

# An introduction to the Research Excellence Framework: A new research evaluation framework for universities in the UK

— A comparison with the status of research evaluation in Japan —

Ryu OHTANI<sup>1\*</sup>, Mariko KAMO<sup>2</sup> and Naoto KOBAYASHI<sup>3</sup>

[Translation from *Synthesiology*, Vol.6, No.2, p.118-125 (2013)]

This article introduces the Research Excellence Framework (REF): a new research evaluation framework developed by the Higher Education Funding Council for England (HEFCE) to evaluate the research activities of higher education institutions in the UK. To promote excellent research so that its potential effect can be maximized, the REF has three criteria for evaluating institutions: output quality, impact, and research environment. Remarkably, although the purpose of the REF is university evaluation, social impact is explicitly included within the evaluation framework. It is proposed that output quality be evaluated by conventional review with the aid of quantitative bibliometric indicators, such as the number of paper citations. In this article, we introduce the REF, from output to impact, and make recommendations for the evaluation of research for universities in Japan.

**Keywords** : REF, research and technology development evaluation, excellence, bibliometrics, impact, research environment, outcome

## 1 Introduction

The National Institute of Advanced Industrial Science and Technology (AIST) continuously conducts *Type 1 Basic Research*<sup>[1]</sup>, *Type 2 Basic Research*<sup>[1]</sup>, and *Product Realization Research*<sup>[1]</sup> on industrial technology to realize the philosophy of “*Full Research in Society, for Society*”<sup>[1]</sup>. Its aim is to contribute to the realization of sustainable development by disseminating research results to society. To promote the broad-ranging research activities from basic research to applied research coherently and continuously is a significant new challenge, where a new research evaluation system is urgently needed.

The Higher Education Funding Council for England (HEFCE) provides basic funds to higher education institutions (HEIs) such as universities in the UK to support research infrastructure including full-time staff and their salaries, facilities, libraries, information systems, and so forth. In effective fund apportion, the HEFCE is currently developing the Research Excellence Framework (REF), which is a new, uniform research evaluation framework.<sup>[2]</sup> The REF is expected to be introduced for the evaluation starting in 2014, which will replace the Research Assessment Exercise (RAE), a research evaluation that has been conducted approximately every five years until now. (For the RAE, see Ref. [3].)

One of the major features of the REF is that it has explicit “impact” indicators as an evaluation criterion in addition to that for “output quality” such as the significance of a research

paper. The HEFCE has proposed a weighting scheme that allocates 20 % to impact relative to 65 % to output quality, indicating that the impact is considered important. Because the budget is allocated selectively based on the evaluation, the introduction of this new framework will significantly influence the research activities at UK universities where the ratio of basic research is relatively high. The fact that an impact criterion was explicitly introduced this time suggests that universities in the UK are expected to generate continuous and strategic ripple effects for society through research.

This article intends to provide an overview of the REF based on the second REF draft<sup>[4]</sup> released in 2009. We believe this new evaluation framework in the UK will be beneficial for universities in Japan, which are currently struggling to establish a new evaluation system. In addition, because the REF’s underlying vision to systematically promote research is quite similar to the strategic research and technology development from basic research to applied research undertaken at AIST (referred to as “*Full Research*”<sup>[1]</sup>), we might also expect to gain substantial insights for the evaluation of research activities at AIST. In this article, we have overviewed the concept of the REF and also discussed some implications for research evaluation in Japan.

## 2 The purpose of evaluation

The crucial point to understand before delving into the details of the REF evaluation framework is its rationale for conducting the evaluation. This is because the evaluation procedures (i.e.,

1. Institute of Geology and Geoinformation, AIST Tsukuba Central 7, 1-1-1 Higashi, Tsukuba 305-8567, Japan \* E-mail: ohtani-ryu@aist.go.jp, 2. Evaluation Department, AIST Tsukuba Central 2, 1-1-1 Umezono, Tsukuba 305-8568, Japan, 3. Center for Research Strategy, Waseda University 513 Wasedatsurumaki-cho, Shinjuku-ku 162-0041, Japan

Original manuscript received December 3, 2012, Revisions received January 8, 2013, Accepted January 16, 2013

methods, area of focus, design of review panel, etc.) are all derived from the rationale. In other words, the question of “what is this evaluation for?” becomes the foundation of the grand design for the entire evaluation.

