £ 14 miljoner (för en 6-årig cykel) (inklusive pilotstudierna), varav huvuddelen avser kostnad för panelerna. Universitetens samlade kostnader uppskattas av HEFCE till £ 40–50 miljoner i direkta kostnader. I detta är inte ev. intäktsbortfall för ”utebliven forskning”, i samband med såväl REF som lärosätenas egna förutvärderingar, inräknat. I och med att man använder i huvudsak nationella experter, kommer ett stort antal personer att vara upptagna med panelarbete under ca ett halvår.

Lärosätena har oftast inte räknat på kostnaderna. Några reaktioner vi fick från dem var att ”arbetet (med REF) måste ändå göras”. En del menade att de ändå skulle gjort det för att komma fram till ”*best-practice*” och att ”kostnaderna är enorma, men det är även fördelarna”. Dock menade flera att det borde finnas enklare och mer kostnadseffektiva sätt att göra utvärderingen på. Exempelvis gjorde University of Reading en grov uppskattning att kostnaderna var £ 0.5 miljoner per år under två år. För den enskilda forskaren krävs inte så mycket merarbete, så länge hen producerar fyra *outputs* av hög kvalitet under en viss tidsperiod. Den enskilde forskaren har inte nödvändigtvis så mycket att vinna på REF eftersom en *submission* sker på institutionsnivå och REF-pengarna kommer till lärosätet centralt. Den mesta hanteringen sköts av ledningen centralt och på institutionsnivå och av centrala, administra-

³⁴ För närmare detaljer se:

<http://www.ref.ac.uk/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf>, s. 6–7.

tiva REF-funktioner. Något som nämndes av flera var känslomässiga kostnader och psykosociala arbetsmiljöproblem som konsekvenser av den interna urvalsprocessen inför REF.

2.9 Konsekvenser av RAE/REF

Frågan är om systemet uppnår sina syften? Ett kort svar är – ja. Vid i stort sett samtliga intervjuer vi gjort har man haft uppfattningen att så är fallet.

När det gäller syftet att premiera excellent forskning är flera faktorer kvalitetsdrivande: det yttre trycket att visa på kvalitet i forskningens resultat, diskussionen om vad som utgör kvalitet, utvärderingen i sig och den urvalsprocess som äger rum på lärosätena inför en RAE/REF. Någon menade att REF troligtvis påverkar systemet mer innan det genomförs 2014 än efter. Samtidigt är det svårt att strikt föra i bevis att det är RAE/REF-systemet som i sig drivit mot excellens i forskningen. Den kvalitetsdrivande effekten har dock planat ut efter återkommande utvärderingar under 25 års tid, effekten var störst i början. Enligt HEFCE har Storbritannien det mest produktiva statligt finansierade forskningssystemet bland G8 länderna. På politisk nivå är man också nöjd med att Storbritannien behållit sin position vid internationella jämförelser av bibliometri och på internationella rankinglistor.

När det gäller syftet att visa på att offentliga forskningsmedel används på ett ansvarsfullt sätt är systemet tillfyllest för de krav som ställs av t.ex. finansdepartementet. Att införa *impact* som en aspekt är ett sätt att ytterligare svara mot skattebetalarnas, näringslivets och politikernas krav på att visa att pengarna till forskning är motiverade.

Som underlag för fördelning av merparten av basanslaget anser de flesta att RAE/REF systemet fungerar väl, eftersom det är robust och premierar kvalitet i prestation. Det finns dock vissa kritiska röster som accepterar utvärderingsmetoden som sådan, men inte anser att den lämpar sig för resursfördelning. Den kommande omvandlingen av REF-resultatet till fördelning av basanslaget är medvetet icke-transparent eftersom HEFCE ska kunna ändra i algoritmen p.g.a. politiska överväganden och för att undvika ”*game-playing*” (vilket systemet dock inte tycks undvika). Fördelningsnyckeln bestäms alltid efter att RAE/REF resultaten kommit in, vilket vi hörde kritik av på lärosätena. RAE/REF-resultaten kan ha olika betydelse för olika discipliner, t.ex. får konst, humaniora och samhällsvetenskap en stor del av sina pengar till forskning via basanslagen, medan t.ex. biovetenskaper i huvudsak finansierar sin forskning med externa medel. Lärosätena uppfattar det som om RAE/REF-resultaten indirekt påverkar framgången för ansökningar till forskningsråden, även om det inte ingår i forskningsrådets instruktioner till sina experter.

Det finns ingen uttalad policy att koncentrera resurserna, men det blir en konsekvens av RAE/REF. Systemet ger inte några drastiska omfördelningseffekter mellan lärosätena, men förstärker ytterligare den befintliga koncentrationen av resurser till de fyra (av totalt drygt 160 st.) mest excellenta lärosätena (Oxford, Cambridge, Imperial College och University College London). Det är samma fyra lärosäten som också erhåller mest forskningsmedel från forskningsråden. Hälften av basanslagen som fördelas på basis av RAE2008 går till dessa fyra. Det finns dock en viss omfördelning av resurser i ”mittskiktet” av lärosäten. RAE/REF förmår även att identifiera små men excellenta miljöer s.k. ”*pockets of excellence*” som då premieras. På HEFCE menade man att det inte är hållbart att alla lärosäten ska vara forskningsintensiva, då sprids resurserna för tunt. Inga regionala hänsyn tas heller i utvärderingssystemet.

REF hindrar inte att forskning av låg kvalitet produceras eftersom lärosätena själva styr över sin interna fördelning. Vare sig HEFCE eller forskningsråden finansierar forskning av låg kvalitet utan det görs i praktiken av studentavgifter. Att skicka in *outputs* från forskning med låg kvalitet till REF straffar sig inte direkt finansiellt, men det spelar roll för ryktet. Andelen forskare och *submissions* per lärosäte publiceras efter REF och används av dem som gör rankinglistor. Även om en stor del av basanslaget baseras på resultatet av REF förefaller prestige, och den därmed sammanhörande attraktionskraften, vara en minst lika viktig drivkraft för lärosätena. På HEFCE menar man att lärosätena genom den externa granskningen i RAE/REF får ett underlag att lägga ned enheter som de egentligen själva visste inte fungerade tillräckligt bra.

Farhågor har rests kring taktisk rekrytering av forskare inför REF. Enligt HEFCE, som för övrigt anser att mobilitet är positivt, sker detta inte i någon större omfattning. Däremot har REF lett till upp-

drivning av lönerna, eftersom forskarna förhandlar om högre lön för att stanna kvar på lärosätet. Det faktum att *output* följer individen medan *impact* tillgodoräknas lärosätet där forskningen bedrevs, gör att dessa två aspekter motverkar varandra, vilket kan påverka mobilitet och taktisk rekrytering.

Mycket tid på lärosätena ägnas åt ”*game-playing*”, dvs. att variera de tre aspekterna *output*, *impact* och *environment* i förhållande till varandra, för att optimera utfallet vad gäller pengar och/eller prestige för egen del. Den frihet lärosätena har att vara taktiska leder till en mycket omfattande och tidsödande selektionsprocess. Några var kritiska till detta och menade att det vore bättre om de istället skulle uppmanas att skicka in all forskning (vilket å andra sidan skulle innebära mer arbete för panelerna).

Systemet driver mot fyra *outputs* av god kvalitet per inrapporteringsperiod, vilket får olika konsekvenser för olika ämnen (beroende på om fyra *outputs* är ett större eller ett mindre antal än normalt). Det leder därmed till ändrade publiceringsmönster, t.ex. att det blir vanligare att skriva artiklar inom det humanistiskt-samhällsvetenskapliga området. Inom andra ämnesområden kan det möjligen leda till en rörelse från korta, uppstyckade artiklar (”*least publishable unit*”) till större, mer övergripande artiklar. En del ansåg dock att den enskilda forskarens drivkraft inte påverkades nämnvärt av RAE/REF, man vill ändå publicera sig i de bästa tidskrifterna. Systemet leder till en allmän volymökning av publikationer, speciellt i slutet av inrapporteringsperioden, vilket orsakar längre ledtid fram till publicering p.g.a. köbildning hos förlag och tidskrifter.

Samarbetsmönster kan påverkas p.g.a. reglerna för tillgodoräknade av *outputs*, där mer omfattande samarbeten (fler än sex personer) och samarbeten mellan forskare inom institutionen kan missgynnas.

Det finns en viss osäkerhet kring hur tvärvetenskaplig forskning ska hanteras för att inte missgynnas av systemet, även om man försökt vidta åtgärder genom panelernas sammansättning och genom samråd mellan paneler vid behov.

Det blir en ökad konkurrens om experter som ska rekryteras till REF, till forskningsrådets ordinarie ansökningsprocesser, till granskning av publikationer och till lärosätenas egna kvalitetsutvärderingar.

Införandet av *impact* som en aspekt i REF leder till en förändrad kultur på lärosätena då de nu bemödar sig om att försöka åstadkomma och spåra konsekvenserna av sin forskning i ett vidare perspektiv.

En annan gynnsam effekt är att RAE/REF sätter igång strategiarbeten på lärosätena och att det skapas forskningsmiljöer. Det är dock inte självklart att det blir synergieffekter mellan universitetets eget strategiarbete och de strategier de måste ha för att komma väl ut i RAE/REF. Lärosätena har ofta vidare strategier än så. Enligt uppgifter vid intervjuerna kan inte RAE/REF till fullo täcka lärosätenas egna behov av att belysa och följa upp sin verksamhet, utan behovet av egna utvärderingar kvarstår.

Vanligtvis har forskare både forskning, undervisning och administration i sitt anställningskontrakt, men det finns en tudelning mellan i huvudsak forskande och i huvudsak undervisande personal som förstärks av REF.

Det finns tecken på att systemet leder till vissa arbetsmiljöproblem, p.g.a. stress och press att tillhöra de utvalda forskare som kan ingå i REF och att inte bli exkluderad för att man underpresterat. Att inkluderas eller exkluderas kan vara avgörande för karriären. Universiteten är skyldiga att ha en ”*code of practice*” för urvalet av forskare. Lärosätena har olika strategier för hur man handskas med dem som underpresterar, en del får extra stöd och ledning, för en del skrivs kontrakten om så att de bara undervisar och administrerar (inte forskar). En del förbättrar sina prestationer markant medan andra väljer att sluta. Fackföreningen UUS (motsvarigheten till SULF) är starkt kritisk till systemet.

3. NATIONELLA KVALITETSGRANSKNINGSSYSTEM AV FORSKNING I ANDRA LÄNDER

3.1 Dessa länder har vi studerat

Prestationsbaserade resurstilldelningssystem har införts i flera länder, och det finns länder som genomför nationella kvalitetsgranskningar av forskningskvaliteten vid universitet och högskolor utan att det är (direkt) kopplat till resurstilldelning. Länderna som beskrivs här är de nordiska länderna³⁵ (förutom Island), Australien³⁶, Nya Zeeland³⁷, Italien³⁸, Flandern³⁹ (en av Belgiens två självstyrande regioner) och Nederländerna⁴⁰. Storbritannien diskuterades i föregående kapitel. Andra länder där nationella kvalitetsgranskningssystem förekommer men som inte ingått i kartläggningen är t.ex. Hongkong⁴¹, Tyskland⁴² och Frankrike⁴³. Tyskland och Nederländerna har nationella system utan att de är kopplade till resurstilldelning.

3.2 Syften och process för att utveckla systemet

Kvalitet, excellens och nytta är ord som ofta återkommer i motiveringar för ett prestationsbaserat resurstilldelningssystem. Exempelvis i Finland är målet ”ett mer kvalitetsinriktat, verkningsfullt, effektivt, mer profilerat och mera internationellt universitetsväsende”. Här kan det noteras att internationalisering uppges vara ett ändamål i sig vilket reflekteras i val av indikatorer (se nedan).

