

From the RAE-able to the REF-able?
A Note on Formative Reactivity in National Research Quality Assessment

“RAE-able” から “REF-able” へ？
——研究評価をめぐる高等教育機関の反応的変容に関する試案的考察

SATO Ikuya, ENDO Takahiro

佐藤 郁哉, 遠藤 貴宏

1. Introduction	85
2. The RAE and “RAE-ability”	87
2.1 Basic Features of the RAE	87
2.2 RAE-ability: Financial and Reputational Significance of the RAE	88
3. Assessment of the Assessment	89
3.1 Major Criticisms of the RAE	89
3.2 Reverse Effects and Sub-optimization	90
4. The Case of “RAE-able Publishability”	91
4.1 Centrality of Research Publication	91
4.2 Publishable but Perishable? Publication Explosion and Premature Publications	91
4.2.1 Excessive Publications	91
4.2.2 Premature Publications	92
4.3 Sub-optimizing Publishability	93
4.3.1 Increasing Predominance of Journal Publications	93
4.3.2 Homogenization of Research Publications	94
4.3.3 What is Good for the Scientist Is not always Good for Science	94
5. Heuristic Conceptual Framework and Issues of Future Research	95
5.1 Formative Reactivity	95
5.2 Unanticipated (but not Unintended) Positive Consequences	96
5.3 Evolutionary Trajectories of Policy Intentions	97
6. Concluding Remarks: From the RAE-ability to the REF-ability?	98
Acknowledgements	99
Bibliography	99
ABSTRACT	104

From the RAE-able to the REF-able?

A Note on Formative Reactivity in National Research Quality Assessment

SATO Ikuya*, ENDO Takahiro**

ABSTRACT

Reactivity to university evaluations is often considered in negative terms, as a bias-generating factor that hinder accurate measurement. Generic terms that are applied to reactive responses—for example, game-playing, gamesmanship, and creative story-telling—also convey negative images. Closer examination of the assessment literature, however, suggests that higher education institutions' adaptive and strategic responses to institutional assessment may occasionally bring about positive consequences. The UK's Research Assessment Exercise provides an interesting case study of reactive responses to assessment and their consequences. On the basis of a literature survey and interview-based fieldwork conducted in the UK, this research note presents a heuristic conceptual framework for analyzing various types of reactivity to assessment. The conceptual framework consists of three sensitizing concepts—unintended consequences, formative reactivity, and the evolutionary trajectories of policy intentions.

KEYWORDS

Research Assessment Exercise; Research Excellence Framework; Unintended Consequences; Measurement Reactivity; Formative Reactivity; Evolutionary Trajectories of Policy Intentions

1. Introduction

It may seem trite to say that the very act of assessing someone often transforms the behaviors of those being assessed, thereby affecting the assessment results. Generally known as “reactivity” in social measurement (Webb et al., 1981; Weiss, 1998, p.149; Singleton and Straits, 2010, pp.132-133, 214-220), this tendency is often considered in negative terms, as a bias-generating factor that is obstructive to “correct” or “accurate” measurement. In the literature on university evaluation as well, reactive or adaptive reactions by higher education institutions (HEIs) and their academics are in many cases regarded as nui-

sances that hamper the precise assessment of research and education at these institutions. Generic terms that are applied to such reactive responses—for example, game-playing, gamesmanship, and creative story-telling (or *sakubun* in Japanese)—also convey negative images. These terms usually refer to obviously deceptive or marginally dubious behaviors that are intended to “look good rather than do better” (Guena and Martin, 2003, p.297).

It has also been pointed out that reactive behaviors of HEIs and their academic staff not only tended to distort the assessment results but also resulted in a number of consequences that were unforeseen and/or unintended by policymakers (see for example,

* Professor, Graduate School of Commerce and Management, Hitotsubashi University

** Associate Professor, Research Institute of Economics and Business, Kobe University

¹ Bahram Bekhradnia, a former director of research at HEFCE (Higher Education Funding Council of England), was recently quoted as saying that “The ideal RAE would be one where people wake up and find they got a score but don't know how they got it” (Jump, 2013a). See also HEFCE (1997, para.36).

Power, 1997; Strathern, 2000; Elton, 2000; Guena, 2001; Lucas, 2006; McNay, 2009).¹ At the most extreme level, such unanticipated ramifications resulted in means-ends inversion or goal displacement,² which ultimately led to reverse effects, i.e., consequences that were diametrically opposite to the original policy intentions.

For example, Kitahara and Hirota (2012) recently made the following observation in relation to consequences of the growing emphasis on university evaluation in Japanese higher education:

[I]t frequently happens that achieving a high grade in the assessment becomes the primary objective of everyday activities at universities, and almost every aspect of university life is geared to the evaluative scores. It is mistaking the means for the end, in the sense that the system and technicalities of assessment end up in defining the very goal of education and research, and even the *raison d'être* of the university (Kitahara and Hirota, 2012, p.26).

The UK's Research Assessment Exercise (RAE) provides an extremely interesting case study of reactive responses to assessment and their consequences, including goal displacement and reverse effects. As we will show shortly, while the RAE has often been hailed as an exemplary model of nation-wide research assessment, it was also found during the audit process to have been fraught with various types of reactive behaviors by HEIs. In fact, the RAE literature, particularly those materials written by its critics, abounds with descriptions and analysis of HEIs' reactive responses to research quality assessments and

the consequent negative effects on research and education.

It is noteworthy in this regard that the term "RAE-able" has been used in the UK's academia to refer to the degree to which specific research outputs or academics could achieve the higher grades in national research assessments, and thereby to the financial and/or reputational returns of their respective HEIs.³ It must be obvious that many of the HEIs' game-playing activities were carried out mainly in order to improve the "RAE-ability" of research outputs, academics, and institutions.

It should be noted at this juncture, however, that HEIs' attempts to improve RAE-ability included not only game-playing activities leading to negative consequences. As will be shown later, HEIs' RAE-ability strategies sometimes also appear to have yielded a number of *positive* consequences either as their intended results or "by-products": these include a clearer focus on research and more structured research strategies at each HEI. It is quite instructive, then, to delve into positive as well as negative ramifications of HEIs' reactive behaviors aimed at improving RAE-ability.

The UK's experiences will provide important lessons for Japanese policymakers (and those who are in charge of implementing university evaluations) in view of the fact that the development of new "objective evaluative criteria" for the institutional assessment of national universities is one of the most pressing issues in the ongoing university reforms of Japanese national universities. "Selectivity and concentration (*sentaku to shuchu*)" constitutes one of the guiding principles of the reform plans, which were first publicized in 2012 (MEXT, 2012). The newly

² "Goal displacement" is used here in a more extended sense than Merton's original formulation. When this idea was first proposed by Merton (1940), "displacement of goals" referred mainly to the conversion of instrumental values into terminal values induced by bureaucrats' ritualistic adherence to formulaic procedural rules. In later usage, this term was used more extensively to refer to the process through which an organization's formal goals were transformed or even subverted. See Perrow (1961; 1986, p.263) and Scott (1967).

³ While there are only occasional references to either "RAE-able" or its kindred terms (e.g., RAE publications, RAE effective) in the RAE literature (see for example, Bence and Oppenheim, 2004, p.61; McNay, 2007, p.207; Lucas, 2006, p.151), we could confirm, through interviews with British academics, that both "RAE-able" and "REF-able" have been used frequently in the UK's HEIs. For seminal discussions on the more general or sweeping tendency of "auditability," see Power (1994, 1997). See also Power (2007).

created evaluative criteria are supposed to be used for the differential allocation of public funds to national universities. If reactivity of some sort is inevitable in any social measurement, those who construct new evaluative criteria should at least anticipate reactive responses to the evaluation and their various ramifications.

The authors of this research note have been conducting literature survey and interview-based field research on the RAE and its successor, i.e., the REF (Research Excellence Framework) in the UK.⁴ This paper provides an interim report of the research primarily on the basis of the literature survey, and presents a heuristic conceptual framework for making sense of different types of assessment reactivity and their consequences. (Results of the interview research will be presented elsewhere.) Drawing on the sociological literature, the proposed analytical framework consists of the following three sensitizing concepts⁵—unintended consequences of social action, formative reactivity, and the evolutionary trajectories of policy intentions.