At the center of REF’s vision lies the concept of identifying, rewarding, and promoting research of “excellence.” More specifically, the REF takes into account the UK’s Government aim and expectations as to the role of research in HEIs: they should produce world-leading research across a wide range of academic fields, which in turn would play an essential role in underpinning social and economic growth; they should not only limit themselves to improving the quality of every form of research but also significantly exploit its full potential impact to the economy and society by effectively sharing research findings, disseminating and applying research results, and promoting productive interactions between HEIs and businesses or other users of research results. All these can be made possible by having excellent research as the foundation.

It is for this policy intention that the REF has been developed. The research evaluation should be designed to identify, reward, and promote excellent research effectively and efficiently to achieve this goal. There is a strategic national objective, rather than merely a “show” to fulfill an ostensible responsibility.

To promote excellent research at HEIs and demonstrate their maximum impact potential in this manner, three distinct research evaluation criteria: output quality, impact, and research environment have been introduced in the REF. These criteria have been accessed in advance through pilot exercises regarding the validity of their evaluation methods.<sup>[5]</sup> In the following sections, we have reviewed several key points in each criterion.

### 3 Evaluation of output quality

In the first evaluation criterion, namely output quality, the highest quality research output in a given research unit—the unit of evaluation roughly equivalent to a university’s department—is selected and evaluated. The HEI under evaluation is required to identify excellent research conducted by the members of the research unit during the evaluation period. The quality will then be accessed through expert review with additional evidence such as citation information and other quantitative indicators of the output.

One of the new features in the REF is the introduction of these bibliometric data to reduce the burden of expert review and improve the transparent process. The HEFCE had initially considered replacing expert review completely with quantitative bibliometric assessment, at least in the fields of science, engineering, and medicine. However, pilot exercises and broad discussions concluded that the method is not sufficiently mature and robust to replace expert review. As a

result, citation data will be provided to inform and supplement expert review of outputs in the science-related fields. For other fields such as arts, humanities, and social sciences, use of quantitative indicators are not recommended.

For an expert review that uses bibliometric data, the information relevant to the submitted outputs will be provided. It will be generated using REF’s data collection system with a standardized process. The information is also available upon the HEI’s request prior to the evaluation, which helps not only reduce the burden for preparation of output submission, but also eliminate discrepancy of information by using consistent dataset. Organization of the information system is another characteristic of the REF.

The primary focus to evaluate output quality is to identify excellent research. It focuses on research of the highest quality and is not intended to conduct a comprehensive evaluation of all research outputs conducted in the HEI under evaluation. This is to reflect the underlying policy that excellent research is being identified, rewarded and promoted. Therefore, even if it were feasible to evaluate all outputs, the REF does not look at those who are not engaged in important, high-quality research.

The REF also requires that the HEI under evaluation select outputs to be evaluated themselves. This is to encourage the HEI to recognize that only institutions under evaluation can identify their own research projects/individuals playing substantive roles. This also helps the HEI cultivate awareness on research management at these institutions.

Grey literature and any output published in a non-standard format are eligible for submission as well as traditional outputs such as refereed papers. These include documents published by the government and academic institutions without being circulated in the general publishing market, confidential reports to government, software, and other such cases. Similarly, research whose citation information is not likely to be available and applied research are eligible for submission. Such research includes projects conducted directly for/with the research users and studies to provide information for public policies. The intention is to give equal and maximum consideration to all high-quality research regardless of the format of outputs, not just to outputs that are measurable by quantitative indicators such as impact factors and/or citation index. The quality of output is evaluated based on the levels of “rigor”, “originality”, and “significance”.