En kompromiss mellan det politiska kravet på en tydligare koppling mellan forskning och innovation och nytta, och akademins krav på utvärdering av forskningskvalitet ledde till Research Quality Framework-systemet (RQF) i Australien, föregångaren till det nuvarande Excellence in Research for Australia (ERA). I Nya Zeeland är målet att belöna och säkerställa excellens i universitetsforskningen. Systemet i Nederländerna utvecklades som en följd av landets autonomireform där det fanns en förväntan på motprestation från lärosätena i form av kvalitetssäkringssystem för både forskning och utbildning.

35 <http://fivu.dk/uddannelse-og-institutioner/videregaende-uddannelse/universiteter/okonomi/basismidler-efter-kvalitet;>

Aftale om fordeling af globaliseringsmidlerne til forskning og udvikling (Danska regering, 5 november 2008); *Aftale om ny model for fordeling af basismidler til universiteterne* (Danska regering, 30 juni 2009).

NOU (Norges offentliga utredningar) 2011:6 Et åpnere forskningssystem. www.minedu.fi; Undervisnings- och kulturministeriet, Finland (2011) Kvalitetsinriktat, internationellt, profilerat och verkningsfullt universitet – ett förslag till finansieringsmodell för universiteten fr.o.m. år 2013.

36 <http://www.arc.gov.au/era/>

37 <http://www.tec.govt.nz/Funding/Fund-finder/Performance-Based-Research-Fund-PBRF-/quality-evaluation/>

38 <http://www.anvur.org/index.php?q=it/homepage&lang=it>

<http://www.metrisnet.eu/metris/index.cfm/report/findByStructureAndCountry/38/44>

Franceschet, M. och Costantini, A. (2011) The first Italian research assessment exercise: a bibliometric perspective, *Journal of Informetrics* 5(2): 275-291.

39 <http://www.ecoom.be/en/services/bof>

<http://www.ewi-vlaanderen.be/ewi/wat-doen-we/programmas-subsidies/financiering-van-onderzoek/bijzondere-onderzoeksfonds>

Presentation av Tim Engels (Centre for R&D Monitoring, University of Antwerpen), i Helsingfors, 2012.

40 http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20091052.pdf,

<http://www.government.nl/documents-and-publications/leaflets/2012/04/17/the-science-system-in-the-netherlands.html>

41 Butler, L. (2010) *Performance-Based Funding for Public Research in Tertiary Education Institutions*, Workshop proceedings, OECD, 2010 Kapitel 2.

42 Times Higher Education 14th February 2013

43 <http://www.aeres-evaluation.com/>

3.3 Vilken instans ansvarar för att genomföra utvärderingarna?

Av de länder som har ingått i kartläggningen är det vanligast att det ansvariga ministeriet eller en underordnad men oberoende myndighet avsvaret för kvalitetsvärderingssystemet. I Australien och i Tyskland är det dock ett forskningsråd som ansvarar för uppdraget.

För de länder (SE, FIN, NO, DK, FLA) som enbart baserar resurstilldelningen på indikatorer är det ministeriet som ansvarar för själva systemet men indikatorerna sammanställs i viss mån av myndigheter och övriga statliga och icke-statliga organisationer. Länder som använder sig av peer review modeller tenderar att ge uppdraget till en oberoende statlig myndighet – HEFCE i Storbritannien, *Tertiary Education Commission* i Nya Zeeland och *National Agency for the Evaluation of Universities and Research Institutes* (ANVUR) i Italien.

Tabell 2. Översikt av olika länders system för prestationsbaserad resurstilldelning.

	Australien	Nederländerna	Nya Zeeland	Storbritannien	Italien
Prestationsbaserad resurstilldelning?	JA	NEJ	JA	JA	JA
Vem ansvarar för utvärderingen?	Australian Research Council på uppdrag av ministeriet	University board som i sin tur kan låta t ex QANU göra den	Tertiary Education Commission	HEFCE på uppdrag av alla fyra länder i Storbritannien	National Agency for the Evaluation of Universities and Research Institutes (ANVUR), på uppdrag av Ministry for Education, University and Research
Vilken nivå används som bas för utvärderingen?	8 disciplinära "research evaluation committees" utvärderar universiteten	Forskningsinstitut och forskargrupper	Varje anställd forskare bedöms i en av 12 peer review paneler.	36 disciplinära utvärderingsenheter, men fördelning på universitetsnivå	University departments (eller "structures"). 14 st "evaluation areas"

	Danmark	Finland	Flandern	Norge
Prestationsbaserad resurstilldelning?	JA (Indikator)	JA (Indikator)	JA (Indikator)	JA (Indikator)
Vem ansvarar för utvärderingen?	Ministeriet för forskning, innovation och vidaregående utbildningar	OMI, Utbildnings- och kulturministeriet.	Flemish Dept. of economy, science and innovation (EWI)	Kunnskapsministeriet
Vilken nivå används som bas för utvärderingen?	Universitet	Universitet	Universitet	Universitet

I Frankrike finns en särskild utvärderingsmyndighet för högre utbildningssektorn, AERES. I Nederländerna är det universitetens eget ansvar att genomföra en kvalitetsutvärdering baserad på *the Standard Evaluation Protocol (SEP)*⁴⁴ vart sjätte år. The Netherlands Organisation for Scientific Research (NWO), The Royal Netherlands Academy of Arts and Science (KNAW) och VSNU (motsvarar SUHF) har gemensamt kommit överens om SEP enligt vilket alla utvärderingar görs. KNAW och NWO ansvarar för att utvärdera sina forskningsinstitut på samma sätt. Ofta läggs genomförandet av utvärderingen ut på den oberoende organisationen Quality Assurance Netherlands Universities (QANU).

3.4 Peer review eller indikatorer?

Vad gäller metodval finns det system som bygger på peer review, system där endast indikatorer ligger till grund för resurstilldelningen och system som bygger på en blandning av dessa två. Det sistnämnda kan ytterligare delas upp i modeller där indikatorer och peer review tillämpas som parallella självständiga spår eller s.k. informerad peer review där indikatorer ingår i underlaget som expertpanelerna baserar sina bedömningar på.

I de nordiska länderna Norge, Danmark, Finland och Sverige används system som baseras enbart på metriska indikatorer, likaså i Flandern.

⁴⁴ http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20091052.pdf

I Storbritannien, Hongkong och Nya Zeeland används huvudsakligen peer review. Vissa paneler i Storbritannien väljer även att ta del av bibliometriska underlag. Systemen i Nederländerna och Tyskland är också peer review-baserade (men används inte för resursfördelning). I Italien är resurstilldelning kopplad till parallella peer review och indikatorspår. Det australiensiska systemet beskrivs bäst som indikator-informerad peer review. Dessutom finns det i Australien och Italien skillnader mellan olika ämnesområden där vissa expertpaneler endast använder sig av peer review för att utvärdera *output* från forskningen medan andra använder sig av bibliometri och andra indikatorer för att informera expertbedömningen.

3.5 Vad som utvärderas

Vad gäller den strukturella nivån som utvärderas skiljer även den sig åt mellan länderna – från den individuella forskaren i Nya Zeeland till lärosätetsnivå i de indikatorbaserade systemen. Resurstilldelning i Storbritannien, Italien och Australien sker på lärosätetsnivå men själva bedömningen är uppdelad ämnesvis varefter resultaten sedan aggregeras till universitetsnivå.

ERA i Australien är en ämnesbaserad utvärderingsmodell, där varje lärosäte rapporterar data för discipliner på fyrsiffernivå.⁴⁵ Om volymen forskning inte är tillräckligt stor på fyrsiffernivå görs analysen på tvåsiffernivå. Som exempel på tvåsiffernivå kan nämnas ”*Physical Sciences*” (02) och på fyrsiffernivå ”*Astronomical and Space Sciences*” (0201). Om volymen forskning är alltför låg även på tvåsiffernivå, kommer ingen ERA-utvärdering att genomföras för det ämnesområdet vid det lärosätet. Ämnesområdena utvärderas i åtta ämneskluster. Sammanlagt redovisas/utvärderas 157 ämnen på fyrsiffernivå och 23 ämnesområden på tvåsiffernivå – alltså fördelade på de åtta klustren. För vissa indikatorer tillämpas åttaåriga redovisningsfönster, för andra tillämpas treåriga perioder. Utvärderingarna genomförs av åtta s.k. *Research Evaluation Committees* (en för varje kluster) under en period av 3–4 månader. I ERA2010 ingick totalt 41 lärosäten. Panelerna består av ”*nationally and internationally recognised researchers with expertise in research evaluation*”. I ERA 2010 användes dessutom, utöver panelerna, sammanlagt 149 externa s.k. expert reviewers som läste *outputs*.

Det nya zeeländska systemet bygger på s.k. *evidensportföljer* för varje forskare. Dessa bedöms i 12 granskarpaneler som täcker 42 *subject areas*. Det finns även två rådgivande organ varav en är expertgrupp för tillämpad forskning för att bevaka just den forskningsgrenen. Det är för oss oklart om de består av användare eller akademiker. Alla universitetsanställda som är ”avsevärt” aktiva i undervisnings- eller forskningsuppgifter ska ingå i utvärderingen. Varje forskare rankas och universitetet som helhet får en sammanvägd poäng.

Det italienska systemet är uppdelat i 14 ämnesområden. Nederländerna avviker något med forskargrupper/forskningsinstitut som basenhet i utvärderingen. Både forskningsinstitutionen och dess forskningsprogram utvärderas i sin helhet, och utvärderingen sker alltså på två nivåer.⁴⁶

3.6 Hur bedömningsprocessen går till

Bedömningsaspekter

De olika bedömningsaspekter samt eventuella viktningar som används i själva utvärderingen i respektive land som använder sig av peer review som metod sammanfattas i tabellen nedan.

⁴⁵ FoR (Field of Research) enligt *Australian and New Zealand Standard Research Classification* (ANZSRC)

⁴⁶ ”The units of evaluation may vary among the participating organisations – the universities, KNAW and NWO. An institute may be defined as a ’group of researchers with an articulated shared mission, operating within one or more research programmes under the same management.’ Standard Evaluation Protocol 2009–2015.

Tabell 3. Aspekter som bedöms i olika länders utvärderingssystem.

Storbritannien	Nya Zeeland	Australien	Nederländerna	Italien
Forskningsoutput (65 %)	Forsknings- kvalitet (60 %)	Forskningsoutput (42 %)	Forskningskvalitet	Forskningskvalitet (50 %)
	Forskningsmiljö (9%)	Forskningensvolym och - aktivitet	Produktivitet	Externa resurser (10 %)
Forskningsmiljö (15 %) <i>inkl. extern finansiering, doktorander</i>	Peer esteem (9 %)	"Recognition" (esteemmått)	Relevans	Egna resurser (5 %)
	Extern finansiering (15 %)			Mobilitet (10 %)
Impact (20%)	Doktorsexamina (25 %)	Research application	Vitalitet och genomförbarhet	Internationalisering(10 %) Improvement (5 %)
				Diverse innovations- och samarbetsmått
				"Third mission"

Indikatorer

Indikatorerna som används i olika system kan syfta till att mäta *output* från forskningssystemet (exempelvis publiceringsvolym, antal avlagda doktorsexamina) eller *input* (t.ex. intäkter från externa forskningsfinansiärer). I några av systemen finns det indikatorer som syftar till att fånga andra dimensioner av kunskapstriangeln än just forskning. Det kan vara mått kopplade till utbildning på grund- och avancerad nivå (exempelvis i Danmark och Flandern) eller till den tredje uppgiften (exempelvis i Italien). Många av indikatorerna mäter volymer, exempelvis antal doktorsexamina och antal publikationer, och kan på det viset sägas vara beskrivande. Andra är mer värderande till sin karaktär, exempelvis citeringar och intäkter från externa finansiärer. Intäkter från externa finansiärer är även ett volymmått men kan anses vara en proxy för kvalitet. Forskningsanslag från externa finansiärer (forskningsråd, European Research Council m.m.) ingår som en indikator i alla system (även i Sveriges nuvarande) utom i Flandern.