2. The RAE and “RAE-ability”⁶

2.1 Basic Features of the RAE

The UK’s Research Assessment Exercise (RAE) is arguably the best known national research assessment system in the world, and many of its ideas have been borrowed and emulated worldwide.⁷ The UK’s first nation-wide research assessment, initiated in

1986, was called “Research Selectively Exercise.” It was renamed the Research Assessment Exercise during the last phase of the second exercise conducted in 1989.⁸ Since then, the RAE was carried out four times (1992, 1996, 2001, 2008) and has been replaced by the Research Excellence Framework, which is due in 2014.

While detailed technicalities of the research performance assessment have changed from RAE to RAE, the stated aim of the RAE has remained almost the same: to assess the research performance and capabilities of departments (or “unit of assessment [UoA]”) of the UK’s HEIs. Since the third assessment exercise in 1992, each department was required to submit a dossier consisting of its “research active” staff members’ research outputs and other documents substantiating the department’s research performance and infrastructure (e.g., research environment, research funding, and number of students). Each research active member was asked to specify up to four outputs that s/he produced during the assessment period. The research outputs and other information were assessed by panel members of the relevant subject areas.

The assessment outcomes were used by four government funding bodies—HEFCE (Higher Education Funding Council for England), SFC (Scottish Funding Council), HEFCW (Higher Education Funding Council for Wales), and DELNI (Department for Employment and Learning, Northern Ireland)—to

⁴ So far, we have conducted formal interviews with more than ten informants. They include one current and one former HEFCE director as well as nine academic staff members of UK’s HEIs. Countless informal conversations with UK’s academics also have provided us with important contextual information that is helpful in understanding the impacts of the RAE and REF on UK universities and their academics.

⁵ What sociologists call sensitizing concepts refer to the conceptual ideas that give “a general sense of reference and guidance in approaching empirical instances” (Blumer, 1954, p.7). They are usually contradistinguished from “definitive concepts,” which, as in the case of operationalized concepts, tend to restrict analytical perspectives by delimiting the connotative and denotative range of concepts (Bulmer, 1979; Becker, 1998, Ch.4; Sato, 2006, pp.94-99).

⁶ The description in this section is largely based on the following sources—HEFCE (1997), RAE (2008, n.d.), McNay (2003), Bence and Oppenheim (2005), Thompson (2005, p.281), Goldfinch and Yamamoto (2012, Ch.4), Brown and Carasso (2013, Ch.4). For a concise summary of the UK’s national research assessment including the REF, see Jump (2013).

⁷ Japanese sources that include references to the RAE and REF include Iwata and Tokuda (2002), Iwata (2005), Yokoyama (2006), Hada (2005, 2010), and Hayashi (2009).

⁸ While the title of the Funding Council’s report on the second exercise was *Report on the 1989 Research Assessment Exercise* (UFC, 1989), all the documents included in the report as annexes (e.g., circular letters, questionnaires, submission forms) included the term “Research Selectivity Exercise” in their titles.

determine the amount of funds they would allocate to their respective HEIs as the “quality-related” component of their block grants. This quality-related or “QR” research fund forms one leg of the UK’s public research funding system called the “dual support system.” The other leg is constituted of the research funds supplied by Research Councils and other government agencies and departments (HEFCE, n.d.).⁹

Panel assessments of each department were expressed in terms of an aggregate score in a five- or seven-point scale from the RAE 1989 to RAE 2001. In the RAE 2008, the aggregated discrete scores were replaced by a “quality profile,” showing the ratio of each department’s research activities falling into five quality levels, ranging from 4* to “unclassified.” The research activities graded as 4* were supposed to have attained the level of “world-leading in terms of originality, significance and rigour.” On the other hand, an unclassified activity was one that “falls below the standard of nationally recognized work. Or work which does not meet the published definition of research” (RAE 2008, n.d.). The same five quality levels were supposed to be employed in the REF 2014.

2.2 RAE-ability: Financial and Reputational Significance of the RAE

For most university departments and HEIs in the UK, significant financial and/or reputational resources were at stake in the assessment results, and to attain higher grades (or quality levels) in the RAE was one of their most serious concerns.

Table 1A through 1C show changes in the funding weights applied to respective grades (or quality levels) of the RAEs.

As can be seen from these tables, while weightings for the QR funding allocations were kept constant for

Table 1A Changes in Funding Weights following the RAE 1996

RAE rating	Funding weights				
	1997/98	1998/99	1999/00	2000/00	2001/01
5*	4.05	4.05	4.05	4.05	4.05
5	3.375	3.375	3.375	3.375	3.375
4	2.25	2.25	2.25	2.25	2.25
3a	1.5	1.5	1.5	1.5	1.5
3b	1	1	1	1	1
2	0	0	0	0	0
1	0	0	0	0	0

Source: Brown and Carroso (2013), pp.54-57

Table 1B Changes in Funding Weights following the RAE 2001

RAE rating	Funding weights						
	2002/3	2003/4	2004/5	2005/6	2006/7	2007/8	2008/9
5*	2.71	3.375	3.375	3.755	3.948	4.036	4.036
5	1.89	2.793	2.793	3.006	3.12	3.175	3.18
4	1	1	1	1	1	1	1
3a	0.31	0	0	0	0	0	0
3b	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0

Source: Brown and Carroso (2013), pp.54-57

Table 1C Changes in Funding Weights following the RAE 2008

RAE rating	Funding weights				
	2009/10	2010/11	2011/12	2012/13	2013/14
4*	7	9	9	3	3
3*	3	3	3	1	1
2*	1	1	0.294	0	0
1*	0	0	0	0	0
Unclassified	0	0	0	0	0

Sources: Brown and Carroso (2013, pp.54-57), HEFCE (2013)

five years following the RAE 1996, a steeper gradient was applied after the RAE 2001. Along with this, funding was significantly reduced or entirely removed from departments or activities assessed with

⁹ Whereas Research Council grants are provided on the basis of the assessment of proposals for specific research projects and programs, allocative decisions for the QR research grants are informed by an assessment of the research performance of HEIs’ departments using their submissions to RAEs. In other words, the dual support system combines a prospective evaluation of research proposals (for Research Councils and other government bodies) and a retrospective evaluation of research performance measured through the RAEs.

¹⁰ The relative ratio between block grants for teaching and those for research varies considerably among different types of universities (Watson, 2005, pp.108-113; Watson, 2013, Ch.11; Goodman, 2013, pp.45-47).

lower grades. After 2012/13, only those activities graded at the top two quality levels were entitled to receive QR funds.¹⁰

The significance of the RAE results is not limited to the financial returns accruing from higher grades. For a substantial number of departments and HEIs, symbolic rewards gained from higher RAE results often far outweighed the financial benefits of QR funding. In fact, since numerical grading system was adopted in the RAE 1989, various news media including the quality press published university research league tables every time the RAE results were made public (Shattock, 1994, p.69). We can also find references to the rankings produced in UK's HEIs league tables in brochures and websites.

It is widely believed that that the prestige or reputational capital accruing from being assigned higher grades in the RAE, and consequent league table ranking, has been transformed into economic capital by means of being awarded increased external funds (Henkel, 1999, pp.110-111; Curran, 2000, p.391). Higher grades and rankings have also been instrumental in attracting international as well as distinguished UK/EU students (cf. Lucas, 2006). On the other hand, it has sometimes happened that those departments that fared poorly in the RAEs were closed or merged with other departments (McNay, 1997, para.2.9; HM Treasury, 2003, pp.88, 92).

In this way, the national research quality assessment that had been coupled with consequent selective research funding worked as an important restructuring device for the UK's higher education system (Shattock, 2009, p.19). It is no wonder that the UK's HEIs sometimes resorted to behaviors that were explicitly aimed at improving their RAE-ability, including game-playing. Such self-interested behaviors at each HEI, collectively, have sometimes led to unanticipated and certainly unintended consequences for the scientific community and the country as a whole.

3. Assessment of the Assessment

3.1 Major Criticisms of the RAE

Almost since its inception, the RAE (and its prede-

cessor the Research Selectivity Exercise) has been subject to intense criticisms—not only from the “losers” but also from the “winners” who performed relatively well in the research assessment—and these criticisms were publicized by the news media.

One of the major criticisms raised against the RAE was the possibility that it tended to stabilize the existing hierarchy among the UK's HEIs. This criticism was closely related to the retrospective nature of the RAE's assessment process. Since the RAE focused on research performance during the assessment period, the so-called “Matthew Effect” (Merton, 1968; Guena, 2001, pp.624-625) became almost inevitable since the resources to be allocated were limited. In fact, those HEIs that have proven their capacity to deliver high quality performances in the past might continue to enjoy a relatively greater share of research funding under the RAE.