### 4 Evaluation of impact

The second criterion, namely impact, evaluates the extent of demonstrable influence built upon excellent research. Here, the term “impact” implies influence on the economy, society, public policy, culture, and quality of life, but does not include intellectual influence on academic communities, which is

evaluated as “output quality”.

It is crucial to understand the concept of “impact” for conducting rigorous evaluation. In the REF, it is necessary for the HEI under evaluation to demonstrate that the impact for which they claim credit was built upon excellent research undertaken by a research unit within. In addition to that, it is required that the unit was actually involved in creating the impact (Figure 1, also see the description). In other words, it is not allowed to claim credit for impact that was developed and exploited by others irrelevant to the research unit even if the impact was originally based on research conducted by the unit.

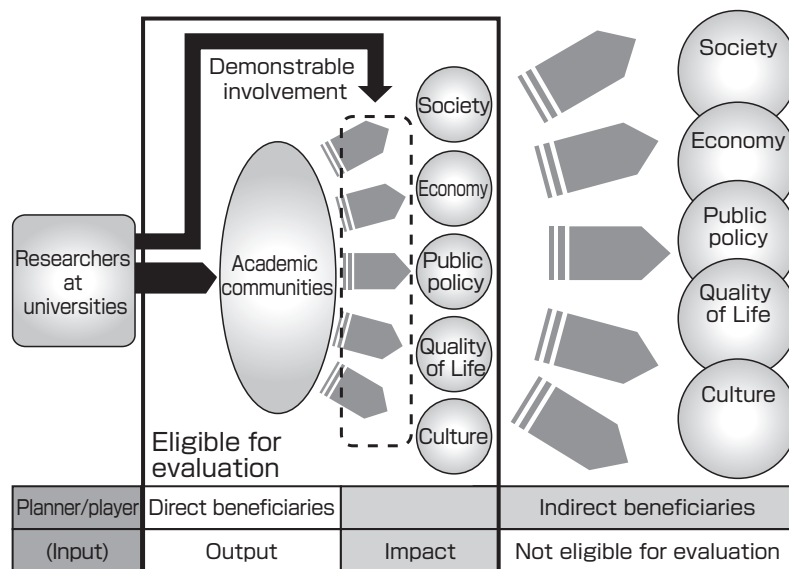
For instance, let us first consider a hypothetical situation where a research unit obtained, compiled, and published a research finding, but did not further attempt to generate a profit by, for example, launching a venture company. Let us now consider that a for-profit company discovered the finding by chance, and consequently generated a vast amount of profit by using it without the unit’s involvement. According to the REF’s definition, this influence of the research finding to the company cannot be regarded as the research unit’s impact. In such a case, only the output quality of the original finding is eligible for evaluation as the achievement of the research unit. Thus, the “impact” defined by the REF is a very specific concept (examples of impact are summarized in Ref. [6]). This is one of the significant differences with Japanese universities’ evaluation where rigorous demonstration and verification for

social/economic impacts are not necessarily required.

In general, it takes considerable time for impact to emerge; an impact that becomes evident during an assessment period is often the result of research during the earlier assessment period. The REF regards the impact as eligible for evaluation as long as it emerges during the assessment period.

The REF states that impact is evaluated at the research unit level because evaluating impact by individual research outputs and/or researchers would not be feasible in practice while evaluating impact by the entire institution (e.g., university) would be too crude. During evaluation, HEIs are required to identify the actions that resulted in impact, and also explain how the impact was exploited and developed.

Since it is generally considered difficult to evaluate impact using quantitative data only, descriptive methods are mainly employed. More specifically, the evaluation is conducted using two descriptive methods—case studies and an impact statement, both of which should comprise narrative explanations with appropriate indicators of impact as supporting evidence. The REF is proposing several standard indicators categorized into the types of impact (shown in Table 1) to reduce the burden of unit’s preparation for evaluation. The extent of impact is assessed by how widely the impacts are seen and how transformative the impacts will be.