Forskarutbildning ingår som en indikator i samtliga system (I, UK, FIN, DK, NO, NZ, FLA), Sveriges nuvarande system undantaget. Antalet doktorander, research fellows och post-docs ingår som mått i Italien. I Norge ingår antal doktorander, medan det i övriga länder är avlagda doktorsexamina som räknas. Antalet doktorander kan ses som ett mått på forskningsmiljön, medan antalet doktorsexamina oftare ses som en indikator på forskningskvalitet.

I Australien används 24 indikatorer som grupperas enligt de fyra rubrikerna i tabellen ovan. Australien är även intressant för att olika aspekter bedöms på olika sätt i olika ämnesområden. Alla indikatorer redovisas (och utvärderas) inte för alla ämnen, ämnesområden och kluster. Peer review av böcker, artiklar etc. ersätter citeringsanalyser för hela *Humanities and Creative Arts* och större delen av *Social, Behavioural and Economic Sciences*. Några undantag är ämnesområdet *Built Environment and Design* inom *Engineering and Environmental Sciences*, där peer review ersätter citeringsanalys, och ämnesområdet *Psychology and Cognitive Sciences* inom *Social, Behavioural and Economic Sciences*, där citeringsanalys ersätter peer review.

Alla nordiska länder använder sig av bibliometriska mått och samtliga använder även andra indikatorer i resurstilldelningen. Sverige är det enda nordiska land som inkluderar citeringar i sina bibliometriska indikatorer, de övriga använder enbart (viktade) publikationsvolymer. Utanför Norden används även citeringsmått i Australien, Italien och Flandern. I Finland, Flandern och Italien ingår mått som försöker fånga dimensionen internationalisering (alternativt mobilitet). Exempelvis genom att mäta andelen utländska forskare och doktorsexamina avlagda av utländska doktorander, antal utresande och inresande forskare, antal publikationer samförfattade med utländska forskare och andel erhållna externa internationella forskningsmedel.

Viktningen i de indikatorbaserade systemen redovisas i tabellen nedan. Intressant att notera är att i Norge används liknande indikatorer men med olika viktningar för tilldelningen till universitet och högskolor, forskningsinstitut respektive sjukvårdssektorn (här redovisas endast för universitets- och högskolesektorn)⁴⁷.

47 NOU 2011:6 Et öppnare forskningssystem

Tabell 4. Viktning i de indikatorbaserade systemen.

Danmark		Norge		Finland		Flandern	
Grundutbildning	(45 %)					Examina på grund- och avanceradnivå	(25 %)
Extern finansiering	(20 %)	EU-medel Medel från norska forskningsrådet	(20 %) (20 %)	Konkurrensutsatt forskningsfinansiering	(27 %)		
Doktorsexamina	(10 %)	Doktorander	(30 %)	Doktorsexamina	(27 %)	Doktorsexamina	(35 %)
Publikationer	(25 %)	Publikationer	(30 %)	Publikationer	(38 %)	Publikationer Citeringar	(18 %) (18 %)
				Internationalisering	(9 %)	Mobilitet och diversitet	(4 %)

Utvärdering av impact, nytta samverkan m.m.

Forskningens nytta och samverkan med samhället/näringsliv är en dimension i de brittiska, australienska, nederländska och italienska systemen. Vidare pågår diskussioner i Nederländerna och Australien om att utvidga och/eller utveckla *impact* som aspekt i bedömningen.

Nyttoaspekten i det befintliga nederländska systemet finns som ett underkriterium till relevans: samhällsrelevans (*”societal impact”*). Det syftar på hur forskning påverkar specifika/enskilda aktörer eller samhällsfunktioner. I Italien används indikatorer som berör näringslivssamverkan såsom patent, inkubatorer, spin-off och konsortia men även indikatorer som rör kulturarv (t.ex. arkeologiska fornninnen och muséer) och aktiviteter som hör till tredje uppgiften.

Några formella beslut om att inkludera *impact* i det australiensiska ERA finns inte. Utvecklingsarbete pågår på det berörda ministeriet och ett antal lärosäten i Australien har på eget initiativ genomfört en pilot av fallstudiemetodiken 2012.⁴⁸ Dimensionen *”research application”* i det nuvarande ERA fångar i viss mån innovationer och samverkan i Australien. Här ingår indikatorer såsom patent och övriga immateriella rättigheter, kliniska riktlinjer och forskningskommunikation. Det numera nedlagda RQF-systemet i Australien bedömde forskarportföljer som kunde kompletteras med s.k. kontextnarrativ för att beskriva *”relevanta kvalitets- och impact anspråk”* för en viss forskargrupping där även *”short supporting statements of research impact by relevant end-users”* kunde ingå. Grundforskning undantogs dock från kravet på utvärdering av *impact*.⁴⁹ Det något diffusa begreppet *”peer esteem”* ingår i det nya zeeländska systemet och var en del av de nedlagda brittiska RAE och australiensiska RQF systemen. *Peer esteem* är dock inte direkt kopplat till *impact* utan är ett sätt att bedöma forskarens eller forskningens genomslag inom, såväl som bortom, de akademiska kretsarna.

3.7 Hur utvärderingen omsätts till resurser

Det bör beaktas att det är stor skillnad mellan länderna avseende dels hur stor andel basanslaget utgör av den totala statliga finansieringen av universitetsforskning, dels hur stor andel av universitetens forskningsintäkter det utgör. Detta innebär att det är svårt att direkt jämföra olika resurstilldelningsmodellens påverkan på forskningssystemet i stort och i synnerhet hur de olika indikatorerna och viktningen av desamma påverkar forskningen. Dessutom kan det vara relevant att ta hänsyn till hur stor del av forskningsintäkterna som är konkurrensutsatta totalt. Anslag som tilldelas via forskningsråd är generellt sett konkurrensutsatta genom peer review-processer och i vissa länder finns det separata bas-

⁴⁸ E-post från Mary Kelly, ARC, 2013-04-24.

⁴⁹ *The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research.* Donovan, C. (2008). In C. L. S. Coryn & M. Scriven (Eds.), *Reforming the evaluation of research.* New Directions for Evaluation, 118, 47–60.

anslag för exempelvis forskarutbildning, innovation, infrastruktur m.m. som inte nödvändigtvis är konkurrensutsatta.

Hur stor andel av basanslaget som är prestationsbaserat varierar också. I Storbritannien är ca 2/3 av basanslaget till grundforskning prestationsbaserat. Det samlade basanslaget till utbildning och forskning i Finland fördelas till 25 procent efter strategiska avgöranden och till 75 procent enligt prestationsmått (varav prestationsmått relaterade till forskning utgör 34 procentenheter av dessa 75 procent). I Flandern är det ungefär 45 procent av de statliga medlen till grundforskning som fördelas på basis av bibliometriska data enligt den s.k. BOF-nyckeln. Enligt norska siffror från 2010 uppgick den resultatbaserade omfördelningen (RBO) som en del av de statligt finansierade FoU-utgifterna under 2009 till 14 procent för den högre utbildnings-, 12 procent för sjukvårds-, och 2 procent för institutsektorn. I Danmark anges siffran vara 2 procent av det totala forskningsstödet 2010, något som kommer att öka framöver. 12 procent av alla medel till universiteten påverkas av utvärderingsresultatet i Nya Zeeland.

I Italien är ungefär 7 procent av basanslaget prestationsbaserat. Resultatet av det tidigare systemet för utvärdering i Italien VTR (Valutazione Triennale della Ricerca 2001–2003) ligger till grund för hälften av fördelningsnyckeln, medan intäkter från EU och konkurrensutsatta medel tillsammans står för resterande hälft.⁵⁰ I Australien används resultaten av ERA som underlag för fördelningen av ett av de sex basanslag som finns där (SRE *Sustainable Research Excellence in Universities Scheme*) och då endast för en viss andel av detta anslag. I jämförelse med t.ex. Storbritannien påverkar ERA endast i tämligen begränsad omfattning resursfördelningen.⁵¹ I Nederländerna och i det föreslagna tyska systemet är inte resultaten av kvalitetsvärderingarna kopplade till tilldelningen av resurserna.

För en uppskattning av kostnaderna för att genomföra utvärderingarna och ett resonemang kring konsekvenser hänvisas till utredningen *Prestationsbaserad resurstilldelning för universitet och högskolor* (U2011-7356-UH) och till Hicks (2012).⁵² Effekterna av fördelningssystemen diskuteras av Butler i en skrift från OECD.⁵³

50 Det nuvarande heter VQR (Valutazione della Qualità della Ricerca) och sträcker sig 2004–2010.

<http://www.metrisnet.eu/metris/index.cfm/report/findByStructureAndCountry/38/44>

51 E-post från Mary Kelly, ARC, 2013-04-24.

52 Hicks, D, (2012) Research Policy 41: 251–261. *Performance-based university research funding systems*.

53 Butler, L. (2010) *Performance-Based Funding for Public Research in Tertiary Education Institutions*, Workshop proceedings, OECD, 2010 Kapitel 2.

4. KVALITETSGRANSKNING AV FORSKNING VID SVENSKA LÄROSÄTEN

Om inte annan källa anges hänvisas till respektive lärosätes utvärdering, se referenser i kapitel 7.

4.1 Syften och process för att utveckla systemet

Sju svenska lärosäten har hittills genomfört egna kvalitetsgranskningar: Göteborgs universitet (GU), Karolinska institutet (KI), Kungl. tekniska högskolan (KTH), Lunds universitet (LU), Sveriges lantbruksuniversitet (SLU), Uppsala universitet (UU) och Örebro universitet (ORU). KTH och UU har genomfört två granskningar.⁵⁴ Dessutom genomför under 2013 flera lärosäten granskningar för första gången. Mittuniversitetet (MiUn), Högskolan i Halmstad (HH) samt Högskolan i Skövde (HiS) arbetar tillsammans med att ta fram riktlinjer för utvärderingsprocessen som pågår under 2013.⁵⁵ Mälardalens högskola (MdH) är i planeringsfasen för en granskning som genomförs under 2014. I samtliga granskningar används peer review som huvudmetod. Några lärosäten, t.ex. GU, avvaktar ett framtida nationellt system och planerar beroende på det en andra utvärdering tidigast 2015.

Ingen av de genomförda utvärderingarna har haft som främsta syfte att omfördela resurser; däremot kan det sägas vara ett indirekt syfte. Ett genomgående övergripande syfte för samtliga lärosätens granskningar har varit att fungera som grund för ett strategiskt utvecklingsarbete. Genom att identifiera styrkor och svagheter i forskningen har lärosätena önskat få underlag för att förbättra den vetenskapliga kvaliteten, exempelvis genom att stödja forskargrupper med potential eller avveckla mindre framgångsrik verksamhet.

