

Chapter 8

The Excellence of IT: Conceptions of Quality in Academic Disciplines

Andrew Abbott

Abstract The chapter approaches the subject matter, the Research Assessment Exercise (RAE), from the perspective of the quasi-ethnographer.

In the UK, as funding councils have started to finance institutions on the basis of a RAE designed to measure 'quality', questions arose as to how such concepts can be measured, whether or how research groups ought to be funded, and to what extent such measurements are subject to manipulation. The chapter addresses the first question on the basis of responses to the RAE by academics. Five dimensions are explored: cognizability (or measurability), units of measurement (or what should be measured), space and time (or their spatiotemporal character), modality (positive and negative aspects affecting perceptions) and content ("what is 'excellence' in research?"). No shared foundations on which to base a RAE could be deduced: "nowhere in these comments is there anything about the substantive content of excellent work nor, indeed, of work at all". This statement is true for both the sciences on the one hand and the humanities and social sciences on the other.

There is a curious anomaly in the study of knowledge and its institutions. While the history and sociology of science have flourished, the history and sociology of the social sciences and humanities have not. There has been a steady but small flow of empirical work. But there is little or nothing corresponding to the great theoretical debates between the Mertonians and their opponents over scientific knowledge and its purported norms.

This difference may reflect the different projects of the two great continents of knowledge. The practice of science assumes that there are answers "out there" in the natural world and that we use theory and empirical investigation to find them. But the humanities and humanistic social sciences (HHSS) are more interested in creating new interpretations of well-known things, or putting together things previously kept apart, or maximally filling a space of possible things to say. These practices do not have the same directional quality as do the sciences, in which old work is necessarily subsumed and rejected by new work. There is no best interpretation of

A. Abbott (✉)

Department of Sociology, University of Chicago, 1126 E. 59th St., Chicago, IL 60637, USA
e-mail: aabbott@uchicago.edu

M. Herbst (ed.), *The Institution of Science and the Science of Institutions*,
Boston Studies in the Philosophy and History of Science 302.

DOI 10.1007/978-94-007-7407-0_8, © Springer Science+Business Media Dordrecht 2014

Moby Dick, nor must we possess such an interpretation before we write a book about Melville. There are to be sure what I (Abbott 2008) called Rankian facts, which are subject to purely empirical investigation: the date of Melville's death, the status of his finances in 1854, and so on. But there is no necessary order or direction to interpretations of his great novel, just as there is no necessary order or direction to HHSS knowledge more broadly.

But despite the surface plausibility of this identification of differing goals, it is an empirical question whether humanists and humanistic social scientists do in fact think they accomplish something different than do scientists. To investigate this question I here use data on disciplinary reactions to an outside threat that challenges disciplines to clarify not only their conceptions of research excellence, but also of their underlying knowledge projects. My chief finding is that the knowledge projects of the sciences and humanities seem not different but similar, and in particular similar in their contentless character.

Method

Our data come from the British Research Assessment Exercise (RAE). Dating from the late 1980s, the RAE uses expert panels to rate every department and research unit in the United Kingdom. Units to be assessed submit a fixed number of papers per person, as well as a unit narrative and various statistics. A concurrent consultation process identifies and appoints expert panels in each field, which read the submitted materials, rate both individuals and units, and publicly report the rating of units and the distribution of the ratings of individuals. These ratings then determine a substantial portion of unit funding.

Disciplinary responses to the 2001 RAE can all be found on the relevant website.¹ They respond to a widely-distributed "invitation to comment" from a committee charged with the post hoc evaluation of RAE 2001. The invitation resulted in verbatim responses to a varying number of fixed questions by a considerable number of voluntary (and hence self-nominated) respondents. There are 87 total responses whose disciplinary origins can be clearly identified: 30 from the humanities, 31 from the social sciences, and 26 from the natural sciences. These 87 responses cover most of the academic disciplines in those areas. Since I aim to contrast patterns of discourse and assumptions across the broad continents of academic knowledge, such breadth is sufficient. As for quantity judgements, the data will support "most, many, some, few", but nothing more detailed.

The data's self-nominated character might seem worrisome; perhaps these are mainly complaints from departments that did badly, getting ratings of 3a or less. This is however not the case. There are 24 responses (of the 87) from departments. Of these only 4 are from departments rated 3a and another 5 from departments rated

4. The rest are rated 5 or in a few cases 5*, the highest rating in RAE 2001. Nor is concentration a problem: the departmental responses are scattered by geographic location, by discipline, and by status of university. The remaining responses are from national bodies (disciplinary or interdisciplinary societies, and councils of chairs or of program directors) or from senior administrators within the three areas here discussed.²

Strategic intent is clear in some responses: for example, interdisciplinary groups often push for their own separate panels in the next RAE. But the majority of these responses are neither rants or complaints. If there is an overriding emotion, it is an exasperated perplexity. Moreover, any biases implicit in self-nomination, emotional reaction, and strategic response do not affect the investigation. My aim is to recover from the language and assumptions of these responses their underlying conception of excellence in research. There is little reason to think that scholars would systematically deform their conceptions of excellence in this setting. They are more likely to have used their customary conceptions, indeed all the more so if they were angry. General differences between the knowledge projects of the humanities and social scientists should still be obvious.³

The more important bias is in ourselves. As scholars we think we already know perfectly well what is 'quality' research or 'excellent' research. Yet it is a discipline's substantive sense of these words that we seek. To disarm such prior beliefs, I global-replaced the word 'excellence' throughout the data with the word IT. I replaced the word 'quality' with the word ITQ. I replaced the word 'excellent' with the word ITLIKE. I shall follow that convention in quotes below, hoping to prevent readers, too, from importing their own knowledge into their understanding of IT.