**Fig. 1 Definition of output and impact: scope of the REF evaluation**

“Academic knowledge” is produced through research activities by researchers in institutions such as universities, and disseminated to various audiences, such as society and the economy, resulting in multiple ripple effects outside academic communities. In this process, university researchers are the “planners and players” of research, academic communities where the academic knowledge is produced and shared are “direct beneficiaries”, and those who benefit from its ripple effects are “indirect beneficiaries”. In the REF, the term “output” corresponds to this academic knowledge and the term “impact” is defined as the ripple effects that were built on excellent research conducted by researchers whose involvement in exploiting and developing the ripple effects are demonstrable (the area within the dashed line). The scope of the REF evaluation is the outputs of the highest quality and the impact defined above (the area enclosed by the thick line). (Created based on Ref. [10].)

**Table 1. Main types of impact suggested in the REF <sup>[4]</sup>**

① Delivering highly skilled people
② Creating new businesses, improving the performance of existing business, or commercializing new products or processes
③ Attracting R&D investment from global businesses
④ Better informed public policy-making or improved public services
⑤ Improved patient care or health outcomes
⑥ Progress towards sustainable development, including environmental sustainability
⑦ Cultural enrichment, including improved public engagement with science and research
⑧ Improved social welfare, social cohesion or national security

(Source: An excerpt of the key points of “Research Excellence Framework – Second consultation on the assessment and funding of research”, pages 41 and 42<sup>[4]</sup>)

There are several issues that face challenges in evaluating impact: time lags, attribution, and corroboration.

“Time lags” refer to the fact that there can be lengthy time lags between the time at which research is conducted and the time at which its impact emerges. To address this challenge, the REF considers the impact of research over a sufficiently long period and sees broadly the entire impact of the research unit rather than corresponding individual impacts to specific outputs.

The term “attribution” refers to the fact that the process of impact to emerge from research is not linear, and there are numerous factors that affect the formation of impact. Therefore, it is necessary to identify what part of the impact is attributed to the activities of the research unit. To address this issue, the REF focuses on the unit’s research-driven contribution to the exploitation of research in realizing impact. In this process, the REF also recognizes the limitation of quantitative metrics to measure the impacts of research. Thus, impacts are “assessed” qualitatively with supporting evidence, which is one of the reasons the REF came to require the narrative explanations for submission.

Finally, the term “corroboration” refers to various challenges associated with the verification of claims made about the impact. In the REF, third party corroboration and expert panels including the user of the research are formed to investigate the credibility of the evidence submitted by the research unit if necessary.

## 5 Evaluation of research environment and overall score allocation

The third evaluation criterion, namely research environment,

assesses the quality of the research unit’s environment: sustainability, contribution toward stimulating higher education research bases, and the extent of the contribution beyond the institution and the field. This criterion evaluates the research unit’s wide range of support activities and its development of research infrastructure to effectively disseminate and apply excellent research on a continuous basis.

The REF allocates the heaviest weight to “output quality” and a relatively heavier weight to impact than to research environment in accordance with the strategic objective of the REF evaluation explained above. Because impact should be built on excellent research, the allocation influences the overall evaluation outcomes by giving a high score to excellent research that realizes impact.

## 6 Evaluation design

The REF does not evaluate individual researchers, or whole institutes; rather, it conducts evaluation “at the level of coherent research units that produces substantive bodies of work.”<sup>[4]</sup> The units are to be defined by the range of substantive and coherent academic activities, whose level of details is suitable for funding and informing the REF of the research management. This is considered to be equivalent to a university’s department in many cases.

Research units under evaluation should provide submissions that include information on research staff, outputs of high quality, details on qualitatively and quantitatively information on impact and research information, and so forth. The evaluation criteria and process will be set to be consistent across all disciplines and review panels although the REF allows some flexibility in terms of disciplinary differences. This would be useful for maintaining consistency in appraisals for funding purposes, and in eliminating unnecessarily complex procedures in preparation of evaluation for review panels.