The RAE assessment process was also criticized for its undue emphasis on basic sciences, at the expense of applied research (McNay, 1997, pp.97-98, 2003, p.52; House of Commons, 2002, EV119; 2004, EV53, para.27, 28; Roberts, 2003 Annex D para.21; HM Treasury, 2003, p.84). This possible bias against applied research was explained partly in terms of its panel members, the majority of whom were academics. The relative ease of measuring research performance by means of purely academic benchmarks (e.g., publications in leading journals) was mentioned as another important source of bias in favor of the basic sciences. For similar reasons, the RAE was said to have discouraged interdisciplinary research, which may fare relatively poorly in established publication venues.

Considerable administrative burdens on HEIs and the heavy workload of senior academics who assumed the roles of panel members or assessors were mentioned as another major drawback of the RAE process. While a panel-based peer review, if properly done, could produce reliable and comparable assessment results, it is also a highly labor-intensive and costly exercise. In the case of the RAE 2008, more than 1,000 panel members assessed the over 210,000 research outputs of some 52,400 researchers re-

turned by 2,344 departments (or UoAs) (RAE 2008, 2009a). Many universities also expended substantial amounts of resources for various measures specifically directed at the RAE—for example, mock RAEs, the appointments or employment of RAE officers, and outsourcing certain jobs to consulting firms. The estimated total cost of RAE 2008 was approximately 47 million pounds (PA Consulting Group, 2008, p.4).¹¹

Undue emphasis on research at the expense of teaching at HEIs has been also attributed to the RAE. If material and symbolic rewards for academics are heavily skewed toward research, many of them may put more emphasis on research than on teaching and other services to students, including pastoral care for undergraduate students. The differentiation in teaching load, including the establishment of teaching- or research-only positions, may also be detrimental to the educational function of the university as a whole.¹²

3.2 Reverse Effects and Sub-optimization

It should be noted at this point that many of the purportedly negative consequences of the RAE and selective research funding policy must have been originally unforeseen by policymakers.¹³ In fact, various ramifications of the RAE, especially negative ones, are often characterized as “unintended consequences” (see for example, Elton, 2000; Guena, 2001; UUK 2003, pp.5, 8; Thompson, 2005, pp.281-282; Lucas, 2006, pp.3, 80; McNay, 2009, p.35). We should also note the well-known fact that virtually any public policy brings about not only expected favorable results but also some sort of unexpected negative consequences. If we use “unintended consequences” merely as a general descriptive term, then, there will be little analytical gain. On the other hand, if we could delve into specific unforeseen conse-

quences and clarify their backgrounds, we might be able to acquire substantial analytical insights (cf. Sieber, 1981, Ch.2).

Among various unforeseen consequences of the RAE, two reputedly negative outcomes merit special attention, since they assume the character of reverse effects or perverse effects, that is, consequences that are in diametrical opposition to the original policy intentions.¹⁴

One possible reverse effect of the RAE and selective research funding is its potentially negative impact on applied research. As has been previously discussed, the RAE has been criticized for its overemphasis on basic sciences at the expense of applied or practitioner-oriented research. It has often been pointed out that the Research Selectivity Exercise was initiated in the late 1980s under the Conservative government led by Margaret Thatcher, in order to make university research more socially useful, and thereby make government expenditures for research publicly accountable. If the exercise actually resulted in the neglect of applied research, we might say that the value for money (VFM) principle has been reversed (McNay, 2009, p.48).

Critics have also argued that the national research assessment have sometimes (though not always) had had detrimental effects on academic research in the UK. According to them, while one of the officially proclaimed policy goals of the RAE was the assurance of research quality, it has sometimes ironically led to a decline in the quality of research in some academic disciplines. In many cases, critics have mentioned the undue emphasis on publishability as the chief immediate cause of the reverse effect. They argued that the pressure to produce a certain number of publications within a fixed period of time have yielded a tendency to aim at instantly publishable research

¹¹ Estimates of the costs for RAE 1996 vary from 27 million to 37 million pounds (Roberts, 2003, para.62). See House of Commons (2002, para.33) and Kinmonth (2005) for criticisms of these estimates.

¹² Other criticisms of the RAE include the following: increasing managerialism in the university, the loss of academic freedom, and the loss of a collegial and co-operative culture among academics.

¹³ In some sense, the stabilization of the established hierarchy among HEIs may have been an exception in this regard. See Swinnerton-Dyer (1991).

¹⁴ See the fifth section of this paper about the precaution we need to take in talking about policy “intentions.”

topics and avoid risky topics and/or approaches. They have also pointed out that such tendencies may eventually lead to the homogenization of research outputs (The next section of this research note will take up this issue of publishability in detail.)

It should be noted here that these two crucial reverse effects, namely the negative impact on applied research and the homogenization of research, were closely related to the HEIs' efforts to improve the RAE-ability of their academic staff members' research outputs. In other words, reactive responses to the assessment appear to be at the root of the reputedly perverse effects of the national research assessment and selective research funding.

It should be self-evident that the HEIs' quest for better RAE-ability was essentially based on self-interest. If the self-seeking behaviors eventually led to detrimental consequences at the national level, this could be regarded as a typical case of sub-optimization (cf. Bekhradnia, 2004, para.22-28).¹⁵ In other words, in such a case, there is an apparent contradiction or mismatch between what is good for each HEI and what is good for society as a whole. Any analysis of the negative consequences arising from reactive responses to university evaluations will have to take account of this possibility.

4. The Case of "RAE-able Publishability"

4.1 Centrality of Research Publication

While both the neglect of applied research and undue emphasis on publishability merit serious attention, this research note focuses on the latter issue because the problems inherent in RAE-able publishability appear to clearly illustrate ways in which each HEI's quest for maximum RAE-ability has led to sub-optimal results for society as a whole. In fact, tactics concerning the publishability of research stand out between various RAE-ability strategies adopted by UK's HEIs.

This preoccupation with publishability can be ex-

plained mainly in terms of the presumed centrality of research publications in the assessment results. While relative weights given to various evaluative elements remained obscure for a long time (Goldfinch and Yamamoto, 2012, p.137), it was generally believed that research outputs (in most case, research publications) was the predominant factor influencing the panels' grading decisions. This belief was partly confirmed when the relative weightings were made explicit in the RAE 2008. While panels had been instructed to allocate at least 50 percent of the weighting assessment to research outputs, most panels and sub-panels assigned more than 70 percent to research outputs.¹⁶

In view of the centrality of the research outputs to the assessment process, it is no wonder that HEIs and their departments made every effort to enhance the publishability of their academics' research activities. In fact, much of the so-called "game-playing" or "gamesmanship" related to the RAE centered around the publishability of research: just as an individual researcher without a strong-enough list of publications tended to disappear from the academic scene and might eventually "perish," a HEI without a strong portfolio of publications might perish as a "research-active" institution. Similarly, a department without a strong list of publications might lose its independent status by being merged with another department.

In this way, the RAE is sometimes said to be responsible for introducing a "publish or perish" culture and mentality into the UK's academia (Bekhradnia, 2009, para.8).

4.2 Publishable but Perishable? Publication Explosion and Premature Publications

4.2.1 Excessive Publications

The publishability strategies adopted by HEIs included not only legitimate strategies but also semi-legitimate and obviously deceptive ones.

¹⁵ While Bekhradnia (2004) refer mainly to the sub-optimizing effects of HEIs' adaptive behaviors on teaching, similar things can be said of the effects on research.

¹⁶ See Research Fortnight (2005) for details of the specific weights given by each panel.

One of the most deceptive publishability strategies was the deliberate misreporting found in the RAE 1989. According to a report by the Universities Funding Council [UFC], the types of misreporting included the following: (1) altering the date of publication to bring a publication within the review period, (2) claiming authorship of edited books, (3) claiming sole authorship of co-authored publications, and (4) the inclusion of publications representing research undertaken in another HEI (UFC, 1989, para.24; see also Bence and Oppenheim, 2005, p.145).