4.2 Vad som utvärderas

Vilken enhet som ligger till grund för lärosätens utvärderingar varierar, men det är antingen institutions- eller forskargruppernivå som utgör en utvärderingsenhet. Det råder stora skillnader i antalet utvärderingsenheter från hundratals s.k. ”*research constellations*”, grupperade under nio medicinska områden (KI), till dryga sjuttio (motsvarande ungefär institutioner i LU:s och UU:s utvärderingar⁵⁶) till knappt femtio (KTH) och knappt fyrtio (ORU). Just skillnaderna i storlek och struktur mellan de utvärderade enheterna, både inom och mellan lärosäten, har av många lyfts fram som ett stort problem som försvårar jämförelser. *Hur* man definierar en utvärderingsenhet och *vem* som gör det synes vara en av de stora utmaningarna.

Det varierar något mellan utvärderingarna i vilken mån de utvärderade enheterna själva gavs möjlighet att välja vilket ämnesområde/utvärderingsenhet de ville höra till. Generellt kan sägas att de rådde stor frihet på den punkten. Vid LU:s granskning bestämde prefekten sin utvärderingsenhet genom det sätt på vilket rapporten skrevs. I vissa institutionsrapporter utlämnades delar av institutionen, andra fick inte den framträdande ställning de borde ha haft, och vissa fick en för stark framtoning.⁵⁷

54 De genomförda granskningarna: Göteborgs universitet: *Research evaluation for development of research* 2010 (RED10); Karolinska institutet: *External Research Assessment* (ERA); Kungl. tekniska högskolan: *Research Assessment Exercise* 2008 respektive 2012 (RAE2008 och RAE2012); Lunds universitet: *Research Quality Assurance for the Future* (RQ08); Sveriges lantbruksuniversitet: *Kvalitet och nytta* (KoN09); Uppsala universitet: *Kvalitet och förnyelse* 2007 respektive 2011 (KoF07 och KoF11); Örebro universitet: *Örebro Research Evaluation* (ÖRE10). Den forskningsutvärdering som genomförs 2013 vid Mittuniversitetet, Högskolan i Halmstad och Högskolan i Skövde går under benämningen *Assessment of Research and Coproduction* (ARC13).

55 En gemensam styrgrupp koordinerar genomförandet men respektive lärosäte styr och ansvarar för sin egen utvärdering. Slutrapport planeras vara färdig i november/december 2013.

56 I UU:s utvärderingar ingick dock mer än 500 forskargrupper.

57 Intern slutrapport från projektledningen avseende forskningsutvärderingen Lund University *Research Quality Assurance for the Future* - RQ08, Bengt Söderström.

I KI:s utvärdering fick varje utvärderingsenhet själv bestämma sin sammansättning. Det resulterade bl.a. i att hundratalet professorer ”were not assessed as constellation leaders because they were either included in another named constellation or were simply not included in the assessment.” Vidare sorterade vissa prefekter aktivt bort svagare forskargrupper i förhoppningen att institutionen skulle falla bättre ut i utvärderingen.⁵⁸

Vid utvärderingarna av UU var syftet inte att utvärdera all forskning utan expertpanelerna hade som uppgift att ”identify particularly strong research, emerging science and opportunities for renewal.” De utvärderade enheterna behövde inte heller rapportera in alla forskningsaktiviteter. Från UU uppskattar man dock att så mycket som 85–90 procent av universitetets forskningsaktiviteter (grupper/projekt etc.) presenterades för panelerna, antingen skriftligt i institutionsunderlagen och/eller genom presentationer vid panelbesöken på plats.⁵⁹

Vid SLU:s utvärdering fick prefekterna själva avgöra tillhörighet på de forskargrupper som utgjorde respektive utvärderingsenhet. Forskargrupperna var disciplinära snarare än organisatoriska, vilket hade fördelen att resultat och återkoppling kunde bli mer detaljerad och specifik. Nackdelen var dock att implementeringen till följd av resultaten försvårades.

Ingår forskarutbildning i utvärderingarna?

MiUN, HH och HiS är de enda lärosäten som explicit inkluderar forskarutbildning i utvärderingen. De enheter som bedriver forskarutbildning ska besvara ett kompletterande raster av frågor med fokus på process och examensmål. Denna information bedöms dock inte av de internationella panelerna, utan utgör tillsammans med den ”ordinarie” granskningen av forskningskvalitet ett underlag för utveckling av forskarutbildningsämnen.⁶⁰ Delar av forskarutbildningen ingår i utvärderingarna i form av avhandlingar (som *output* t.ex. vid SLU) men det sker inte en systematisk utvärdering av själva utbildningen som sådan. I de flesta fall kommenterar expertpanelerna olika aspekter av forskarutbildningen när forskningsmiljön bedöms, men det sker utan synbar systematik. Antalet doktorander och doktorsexamina är indikatorer som oftast ingår i bedömningsaspekten forskningsmiljö.⁶¹

4.3 Hur bedömningsprocessen går till

Bedömningsprocessen har varit förhållandevis lika vid alla lärosäten. De lärosäten som genomfört utvärderingar senare har ofta hänvisat till tidigare utvärderingar, vilket delvis kan förklara att man finner många likheter emellan dem. Processen inleds med att en intern projektorganisation skapas. Bedömningsaspekter och kriterier beslutas och underlag samlas in. Peer review paneler granskar underlagen och genomför platsbesök.⁶² En slutrapport sammanställs med panelernas bedömningar.

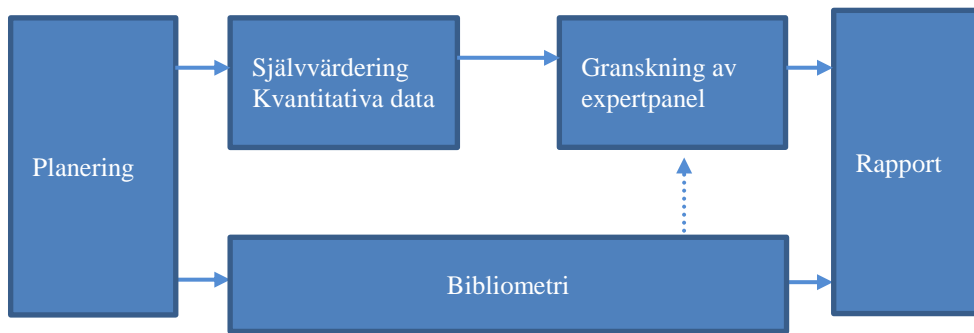
58 KIs utvärdering, s 6.

59 Korrespondens med Per Andersson, utbildningsledare Uppsala universitet.

60 Korrespondens med Jonas Harvard, Mittuniversitetet.

61 Några lärosäten hänvisar till separata uppföljnings- och utvecklingsprojekt med fokus på grundutbildning, forskarutbildning och/eller forskarskolor, t.ex. *Education Quality 11* (EQ11) vid Lunds universitet, BLUE11 vid Göteborgs universitet samt liknande projekt vid Uppsala universitet och KTH.

62 I LU:s och SLU:s fall skiljde sig platsbesöken åt något jämfört med de andra mer traditionella, där hela panelen träffar varje utvärderingsenhet på plats vid respektive lärosäte. I LU:s fall var det endast ordförande och vice ordförande i panelerna som genomförde platsbesök. I SLU:s fall tillkallades alla utvärderade forskargrupper till Ultuna i Uppsala för intervjuer för att undvika ett intensivt resande över hela Sverige.



Figur 2. Skiss över hur bedömningsprocessen i de flesta fall gått till vid lärosätena. Den streckade pilen visar att bibliometri i ett fåtal fall direkt utgjort underlag för expertpanelernas bedömning. I de flesta fall har dock bibliometriska data samlats in parallellt.⁶³

Underlag för bedömningen

I samtliga granskningar ingår självvärderingar av något slag samt en kvantitativ del med uppgifter om personal, antal publikationer, andra typer av *outputs* (t.ex. patent), antal samarbeten, storleken på externfinansiering, antal doktorsexamina mm.

Bibliometri har i vissa fall använts som underlag till panelerna (KTH:s andra utvärdering och SLU) och i andra som ett parallellt underlag, i vissa fall genomförd efter avslutad peer review granskning, med syfte att stämma av mot panelernas bedömningar (GU, KTH:s första utvärdering, UU, ORU). I LU:s fall genomfördes en bibliometrisk studie först efter önskemål från panelerna.

KTH:s andra utvärdering, RAE12, inkluderade som första lärosäte aspekten ”*economic and social impact on the research performed.*” Inspirerade av REF14 i Storbritannien fick varje utvärderingsenhet skicka in en s.k. *impact statement* tillsammans med högst två fallstudier. Sammanlagt inkom 94 fallstudier. På liknande sätt planerar även MiUn, HH och HiS att inkludera fallstudier i sina granskningar. Liksom i RAE12 får varje enhet skicka in högst två stycken och gränsen är max 4 sidor per fallstudie. Till skillnad från både REF14 i Storbritannien och RAE12 vid KTH inkluderar de senare fallstudierna både vetenskaplig och utomvetenskaplig påverkan/nytta.

Bedömningsaspekter⁶⁴

Det råder stor överensstämmelse om vilka aspekter som granskas i lärosätenas utvärderingar. Alla utvärderingar utom KI och ORU har en aspekt som explicit heter *vetenskaplig kvalitet*. Tre av sju har *originalitet* som en utvärderingsaspekt (GU: unikheter; KI: originalitet och *impact*; ORU: huruvida forskningen är nydanande). *Relevans* är en aspekt som återfinns hos tre (GU, LU och SLU). SLU skiljer sig något åt, som redan nämnts, i och med att bedömning av nytta (*impact*) ingår som ett tydligt syfte. Således har SLU en aspekt som heter just *relevans och genomslag (nytta)*. I KTH:s andra granskning 2012 har man lagt till *impact and engagement with society*. I utvärderingen vid MiUn, HH och HiS finns den snarlika aspekten *impact, engagement and co-operation with society*. *Vitalitet* är en aspekt som används explicit i fyra utvärderingar (GU, KTH, LU och ORU). *Strategi och potential* är också vanligt förekommande (KI, KTH, SLU, UU, ORU). *Organisatorisk kapacitet* återfinns i tre utvärderingar (GU, LU och ORU). I tabellen nedan finns lärosätenas bedömningsaspekter sammanställda.

⁶³ Figuren är en något modifierad variant av en figur som presenterades i den utvärdering som gjordes vid Göteborgs universitet 2010. RED10

⁶⁴ I många utvärderingar kallas dessa aspekter för ”*general assessment criteria*” eller motsvarande. Även det som i denna rapport benämns som bedömningsskalor, betecknas ibland som kriterier.

Tabell 5. Sammanställning över bedömningsaspekter i lärosätenas utvärderingar.

GU	Quality; Productivity; Uniqueness; Relevance; Organisational capacity; Interactive vitality
KI	Originality and impact; Productivity and potential for continued excellence; Research infrastructure
KTH	RAE08: Scientific quality (basic research); Applied research quality; Scholarship; Vitality and potential; Strategy RAE12: Quality of research output; Impact and Engagement with Society; Research environment
LU	Quality; Productivity; Relevance; Vitality and Organisational capacity
SLU	Scientific quality; Recognition and leadership; Relevance and Impact; Strategy and Potential
UU	KoF07: General assessment of the department/unit; Quality of research; Research environment and infrastructure; Networks and collaborations; Opportunities for renewal and emerging science; Actions for successful development. KoF11: General assessment of the department/unit; Quality of research; Research environment and infrastructure; Networks and collaborations; Opportunities for renewal and emerging science; Actions for successful development; Effects of the KoF07-evaluation; Other issues
ORU	Vetenskaplig produktion; Produktivitet och kvalitet; Huruvida forskningen är nydanande och tillämpbar rent vetenskapligt och för samhälle/näringsliv; Forskningsorganisatorisk sammansättning; Forskningsområdets aktuella fokus, utvecklingspotential och möjlighet till förnyelse.
MiUn/HH/HiS	Scientific results; Research environment and infrastructure; Impact, engagement, and co-operation with society; Self-assessment and future development

Alla lärosäten har haft en bedömningsskala som varierar mellan fyra och sex steg. SLU har en sexgradig för vetenskaplig kvalitet medan bedömningen av nytta (*impact*) sker enligt fyra kriterier utan betygsättning. I KTH:s andra utvärdering bedömdes den vetenskapliga kvaliteten efter en sexgradig skala, från världsledande till under standard, medan aspekterna *impact* och forskningsmiljö bedömdes enligt tre nivåer.