## 7 Relevance to outcome-based evaluation and university evaluation in Japan

AIST began incorporating outcome-based evaluation into the evaluation of research units in fiscal year 2005. This is an evaluation method in which the objective, with regard to the outcome of the research to be produced for the social contribution in the future, is explained along with the scenario and roadmap on how to get there.<sup>[7][8]</sup> Initially, there was concern about the feasibility of the evaluation method; however, after more than seven years of experience, the method is well-established today, and it seems that it has had a significant effect.<sup>[9]</sup> What is unique about this evaluation is that an evaluation committee, which comprises of peers and stakeholders (e.g., people from universities, industry, media,

government, independent administrative institutions, etc.), assesses whether the target for future outcomes and the path and methods to get there are reasonable. Although outcomes and impacts are closely related and partially overlap, they are not necessarily the same. According to Hirasawa,<sup>[10]</sup> outcome in the context of research and development refers to the “intrinsic or substantive aspects of the research and development results,” and in the context of policies it is the intrinsic substance aligned to the purpose of the policy. For example, if there is a policy for disease prevention, then the intrinsic substance refers to, for example, trends in the number of patients. On the other hand, impact is positioned as “the ripple effect caused by an indirect result of the research after it has left the research and development sector” in the context of research and development, and as “the ripple effect other than intended results” in the context of policies. When outcome is defined as an intended result, while impact is defined as a ripple effect excluding intended results, it becomes clearly understandable that the cases of research and development—such as the ones at AIST—conducted after setting a clear objective, and the cases of research like the ones conducted with basic university funds to pursue “truth,” have different aspects for evaluation: outcome and impact, respectively.

Meanwhile, under university evaluation (for national universities) in Japan, the National Institution for Academic Degrees and University Evaluation (NIAD-UE) has been assessing the state of education research in the medium-term target period.<sup>[11]</sup> The “Analysis of the State of Undergraduate and Graduate Schools” in Japan, which is roughly equivalent to the REF in the UK, corresponds to the output evaluation in the REF. As part of this research achievement analysis results, there is a section for the university to describe the social, economic, and cultural significance of the research in addition to its academic significance. The university is meant to explain the social, economic, and cultural contribution excellence of the research achievement along with objective indicators. Therefore, it is not that they do not consider the impact of research achievement, although we cannot say that it has incorporated a definite impact evaluation. It is interesting to note that lately the NIAD-UE has considered revising the method for evaluating national universities in the second mid-term period (fiscal year 2010 to 2015).<sup>[12]</sup> In particular, since it is currently at the university’s undergraduate and graduate departments’ discretion to determine how much of the support documents are to be devoted for evaluating academic significance as opposed to social, economic, and cultural significance, discussions are underway to enable the university to submit support documents for the evaluation of both areas of significance, rather than just one.

We believe that the largest difference between the REF and Japan’s university evaluation lies in its utilization. Whereas the results of REF are clearly reflected in changes

in the budget to individual universities, Japan’s university evaluation has a major defect in that the results of the evaluation are not clearly reflected.

## 8 Conclusion

Thus far, we have provided an overview of the REF currently being developed. The fact that the UK has specifically introduced the concept of “impact” to evaluate universities where basic research is primarily conducted, and allocated a major significance to evaluating the dissemination of research results to society is quite characteristic.<sup>(note)</sup> It seems that in the UK,<sup>[13]</sup> the recognition is widespread that universities focusing on research with abundant entrepreneurial spirit to develop international competitiveness are indispensable to the prosperity and well-being of the state. Furthermore, it can be said that the UK is also focusing on promoting a new intellectual industry based even more on university knowledge to compensate for the weakness in manufacturing industry. From these reasons, it is likely that the UK is putting a great effort into the evaluation and the reform of its methods to produce tangible benefits.

This trend is commonly observed in developed countries that are striving to promote innovation, enhance international competitiveness, and build a sustainable society through research and development. This type of impact evaluation might become a useful tool at public research institutions in Japan such as AIST.