The author of the report pointed out that the UFC (and its predecessor the UGC [University Grants Committee])¹⁷ was probably naïve enough to assume that such dubious practices would not occur. In other words, for the Funding Council, such a deceptive response was one of the reactive responses that had not been anticipated. A similar thing can be said of the HEIs' (mis) interpretation of the publication list or publication count submitted by each unit as a requirement of the RAEs 1989 and 1992. Many HEIs interpreted this as a message confirming the high importance that had been assigned to the quantity of research conducted. This misconception resulted in the "publication explosion," or phenomenal increase in the quantity of publications (HEFCE, 1997, para.30; McNay, 2011, p.52), or "premature and excessive publication" (Richards, 2001; Jump, 2013). Publication count was dropped from the required information to be submitted to the assessment panels effective with the fourth exercise in 1996. Nonetheless, it has been pointed out that both HEIs and academics tended to regard the specification of "up to" four research outputs not as the upper limit, but as a production quota of "at least" four publications (Talib, 2000, p.42; Richards, 2000; Piercy, 2000; UNIVERSITAS, 2001, p.42).¹⁸

A number of critics have also suggested the possibility that the RAE led to the proliferation of academic journals, many of which were in-house journals (Talib, 2000, pp.44-45; Elton 2000, p.276; Piercy, 2000, p.29; Bessant et al., 2003, p.65; Macdonald and Kam, 2007, pp.648-649).¹⁹ It is quite plausible that even if publication counts were no longer required in the RAE submission, newly established journals endure and continue to be a major factor leading to the publication explosion.

4.2.2 Premature Publications

It should be apparent that an increased quantity of publications does not necessarily mean an improvement in the quality of these publications. Critics have pointed out that the excessive number of publications has led to an overall decline in the quality of research publications—most evident in the case of premature publications (McNay, 1997, para.4.47, 5.21; Henkel, 1999, p.119; Mynott, 1999, pp.129-130; Walford, 2000; Bence and Oppenheim, 2004). It has been reported that authors of academic papers have tended to decline editors' requests to revise their manuscripts, in order to meet their publication quota within an assessment period. This tendency was especially prominent in the run-up to the submission deadline for the RAE. Some critics also reported that a flood of book manuscripts and publication proposals clogged the pipelines of British and American publishers around the RAE's deadline (Myott, 1999; Richards, 2000; Kinmonth, 2005, p.164).²⁰

The accelerated tendencies of salami publications were also attributed to the HEIs' and their academics' preoccupation with RAE-able publishability (Smith, 1998; Elton, 2000, p.276; Talib, 2000, p.45; see also Taylor, 2001). This practice consisted of publishing essentially the same content in different

¹⁷ See Shattock (1994) and Kogan and Hanney (2000) for accounts of history of the UK bodies responsible for the allocation of public funds to HEIs.

¹⁸ A number of reports have noted the cyclical publication pattern of UK researchers. See for example, UUK (2003, p.11) and Shattock (2009, p.19).

¹⁹ Even if a considerable proportion of the new journals are essentially in-house journals published by specific universities, they provide researchers with outlets for "refereed articles."

²⁰ The director of a major American university press also commented on the problems caused by the flood of manuscripts from the UK in the run-up to the RAE deadline (personal interview, 2010).

papers. Some also attempted to increase the number of papers they published by slicing the same data set into small chunks, each of which was published as a separate paper.

Articles and books written in haste and without sufficient revisions, so as to meet a submission deadline, may have relatively short lifespans. They may also have little readership (Corner, 2009). The establishment of new specialist journals as well as the increased quantity of published outputs (Talib, 2000, p. 44-45; McNay, 2011, p.52; See also McNay, 2008) will further lead to the decrease in readership (Piercy, 2000; Bence and Oppenheim, 2005, p. 151; Goodman, 2013, p.48). In other words, such research outputs may certainly be publishable, but at the same time, they may be perishable as a medium of scholarly communication.

4.3 Sub-optimizing Publishability

4.3.1 Increasing Predominance of Journal Publications

In the case of certain academic disciplines, publishability strategies adopted by HEIs' and their academics have led not only to an increase in the total quantity of research publications, but also to the transformation of other aspects of publications including pre-

ferred publication media and content. This tendency can be glimpsed from the increasing ratio of journal publications to all research outputs submitted to the RAE, as shown in Table 2.

As can be seen from this table, the proportionate share of journal articles relative to the total number of outputs jumped by approximately 8 percent from the RAE 1996 to the RAE 2001. It further increased and reached more than 75 percent in the RAE 2008. In other words, more than three out of every four research outputs submitted to the last RAE were journal articles. The increasing predominance of journal articles in the research outputs submitted to the RAE suggests that "RAE-able publishability" was increasingly defined in terms of journal publications.²¹ In order for an academic to earn the higher quality evaluation, it was not enough that s/he published it as a journal article. S/he had to publish it in a leading, peer-reviewed journal with a high impact factor, preferably in a US- or UK-based top-notch journal. Moreover, s/he had to publish such articles on a relatively regular basis, in order to meet the production quota of "four publications" within each assessment period. (Since not all manuscripts sent to journals are accepted, the academic will have to accumulate a backlog of some number of manuscripts, each of which is at a certain stage of the submission process leading to eventual publication or rejection.)

In this way, the requirements of "RAE-able publishability" have dictated almost every aspect of academics' research behavior in certain disciplines. After all, their activities were supposed to evolve around the publication of research findings in the form of journal articles.²² This meant that other publi-

Table 2 Number of Journal Publications, RAE 1996, 2001, 2008

	RAE 1996	RAE 2001	RAE 2008
No. of Items	132,077	143,362	162,320
% of Total Submissions	62%	69.7%	75.3%

Sources: Bence and Oppenheim (2005, p.151), RAE 2008 (2009, Annex C)

²¹ It is ironical, in this regard, that the Guidance on Submissions for the RAE contain the following clarification: "All types of research and all forms of research output shall be assessed on a fair and equal basis" (RAE 2008, 2005, para.12). Journal articles as preferred publication media for assessment can be partly explained in terms of their quantifiability and assumed comparability, both of which also lighten the workload of panel members and assessors.

²² There are apparently considerable variations among different academic disciplines. For example, in the case of the Business and Management subpanel, the proportionate share of journal articles in the total outputs documented in submissions to RAEs 1996, 2001 and 2008 was 69%, 80%, and 92% respectively (Bence and Oppenheim, 2004; Willmott, 2011; Mingers and Willmott, 2013). On the other hand, in the case of the History subpanel, journal articles accounted for just a little more than 35% of the outputs included in submissions to the RAE 2008, while books occupied as much as 24%. Book chapters also accounted for approximately 30% (RAE 2008, 2009, pp.5-6). Detailed analysis of these disciplinary variations will provide important information in understanding the process in which HEIs' formative reactivity affects the quantity and quality of research.

cations such as monographs, chapters in edited books, textbooks and conference papers were not given much weight (Thompson, 2005, pp.280-285). Moreover, activities that were not closely linked with journal publication (e.g., collaboration with business firms, public lecture, editorship and reviewership) might have been slighted by academics themselves, or discouraged by their departments.

4.3.2 Homogenization of Research Publications

It has been pointed out that the imperative of RAE-able publishability has sometimes led to the homogenization of research regarding such matters as theoretical perspectives, methodologies and research topics (Whitley, 2007; Shattock, 2009; UUK, 2009, p. 5; Lambert Review, 2003, p.84; McNay, 2011, p.53; Lee, Pham, and Gu, 2013).²³ In other words, isomorphism in the publication form has resulted in isomorphism in the content of publications. In many cases, homogenization of the content of journal articles has been explained in terms of two factors: (1) the need for immediate publication, and (2) the concentration of manuscript submissions to a limited number of leading journals.

As already mentioned, the UK's HEI-based (research-active) academics in general are expected to produce at least four publications within each assessment period. For a manuscript to be accepted, it is much safer to report predictable results that conform closely to existing beliefs. Conservatism was frequently cited as one of the serious problems of editorial peer review: editors and reviewers not only tend to have vested interests in established paradigms but are also likely to be constrained by the worldview of existing theoretical paradigms (ABRC, 1990, para. 4.47-4.51; Weller, 2001, pp.96-100, Ch. 7; Biagioli, 2002; Schatz, 2004, Ch.3; Waters 2004, Part II; Nightingale and Scott, 2007, p.547; Harley and

Acord, 2011, pp.16-26). When this conservative tendency was combined with the pressure to produce a certain number of research outputs within a limited period, the likely result would have been the homogenization of research.