Having made this replacement I seek to infer the meaning of IT, ITQ, and ITLIKE from statements made about them. For example, the phrase "a work of lower IT" implies that IT is a quality that inheres in works and that has a rank order, just as the sentence "only the ITQ of recent publications should be measured" implies that ITQ can be lost by an individual over the life course. Neither of these statements, it should be noted, gives any information about the substantive content or purpose of IT.

The Excellence of IT

I consider the meanings of IT and ITQ in these documents under five headings. The first of these is cognizability: how can IT be known? The second is units: of what

²Of responses from administrators, I used only those from deans or heads of schools, since these can be attributed to the three subject continents. Of those from societies, some are based on surveys, some on consultation with membership or with committees, and some on the opinion of whatever society officer chose to respond.

³I should underscore that I am not interested in explaining why respondents thought this or that about the RAE, in finding 'causes' of their attitudes like RAE rating received, type of institution, funding, etc. Nor would knowing such 'causes' change my analysis, which is addressed to implicit conceptions of knowledge.

¹The responses can be found at www.ra-review.ac.uk/invie/resp/resp/subject.asp. Since the RAE system is changing fundamentally, there will probably be no such response to the 2008 exercise.

kinds of things is IT a property? The third is spatiotemporal character: where is IT? where might IT appear? When? Does IT endure? The fourth concerns the issue of modality: is IT necessary or merely possible? absolute or relative? always positive or sometimes negative? Finally what is IT substantively? Do the sciences say that IT work "extends our understanding of the natural world", while humanists think IT work "deepens our interpretations of human action and values"?

Cognizability—Metrics and Experts

Two central facts emerge about how IT and ITQ are known. First, IT and ITQ cannot be measured. Second, IT and ITQ can be recognized only by experts: "The expert and the peer are [...] one and the same" <#12>⁴.

Humanists are blunt about this: "We reject", says one group, "any approach that relies solely on the dubious notion of objective data" <#10>. Such measures, says another response, have "no obvious pertinence to arts and humanities research" <#11>. These denials are however limited to the arts and humanities: an algorithm based on metrics "may be quite properly useful in science and engineering" <#12>, remarks one group.

By contrast, the statisticians and some of the social scientists feel that measuring IT by metrics is feasible and preferable. The Royal Statistical Society says that "Good practice, which provides academic objectivity and possesses [virtues like rigor, fairness, clarity, and so on] is well understood, and indeed commonly taught by statisticians [...]"^{<#19>}. Or again, "where hard metrics can be used to inform expert judgment, panels should be formally obliged to use them" <#6>⁴.

But there is disagreement. "The notion that such metrics are 'totally objective' is clearly a fallacy as virtually all metrics are subject to different interpretations" <#9>, says one group. "The work of the sociologists of science [...] has repeatedly demonstrated the underlying subjectivity behind such measures" <#2>, says another. Indeed, there is some overt hostility. The heads of anthropology departments say "monographs of fundamental importance to the continuation of the discipline do not necessarily rate high in bibliometric measures, while meretricious works may do so" <#22>. Another group is blunter: "Garbage in, garbage out" <#26>. As in the humanities, suspicion of colleagues surfaces: "Citation enhancement is a known sport", remarks the LSE Policy group <#15>. Others argue that citation analysis "will largely reflect the existing disciplinary nature of knowledge while there is no a priori reason why such organization will correspond to future needs" <#17>.

The mixed verdict of the social scientists disappears in the sciences, which are—like the humanities—overwhelmingly negative. There are to be sure some radical objectivists. "The system should be measured using metrics, the majority of which

are comparable and objective fact" <#7>. Or "The assessment must be automatic, leaving no room for subjective assessment" <#13>. But these are rare drops in a sea of rejection: "[S]ome objective data must be fed into the consideration of research strength, but it is inappropriate to judge research purely on the basis of metrics" <#3>; "No single metric can measure the ITQ of a piece of work" <#11>. As elsewhere, there is suspicion of colleagues: "care must be taken to avoid the negative consequences of 'metric-driven' behavior" <#9>.

In summary, the vast majority from all fields believe that IT and ITQ cannot be measured by objective indicators. Such indicators are irrelevant. If relevant then they are unfeasible. If feasible, then they are wrong. Disagreement comes only from those who invented those indicators or whose daily practice involves them. To some extent this immeasurability is attributed to games-playing and manipulation; to some extent metrics are thought useful in other fields or as minor adjuncts to other forms of judgment. But the main and decisive fact is rejection.

While IT and ITQ cannot be measured, they can nonetheless be recognized by members of the group itself. This fact is accepted from the humanities ("If you want a sound assessment, it makes sense to ask the experts" <#4>), to the social sciences ("expert peers should carry the main burden of judgments of ITQ" <#19>), to the natural sciences ("lay people cannot judge research IT because by its very nature it exists at the upper limits of even expert knowledge" <#17>).