Meanwhile, there are debates on the introduction of an “impact” criterion for university evaluation.<sup>[14][15]</sup> The weighting scheme in the second draft that allocated a weight of 25 % to impact and 60 % to output, was revised in the final draft in 2011 to a weight of 20 % to impact and 65 % to output, based on the results of pilot exercises and feedback.<sup>[16]</sup> It will be interesting to see how this will be put into practice in the future.

## Acknowledgment

Dr. Osamu Nakamura, Director of AIST Chugoku, and Dr. Chikao Kurimoto, Principal Evaluator, conducted a detailed peer review and provided valuable feedback to us. We also obtained valuable information regarding university evaluation from Associate Professor Takayuki Hayashi at the National Institution for Academic Degrees and University Evaluation. This article was based on the presentation at the 2nd research evaluation seminar held by the AIST Evaluation Department on March 22, 2012. We would like to take this opportunity to indicate our greatest appreciation to the officials of the Evaluation Department who provided us with this invaluable opportunity.

Note: The Research Council UK (RC) also provides separate

research funding for UK universities (approximately half of the budget for basic research); however, they provide funds by evaluating each project or program independently. While they also have “impact” indicators as an important evaluation criterion, they look at expected future impact from the research results rather than actual impact.

## References

- [1] Editorial Board of *Synthesiology*: “Messages from the editorial board”, *Synthesiology*, 3, i (2010).
- [2] <http://www.ref.ac.uk/>
- [3] S. Iwata: “Toward RAE 2008: research assessment exercise in the UK.” *Research on Academic Degrees and University Evaluation*, 5, National Institution for Academic Degrees and University Evaluation (2007).
- [4] [http://www.hefce.ac.uk/media/hefce1/pubs/hefce/2009/0938/09\\_38.pdf](http://www.hefce.ac.uk/media/hefce1/pubs/hefce/2009/0938/09_38.pdf)
- [5] <http://www.ref.ac.uk/background/>
- [6] <http://www.ref.ac.uk/background/pilot/>
- [7] National Institute of Advanced Industrial Science and Technology Research Evaluation Committee: *Research and Development Evaluation at AIST (Mid-term Summary)* (2004).  
<http://unit.aist.go.jp/plan/research-evaluation/report/report.pdf>
- [8] O. Nakamura, O. Nakamura, M. Takagi Sawada, S. Kosaka, M. Koyanagi, I. Matsunaga, K. Mizuno and N. Kobayashi: Strategic evaluation of research and development in Japan's public research institute, *New Directions for Evaluation*, 118, 25-36 (2008).
- [9] O. Nakamura, S. Ito, K. Matsuzaki, H. Adachi, T. Kado and S. Oka: Using roadmaps for evaluating strategic research and development: lessons from Japan's Institute for Advanced Industrial Science and Technology, *Research Evaluation*, 17, 265-271 (2008).
- [10] R. Hirasawa: Framework of performance measurement and the definition of outcome, *Final reports on “Outcome and Impact Evaluation System for Research and Development”* (2006).  
<http://scfdb.tokyo.jst.go.jp/pdf/20051270/2005/200512702005r.pdf>
- [11] National Institution for Academic Degrees and University Evaluation website:  
[http://www.niad.ac.jp/n\\_hyouka/kokuritsu/index.html](http://www.niad.ac.jp/n_hyouka/kokuritsu/index.html)
- [12] National University Education and Research Evaluation Committee document:  
[http://www.niad.ac.jp/n\\_kikou/shokaigi/hyouka/kokuritsu/\\_icsFiles/afieldfile/2012/06/21/no6\\_9\\_H240619siryou2-3.pdf](http://www.niad.ac.jp/n_kikou/shokaigi/hyouka/kokuritsu/_icsFiles/afieldfile/2012/06/21/no6_9_H240619siryou2-3.pdf)
- [13] G. Roberts: “Igrisu ni okeru kenkyu hyoka no kaikaku (reform of university evaluation in the UK).” *Shinjidai o kirihiraku daigaku hyoka: Nippon to igirisu (university evaluation to open up a new era: Japan and the UK)*, Y. Hada ed., 233-251, Toshindo (2005).
- [14] N. Gilbert: Plans for UK research assessment revealed, *Nature*, doi: 10.1038/news.2009.933 (2009).
- [15] N. Gilbert: UK science will be judged on impact, *Nature*, doi: 10.1038/468357a (2010).
- [16] <http://www.hefce.ac.uk/news/newsarchive/2011/name/62310/en.html>