Critics also mentioned contributors' tendencies to submit their manuscripts to a limited number of leading journals as another important factor leading to the homogenization of research publications. It was generally believed that to be RAE-able the researcher had to publish in leading journals with high-impact factors. While some mainstream quality journals cover broader research areas than specialist journals (Macdonald and Kam, 2007, p.648), academic journals generally have their own distinctive identities and editorial policies. It is no wonder that manuscripts conforming to the distinctive characteristics of mainstream journals would have a better chance of being accepted for publication.

On the other hand, contributors with manuscripts addressing a new field of research and including unorthodox theoretical and/or methodological ideas might find getting published in leading journals quite difficult. While they could eventually find their homes in newly created or specialist journals, those journals generally have a low impact factor, and are regarded as having less RAE-ability (Henkel, 1999, p.118). Departments, therefore, will not encourage submissions to journals with low impact factor scores.

In this way, the imperatives of RAE-ability were not only defined by the inferred importance of publishing manuscripts, but they also tended to determine what would be published, and eventually what subjects would be researched.

4.3.3 What is Good for the Scientist Is not always Good for Science

In some respects, HEIs' and their academics' pursuits of RAE-able publishability may have resulted in

²³ Michael Rees, chairman of the British Medical Association's medical academic staff committee, once stated that a decrease in the number of clinical investigations in the UK could be explained in terms of the impact factors of journals: papers reporting laboratory-based, rather than clinically based findings, and research tending to be published in journals with high impact factors (Brown, 2007, p.564). For similar changes in academic disciplines caused by universities' management strategies, see Lucas (2006).

the improved “quality” of published research outputs in the eyes of journal reviewers and RAE panel members. Yet, it is another question whether the ideas and findings included in the outputs could contribute to the advancement of academic knowledge. Neither is it clear whether all or most of the RAE-able publications have wider social relevance.

In fact, it has been pointed out that the proliferation of in-house and specialist journals tends to delimit the readership of journal publications (Piercy, 2000; Corner, 2009; Bence and Oppenheim, 2005, p.151; Goodman, 2013, p.48). Just as in the case of vanity publishing in Japan, in some cases would-be authors of journal publications far outnumbered potential readers of the articles. Many of the articles addressed very narrowly focused issues by means of esoteric theoretical frameworks and research methods. Such articles were often full of jargon that was almost unintelligible not only to lay readers but also to those who specialized in the same academic disciplines.

Armstrong (1982) once pointed out that undue emphasis on publishability sometimes lead to a disjunction between “what is good for the scientist and what is good for science” (Armstrong 1982, p.88). The arguments in this section suggest that in the case of the reactive responses to the RAE, they have sometimes led to a conflict between what is good for HEIs and their academics and what is good for science and society as a whole (Nightingale and Scott, 2007, p.547). In fact, HEIs’ and their academics’ behaviors, as driven by the imperatives of RAE-able publishability, appear to be a typical case of sub-optimization in which the pursuits of local interests result in the failure of these behaviors to benefit the system as a whole.

It should be noted, however, that undue emphasis on RAE-able publishability is only one of the HEIs’ responses to the RAE audit process. Reactive responses to the RAE may occasionally have brought about unanticipated *positive* consequences.

5. Heuristic Conceptual Framework and Issues of Future Research

5.1 Formative Reactivity

In an article titled “Becoming PBRF-able,” Middleton (2009) pointed out that “the impact of research assessment runs deeper than mere measurement of ‘what is already there’; such processes are productive, or formative.”²⁴ PBRF is the acronym for the Performance Based Research Fund, the national research funding scheme in New Zealand. By “PBRF-able,” Middleton refers to the process in which activities and research outputs of academics in New Zealand tend to be molded into the patterns that are expected to attain favorable assessment results in that country’s funding scheme. In this sense, the PBRF not only measures research performance, but also forms and shapes to a considerable extent what is to be measured.

Similar observations can be made of the UK’s Research Assessment Exercise. The descriptions provided in preceding sections, albeit admittedly sketchy, should be sufficient to show that the RAE is not a neutral measurement device that faithfully gauges “what is already there.” The RAE audit process has often functioned as an important causal agent in forming and transforming research outputs in the UK, and so eventually the volume and quality of research activities.

The foregoing arguments suggest that the formative impacts of the RAE arose largely from HEIs’ reactive responses that have been aimed mainly at improving RAE-ability. This type of reactive response to research assessment may be called “formative reactivity.”

Formative reactivity is apparently distinct from what is generally known as measurement reactivity, a typical example of which is “game-playing” or “gamesmanship” in the RAE. Those who are involved in adopting game-playing behavior in the assessment process intend to conceal the true values (research performances) of what is being measured,

²⁴ For similar arguments, see Besley (2009, *passim*) and Goldfinch and Yamamoto (2012, p.315).

and try to “look good rather than do better” (Guena and Martin, 2003, p.297). On the other hand, HEIs and their departments showing formative reactivity may attempt to actually do better, namely, they may encourage academics to improve their research performances—at least as far as the RAE-ability or publishability of research is concerned.

It appears to us that the two types of assessment reactivity (i.e., measurement and formative reactivity) are not always clearly distinguished from each other in the RAE literature. Lumping various kinds of strategic responses to the RAE together, as “game-playing” or “gamesmanship,” will lead policymakers as well as critics to overlook the formative and transformative potential of research assessment. On the other hand, making an analytical distinction between the two types of reactivity will make it possible to appreciate the potentiality of the assessment process as an important change agent.

An important research issue, then, is to disentangle the two types of reactivity in the HEIs’ adaptive and strategic responses to the RAE. It is also important to delve into the process during which each type of reactive response leads to various types of positive and negative consequences for research and education.²⁵ Undertaking this area of research will ultimately contribute to better understanding the significance of evaluation and “benchmarking” in higher education and research.

5.2 Unanticipated (but not Unintended) Positive Consequences

It should be obvious that formative reactivity to the RAE, in many cases, spans a longer period of time than measurement reactivity, and often has far greater impacts on HEIs’ organizational structures and processes. In fact, measurement reactivity, typified by the misreporting of publication information, takes place for a relatively short period of time, closely related to each assessment, and is not likely to have an

impact on HEI’s core activities. On the other hand, formative reactivity may affect crucial aspects of HEIs for decades.

Increased emphasis on publishability is a typical example of formative reactivity to the RAE that could have long-lasting influences on the HEIs and their academics. It should be noted that while the HEIs’ adaptive responses aimed at sub-optimizing publishability tend to result in negative consequences, other strategic responses may bring about beneficial results both to HEIs and to society as a whole.

For example, it has been reported that the RAE has encouraged UK’s HEIs and their departments to adopt more structured approaches to research activities (McNay, 1997; Lucas, 2006; Thomas, 2007, p.43; PA Consulting Group, 2008, p.9). These approaches include such measures as meritocratic personnel policies, increased emphasis on graduate education, and the designation of internal research funds. Such measures may occasionally have led to excessive and premature publications. Yet, they must also have facilitated proper completion of research and lead to wider dissemination of research outcomes in the form of articles and books (Curran, 2000, pp.391-393; Guena and Martin, 2003, pp.295-300). These strategic research management policies must also have encouraged UK’s academics to have a clearer focus on research, and adopt an output-oriented stance toward their own and others’ research activities.

In this way, HEIs’ strategic responses aimed at improved RAE-ability may have been, to a certain extent, instrumental in improving the quantity and quality of university research in the UK. It appears that the accounts of such possible positive consequences of the RAE are relatively few, as compared to the stories about its negative consequences, especially as far as media reports on the RAE are concerned. In fact, stories about such matters as the poaching of star researchers, the predicaments of academics who fared poorly in the research assessment (or mock as-

²⁵ A strand of research on “audit society” (Power, 1994, 1997) and “audit culture” (Strathern, 2000) includes a number of important ideas and concepts for analyzing the differences between the two types of reactivity. For example, what Power (1997, pp.95-97) calls decoupling is mainly concerned with measurement reactivity, while what he calls colonization (Power, 1997, pp.97-98) is concerned with formative reactivity.

assessment), and scandalously deceptive game-playing activities by HEIs must have been far more newsworthy than stories about any positive consequences stemming from the national research quality assessment.²⁶ In learning from the UK's experiences then, we need to identify the various consequences of the RAE carefully. We should also examine the causal linkages between those consequences and the strategic responses by HEIs in relation to the research assessment implementation processes.