But only true experts can recognize experts. Even closely allied fields are considered inept. Architects complain that the built environment panel "had little appreciation of architecture as a discipline [...] nor of creative design output or of book output that crossed the technology/history divide" <#7>. Economic history, women's studies, and criminology make similar complaints. Even within mathematics this "outsiders" issue is raised: "grades awarded to some departments reflect the uneven distribution of disciplines [i.e., subfields of mathematics] across the panel and a formulaic approach to grading rather than the intrinsic ITQ of the work being assessed" <#22>.

All kinds of outsiders are rejected: practitioners ("users and practitioners [...] are important constituencies, but final decision must rest with panels composed predominantly of peer academics" <#18>), financial experts ("The inclusion of financial experts appears only to be justified if there is an explicit need to judge value for money rather than absolute research IT" <#17>), non-UK disciplinary experts ("Many of them, despite their eminence, are not familiar with the mindset of the UK civil service [...] It is questionable if they contributed anything to the process" <#29>), even critics or audiences ("practitioner researchers' should identify the peers of the work, the sort of practitioners who will most immediately understand and appreciate what this piece is doing, explain what one's peers will find original and striking about the piece [...] <#18>).

In short, IT cannot be measured, but IT can be recognized by experts, who are defined as the members of the group whose ITNESS is at issue. This is a totemic religious system in which only clan members can recognize the totem.

⁴ <#12> means the twelfth humanities response in our list of respondents. <#3* > and <#6* > are the equivalents for social and natural sciences. Since the quotes are retrievable via search algorithm from the on-line data, we save space by omitting a list of respondents.

Units

ITness is implicitly predicated of several things: works, research in general, individuals, groups (research groups, departments, centers), and even whole disciplines. There is also explicit discussion of the proper unit in which to conceive of ITness and of who among scholars is capable of ITness.

IT is sometimes explicitly predicated of works and research. The latter is more common: "research ITQ in the humanities" <h1>. Sometimes it is not research that has ITQ, but "output", as in "Research income measures input and cannot verify the ITQ of output" <h3>. Only once in a while do we hear of highly specific IT as in the phrase "accurate measurement of research ITQ at the level of specific contributions" <ss11>. The idea that primarily works have ITness pervades the discussion of the ITness of individuals. Thus, the common belief that highly-rated departments or individuals may rest on their laurels logically contains the assumption that a given individual can produce works of varying IT throughout the life course, which in turn implies that the true locus of IT is the work, not the individual.

That point about implicit units made, however, the majority of responses explicitly locate IT either in individuals or in groups of one sort or another, rather than in works. "Assessment could be at Unit of Assessment [UoA] level, but measure of IT should be at individual or research group level" <ns8>. A group of humanists speaks of "the ITQ or range of an individual's or a department's research" <h14>. Most commonly, the explicit unit of IT level is the individual. For the economists, for example: "The ITQ of a researcher is a latent signal to be extracted from observable outcomes in the noisy outcome that constitutes creative research" <ss11>. Or from the humanities—"The lone researcher who may [...] produce internationally recognized publication [...]". <h30>. And from the historians "The ITQ of research in any institution of higher education is ultimately dependent on the ITQ of its individual researchers" <h26>. Even the physicists, with their gargantuan team experiments, recognize an individual level: "where possible, [assessments] should be made at the individual level. In subjects where papers have many authors, the individual should be encouraged to state their [sic] role in the research" <ns10>. All of these phrasings assume that IT is something that finds its place in people, and most often that it is an enduring quality of people.

By contrast, there are respondents—chiefly but not exclusively from the natural sciences—whose language locates IT in groups. For example, "[there are] nodes of IT within the units" <ns15>. Some respondents are adamant that the group level is the real or proper one. "If individuals are assessed we could lose synergy created by larger groups and units of assessment" <ns17>. Or "Assessment of individuals would be unwieldy and invidious and assessment at the level of institutions would conceal centres of IT [...]". <ss19>.

By contrast, others are quite hostile to group-level concepts of IT. "While collaborative work is practiced by many Arts and Humanities Scholars, the bulk of their work is and will be for the foreseeable future done in a single scholar mode which experience has shown yields large amounts of ITlike output [...]". <h27>. Still others favor a billevel concept: "selected publications should remain a key diagnostic of

individual calibre, but [...] the final departmental scores should also reflect to some degree the overall research culture of research groups and the department" <ss25>.

These various loci of IT are clearly not exclusive. But overall it seems that IT exists in certain people as a potential, perhaps even as a realized potential. It may in some fields be located in groups or "research cultures". But in the last analysis IT is directly manifest only in work itself.

Space and Time

As we have seen earlier, units of IT like individuals and departments are embedded in a larger community of experts who are capable of judging a particular type of IT. This larger community—implicit throughout—is a 'discipline'. The detailed properties of disciplines are never given; what a discipline shares is simply its ability to judge the IT of its own work. A discipline, that is, is a type of IT.

This exclusivism is evident in the complaints mentioned earlier about the inability of outsiders to judge the IT of research or researchers, which arise most often in discussions of 'interdisciplinarity'. "There must be a more clearly worked out system of inter-panel assessments for genuinely interdisciplinary groups" <h20>, says one design department Faculty in women's studies resent being rated in disciplines "where the specificities of their research cannot gain full recognition" <ss19>. The History at Universities Defence Group was blunt: "Interdisciplinary research and multidisciplinary research should be encouraged, but the tail must not wag the dog" <h26> 5.