Indikatorer

Förutom bibliometriska data över publikationer som används i alla granskningar på ett eller annat sätt, sammanställdes också kvantitativ information om personalsammansättningen. I övrigt är de indikatorer lärosätena vanligtvis använder för att bedöma forskningskvaliteten avsedda att mäta *output* från forskningssystemet (exempelvis publiceringsvolym, antal avlagda doktorsexamina, antal samarbeten) eller *input* (t.ex. intäkter från externa forskningsfinansierare). Indikatorer rörande *impact*, nytta och samverkan inkluderar vanligtvis samarbeten med externa parter (näringsliv och samhälle), mobilitet samt kommunikation och andra externa aktiviteter. För detaljer kring de indikatorer som används hänvisas till respektive lärosätes utvärdering.

Utvärdering av utomakademisk impact, nytta och samverkan

I KTH:s första granskning fanns ett uttalat syfte att stimulera entreprenörskap och innovation och i den andra granskningen att stimulera spridning och användning av forskningsresultat i samverkan med näringsliv och samhälle. I den andra utvärderingen, som nämns ovan, fick varje enhet skicka in en *impact statement* och upp till två fallstudier per enhet.

Ett liknande förfarande sker vid den pågående utvärderingen vid MiUn, HH och HiS, men där inkluderas även akademisk *impact* i fallstudierna. Samma utvärdering är också den enda som explicit inkluderar samproduktion med offentliga och privata aktörer (*co-production*). I vissa utvärderingar, t.ex. LU:s, ingick delar av *impact* som en del av aspekten ”*relevance*” (*scientific, social and socio-economic significance*).

SLU var den enda utvärdering som förutom att bedöma vetenskaplig kvalitet och relevans, även *separat* bedömde nytta ur ett användarperspektiv utförd av SLU:s intressenter (näringsliv, myndigheter, intresseorganisationer, m.fl.). Detta gjordes genom intervjuer med fem separata intressentpaneler. Intervjuerna med intressenterna genomfördes av en konsult.

4.4 Expertpanelerna

Det vanligaste förfarandet för att rekrytera experter till panelerna var att utvärderingsenheterna själva fick komma med förslag på experter. Beslut om vilka som slutligen skulle ingå fattades på ledningsnivå.⁶⁵ Antalet paneler varierar i granskningarna från som minst sju paneler (ORU) till som mest 25 (UU 2011). Panelernas storlek varierade från som minst ca fyra till som mest ca 11 personer. Majoriteten av de sakkunniga i panelerna var verksamma utanför Sverige. I LU:s fall var samtliga utländska, i ORU:s granskning var granskarna företrädesvis från Sverige och de nordiska länderna. I KI:s granskning var alla verksamma utanför Sverige utom vice ordförande som verkade i Sverige. I UU:s båda utvärderingar ingick i alla paneler en expert verksam i Sverige samt en så kallad panelguide från UU. Även KTH:s utvärderingar hade en KTH-koordinator med i varje panel.

Endast SLU och KTH inkluderade användareexperter i sina utvärderingar, vilket (delvis) förklaras av att dessa två lärosäten hade samhällsrelevans/nytta som en tydlig separat bedömningsaspekt. KTH hade vissa panelmedlemmar med industribakgrund. SLU hade förutom ett par användareexperter i panelerna för vetenskaplig kvalitet, som representerade svenska intressenter, även fem separata intressentpaneler med mellan fem och åtta personer vardera. I den granskning som påbörjats av MiUn, HH och HiS ska en generalistpanel tillsättas med representanter från samhälle och näringsliv.

4.5 Tekniska lösningar för inrapportering

De flesta lärosäten inrättade en särskild projektorganisation för sina respektive granskningar. I de flesta fall var vice-rector eller motsvarande ansvarig. Webbaserade projektportaler skapades för ändamålet för att samla in och distribuera information.⁶⁶ Man informerade via interna och externa hemsidor.

4.6 Hur utvärderingen omsätts till resurser

Inget lärosäte har haft som uttalat syfte att omsätta granskningarnas resultat till resursfördelning. Där- emot har så blivit fallet i och med att i princip alla lärosäten omfördelat medel man hållit inne. Det har rört sig om i storleksordningen 30–85 miljoner kronor per år. Generellt har det rört sig om strategiska satsningar på både starka forskningsmiljöer och svagare områden med förbättringspotential, i enlighet med resultaten i utvärderingen. Det har också förekommit särskilda satsningar på t.ex. jämställdhet (SLU och LU). Vid UU fick fakulteterna komma med prioriteringsförslag istället för att ledningen ”automatiskt” delade ut pengar till dem som rankats högst. Vid SLU tillsattes 18 nya s.k. samverkanslektorat som ett resultat av utvärderingen.

4.7 Kostnader för framtagning, implementering och drift av systemet

Kostnaderna för lärosätenas utvärderingar varierar men ligger någonstans runt 10–22 miljoner kr, inkluderat det arbete som institutionerna själva lägger på att rapportera in resultat etc.⁶⁷ I ORU:s fall var kostnaden avsevärt lägre, ca 1,5 miljoner kr för de direkta kostnaderna, beroende på det färre antalet

⁶⁵ Det saknas information om hur denna process gick till mer precist i flera utvärderingar.

⁶⁶ Flera lärosäten använde sig av Designtech's *Project-Coordinator* ® X.

⁶⁷ UU beräknade de direkta kostnader till c:a 6,5 miljoner kr (KoF07) respektive 7,5 miljoner kr (KoF11), varav c:a 75 procent var kostnader för expertpanelerna (resor, hotell, måltider, arvoden etc.). Kostnader för arbetstid uppskattades till det dubbla, dvs. ca 13 miljoner kr (KoF07) respektive 15 miljoner kr (KoF11). Även LU beräknade de direkta kostnaderna till ca 7 miljoner kr och de indirekta kostnaderna, dvs. det arbete som lärosätet lade på att rapportera in resultat etc. till ungefär lika mycket. För SLU uppgick kostnaderna till 10 miljoner kr och för KTH 6 miljoner för de direkta kostnaderna och ungefär lika mycket till för de indirekta.

panelmedlemmar (30 fördelade på sju expertpaneler, att jämföra med t.ex. UU som hade ca 200 experter fördelade på 25 paneler i KoF11). Den utvärdering som genomförs vid MiUn, HH och HiS delfinansieras av KK-stiftelsen med högst två miljoner vardera till varje lärosäte, dock högst hälften av kostnaden.⁶⁸

4.8 Konsekvenser

Samtliga lärosäten som genomfört granskningar vittnar om att det varit lärorika men arbetsamma processer. Många har arbetat under ett mycket snävt tidsschema, vilket satt stor press framför allt på nyckelpersoner inom lärosätet. Många vittnar om att det tagit längre tid än beräknat att rekrytera experter till panelerna.

Som framgått har en av de stora utmaningarna varit *hur* man definierar en utvärderingsenhet och *vem* som gör det. Det har varit problematiskt att de utvärderade enheterna har skilt sig åt vad gäller exempelvis storlek och förutsättningar. När de utvärderade enheterna inte sammanfallit med lärosätets organisatoriska enheter har det försvårat implementeringen av eventuella åtgärder till följd av utvärderingens resultat. Vidare har det gjort jämförelser inom lärosätet i princip omöjliga.⁶⁹

Ytterligare en omständighet som försvårat jämförelser har varit att inte alla forskningsaktiviteter har inkluderats i utvärderingarna. Vissa lärosäten har haft som princip att inkludera allt, medan andra utvärderingar har byggt mer eller mindre på frivillig medverkan.

Något som är viktigt att understryka är vikten av att från början klargöra både syfte och hur man avser hantera resultaten. Det har inte varit ovanligt att de utvärderade enheterna haft orimliga förväntningar på vad man kunde förvänta sig efter avslutad utvärdering i termer av extraresurser och liknande, vilket i sin tur lett till viss frustration och besvikelse.

Det var endast två lärosäten (KTH och KI) som hade bibliometri som direkt underlag till panelerna. Övriga genomförde bibliometriska analyser parallellt med expertgranskningen för att stämma av resultaten mot varandra. I efterhand⁷⁰ var övriga lärosäten (GU, LU, SLU, UU och ORU) överens om att bibliometri bör användas som ett av flera underlag till panelerna och inte som eget underlag. Panelerna gör ändå sina egna bibliometriska analyser och det är olyckligt eftersom dessa inte kvalitetsgranskas på sedvanligt sätt. Att analysera bibliometriska sammanställningar kräver dessutom särskild kompetens för att kunna göra välgrundade slutsatser och analyser.

68 ”All forskning vid Högskolan utvärderas nästa år i ARC13”.

<http://www.hh.se/pressmedia/aktuellt/nyheter/nyhetsarkiv/nyheter/allforskningvidhogscolanutvarderasnastaariarc13.65437331.html>

69 Jämförelser *mellan* lärosäten är naturligtvis ännu svårare, givet de olika uppläggen. Detta har emellertid inte heller varit något syfte för utvärderingarna.

70 Synpunkterna framfördes vid ett SUHF-seminarium i Uppsala 4 december 2012.

5. SLUTSATSER

Arbetsgruppens huvudsakliga slutsatser efter kartläggningen av kvalitetsgranskningssystem i Storbritannien, andra länder och på svenska lärosäten sammanfattas nedan.

- Syftet med att införa ett kvalitetsgranskningssystem bör vara tydligt uttryckt och styra vilken modell som används. Det är inte lämpligt att importera något annat lands system direkt, hänsyn bör tas till aktuella förutsättningar.
- Det är att rekommendera att starta med en så enkel och okomplicerad modell som möjligt, som inte kräver mer resurser i fråga om tid och pengar för alla inblandade än vad som är absolut nödvändigt för att uppnå ett tillräckligt gott resultat. En viktig princip bör vara att fördelningen av basanslaget är transparent.
- Kommunikation och konsultation med relevanta aktörer är viktigt för att säkerställa ett system som ger avsedd effekt, liksom att genomföra pilotförsök. Utvecklingsprocessen av ett nytt system måste få ta tid. Det är också av största vikt att göra en konsekvensanalys av ett förslag till system innan det genomförs, dels för att identifiera hur det kan komma att påverka lärosätena, dels för att identifiera eventuella oönskade systemeffekter.
- Prestigen kan i en del fall vara lika viktig som den direkta medelstilldelningen. Ett styrande incitament kan därför kanske uppnås utan att så stor del av basanslagen fördelas på basis av resultatet av kvalitetsgranskning. Arbetet inför en utvärdering kan vara minst lika kvalitetsdrivande som själva utfallet.
- Man bör också ha i åtanke att utformningen av ett system kan påverka beteenden t.ex. ifråga om mönster för samarbeten, publicering samt inställningen till betydelsen av forskningsresultatens samhällsrelevans och nytta.
- Nationella kvalitetsgranskningssystem av forskning sätter igång strategiarbeten på lärosätena, men det brittiska systemet visar att de inte nödvändigtvis ersätter lärosätenas behov av egna uppföljningar och utvärderingar av sin verksamhet.
- I Sverige måste vi rimligtvis använda oss av internationella vetenskapliga experter för att undvika jäv. Internationella experter kan inte förväntas lägga ned lika mycket arbete som om man anlitar nationella experter. Panelerna bör få tydliga direktiv och planering kan i huvudsak ske tillsammans med ordförandena, resten av experterna kan involveras i ett senare skede för att förenkla processen.
- Bibliometri kan användas som ett (av flera) underlag för bedömning av vetenskaplig kvalitet inom flera områden som naturvetenskap, teknik och medicin, men troligtvis inte inom alla ämnesområden.
- Bedömning av forskningens relevans och nytta kan möjligtvis ske på lärosätetsnivå istället för per disciplinär utvärderingsenhet. Fallstudier verkar vara en metod som fungerar för alla ämnesområden. Dock verkar det inte vara rimligt att kräva att varje enskild forskare ska kunna visa på forskningsresultatens vidare samhällspåverkan.
- Ett smidigt inrapporteringssystem minskar arbetsbördan för alla inblandade. Förutsättningarna är relativt goda i Sverige tack vare ansöknings- och återrapporteringssystemet Prisma som utvecklas gemensamt av Vetenskapsrådet, FAS och Formas. Systemet kommer att vara öppet för fler aktörer att ansluta sig till.