It seems certain that many of the HEIs' strategic responses had not been clearly envisaged by policy-makers when they first contemplated the research assessment process. In fact, as previously pointed out, while the stated aim of the RAE remained almost the same, specific technicalities of the assessment process changed from RAE to RAE, and many of these changes were made in response to HEIs' reactive responses, sometimes almost in the manner of a game of cat and mouse.

The improved quantity and quality of university research arising from HEIs' reactive responses, then, could be characterized as unanticipated positive consequences of the RAE.²⁷ On the other hand, one could say that the RAE succeeded in attaining its intended effects, if the improvement of research quality or quality assurance had actually been included in the original policy goals of the national research assessment. In employing the idea of unintended consequences of social action to the analysis of the assessment reactivity, then, we should make a distinction between unanticipated and unintended consequences.

5.3 Evolutionary Trajectories of Policy Intentions

We should note, at this juncture, that it is not al-

ways easy to identify widely agreed-upon explicit goals or "intentions" in public policies. In fact, as students of unintended consequences have repeatedly pointed out, social action does not always involve clear-cut, explicit purposes (Merton, 1936, p.896; Sieber, 1981, p.12; Corwin, 1981, p.xv). Even in the case of public policy, for which an explicit and formal statement of purpose is assumed to be required, it is sometimes quite difficult to identify its original intention (Weiss, 1980; March, 2009). Policy intentions or policy goals are often left vague, "so as to afford leeway in action or to gain consensus among key participants and supporters" (Sieber, 1981, p.12). Moreover, stated, as well as unstated policy goals may change over the course of time through the process of trial-and-error, and due to environmental changes such as a shifting political climate and budget conditions.

The same can be said of the assumed policy goals of the RAE. The literature on research selectivity and national research assessment shows that these programs have complicated roots in policy discourses that are not easily disentangled (Swinnerton-Dyer, 1991; Shattock, 1994, pp.67-70; Kogan and Hanney, 2000, pp.96-108; McNay, 2009, pp.38-49; Adams and Gurney, 2010; Brown and Carasso, 2013, pp.43-46). Various organizations and people (e.g., politicians, government officials, executive members and officers of the funding bodies, members of learned societies) are involved in contemplating, making, implementing and evaluating the research assessment and funding policies. The literature also suggests that the goals of performance-based funding policies include at least some of the following at different points in time: a reasonable allocation of budget cut, effective use of limited research funds through increased differentiation among HEIs, quality assurance, and

²⁶ It is somehow ironic that the very news media that constructed a university league table on the basis of RAE results also frequently carried articles criticizing the purportedly negative consequences of the national research assessment, some of which were closely related to the league table's impacts.

²⁷ It is crucial to distinguish between the effect of the RAE's audit process and the effect of selective funding that was supposed to be informed by the assessment results. Some critics pointed out the possibility that the long-term improvements in research quality and management were caused mainly by managerial necessity and selective research funding rather than by adaptive or strategic responses to the evaluative process of the RAE (Thomas, 2007, p.46; McNay, 2009, p.45).

funding accountability. Environmental changes (e.g., regime changes, changing budget conditions, ministerial changes and changes of key officers) may also have impacted the shifting goals or shifting weights and priorities among various goals. It has been also pointed out that the RAE's functions evolved over time (Roberts, 2003, para.71; PA Consulting Group, 2008, p.9).²⁸

In considering the idea of unintended consequences in the analysis of the RAE, therefore, we should delve into the intricate details of the policy's "intentions" by asking the following question: "Who intended what kinds of goals, at what point in time, and under what circumstances?" In other words, we need to examine the evolutionary trajectories of policy goals and intentions carefully. Such an analysis will be especially important in learning from the UK's experience so as to achieve a closer alignment between policy goals and specific assessment procedures.

6. Concluding Remarks: From the RAE-ability to the REF-ability?

The UK's experiences with national research assessment and selective research funding are instructive to other countries, including Japan, in that the UK's HEIs and their academics are now undergoing a new evaluative exercise—Research Excellence Framework (REF). With the introduction this new evaluation, "REF-able" (pronounced [refəbl]) has taken over "RAE-able," and has become a commonly used word in the parlance of British academia. Addressing the issue of what has (and has not) changed

in the transition from RAE-ability to REF-ability will provide important insights for understanding reactivity in university evaluation.

One of the most crucial changes from RAE-ability to REF-ability is concerned with the introduction of an evaluative element concerning the policy/practice impact of academic research, or simply the "impact element." The introduction of the impact element into the REF was arguably the most drastic and most controversial changes in the history of national research assessment in the UK (Bekhradnia 2009, para. 15; Fernadez 2009; Shepherd 2009; Pettigrew, 2011; Goldfinch and Yamamoto, 2012, p.160; Oancea, 2013; *Nature*, 2013; Owen, 2013).²⁹ For the first time in the history of national research quality assessment in the UK, non-academic research impacts are to be graded in their own right as a separate element.³⁰ Whereas practical impacts may have been taken into account by panels in their assessment of the overall quality of submitted research in previous RAEs, in the REF 2014, the impact is supposed be assessed as a separate element being given as much as 20 percent weight: the other two weighted elements are research quality (65 percent) and peer esteem (15 percent).

As previously pointed out, undue emphasis on basic science at the expense of applied research has been one of the most significant criticisms raised against the RAE. The introduction of the impact element, therefore, could be regarded as an important move to redress this tendency. Paying proper attention to the practical relevance of academic research

²⁸ According to Roberts (2003), the RAE was essentially a "quality assurance policy" and its original function was "driving up standards through reputational incentives." On the other hand, the report of the PA Consulting Group (2008) argued that RAE 1992 was implemented to "grade research departments across the UK for the purpose of selective funding allocation." Differences in these two interpretations appear to attest to some ambiguity in policy goals.

²⁹ Both the REF itself and its impact element have complex roots in policy discourses (Oancea, 2013). When the idea of the REF was initially announced in the pre-Budget report of December 2006 (HM Treasury, 2006) without any consultation with the Education Department and the Funding Council (Brown and Carosso, 2013, pp.53-54), the REF was supposed to be based on bibliometrics and was therefore a very different system from the RAE. Yet, through consultations and political negotiations, the REF eventually became a modified version of the RAE instead of its total replacement. As a result, the core features of the RAE, such as research assessment by peer panels and the assessment of subject areas rather than individual researchers, remained almost intact (Goldfinch and Yamamoto 2012, pp.157-163; Bekhradnia, 2009, para. 6).

³⁰ We should note, however, that research that is eligible for impact assessment must be underpinned by "excellent" research outputs that can be graded 2* (internationally recognized) or above, in terms of academic quality. In other words, academic excellence and non-academic impact are not totally separate criteria.

may also serve as a corrective countermeasure to the earlier excessive emphasis on publishability, which has tended to delimit academics' attention solely to academic audiences. In other words, the inclusion of impact elements in the assessment may work as an antidote to the sub-optimizing behaviors of HEIs: they may attempt to find a way to make what is good for science also good for society.

On the other hand, however, if the reactive responses of HEIs and departments to the impact assessment overstep a certain limit, the result may be another formative reactivity of a negative kind. For example, it may sometimes happen that the assessment of impacts ends up with an emphasis on immediate applicability that is readily assessable, or "REF-able." In such a case, longer-term research relevance, which is not readily subject to assessment, may be overlooked or slighted. In other words, in such a case, what is good for the HEIs and society in the short term is not good for science and society in the long term.

A recent article in the *Times Higher Education* (Jump, 2013b) reported that a pro vice-chancellor of a university predicted that British universities will begin gathering impact case studies in real time so as to make future submissions easier. Similarly, a survey sent by a venture firm to department heads around the UK included a question asking if they were interested in "embedding impact measurement into everyday practice."

Just as the exclusive focus on publishability tends to delimit the scope of research activities, excessive preoccupation with immediate applicability will narrow the research scope. Above all, such research practices will sterilize the fertile soil from which serendipitous ideas produce fruit. Anyone who is in charge of making research policy should not forget the fact that serendipity is nothing but the most important unanticipated consequences of the human undertaking that is called "research" or "science" (Merton, 1957; Merton and Barber, 2004).

Acknowledgements

The research for this note was partly supported by

JSPS KAKENHI Grant Numbers 23530618 and 24243074.