Exclusivism is also underscored by complaints that panel experts were generally unable to comprehend the details of specialized work. We hear this from scholars in French language, from geographers, from psychologists, from mathematicians. Certainly the majority agree with the Leicester archaeologists who remark "creating broader subject panels would weaken the reliability and credibility of peer review" <ss25>.

Inside disciplines, the spatiality of IT varies somewhat between the humanities and the natural sciences. Division of labor is unmentioned (probably because taken for granted) in the natural science responses. But it figures occasionally—and negatively—in the humanities, where it is perceived as alien to the humanistic style of research: "Many fields, especially in the humanities, do not require concentration of people to enable high ITQ research. Lone scholars, or small groups of researchers frequently make significant contributions" <h20>.

Explicit discussion of IT differences between units within a discipline is largely limited to the sciences, where it is taken for granted. There is much explicit elitism:

⁵ At the same time, many respondents saw themselves as potentially subject to more than one 'discipline'. The disciplines, that is, may make up not a set of exclusive categories, but rather a system of tolerances, in which many scholars fit under several panels and each panel could potentially assess several such different subgroups of scholars.

"The absolute number of 'high ITQ researchers', rather than the mean, matters in research, since 'extreme values' are what really makes the difference" <ss12>, and "weaker or developing groups should be encouraged not to apply for expert review, and to bid for funding for a small pot of money ring-fenced for developing research capacity" <ns16>.

Conceptions of the distribution of IT change when money is involved. Should IT be conceived as absolute or as relative to available resources? This is not an issue for humanists, who get little money in any case. Among social scientists, it engages only the economists, who predictably argued that "marginal research productivity should be equated across all units so that the impact of the next pound of funding is the same everywhere" <ss11>. But the issue divides the natural sciences. Where one group holds "In essence, exceptional 'value for money' has been achieved in the short-term by jeopardizing the longer-term health of the university research infrastructure" <ns1>, another says "the criteria for judging performance should, in addition, include value for money" <ns5>.

A particular distributional puzzle is the department that loses major funds because of minor ITQ differences. One such victim complains at length, but a very highly rated department voices the same issue, speaking of the "large funding impact a relatively small change in submission can make" <ns8>. At the same time there are complainants throughout the humanities and sporadically elsewhere that large differences in rating had had *no* consequences (except perhaps smaller declines than worse-rated departments had). Thus the relation of funding to the distribution of IT was felt to be insufficiently elastic by one group and excessively elastic by the other.

Across the disciplines, one might have expected that money comparisons would produce explicit statements about the differing knowledge projects of the humanities versus the sciences. But the comparisons are contentless. The natural scientists—being much richer—are of course silent on this matter. Many humanists and social scientists, however, argue that the humanities are excellent value for money, being nearly costless ("Many of the very best research projects are run on limited funds" remarked one group <ss5>). And social scientists criticize cross-field comparison on methodological grounds. There is, says one group, "no empirical evidence [...] that research IT can be measured on a scale that allows meaningful between-discipline comparisons to be made" <ss4>. Even the economists say present scores "are more or less worthless for cross-subject comparisons" <ss11>. But—most important—there is nothing at all about the differing substance of the values being pursued in the differing disciplines.

Like its distribution in disciplinary space, IT has also a distribution in time, both in the life cycles of individuals and in those of disciplines. Moreover, this distribution has a dynamic aspect. All of these themes are central among the responses, but, like the spatial responses, produced no substantive discussion of knowledge projects.

Nearly all responses agree that the IT of departments changes significantly over time. This is implied by the uniform opposition to the proposal of ratings based on general "track record" and infrastructural strength rather than on specific ratings of actual work. One judicious respondent puts it simply: "historic performance in

previous assessment cycles [...] is not a guarantee of future achievement" <h3>. This theme repeats endlessly. Humanists think historical ratings have "obvious in-built tendencies to intellectual conservatism, the privileging of established fields at the cost of developing ones, and the entrenching of institutional elitism" <h7>. Social scientists think they would lead to "elitism and disciplinary sclerosis at one end, stagnation and decline at the other" <ss28>. Natural scientists think historical ratings would produce "stagnation with no motivation for improvement" <ns12> or again that they would lead to "complacency on the part of higher-rated departments and a lack of support for developing fields [...] " <ns17>.

Only a few see problems with this insistent focus on recent ITQ. One group wrote: "It is unfair for the distribution of research funding to be overly influenced by past performance; on the other hand, top-rated departments must secure appropriate reward for IT" <h26>. On the other hand, another said ratings "need to reflect the current situation, or at least the recent past: [...] Otherwise there is no incentive to improve" <ns8>. This widespread argument that "otherwise there is no incentive to improve" implies that IT is driven partly or even largely by external incentives. Indeed, nowhere in this data is a statement of the form "We aim at great work whether you fund it or not, but it will be easier if you fund it". To be sure, the whole rhetoric of the RAE militated against such a statement, but its complete absence is surprising nonetheless.