6. FRÅGOR SOM BEHÖVER UTREDAS VIDARE INFÖR FÖRSLAG TILL ETT SVENSKT SYSTEM

Med utgångspunkt i regeringsuppdraget ”Uppdrag att utreda och lämna förslag till en modell för resursfördelning till universitet och högskolor som innefattar kollegial bedömning av forskningens kvalitet och relevans” (U2013/1700/F) dnr 111-2013-2496 och lärdomarna från förstudien, har arbetsgruppen identifierat ett antal frågor som behöver utredas vidare. För varje fråga finns åtskilliga detaljer, som inte specificeras här, men som behöver utredas i det kommande projekt för att lösa regeringsuppdraget.

Den övergripande frågan

Hur bör ett system vara konstruerat för att på ett så enkelt sätt som möjligt (i termer av begriplighet, krav på resurser, tid och kostnader för alla inblandade) premiera och driva kvalitet och prestation i forskning och samtidigt kunna utgöra ett robust underlag för resursfördelning?

Vad ska utvärderas?

Ska all universitetsbaserad forskning omfattas av utvärdering eller kan ett urval ske? Hur ska i så fall urvalet gå till? Hur avgränsningen av utvärderingsenheter ska göras är en viktig fråga. Kan avgränsningen göras så att befintlig statistik kan användas? Finns det vissa underlag som kan redovisas på lärosätetsnivå?

Hur ska bedömningen gå till?

Bör samma modell användas för utvärdering av alla ämnesområden, som i Storbritannien – eller kan man ha olika metodik och underlag som i t.ex. Australien?

Hur ska expertpanelerna vara sammansatta? Detta hänger bl.a. samman med hur enheterna för utvärdering avgränsas. Hur ska medlemmarna rekryteras och vad ska de leverera? Rollerna för vetenskapliga experter respektive användareexperter behöver förtydligas.

Vilka underlag ska panelerna ha för sin bedömning av vetenskaplig kvalitet respektive forskningens relevans och nytta? Hur ska underlagen vara utformade för att svara mot systemets syfte och förutsättningarna att samla in dem? Ska experter bedöma forskningens kvalitet genom att faktiskt läsa publikationer/motsvarande eller kan det göras på annat sätt? Vad kan redovisas med hjälp av statistiska data, kvantitativa indikatorer och bibliometri? Hur kan olika indikatorer användas – fristående eller som underlag för bedömning av paneler? Vad redovisas bäst i berättande text? Att bedöma forskningens relevans och nytta är en relativt ny företeelse i Sverige och kräver ett särskilt utvecklingsarbete.

Vilka pilotstudier bör och hinner genomföras?

Hur kan Prisma användas som inrapporteringsystem? Vilka anpassningar krävs?

Resursfördelning baserat på utvärderingsresultatet

På vilka sätt kan utvärderingens resultat kopplas till beslut om resursfördelning? Hur ska ämnesområden vägas mot varandra? Hur stor andel av fördelningen ska bero på olika aspekter av forskningens kvalitet och relevans?

Förankring, samråd och samverkan

Kommunikation och konsultation med sektorn bör ske, men när, var, hur och med vilka instanser bör det ske?

Samordning bör ske med VINNOVA:s uppdrag N2013/1162/FIN om att ”utforma metoder och kriterier för bedömning av prestation och kvalitet i läorsätenas samverkan med det omgivande samhället, i termer av relevans och nyttiggörande av forskningsbaserad kunskap”.

Samordning bör också ske med regeringens uppdrag till Vetenskapsrådet (U2012/6798/F, dnr 111-2012-7079) om att ”utveckla och föreslå en modell för kvalitetsutvärdering av klinisk forskning som finansieras helt eller delvis via s.k. ALF-medel”.

Finns det samordningsvinster med att synkronisera Vetenskapsrådets förslag till modell och Universitetskanslersämbetets uppdrag att utvärdera forskarutbildning?

Konsekvensanalys

Vilka konsekvenser (önskade och oönskade) kan man förutse om det utarbetade förslaget till modell införs? Hur mycket driver modellen forskningens kvalitet, produktion och relevans inom högskolan? Vilka kostnader kommer förslaget att innebära?

Organisation

Vilken instans (t.ex. motsvarande HEFCE i Storbritannien) bör förslagsvis ha ansvar för genomförandet?

Ytterligare kartläggning

En grundligare kartläggning bör göras av Australiens system och ev. ytterligare länder.

Implementeringen av och resultaten från REF2014 bör följas.

7. REFERENSER

Kapitel 1 Inledning

Prestationsbaserad resurstilldelning för universitet och högskolor (U2011/7356/UH). Ds 2011.

Vetenskapsrådet PM 2011 *Insatser för att stärka Sverige som forskningsnation (dnr 311-2011-1306)*

Vetenskapsrådets remissvar 2012 (dnr 112-2012-6) på utredningen *Prestationsbaserad resurstilldelning för universitet och högskolor (U2011/7356/UH)*

Vetenskapsrådet PM 2012 (dnr 111-2012-6066) *Utformning av ett system för peer review utvärdering som underlag för prestationsbaserad resurstilldelning för universitet och högskolor*

Regeringens proposition *Forskning och Innovation 2012/13:30*

Regeringsuppdrag (2013) *Uppdrag att utreda och lämna förslag till en modell för resursfördelning till universitet och högskolor som innefattar kollegial bedömning av forskningens kvalitet och relevans. U2013/1700/F. Dnr 111-2013-2496.*

Kapitel 2 REF2014 i Storbritannien

Dokumentation om REF finns tillgänglig på www.ref.ac.uk

<http://www.ref.ac.uk/background/bibliometrics/>

<http://www.ref.ac.uk/background/pilot/>

<https://www.researchfish.com/>

<http://www.rcuk.ac.uk/research/researchoutcomes/pages/ROS.aspx>

<http://www.ref.ac.uk/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf>

<http://www.hesa.ac.uk>

HEFCE (2009/38) *Research Excellence Framework: Second consultation on the assessment and funding of research.*

HEFCE (2009/39) *Report on the pilot exercise to develop bibliometric indicators.*

Technopolis (2009) *First report on 'lessons learned' from the bibliometrics pilot exercise: data collection.*

Technopolis (2009) *Second report on 'lessons learned' from the bibliometrics pilot exercise: outcomes.*

RAND Europe (Grant, J., Brutscher, P-B., Kirk, S., Butler, L. & Wooding, S.) (December 2009): *Capturing Research Impacts – A review of International Practice.* Prepared for the Higher Education Funding for Council for England.

HEFCE 2010. *Guide to Funding – How HEFCE allocates its funds.*

HEFCE (March 2010). *Impact Pilot Exercise. Guidance to expert panels.*

Research Excellence Framework (2010) *Impact pilot exercise. Findings of the expert panels.*

Technopolis (November 2010) *REF Research Impact Pilot Exercise. Lessons-Learned Project: Feedback on Pilot Submissions, Final report.*

HEFCE (December 2010). *Workshops on the impacts of research in the practice-based creative and performing arts, the humanities and social sciences.*

HEFCE REF01.2010. REF2014 *Units of assessment and recruitment of expert panels.*

Research Excellence Framework (2010) *Codes of Practice on the selection of staff.*

HEFCE (2011, updated 2012) REF02.2011. REF2014 *Assessment framework and guidance on submissions.*

HEFCE REF01.2012. REF2014 *Panel criteria and working methods.*

Vetenskapsrådet PM 2012 Forskningsinformation 3.0 (dnr 111-2012-203)

Kapitel 3 Nationella kvalitetsgranskningssystem i andra länder

Australien:

<http://www.arc.gov.au/era/>

Australian and New Zealand Standard Research Classification (ANZSRC).

Donovan, C. (2008). The Australian Research Quality Framework: A live experiment in capturing the social, economic, environmental, and cultural returns of publicly funded research. In: C.L.S. Coryn & M. Scriven (Eds.), *Reforming the evaluation of research. New Directions for Evaluation*, 118, 47-60.

Mejlkorrespondens med Mary Kelly, ARC, 2013.

Danmark:

<http://fivu.dk/uddannelse-og-institutioner/videregaende-uddannelse/universiteter/okonomi/basismidler-efter-kvalitet>

Aftale om fordeling af globaliseringsmidlerne til forskning og udvikling. I: Aftaler om Finansieringsloven for 2009. (Finansministeriet Danmark, 5 november 2008)

Aftale mellem regeringen (Venstre og Det Konservative Folkeparti), Socialdemokraterne, Dansk Folkeparti og Det Radikale Venstre om ny model for fordeling af basismidler til universiteterne (30 juni 2009)

Finland:

www.minedu.fi

Undervisnings- och kulturministeriet, Finland (2011). Kvalitetsinriktat, internationellt, profilerat och verkningsfullt universitet – ett förslag till finansieringsmodell för universiteten fr.o.m. år 2013. ([På finska med svensk sammanfattning](#))

Flandern:

<http://www.ecoom.be/en/services/bof>

<http://www.ewi-vlaanderen.be/ewi/wat-doen-we/programmas-subsidies/financiering-van-onderzoek/bijzondere-onderzoeksfondsen>

Presentation av Tim Engels (Centre for R&D Monitoring, University of Antwerp), i Helsingfors, 2012.

Italien:

<http://www.anvur.org/index.php?q=it/homepage&lang=it>

VTR (Valutazione Triennale della Ricerca 2001-2003), VQR (Valutazione della Qualità della Ricerca) 2004-2010.

<http://www.metrisnet.eu/metris/index.cfm/report/findByStructureAndCountry/38/44>

Franceschet, M. och Costantini, A. (2011) *The first Italian research assessment exercise: a bibliometric perspective*, Journal of Informetrics 5(2): 275–291.

Nederländerna:

Standard Evaluation Protocol 2009-2015

http://www.knaw.nl/Content/Internet_KNAW/publicaties/pdf/20091052.pdf

The science system in the Netherlands. An organisational overview (2012).

<http://www.government.nl/documents-and-publications/leaflets/2012/04/17/the-science-system-in-the-netherlands.html>

Norge:

NOU (Norges offentliga utredningar) 2011:6 Et åpnere forskningssystem, (Kapitel 5.2 Resultatbasert forskningssystemer).

Nya Zeeland:

Tertiary Education Commission <http://www.tec.govt.nz/Funding/Fund-finder/Performance-Based-Research-Fund-PBRF-quality-evaluation/>

Övrigt:

<http://www.aeres-evaluation.com/>

Butler, L. (2010). Performance-Based Funding for Public Research in Tertiary Education Institutions, Workshop proceedings, OECD, 2010.