This research was also supported by a research grant from the Japan Center for Economic Research. The first author of this paper is very grateful to the Nissan Institute of Japanese Studies at Oxford University and St Antony's College for their help in conducting this research in the UK. The author would like to express his deepest gratitude to Professors Takehiko Kariya and Ian Neary of Oxford University for hosting his stay at the UK. This paper profited from our interviews and/or conversations with a number of British and Japanese informants, primarily academics and policy makers. They include (in alphabetical order): Mr. Bahram Bekhradnia, Professor Rick Delbridge, Dr. Tim Edwards, Dr. Shaun Goldfinch, Professor Roger Goodman, Dr. Takayuki Hayashi, Professor Tatsuo Kawashima, Dr. Robin Klimecki, Professor Ian McNay, Professor Alice Oancea, Mr. David Sweeney, Professor Robyn Thomas, Professor David Watson, Professor Hugh Willmott, and Professor Kiyoshi Yamamoto. The authors would also like to extend their appreciation to the participants of a Cardiff Organisation Research Group Seminar and Bristol University Management Seminar who provided helpful comments. We are greatly indebted to the insightful comments and advice of these academics and an anonymous reviewer of *Research on Academic Degrees and University Evaluation*. We are solely responsible for any factual errors or misinterpretations that may be included in this research note.

Bibliography

- ABRC (Advisory Board for the Research Councils) (1990) *Peer Review*. Advisory Board for the Research Councils
- Adams, J. & Gurney, K. (2010) "Funding Selectivity, Concentration and Excellence." HEPI (Higher Education Policy Institute) (<http://www.hepi.ac.uk/455-1793/Funding-selectivity,-concentration-and-excellence---how-good-is-the-UK%27s-research.html>)
- Armstrong, J. (1982) "Research on Scientific Journals." *Journal of Forecasting*. 1, 83-104.

- Becker, H. (1998) *Tricks of the Trade*. University of Chicago Press
- Bekhradnia, B. (2004) "Government, Funding Council and Universities: How Should They Relate?" HEPI (<http://www.hepi.ac.uk/466-1122/Government,-Funding-Council-and-Universities---How-Should-They-Relate.html>)
- Bekhradnia, B. (2007) "Evaluating and Funding Research through the Proposed 'Research Excellence Framework'." HEPI (<http://www.hepi.ac.uk/466-1324/Evaluating-and-funding-research-through-the-proposed-Research-Excellence-Framework.html>)
- Bence, V. & Oppenheim, C. (2004a) "The Role of Academic Journal Publications in the UK Research Assessment Exercise." *Learned Publishing*, 17, 53-68.
- Bence, V. & Oppenheim, C. (2004b). "A Comparison of Journal Submissions to the UK's Research Assessment Exercises 1996 and 2001 for UoA 43 (Business and Management Studies)." *European Management Journal*, 22(4), 402-417.
- Bence, V. & Oppenheim, C. (2005) "The Evolution of the UK's Research Assessment Exercise: Publications, Performance and Perceptions." *Journal of Educational Administration and History*, 37 (2), 137-155.
- Besley, T. (2009) *Assessing the Quality of Educational Research in Higher Education*. Sense Publishers
- Biagioli, M. (2002) "From Book Censorship to Academic Peer Review," *Emergences*, 12, 11-45.
- Blumer, H. (1954) "What is Wrong with Social Theory?" *American Sociological Review*, 19(1), 3-10.
- Brown, H. (2007) "How Impact Factors Changed Medical Publishing: And Science." *BMJ: British Medical Journal*, 334, 561-564.
- Brown, R. and Carasso, H. (2013) *Everything for Sale?* Routledge
- Bulmer, M. (1979) "Concepts in the Analysis of Qualitative Data." *Sociological Review*, 27(4), 651-677.
- Connor, S. (2010) "Nobel Laureates: Don't Put Money before Science." *The Independent*, January 7, 2010.
- Corner, J. (2009) "A Very Peculiar Practice." *Times Higher Education* (June 11, 2009)
- Corwin, R.G. (1981) "Foreword." In Sieber, S. *Fatal Remedies*. Plenum
- Curran, P. (2000) "Competition in UK Higher Education." *Higher Education Quarterly*, 54(4), 386-410.
- Elton, L. (2000) "The UK Research Assessment Exercise: Unintended Consequences." *Higher Education Quarterly*, 54(3), 274-283.
- Fernandez-Armesto, F. (2009) "Poisonous Impact." *Times Higher Education*, December 3, 2009.
- Goldfinch, S. & Yamamoto, K. (2012) *Prometheus Assessed?* Chandos Publishing
- Goodman, R. (2013) "Changing Roles of the State and the Market." In Goodman, R., Kariya, T. and Taylor, J. eds. *Higher Education and the State*. Symposium Books
- Guená, A. (2001) "The Changing Rationale for European University Research Funding: Are There Negative Unintended Consequences?" *Journal of Economic Issues*, 35(3), 607-632.
- Guená, A. & Martin, B.R. (2003) "University Research Evaluation and Funding: An International Comparison." *Minerva*, 41, 277-304.
- Hada, Y. ed. (2005) *University Evaluation for the New Age*. Toshindo (in Japanese)
- Hada, Y. (2010) "Higher Education System in the UK and Foundation Forms of Universities." In *Reports on Foundation Forms of Universities*. Center for National University Finance and Management
- Hayashi, T. (2009). "University Research Evaluation Reforms and Governance of Science." *Kenkyu, Gijutsu, Keikaku*, 24(3), 231-242. (in Japanese)
- Harley, S. (2000) "Accountants Divided: Research Selectivity and Academic Accounting Labour in UK Universities." *Critical Perspectives on Accounting*, 11(5), 549-582.
- HEFCE (1997) "The Impact of the 1992 Research Assessment Exercise on Higher Education Institutions in England." HEFCE 1997 M6/97.
- HEFCE (n.d.) "How We Fund Research." (<https://www.hefce.ac.uk/whatwedo/rsrch/howfundr/>)

- HEFCE (2013) "Quality-Related Funding Data 2013-14 (An example of the calculation of quality-weighted volume used in the 2013-14 mainstream QR funding method)." (<http://www.hefce.ac.uk/data/year/2013/quality-relatedfundingdata2013-14/>)
- Henkel, M. (1999) "The Modernisation of Research Evaluation: The Case of the UK." *Higher Education*, 38, 105-122.
- Iwata, S. (2005) "From RAE2001 to RAE2008." *Research on Academic Degrees and University Evaluation*, 3, 137-198. (in Japanese)
- Iwata, S. & Tokuda, M. (2002) "Assessment Criteria and Working Methods in Panels for Research Assessment Exercise 2001 in the UK." *Research in University Evaluation*, 2: 57-172. (in Japanese)
- Jump, P. (2013a) "Evolution of the REF." *Times Higher Education* (October 17, 2013).
- Jump, P. (2013b) "REF More Burdensome than RAE, Pro V-Cs State." *Times Higher Education* (December 12, 2013).
- Kitahara, K. & Hirota, T. (2012) "How Should We Think about Higher Education and Research?" In Kitahara, K. et al. *Research Reports on University Evaluation*. (in Japanese) MEXT
- Kinmonth, E. (2005) "Problems of Research Assessment in the UK." In Hada, Y. ed. *University Evaluation for the New Age*. Toshindo (in Japanese)
- Kogan, M. & Hanney, S. (2000) *Reforming Higher Education*. Jessica Kingsley Publishers
- Lee, F., Xuan P., & Gu, G. (2013) "The UK Research Assessment Exercise and the Narrowing of UK Economics." *Cambridge Journal of Economics*. 37, 693-717.
- Lucas, L. (2006) *The Research Game in Academic Life*. Open University Press
- Macdonald, S. & Kam, J. (2007) "Ring a Ring o' Roses: Quality Journals and Gamesmanship in Management Studies." *Journal of Management Studies*, 44(4), 640-655.
- McNay, I. (1997) "The Impact of the 1992 RAE on Institutional and Individual Behavior in English Higher Education: The Evidence from a Research Project." HEFCE
- McNay, I. (2003) "Assessing the Assessment: An Analysis of the UK Research Assessment Exercise, 2001, and Its Outcomes, with Special Reference to Research in Education." *Science and Public Policy*, 30(1), 47-54.
- McNay, I. (2007) "Research Assessment: Researcher Autonomy." In Kayrooz, C. & Akerlind, G.S. eds. *Autonomy in Social Science Research*. JAI Press
- McNay, I. (2008). "The Crisis in Higher Education: The Views of Academic Professionals on Policy, Leadership Values and Operational Practices." *Higher Education Review*, 40(2), 3-25.
- McNay, I. (2009) "Research Quality Assessment: Objectives, Approaches, Responses and Consequences." In Brew, A. & Lucas, L. eds. *Academic Research and Researchers*. Open University Press
- McNay, I. (2011) "Research Assessment: Work in Progress, or *la lotta continua*." In Saunders, M., Trowler, P. & Bamber, V. eds. *Reconceptualising Evaluation in Higher Education*. Open University Press
- March, J. (2009) *Primer on Decision Making*. Free Press
- Merton, R. K. (1936) "The Unanticipated Consequences of Purposive Social Action." *American Sociological Review*, 1(6), 894-904.
- Merton, R.K. (1940) "Bureaucratic Structure and Personality." *Social Forces*, 18(4):560-568.
- Merton, R.K. (1957) "The Bearing of Empirical Research on Sociological Theory." In *Social Theory and Social Structure*. Free Press
- Merton, R.K. (1968) "Mathew Effect in Science." *Science*, 159, 56-63.
- Merton, R.K. & Barber, E.G. (2006) *The Travels and Adventures of Serendipity*. Princeton University Press
- MEXT (Ministry of Education, Culture, Sports, Science and Technology) (2012) "University Reform Plans." MEXT (in Japanese)
- Middleton, S. (2009) "Becoming PBRF-able: Re-