About the speed of these various changes, however, there is disagreement. A few responses argue that IT changes slowly. "Changes in research strength are, on the whole, slow" <h4>; "Institutions go down as well as up, albeit slowly" <ss29>. Usually, however, slowness is attributed to the monograph-writing process, rather than being derived from an actual theory of slow change in IT. Many social science and humanities responses want longer review periods to adjust to this longer rhythm of production. The historians remark that "some of the greatest and most influential works of historical scholarship have taken years to conceive and research" <h26>. Many responses suggest, explicitly or implicitly, that faster production means lower IT, a danger for younger scholars in particular. One humanist mentions young colleagues who "rush out a series of articles rather than deepen or broaden their work into a first-rate book-length publication" <h14>. A natural science group note that the "pressure to produce results quickly [is] especially felt by younger staff" <ns11>. In the natural sciences a different concern for the long-run emerged in responses that made an implicit distinction between what we might call merely IT research—safe, solid, and predictable—and truly IT research.

"Much of [speculative research] will fail. It can only be assessed on a very long timescale and it cannot be assessed on the simple basis of the number or short-term impact of individual papers [...] Truly outstanding science, [...] is so unpredictable and can take so long to gestate that it is almost impossible to envisage an assessment process that would work" <ns1>.

This distinction between different temporal types of IT perhaps corresponds to those humanistic comments that saw rapid production as an invitation to lesser IT. Yet neither set of comments links this difference in rates to any differing substantial content of IT or to differences in knowledge projects.

But while some responses emphasized the long-term nature of true IT, the more common response insisted on sharp and sudden changes. From the humanities we hear of "the leaps that many UoAs achieved from 3 to 5 as a result of institutional investment" <#20>. From the social sciences, "Longer [intervals] than about 6 years might fail to pick up substantial changes in relative position" <#13>. From the natural sciences: "Individuals, groups and departments can improve or indeed deteriorate markedly over a period of perhaps 5 years" <#2>.

Discussion of change in ratings revealed further temporal themes. That no one wants purely prospective ratings clearly indicates that no one believes (or wants to believe) that IT can be predicted accurately. All the same, the majority of those who discuss the issue want some prospective component and many of those want that component to dominate. This strong orientation to the future ties to the theme of sudden, sharp changes in IT and to the near uniform rejection of historical ratings. Many believe that support for past performance is inherently problematic. On the one hand is self-fulfilling prophecy. "If a panel awards a 5* [...] then the Department will get resources to allow them to perform well in the future [...]" <#4>. On the other hand is the expectation of complacency and consequent decline. These contradictory arguments perhaps arose in a general hostility to disciplinary elites, since either way their IT was suspect (either because the money helped or because it made them complacent and lazy; oddly, members of those elites are among the respondents arguing these very points).

So we see that IT changes, that those changes are sharp and continual, that IT responds strongly to monetary support. Indeed, constancy of IT is to some extent seen as impossible or even dangerous. A few disagree, discerning a "true IT" whose underlying rhythm is slower and independent.

Modality

We have seen so far how IT can be known, what are IT's units, and how IT is distributed in space and time. We turn now to the modality of IT, IT's positive and negative possibilities. On the positive side are things like unsuspected loci for IT, developing and nurturing IT in persons, the potentiality for IT in all scholars, and the expression of untrammelled IT. On the negative side are the corruption of IT and the manifold means of counterfeiting IT.

The notion that IT can be found in unsuspected places appears in some but not many responses, scattered across the three areas. A humanities association speaks of "institutions [...] that have demonstrated research IT and distinctiveness, often with limited investment" <#3>. A social science group tells us "in HSS, where research is cheap, it is very important to recognize what [institutions] are doing to develop research 'against the odds'" <#2>. Even among natural scientists we hear this belief occasionally. In a biochemical society survey only 34% of survey respondents wanted the field to take a core and periphery shape, suggesting a strong belief in the periphery.

A more common theme is the potentiality for IT in individuals and the need to recognize and develop it early. Even in the humanities, we hear, assessors must have "familiarity with the work of the departments they are assessing, including the work of younger scholars" <#4>. Among the social sciences, anthropologists fear that "the importance of young researchers is likely to be overlooked" <#30>. In the natural sciences this issue is a central concern: "The wider research environment [must be] funded on a scale sufficient to ensure that strong people can flourish early in their career, wherever they find themselves" <#24>. This echoes the earlier complaint that the RAE pressures the young towards non-IT-LIKE work, suggesting a surprising vulnerability of those with potential IT.

An occasional undertone in the HSS documents is a belief that in some sense everyone in a discipline can have IT, evident in widespread concerns for fairness. Each institution should be assessed in the same way, says one superb humanities department: "This is the only equitable way of proceeding, but should be sensitive to institutional differences which affect the kind and range of activity" <#8>. Some argue that IT should be measured not in the abstract or absolute, but against the actual aims of a department or group: "Units should be judged on the basis of their own statements of priority research themes" <#13>. All of these comments breathe a sentiment of IT as a potential in everyone. IT is a matter of achieving what you set out to do. IT therefore lies within the grasp of every group in the system.