## Authors

### Ryu OHTANI

Ryu Ohtani received Ph.D. in science from the University of Tokyo in 1999. Afterwards he joined the Agency of Industrial Science and Technology (currently the National Institute of Advanced Industrial Science and Technology: AIST) at the Ministry of International Trade and Industry where he conducted research on earthquake science and GPS meteorology using precise geodetic measurement. He was a visiting researcher at Stanford University from 2003 to 2005 and received Program Officer Certificate in research and development evaluation from the Program Officer Academy at the Japan Science and Technology Agency (JST) in 2012. He is currently a senior researcher in the institute of Geology and Geoinformation at AIST. In this paper, he wrote part of section 1, sections 2 through 6, and part of section 8 based on the chapters “Executive summary” through “Assessing the impact of research” in “Research Excellence Framework – Second consultation on the assessment and funding of research,” on which this paper is based.



### Mariko KAMO

Completed the master's program at Tohoku University's Graduate School of Science in March, 2009. Joined AIST the following April. After working in the Evaluation Department, joined the Technology Promotion Division of the Industrial Science and Technology Policy and Environment Bureau at the Ministry of Economy, Trade and Industry in April, 2011. She specializes in research evaluation theory. She translated the chapters “The research environment” through “Assessing socio-economic impact” in “Research Excellence Framework – Second consultation on the assessment and funding of research,” on which this paper is based.



### Naoto KOBAYASHI

Completed the doctoral course at the Graduate School of Engineering, Kyoto University in 1978. (Ph.D.) Joined the Electrotechnical Laboratory (ETL), Agency of Industrial Science and Technology (AIST), Ministry of International Trade and Industry (MITI) in 1978, and worked as director of Planning Division and director of Quantum Radiation Division from 1997. Director of Photonics Research Institute, AIST in 2001; vice president and director of Evaluation Department, AIST in 2003; vice president, director of Safety and Environmental Management, and director of General Affairs Headquarters in 2007. Professor of Center for Research Strategy, Waseda University since April, 2009. Specialties are optical device engineering, semiconductor material engineering, quantum beam engineering, and theories of research strategy and evaluation. For this paper, built the basic concept, considered the overall strategy formation, and discussed the synthetic research evaluation. In this paper, he discussed the current state of program evaluation in Japan based on the translation



of the chapters “Overall assessment outcomes” through “Accountability burden” in “Research Excellence Framework – Second consultation on the assessment and funding of research.” He also wrote part of section 1, section 7, and part of section 8, and supervised the paper.

---

## Discussions with Reviewers

### 1 Overall

**Comment (Osamu Nakamura, AIST Chugoku; Chikao Kurimoto, Evaluation Department, AIST)**

With its main focus on providing recommendations for the evaluation system in Japan by introducing a new university evaluation framework in the UK, this is a timely article for considering the evaluation of synthetic approaches that helps to promote research on socio-economic innovations. Therefore, we recommend this article for publication in *Synthesiology* to widely spread the debate.

### 2 Comparison with the evaluation in Japan

**Comment (Osamu Nakamura)**

The overall structure is good. You have devoted a large segment towards introducing the REF in a straightforward manner. But, how about further exploring the discussion regarding the comparison with the Japanese evaluation system in practice (especially university evaluation)? More specifically, I would recommend that you describe the reality of university evaluation in Japan in more detail in section 7 to further highlight the differences from the REF case introduced here.