- search Assessment and Education in New Zealand." In Besley, T. ed. *Assessing the Quality of Educational Research in Higher Education*. Sense Publishers
- Mingers, J. & Willmott, H. (2013) "Taylorizing Business School Research: On the 'One Best Way' Performative Effects of Journal Ranking Lists." *Human Relations*, 66(8), 1051-1073.
- Myott, J. (1999) "Publishing: The View from Cambridge University Press." *History of the Human Sciences*, 12(4), 127-131.
- Nature (2013) "Impact (specials)." *Nature*, 502.
- Nightingale, P. & Scott, A. (2007). "Peer Review and the Relevance Gap: Ten Suggestions for Policymakers." *Science and Public Policy*, 34(8), 543-553.
- Oancea, A. (2013) "Buzzwords and Values: The Prominence of 'Impact' in UK Research Policy and Governance." *Research Trends*. (<http://www.resdeartrends.com/issue-33-june-2013/buzzwords-and-values/>)
- PA Consulting Group (2008) *RAE 2008 Accountability Review*. PA Consulting Group
- Pettigrew, A. (2011) "Scholarship with Impact." *British Journal of Management*. 22, 347-354.
- Piercy, N. (2000) "Why It is Fundamentally Stupid for a Business School to Try to Improve Its Research Assessment Exercise Score." *European Journal of Marketing*, 34(1/2), 27-35.
- Power, M. (1994) *The Audit Explosion*. Demos
- Power, M. (1997) *The Audit Society*. Oxford University Press
- Power, M. (2007) "Organizations and Auditability: A Theory." Paper presented at Lancaster University. <http://www.lums.lancs.ac.uk/events/owt/10300/>
- RAE 2008 (2009a) *RAE 2008 Manager's Report*. RAE 2008
- RAE 2008 (2009b) "RAE 2008 Sub-panel 62 History Subject Overview Report." RAE 2008
- RAE 2008 (n.d.) "About the RAE 2008" <http://www.rae.ac.uk/aboutus/>
- RAE 2008 (n.d.) "Quality Profiles." <http://www.rae.ac.uk/aboutus/quality.asp>
- Research Fortnight (2005) "Making Sense of the Panel Guidance." *Research Fortnight* (July 20, 2005)
- Richards, H. (2000) "Research Exercise Hits Publishers." *Guardian* (August 1, 2000).
- Richards, H. (2001) "Seeking the Best." *Guardian*, January 9, 2001.
- Roberts, G. (2003) "Review of Research Assessment. Report by Sir Gareth Roberts to the UK Funding Bodies. Issued for Consultation May 2003," RA Review
- Sato, I. (2006) *Fieldwork*. Shin-yosha. (in Japanese)
- Schatz, D. (2004) *Peer Review*. Rowman & Littlefield
- Shepherd, J. (2009) "Humanities Research Threatened by Demands for 'Economic Impact'." *Guardian*, October 13, 2009.
- Shattock, M. (1994) *The UGC and the Management of British Universities*. Open University Press
- Shattock, M. (2009) "The Impact of the UK Research Assessment Exercise," *International Higher Education*. 56,18-20.
- Sieber, S. (1981) *Fatal Remedies*. Plenum
- Singleton, R.A. & Straits, B.C. (2010) *Approaches to Social Research*. Oxford University Press
- Strathern, M. ed. (2000) *Audit Cultures*. Routledge
- Smith, D. (1998) "Gold Standard Guarantees Mediocrity." *Guardian*, May 18, 1998.
- Swinnerton-Dyer, P. (1991) "Policy on Higher Education and Research: The Rede Lecture 1991." *Higher Education Quarterly*, 45(3), 204-218.
- Talib, A.A. (2013) "The RAE and Publications: A View of Journal Editors." *Higher Education Review*, 33(1), 32-47.
- Taylor, J. (2001) "The Impact of Performance Indicators on the Work of University Academics: Evidence from Australian Universities." *Higher Education Quarterly*, 55(1), 42-61.
- Thomas, E. (2007) "National Research Assessment in Higher Education." In Burgh H.D., Black, J. & Fazackerley, A. eds. *Can the Prizes Still Glitter?* University of Buckingham Press
- UFC (Universities Funding Council) (1989) *The 1989 Research Assessment Exercise*, Universities Funding Council

- UNIVERSITAS (2003) *Operational Review of the Research Assessment Exercise 2001*, UNIVERSITAS
- UUK (Universities UK) (2003) *Funding Research Diversity*. Universities UK
- Walford, L. (2000) "The Research Assessment Exercise: Its Effect on Scholarly Journal Publishing." *Learned Publishing*, 13, 49-52.
- Waters, L. (2004) *Enemies of Promise*. Prickly Paradigm Press
- Watson, D. (2013) "Leading the British Universities Today: Your Fate in Whose Hands?" In Canllender, C. & Scott, P. eds. *Browne and Beyond*. Institute of Education
- Watson, D. & Maddison, E. (2005) *Managing Institutional Self-Study*. Open University Press
- Webb, E.J., Campbell, D.T., Schwarts, R.D., & Sechrest, L. (1981) *Unobtrusive Measures*. Rand McNally
- Weiss, C.H. (1980) "Knowledge Creep and Decision Accretion." *Knowledge: Creation, Diffusion, Utilization*, 1 (3), 381-404.
- Weller, A. (2001) *Editorial Peer Review*. Assist
- Willmott, H. (2011) "Journal List Fetishism and the Perversion of Scholarship: Reactivity and the ABS List." *Organization*, 18(4), 429-442.
- Yokoyama, K. (2006) "The Effect of 2001 Research Assessment Exercise on Organizational Culture in English Universities." *Research in Higher Education*, 37, 79-96. (in Japanese)

(受稿日 平成26年 1 月28日)

(受理日 平成26年 7 月10日)

[要 旨]

“RAE-able” から “REF-able” へ？
——研究評価をめぐる高等教育機関の反応的変容に関する試案的考察

佐藤 郁哉*, 遠藤 貴宏**

評価活動に対する大学側の反応 (reactivity) については、ともすればその否定的な側面、すなわち正確な測定と評価に対する阻害因としての側面がクローズアップされてきた。反応的な対応に対して適用されてきた用語 (たとえば、ゲームプレイング、駆け引き、作文など) の多くもまた、ネガティブな意味合いが濃いものである。しかしながら、大学評価に関する文献を検討してみると、評価活動に対する高等教育機関の適応的ないし戦略的な対応がむしろポジティブな結果をもたらす場合もあったという可能性が浮かび上がってくる。これらの問題について考察していく上で、英国における研究評価事業 (Research Assessment Exercise) の事例はきわめて興味深い知見を提供している。本稿は、英国で実施した文献サーベイとインタビュー調査にもとづいて、評価に対する大学の対応について分析する上での分析的枠組みを提示する。その分析フレームでは、次の3つの概念が軸となる——意図せざる結果、反応的変容、政策意図の変遷軌道。

* 一橋大学大学院 商学研究科 教授

** 神戸大学 経済経営研究所 専任講師