A final positive possibility is research unlimited by any constraints—external or self-imposed, financial or intellectual. Throughout, this is called "blue skies research", a term that did not appear in the invitation document, but that was nonetheless widely used. Both the term itself and its usage identified blue skies research as the ideal of IT. In the humanities, the term "blue-skies" was often made equivalent to "long-term research". "We have had to become responsive and reactive to invitations from editors and publishers rather than 'indulge' in 'blue skies' projects [...] the outcome of which will not be assessable for some time" <#20>. Indeed, in a way blue skies stand for disciplinary dominance, as in "The capacity for blue skies research and respect for the primacy of academic values must be maintained" <#2>. Blue-sky could also mean novel, as in "breadth of representation is important to ensure that blue-sky research or newly developing areas can also be assessed on equal par with more established topic areas within the discipline" <#9>. One group simply listed the ideal qualities of research: "Novel, entirely blue-skies, flexible, long-term, freely-defined, curiosity-driven research [...] does not fit neatly into this sort of process" <#1>.

These then are the positive possibilities of IT. But there is equally a negative side. IT can be corrupted. IT can be counterfeited. Moreover, the agents of these things are the colleagues hitherto so positively presented as the only legitimate assessors of IT, the researchers who are all capable of IT, the scholars who will take the long view in order to be IT, and so on.

To begin with corruption. To the ideal of blue skies is opposed the equally widespread idea that IT is driven by incentives and indeed by money. To the ideal of a discipline that controls and dominates the direction of knowledge is opposed the discipline that runs after funding. "Special interest groups [could] get together

and lobby effectively to steer money in the direction of their own pet projects thus starving other potentially IT/LIKE researchers who have been less quick to see how to play the game" <MS9>. The system "will lead to researchers taking 'safe' rather than innovative paths in research" <MS3>. "[I] would encourage number-chasing at the expense of genuine value, and originality would be the casualty" <MS4>. Another response spoke of avoiding metrics which "could be seen to follow what was in vogue and for which it may be easier to command grant income" <MS18>.

These statements are not a vote of confidence in colleagues. Quite the contrary, they portray an IT subject to the whims of funders and fashions, driven by modes of assessment that coerce colleagues too weak to resist them. This despair is only worse when we look at the topic of gamesmanship itself. There are so many comments about gamesmanship—an issue which was explicitly raised in the letter of invitation—that one cannot begin to quote them, only to sketch the main themes. The underlying conclusion is that nearly everyone believes that IT can be faked and fairly easily faked. One can pretend that second-rate research-active staff are not really research-active and leave them out of the exercise altogether. One can lump weak staff into a hodge-podge in one unit of assessment in order to look better in the others. One can hire excellent faculty in anticipation of the RAE. One can create what one response called "spurious research clusters" <SS12>. Some responses feared "citation cartels" of various sorts.

Gamesplaying was discussed in nearly all responses. Generally it was seen to be the strategy of chairs, deans, and even vice-chancellors, rather than of individual professionals. It was also seen to be largely opaque: If we use metrics, then there will be gamesplaying, and so on. But it was a pervasive fear. Although one's colleagues might not actually play the game themselves, they would allow it on their behalf. Although the IT of individuals might be relatively pure, that of groups was easily counterfeited. Given the pervasive idealism about other aspects of knowledge, this cynicism about colleagues is doubly surprising.

Content

Our final and central issue is the content of IT. The question "what is IT in research" was posed directly in the letter of invitation, and most responses addressed it.

Most humanities responses define IT as whatever the current experts or panels say it is, "the best work of the current generation", in the words of one response. Many responses explicitly underscore the diversity of definitions across disciplines, insist on locally-defined criteria, and refer questions of impact largely to impact on disciplinary knowledge or practice. Several also spoke of the evolution and dynamism of definitions, one noting that "traditional publication is out of date by the time it is printed". This points to the second major humanistic criterion for IT, voiced variously as creativity, originality, or innovation, but often as simple novelty. A few responses mention a third, more general criterion: "advances our knowledge and understanding" or "makes a substantial contribution".

In summary, IT work in humanities is work that the discipline thinks is new and that perhaps advances some knowledge project, presumably disciplinary. The definition is thus structural and progressive. It is purely formal, having no substantive content whatsoever: nothing about beauty, justice, or other values, nothing about complexity or subtlety of interpretation, nothing about understandings of human existence, the human condition, the human project. IT work is simply whatever the discipline thinks is new and advances the discipline's (unstated) knowledge project. The social sciences evince much the same pattern. Again disciplinary definition and newness are central attributes of IT. To these is added one new theme, the theme of application in the real world, which among social scientists is mentioned as often as disciplinary and novelty. And advancement of knowledge is more often mentioned.

Among the social scientists, one response does set a substantive criterion for IT, although rejecting it at once:

"One useful distinction is between formal and substantive criteria for research IT. Formal criteria (that good research should present an accurate picture of reality and help to develop theory in order to understand that reality) are unlikely to generate controversy" <SS17>.

Here at last is a substantive statement; social research aims to understand social reality (note, however, that the comment has the labels of form and substance backwards). But the response then argues that the only audience for social research that is itself stable enough to sustain consistent IT criteria over time is the discipline. Thus at any given time, the disciplines must set the actual criteria for IT: "IT in contribution to future research is thus suggested as a primary criterion of research IT that is most likely to be 'future-proof'". The respondent thus circles back to IT being simply whatever the discipline says it is.

In the natural sciences, this picture repeats with slightly different emphases. Novelty continues to be central, but applicability in practice now equals it as does advancement of knowledge. Discipline-definition recedes somewhat, but the advancement achieved is clearly understood as disciplinary advancement.