**Answer (Naoto Kobayashi)**

Based on your comment, we have added the following:

It is interesting to note that lately the NIAD-UE has considered revising the method for evaluating national universities in the second mid-term period (fiscal year 2010 to 2015).<sup>[12]</sup> In particular, since it is currently at the university’s undergraduate and graduate departments’ discretion to determine how much of the support documents are to be devoted for evaluating academic significance as opposed to social, economic, and cultural significance, discussions are underway to enable the university to submit support documents for the evaluation of both areas of significance, rather than just one.

We believe that the largest difference between the REF and Japan’s university evaluation lies in its utilization. Whereas the results of REF are clearly reflected in changes in the budget to individual universities, Japan’s university evaluation has a major defect in that the results of the evaluation are not clearly reflected.

In addition, you might ask why “impact” is vigorously introduced this way into the REF evaluation in the UK. It is probably because the university is motivated to proactively use the evaluation results for themselves by readily responding to requests from society and the government to contribute to the social value through these research results.

### 3 Eligible assessment period

**Question (Chikao Kurimoto)**

The paper does not mention the eligible assessment period (in terms of years) for the REF to be implemented in 2014. I suppose that research outputs are those generated within the eligible period (probably a period of multiple years), but research that has impacted society would probably be those completed prior to the eligible period. I do not think there will be a problem as long as the entity being evaluated (equivalent to a department in this

case, assuming that it is in the scale of a research unit at AIST), is maintaining the research elements linking past research to current research. However, there may be difficulties in checking the realities of the research and its impact, and examining the impact after years have passed. What are your thoughts on this?

**Answer (Naoto Kobayashi)**

As you pointed out, the assessment period for the REF is for multiple years. For output, I believe it is the period between the time of the last evaluation and the deadline for submitting assessment documents for the current evaluation. Meanwhile, in the case of impact, while the period in which the eligible impact is made is the same assessment period as for the output, it seems that the research results that have made this impact are expected to have been generated up to approximately 15 years prior to this period. In fact, the pilot study of impact conducted in 2010 specified the period in which the impact materialized to be between 2005 and 2009, and the original research results that generated the impact to have been created after 1993. This information has been added to the paper.

As a note, our direct and indirect interviews with some officials who were involved in this pilot study indicated that it is extremely difficult to go back to research results obtained quite a long time ago, and the organization and storage of documents will become an important consideration in the future.

Referenced URL:

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

REF impact pilot: revised case study template and guidance, July, 2010

### 4 Explanation of the figure

**Question (Osamu Nakamura)**

Regarding figure 1, it would be nice to provide a detailed explanation as to what the figure indicates. Since this diagram is the essence of the REF structure, I would like you to provide an easy-to-understand explanation.

**Answer (Ryu Ohtani)**

We have added a detailed explanation to the caption of the figure as shown below:

“Academic knowledge” is produced through research activities by researchers in institutions such as universities, and disseminated to various audiences, such as society and the economy, resulting in multiple ripple effects outside academic communities. In this process, university researchers are the “planners and players” of research, academic communities where the academic knowledge is produced and shared are “direct beneficiaries”, and those who benefit from its ripple effects are “indirect beneficiaries”. In the REF, the term “output” corresponds to this academic knowledge and the term “impact” is defined as the ripple effects that were built on excellent research conducted by researchers whose involvement in exploiting and developing the ripple effects are demonstrable (the area within the dashed line). The scope of the REF evaluation is the outputs of the highest quality and the impact defined above (the area enclosed by the thick line). (Created based on Ref. [10].)

### 5 Utilization of impact evaluation indicators

**Comment (Osamu Nakamura)**

The types of impact are cited (Table 1). I recommend you to describe that it should be utilized to a great extent, as I think it provides specific and rich information that would be useful in Japan as well.

**Answer (Ryu Ohtani)**

We included a statement as recommended in the paper.