Hicks, D. (2012). Performance-based university research funding systems. Research Policy 41, 251–261.

Times Higher Education 14th February 2013.

Prestationsbaserad resurstilldelning för universitet och högskolor (U2011/7356/UH). Ds 2011.

Kapitel 4 Kvalitetsgranskning av forskning vid svenska lärosäten

Göteborgs universitet:

Research evaluation for development of research 2010 (RED10)
<https://gupea.ub.gu.se/handle/2077/24885>

Karolinska institutet:

External Research Assessment (ERA10) <http://ki.se/ki/jsp/polopoly.jsp?d=36227&l=en>

Kungl. tekniska högskolan:

Research Assessment Exercise 2008 respektive 2012 (RAE2008 och RAE2012)
<http://www.kth.se/en/forskning/rae/2008>
<http://www.kth.se/en/forskning/rae>

Lunds universitet:

Research Quality Assurance for the Future (RQ08)
<http://www5.lu.se/anstaellid/forska/utv-rdering-av-forskning---rq08>

Sveriges lantbruksuniversitet:

Kvalitet och nytta (KoN09)
http://www.slu.se/Documents/externwebben/overgripande-slu-dokument/Kvalitet%20och%20Nytta/KoN_A_webb.pdf

Uppsala universitet:

Kvalitet och förnyelse 2007 respektive 2011 (KOF07 och KOF11)
<http://uu.diva-portal.org/smash/record.jsf?searchId=2&pid=diva2:43034>
<http://uu.diva-portal.org/smash/record.jsf?pid=diva2:461235>

Örebro universitet:

Örebro Research Evaluation (ÖRE10).
<http://www.oru.se/Om-universitetet/Kvalitet/Forskningsutvardering-ORE2010/>

Övrigt

Intern slutrapport från projektledningen avseende forskningsutvärderingen Lund University Research Quality Assurance for the Future - RQ08, Bengt Söderström

Minnesanteckningar från ett SUHF-seminarium för utbyte och uppsamling av erfarenheter av egeniniterade forskningsutvärderingar, Uppsala 4 december 2012

Mittuniversitetet, Handlingsplan för Mittuniversitetets forskningsutvärdering 2013, ARC13: processbeskrivning, organisation och ansvarsfördelning [okänt datum].

Mittuniversitetet, ARC13 – Evaluation Package [okänt datum]

All forskning vid Högskolan utvärderas nästa år i ARC13.
<http://www.hh.se/pressmedia/aktuellt/nyheter/nyhetsarkiv/nyheter/allforskningvidhogskolanutvarderas-nastaariarc13.65437331.html>

Mejlkorrespondens:

Per Andersson, utbildningsledare Uppsala universitet.

Jonas Harvard, Mittuniversitetet.

Andreas Boberg, Utbildnings- och forskningssektionen, Mälardalens Högskola.

Kapitel 5 Slutsatser

Inga referenser

Kapitel 6 Frågor som behöver utredas vidare inför förslag till ett svenskt system

Regeringsuppdrag (2012). Uppdrag i fråga om modell för kvalitetsutvärdering av klinisk forskning finansierad av vissa medel. U2012/6789/F. Dnr 111-2012-7079.

Regeringsuppdrag (2013) Uppdrag att utreda och lämna förslag till en modell för resursfördelning till universitet och högskolor som innefattar kollegial bedömning av forskningens kvalitet och relevans. U2013/1700/F. Dnr 111-2013-2496.

Ändring av regleringsbrev för budgetåret 2013 avseende Verket för innovationssystem inom utgiftsområde 24 Näringsliv. (2013) N2013/1162/FIN

8. INTERVJUER VID STUDIEBESÖK I STORBRIANNIEN 5–15 FEBRUARI 2013

University of Warwick, Coventry 5 februari

Prof Lawrence S. Young Pro Vice-Chancellor for Research (Life Sciences and Medicine) and Capital Development

Dr Giles Carden, Director, Strategic Planning and Analysis

Deltagare från Sverige:

Sara Monaco och John Tumpane

Department of Business, Innovations and Skills, London 5 februari

Carolyn Reeve, Assistant Director, Economic Impact Team

Paul Williams, Research councils

Rosa Fernandez, Research economics

Deltagare från Sverige:

Carl Jacobsson, Maud Quist, Peter Schilling och Sten Söderberg

Medical Research Council, London 6 februari

Ian Viney PhD, Head of Strategic Evaluation

Deltagare från Sverige:

Sara Monaco, Peter Schilling och Sten Söderberg

Quality Assurance Agency for Higher Education (QAA), London 6 februari

Anthony McClaran, Chief Executive

Janet Bohrer, Assistant Director, Research, Development and Partnerships

Harriet Barnes, Development Officer, Research, Development and Partnerships

Deltagare från Sverige:

Carl Jacobsson, Sara Monaco, Maud Quist, Peter Schilling, Sten Söderberg och John Tumpane

Higher Education Funding Council for England (HEFCE), London 7 februari

Sir Alan Langlands, Chief Executive

David Sweeney, Director (Research, Innovation and Skills)

Deltagare från Sverige:

Peter Eriksson, Carl Jacobsson, Maria Landgren, Sara Monaco, Maud Quist, Peter Schilling, Sten Söderberg och John Tumpane

Higher Education Funding Council for England (HEFCE), Bristol 8 februari

Paul Hubbard, Head of Research Policy

Graeme Rosenberg, REF Manager

Vicky Jones, Deputy REF Manager

Deltagare från Sverige:

Peter Eriksson, Carl Jacobsson, Maria Landgren, Sara Monaco, Maud Quist, Peter Schilling, Sten Söderberg och John Tumpane

University of Bristol, Bristol 11 februari

Professor Eric Thomas, Vice-chancellor University of Bristol, Universities UK President and Chair

Chris Hale, Policy Advisor

Professor Guy Orpen, Pro Vice-Chancellor, Research and enterprise

Lesley Dinsdale, Head of Research and Enterprise Policy
Mervyn Miles, Professor, Director, Nanoscience & Quantum Information Centre, H.H. Wills Physics
Laboratory

Deltagare från Sverige:

Sara Monaco, Maud Quist och Erik Roos

Universities UK, Bristol 11 februari

Professor Eric Thomas, Vice-Chancellor University of Bristol, Universities UK President and Chair

Deltagare från Sverige:

Maud Quist och Erik Roos

University of Exeter, Exeter 11 februari

Dr Michael Wykes, Policy, Impact, and Performance Manager

Dr Matt Baker, Research Evaluations and Communications Manager

Sean Fielding, Director of Research and Knowledge Transfer

Professor Debra Myhill, Associate Dean for Research and Knowledge Transfer in the College of So-
cial Sciences

Dr John Plunkett, Associate Professor English

Deltagare från Sverige:

Sten Söderberg och John Tumpane

Research Councils UK, Swindon 12 februari

Biotechnology and Biological Sciences Research Council

Mari Williams, Deputy Director, Corporate Policy and Strategy

Engineering and Physics Research Council

Sue Smart, Head of Performance and Evaluation

Arts and Humanities Research Council

Christopher Walker, Evaluation Manager

Deltagare från Sverige:

Sara Monaco, Rolf Nilsson, Maud Quist, Erik Roos, Sten Söderberg och John Tumpane

University of Reading, Möte på Paddington Station 6 februari

Prof. Christine Williams Pro Vice-Chancellor Research and Innovation.

Deltagare från Sverige:

Carl Jacobsson, Maud Quist och John Tumpane

University of Reading, Reading 14 februari

Hugo Tucker, Dean of the Faculty of Arts, Humanities and Social Science

George Marston, Dean of the Faculty of Science

Rachel Stewart, Acting Director of Planning Support

Nathan Helsby, Planning Support Manager

Emily Hake, PA to Pro-Vice-Chancellor Research and Innovation

Bonhi Bhattacharya, REF intern

Deltagare från Sverige:

Erik Roos, Sten Söderberg och John Tumpane

Imperial College, London 14 februari

Professor Donal Bradley, Pro Rector Research

Miss Josie Lewis-Gibbs, Senior Planning Officer, Strategic Planning Division

Dr Elise Laird, Research Impact Officer

Professor Joanna Haigh, Head of Department & Professor of Atmospheric Physics
Dr Sarah Perkins, Head of Research Strategy, Faculty of Medicine Centre (seconded by HEFCE to write the criteria for impact case studies)
Dr Caroline Small, Honorary Senior Lecturer, Faculty of Medicine
Dr Chrystalla Orphanides, Departmental Research Co-ordinator, Department of Medicine
Deltagare från Sverige:
Sara Monaco, Rolf Nilsson och Maud Quist

University of Ulster, Belfast 15 februari

Professor Hugh McKenna, Pro Vice-Chancellor Research and Innovation, (även Chair Sub-panel 3 Allied Health Professions, Dentistry, Nursing and Pharmacy)
Dr Marie Patton, Director of the Research Office
Mrs Natalie Dallat, Administrative Manager (Research Policy)
Prof Alastair Adair (Building and Environment)
Prof Bernie Hannigan (Health, Immunology), Chief Scientific Advisor, Chief Medical Officer Group, Department of Health Social Services and Public Safety (Northern Ireland)
Prof Paul Seawright (Art & Design)
Prof Richard Barnett, Vice Chancellor University of Ulster
Cathy Gormley-Heenan, Research Institute Director, Social Sciences
Jim McLaughlin, Research Institute Director, Engineering
Sandy Steacy, Research Institute Director, Environmental Sciences
Bryan Scotney, Research Institute Director, Computer Science Research
Susan Hodgett, Social Sciences, även Deputy Chair (vice ordf.) för REF Subpanel 27 Area Studies
Deltagare från Sverige:
Sten Söderberg och John Tumpane

APPENDIX 1 INTERVJUGUIDE

Följande frågebatteri användes som underlag för semistrukturerade intervjuer vid studiebesöket i Storbritannien och skickades ut i förväg till aktuella instanser.

QUESTIONS to BIS and HEFCE

1) Objectives and development of the system

- a) Describe the process in which the REF was developed. Why was a quality assessment system of research introduced and why was it necessary to change it (from RAE to REF)?
- b) Are there any other purposes than resource allocation e.g. quality, streamlining/profiling of HEIs, promoting diversity, equality, etc.?
- c) What is the main purpose of the system (to reward the best or sort out the worst - or both)? Is it primarily a quality assurance exercise or a funding allocation exercise?
- d) At the policy level is the REF system designed to complement other means of allocating funding? (e.g. teaching funding, research funding through research councils etc.) Are there any synergic effects or conflicting incentives?
- e) What are the intended and unintended consequences/outcomes of RAE/REF? Are there also examples of unrealised/unfulfilled expected outcomes?
- f) What are the consequences on the local/HEI level? Do you know how the RAE/REF outcomes have affected e.g. local development work, local resource allocation etc.?

2) Units of assessment

- a) Why are units of assessment used as the basis for the evaluation and not another level (e.g. researcher level, university level, etc.)? And how is this aggregated to funding at the university level?
- b) How are the units of assessment defined and decided upon? (*Units of assessment, there are 36 in REF14*). Why are they disciplinary, are they approx. equal in volume, why 36 in REF and 67 in RAE. Are there restrictions in order to improve the cost-benefit of the assessment exercise? Is it more important that they have the right volume or a more correct disciplinary boundary?
- c) How is interdisciplinary research handled within RAE/REF? How is research collaboration between universities handled within RAE/REF?
- d) How small volumes are relevant for submission?
- e) Is there any kind of measure of efficiency, i.e. output in relation to staff and resources?
- f) REF covers only a selection of research outputs- and impact and environment are assessed only as regards the submitted material. How much of the total share of research in HEIs is estimated to be covered by REF?
- g) How representative is that selection to all research at a particular university included in RAE/REF (and who decides, HEIs or HEFCE)? How representative of UK university research is that selection? (Is it appropriate only to assess the best research as chosen by the universities?) Is the total share of HEI university research evaluated in RAE/REF or only the share financed by HEFCE?
- g) To what extent do the assessments cover retroactive aspects (tangible results), and to what extent do they cover future potential?