There is, however, another new aspect of the natural science responses. Two entries specifically invoke the criterion of fertility in the production of future disciplinary knowledge. One entry remarks:

"IT is probably only something which is recognized long after a work is published; and almost certainly not something done especially with a 'research assessment exercise' in mind, since in many areas this simply produces a spate of overinflated and under-prepared publications" <MS4>.

Several responses echo this distinction between hugely influential, paradigm-shifting work (and the high risk of attempting it) and routine, everyday science. The responses clearly believe that the former is extremely rare.

Thus across the three areas, the constant theme is that excellence—let us use the word itself at last—is defined by a given discipline at any moment, and that, in particular, it is whatever that discipline will at that time find new and innovative. In some sense, innovation is thought to advance the larger project of the discipline, but there is no mention, with the one exception given, of what this project actually

is. In the social and natural sciences, a further criterion of excellence seems to be applicability in the real world. If we reread the humanists' responses with that theme in mind, we will find a few similar claims; for example, "Less quantifiable but no less significant is history's contribution to the quality of the nation's life and the culture of the nation's citizens in general"⁴ (427).

Theorizing Knowledge Projects

In short, these data taken together tell us absolutely nothing about our original problem. We had hoped to uncover the difference between ideal conceptions of knowledge in the two great families of knowledge—the humanities and their social scientific allies on the one side, and the natural sciences and their social scientific allies on the other. But there are no clear differences. There are, to be sure, some indications of difference. The humanities are more likely to think that excellence is decentralized than are the natural sciences. The humanities also customarily think change in excellence is slower than do the natural sciences. But all the same, the humanities recognize fast change and the sciences retain their faith in the slow pace of major discovery. The natural sciences, finally, are more oriented to advancement of a corpus and to applicability than are the humanities.

But while these differences are interesting, they are both few and unsurprising. What is more surprising is the overwhelming similarity. Most qualities of excellence are shared across the two great continents of academia: from how it is known, to who has it, to where it is, to its continual change, to its transience and potential, to its tendency to corruption, and finally to its identity with the new and the creative.

We underscore two particularly important themes within this broad similarity. First, as we just saw, nowhere in these comments is there anything about the substantive content of excellent work nor, indeed, of work at all. Aside from the brief remark of one social scientist, the projects of these disciplines are all unspecified, even unnamed. Although a broad difference in the knowledge projects of the humanities and the natural sciences remains an intriguing possibility, there is no evidence about those projects in this data, and a distinct suggestion that there is no difference, since the conception of excellences is purely formal at this point in history.

The second particularly noteworthy similarity is the pervasive normative dualism. On the one hand things are good. One's colleagues are those who know excellence. Excellence is protean and tolerance-governed. The potential for excellence pervades the future, the young people, and the research periphery. This is a very hopeful picture. But on the other hand, excellence can be sharply discontinuous and falsely concentrated. One's most noted colleagues are probably resting on their laurels, their current achievements the result of self-fulfilling prophecies. Excellence can be corrupted in a dozen ways, and indeed one's colleagues are at best willful beneficiaries of their administrators' manipulations and at worst game-players themselves.

The broad similarity of responses across the disciplines has three possible sources. While this data cannot adjudicate between them, they do however suggest that one of those theories is more plausible than the two others. These theories involve first the RAE context, second the larger social structure of academic knowledge, and third the cultural structure of communities of scholars.

A first reason for this similarity would be the RAE framework itself. Perhaps that framework forces all respondents into the same stances, even in a data-elicitation setting where we might expect disciplinary peculiarities to figure prominently. The strong and consistent negative themes in this data—above all, the suspicion of peers—probably have roots in the RAE itself, which was a divide-and-rule project aiming to induce academia's cooperation in cutting its own throat. This is exactly the mechanism Michael Burawoy (1979) analyzed in *Manufacturing Consent*, in which higher management secures surplus profit via rules which harness inter-worker rivalries to management's advantage. It is therefore hardly surprising that there should be a kind of tacit collusion against the government 'boss'.

A second possible explanation traces the similarity of responses to disciplinary social structure. In a demographically mature academic employment system, advancing careers require interinstitution mobility, which in turn requires currency convertible across institutions. Only research (rather than teaching or local service) can provide such convertibility, and the desire to manufacture research currency produces a research inflation analogous to any other form of credential inflation. Moreover, the rapid exhaustion of many over-researched fields promotes relabeling, theory-churning, and rediscovery of the wheel in an attempt to produce artificial novelty.⁶ On this argument, the similarity of employment and career structure across the two major areas of academia drives a similarity in incentives and behavior patterns that in turn creates the broadly similar research habits uncovered by the responses here analyzed.

These two arguments can account for much of the broad similarity of the responses across the disciplines. In particular, these arguments seem to explain the negative side of the normative dualism noted earlier. What they don't explain is the positive side of the dualism and the prevalence of unnamed knowledge projects. A third theory would address those findings more directly. It looks not to the social structure of academic communities, but their cultural structure. Indeed, it focuses on the normative commitments of scholars, on patterns of values rather than of social structure.

There are many hints suggesting such a theory. There was for example the considerable idealism uncovered by the analysis of modalities of excellence. Disciplines orient to the future. They worry about transiency and the perpetual need for renovation. They believe in the possibility of excellence among those of low status—the peripheral, the young, the unexpected. And they all use the romantic metaphor of 'blue skies'.