3) The assessment process

Criteria and data

- a) How is scientific quality assessed?
- b) How is future research potential assessed?
- c) How is impact on society assessed?
- d) How is research environment assessed?

- e) Are indicators of research volume and activity collected and assessed (as in the Australian ERA)? If not, why not?
- f) Are indicators of recognition (esteem measures) collected and assessed (as in the Australian ERA)? If not, why not?
- e) What bibliometric measurements and metric indicators can be used? To what extent do the panels take the bibliometric measurements into consideration? Is there a problem in terms of equivalence that some units of assessment/sub-panels will use bibliometric measurements while others will not?
- f) How are indicators used as a basis for resource allocation - as an "independent track" or as part of other data assessed by experts?
- g) How were aspects/criteria/indicators developed? Describe the process.
- h) Are the same type of data used for all subject areas?
- i) Are there problems with calibration/comparison of indicators between subject areas?
- j) Is there a possibility to complement data after the final date of submission? How? If yes, who can initiate such completions?
- k) Are there site visits? If yes, who can initiate such site visits?

Expert panels and their evaluation

- a) How are the panel experts nominated and by whom? Who makes the final selection/decision on which experts to be included? Can you reflect on REFs use of mainly domestic experts, as opposed to international experts? How are the panels of experts composed? Are there other members than scientific experts? Stakeholders/industry representatives/representatives of the wider society? (Användarexperter)? Describe the role of the various stakeholders (panel chairs, subpanel chairs etc.)
- b) What introduction do the assessors/panel experts get? Who is responsible for the introduction?
- c) Is there a possibility for the panels to ask for further data/explanations etc. if deemed necessary? Can they ask for personal interviews/site visits? Why/why not?
- d) What remuneration do assessors/expert panels receive?
- e) How is the final assessment/result communicated/distributed? What is communicated? Are there only judgements/scores or also recommendations?
- f) If the evaluations are made on a detailed subject level - how are results/assessments aggregated/weighted/linked to higher levels, such as subject level, cluster level or even university level?
- g) How is scientific quality and research impact on society and other aspects weighted? Are there separate scores of the two aspects, or an overall rating?
- h) Who has the final say as regards the results?

4) How do evaluation results inform the distribution of resources?

- a) What proportion of the block grant for research is subject to redistribution?
- b) How redistributive is the system between individual universities?
- c) How redistributive is the system between individual subjects (Main panels och UoAs)?
- d) How are subject areas weighted against each other in the resource allocation process?
- e) How is scientific quality weighted against research impact on society and other aspects in the resource allocation process?

5) Period of evaluation

- a) Do HEIs report and submit data annually that gradually accumulates or is data gathered at one single point (29 Nov 2013)? Why/why not? How did you decide on what research timeframe is allowed for submission?

6) Cost of implementation and operation of the system

- a) What is the estimated cost to implement and operate a performance-based resource allocation system of peer review evaluation, including the costs at participating HEIs?
- b) Which technical system is used to collect data? What was the cost of designing, building and servicing this system? Collaboration with Research Councils UK? Relation to www.researchfish.com?

7) Objections

- a) What were the main problems in RAE and how have you solved them in REF?
- b) What are the advantages and disadvantages of REF in comparison to RAE?
- c) What are the advantages and disadvantages of RAE in comparison to REF?
- d) What is the main critique?

QUESTIONS TO UNIVERSITY PRO VICE-CHANCELLORS RESPONSIBLE FOR RESEARCH AND INNOVATION

1) Organization and strategy of RAE/REF at the university

- a) What is the university strategy for the REF?
- b) How does the university prepare for REF?
- c) How do you organize REF within the university?
- d) Differences between subject areas/faculties?
- e) Can you estimate the cost of REF to the university (time, money, resources)? How do you look upon the costs in relation to the benefits?
- f) How is interdisciplinary research at your university handled within RAE/REF? How is research collaboration between you and other universities handled?
- g) REF covers only a selection of research outputs - and impact and environment are assessed only as regards the submitted material. How much of the total share of research in your university is estimated to be covered by REF? How representative is that selection to all research at your university?
- h) Do you work with the research councils in preparation for REF/RAE – e.g., to identify researchers/units of assessment?

2) Relationships between RAE/REF and other evaluations

- a) Does the university carry out its own evaluations? How do these compare/differ to REF? What is the relation between these?
- b) Does the university take part in evaluations conducted by the research councils? If so, what are the relations between these evaluations and the REF?
- c) Has the RAE/REF reduced the need for other (e.g., in-house) evaluations? Why/why not?

3) Effects of the RAE/REF for your university

- a) How has the outcome of RAE affected the university? In what ways have the RAE/REF outcomes affected e.g. local development work, local resource allocation etc?
- b) How is the block grant for research distributed within the university? According to REF-/RAE-results?

4) Comparisons between RAE and REF

- a) Would the outcome for your university be better under REF or RAE, why?
- b) Advantages and disadvantages of REF compared to the RAE, and vice versa, for your university and more generally?

5) General views on the RAE/REF-systems

- a) What are the main advantages and disadvantages of the REF from the point of view of your university? How accepted and legitimate would you say the system is at your university?
- b) Do you have any views on the impact measurement, in comparison to the measurement of scientific quality?
- c) What are the intended/unintended [secondary] consequences (positive or negative) of REF (RAE)? Are there any unfulfilled [secondary] goals of REF (RAE)?
- d) Is there anything that you would you change in REF to make it better in the future?

SPECIFIC QUESTIONS TO UNIVERSITY OFFICIALS WHO HAVE BEEN (RAE)/ARE RESPONSIBLE (REF) FOR COORDINATING THE SUBMISSION OF INFORMATION FROM INSTITUTIONS/DEPARTMENTS TO HEFCE

- a) How does the university decide what units of assessment it will enter in for evaluation? Does the university have research in certain areas but does not enter these in any REF unit of assessment, if so - why?
- b) How does the university decide which researchers to enter in REF?
- c) How are items of research for the evaluation selected and how do you ensure that they adhere to the guidelines?
- d) How do you choose the four outputs for each researcher?
- e) Estimation of time and resource consumption (cf. the size of HEI, outcome)
- f) When did the preparation for the REF14 start?
- g) How do you plan and organize the process within the university?
- h) Is it easier for some subject areas/faculties to prepare for and deliver data to REF?
- i) Administrative systems and support staff?
- j) What were the main problems in RAE and how have you solved them in the REF?
- k) What are the advantages and disadvantages of RAE/REF, and vice versa?

SPECIFIC QUESTIONS TO DEANS AND/OR DEPARTMENT HEADS AND/OR INDIVIDUAL RESEARCHERS

- a) Estimate the time put into RAE/REF by your staff/yourself.
- b) What were the main problems in RAE and how have you solved them in REF?
- c) What are the advantages and disadvantages of RAE/REF, and vice versa?
- d) Is REF equally well designed for all subjects and universities (in respect to size, scholarly tradition, etc.)

SPECIFIC QUESTIONS TO PERSONS WHO HAVE BEEN INVOLVED IN RAE AND OR THE REF PILOT STUDIES AS ASSESSORS

- a) Multi-disciplinary units of assessments, how are they handled?
- b) Compare RAE with REF from the point of view of an assessor – differences?
- c) How much time and effort does it involve on your part? Is this equal for all assessors?
- d) Why do you participate in REF? (is it a merit, well-paid, educational, etc.)
- e) What are the advantages/disadvantages of being member of a REF panel?
- f) How do you handle conflicts of interest?
- g) Your reflections on REF's use of mainly domestic experts, as opposed to international experts?

TILLÄGGSFRÅGOR

Dessa frågor användes av gruppen för att komplettera frågorna ovan på plats vid studiebesöket i Storbritannien.

- a) Hur stor frihet har lärosätena att INTE rapportera in sådan forskning som uppfyller minsta-volymp-kriterierna (om de nu inte skulle vilja göra det)?
- b) Hur stor andel av forskningen vid respektive lärosäte, inom respektive UoA (och totalt) faller utanför minsta-volymp-kriterierna och utvärderas inte av det skälet?
- c) Gäller rapporteringskravet all forskning som uppfyller minsta-volymp-kriterierna oavsett finansiering? Eller enbart den HEFCE-finansierade forskningen? Ingår t.ex. uppdragsforskning som till 100 procent är finansierad av industrin (om det nu finns sådan forskning)?
- d) Panelerna har i uppgift att klassificera all forskning som lärosätena har redovisat – de kan inte avstå från gradering och betygssättning. Däremot kan de ange ”unclassified”, ett betyg som dock knappast ligger i lärosätenas intresse att få. Vi skulle behöva få en uppfattning av andelen ”oklassificerad” forskning både totalt, per lärosäte, och per UoA
- e) Hur stor legitimitet som panelernas sammansättning och ämnesbredd anses ha på lärosätena?
- f) Samlas bibliometriska uppgifter in årligen och/eller i anslutning till respektive REF-utvärdering vart sjätte år.
- g) Vilken är kopplingen mellan grundutbildning/forskarutbildning och forskning i termer av hur resurser fördelas? Separata anslag? (jmf basanslag som ska gå till forskning och forskarutbildning.)
- h) Uppmuntras forskningsanknytning – utbildningsanknytning genom särskilda styrinstrument/resurser?
 - i) Avgränsning och urval, aggregering och viktning, för rättvisande bild
 - j) Aspekter och kriterier
 - k) Utvärdering av relevans och nytta för samhället
 - l) Tvärvetenskap
 - m) Peer review-panelernas sammansättning
 - n) Omsättning av utvärderingens resultat till resurser
 - o) Kostnadseffektivt och kvalitetsdrivande
 - p) Undvika ev. oönskade konsekvenser

APPENDIX 2 UNITS OF ASSESSMENT I REF2014 I STORBRITANNIEN

Huvudpanel A

1. Clinical Medicine
2. Public Health, Health Services and Primary Care
3. Allied Health Professions, Dentistry, Nursing and Pharmacy
4. Psychology, Psychiatry and Neuroscience
5. Biological Sciences
6. Agriculture, Veterinary and Food Science

Huvudpanel B

7. Earth Systems and Environmental Sciences
8. Chemistry
9. Physics
10. Mathematical Sciences
11. Computer Science and Informatics
12. Aeronautical, Mechanical, Chemical and Manufacturing Engineering
13. Electrical and Electronic Engineering, Metallurgy and Materials
14. Civil and Construction Engineering
15. General Engineering

Huvudpanel C

16. Architecture, Built Environment and Planning
17. Geography, Environmental Studies and Archaeology
18. Economics and Econometrics
19. Business and Management Studies
20. Law
21. Politics and International Studies
22. Social Work and Social Policy
23. Sociology
24. Anthropology and Development Studies
25. Education
26. Sports-Related Studies

Huvudpanel D

27. Area studies
28. Modern Languages
29. English Language and Literature
30. History
31. Classics
32. Philosophy
33. Theology and Religious Studies
34. Art and Design: History, Practice and Theory
35. Music, Drama, Dance and Performing Arts
36. Communication, Cultural and Media Studies, Library and Information Management