Following these hints, a cultural theory of disciplinary similarity begins with the recognition that the disciplines evolve in a world that is simultaneously one of

⁶For an extended analysis, see: Abbott (2001).

scarcity and of abundance. Because our social theories have roots in political economy, the former of these is much more familiar to us. We are used to thinking about scarcity, both in social structural terms (PhDs competing for scarce jobs, interdisciplinary competition for funds and talent, division of labor) and in cultural terms (exhaustion of research areas, declining productivity of paradigms, etc.). All of these things we understand under the sign of scarcity.

Yet abundance may be more important to the life of disciplines than is scarcity. Specialization arises more from the overwhelming complexity and extent of knowledge than from open competition over scarce topics. The glut of knowledge on the internet echoes an earlier glut that has existed since the 1960s if not the 1920s (Abbott 2011). If we consider the various oppositions that pervade the social sciences—positivism versus interpretation, narrative versus analysis, and so on—it is clear that these seven or eight distinctions can be crossed to generate dozens if not hundreds of possible “disciplinary positions”, only a handful of which have actually been tried as disciplines (Abbott 2004). Not only is our knowledge of the world immensely diverse, we have also only begun to explore the different ways of producing such knowledge.

When we social actors are confronted with abundance, we employ a number of strategies. The simplest is to abandon ourselves to randomness—the strategy of many surfers of cable TV and the internet, as it is of that school of Bible readers who “let God choose” what they are going to read. Another strategy is to subdue abundance through organization, subdivision, indexing, and the other strategies of mapping and abstraction. These enable us to ignore certain areas altogether: “I’m not good at mathematics”, “I don’t like to travel in countries where I don’t speak the language”, etc. Along with this selective ignorance goes specialized knowledge. Organization permits us to know much about some areas even as we know little about others, except as they are fractally related to areas we know well (Abbott 2001).⁷

In between these strategies of selective ignorance and specialized knowledge is that of canon—the somewhat arbitrary selection of certain particular things and people as emblems, representatives, or paradigms. Hence Max Weber dominates our social theory courses even though many of the ideas now attributed to him were borrowed from others. Weber-as-canon cancels the obligation to read each of the dozens of slightly different but equally distinguished theories of society. Yet another control strategy is the screening of abundance for quality or authenticity or righteousness or some other characteristic that justifies selective ignorance within a particular organized field. So we decide that only books published by this or that press are worth reading. Or only women can write gender theory. Or only quantitative analysis can properly identify the impact of social class on education. Although these statements all have their theoretical justifications, it may well be that their main function is actually to reduce the amount of work to be read and comprehended.

⁷Note that this mechanism could explain division of labor as well as does the scarcity/competition argument favored by Durkheim and many others.

All these various strategies turn abundance into scarcity, which we understand much better. For if only certain presses are worth reading, we can have a competition for who gets to publish with them. If only quantitative analysis is worth doing, we can have a competition to find the best form of quantitative analysis. If each discipline knows only certain things, then we can have a competition between them for resources, students, and so on. We find all of this not only more comprehensible, but also right and proper—things as they should be.

In summary, the disciplines may well originate in an attempt to constitute coherent knowledge communities in the enormous chaos of things to be known. Far from being a division of labor forced by intensive competition in a crowded turf, they are in fact like small communities of whales in the immense ocean, banding together for companionship. To that end, they select canons of common work. By agreeing to restrict “classical sociological theory” to Marx, Durkheim, and Weber, sociologists create a common set of language, allusions, problems, and methods. Similarly they organize themselves into subdisciplines, ranked departments, elite publishers and journals, and so on, all in order to recreate a more familiar world of scarcity amid the frightening abysses of abundance.

Relations between the disciplines so constituted are maintained by “interdisciplinary”. Those who have practiced interdisciplinary know that disciplines comprehend each other only with great difficulty, for the process of canon-formation and structuration through which disciplines create intellectual coherence leads to quite idiosyncratic habits. Like all clan systems, therefore, the academic system has kinship rules. These are the inverse of normal kinship rules in modern societies. The full children of interdisciplinary collaboration are outcasts from the core system and usually become parts of the hybrid, problem-driven world of application. (Occasionally, they are part of new discipline formation.) By contrast, the products of what we might call the morgantic form of academic marriage—the extensive borrowing of a set of ideas, canonical works, methods, etc., from a discipline by a member of another discipline—become legitimate descendants of the borrower’s discipline, which thus does not have to breach disciplinary boundaries.⁸ These rules have kept the disciplines continuously in existence over a long period as relatively independent lineages but at the same time prevent their ever completely losing contact with one another.

The creation of a reassuring scarcity via disciplinization (a cultural structure in the ecology of possible knowledge) in turn produces the social structures of internal status hierarchies that drive the careerism noted throughout this data. These hierarchies are dominated by “nearness to the center” of the discipline. This center is indicated by the (arbitrary) canon of works, methods, theories, and so on. In any given short run, this canon is reified into a fixed hierarchy that enacts the scarcity system that provides the incentives and directions that push careerism.

⁸On the argument made here, boundaries are not so much about conflict with competitors as they are protections against the leakage of scholarship into the endless chaos of “all possible knowledge”. This is precisely the reverse mechanism of that posited in Gieryn’s celebrated article (Gieryn 1983).

But participants know this fixed intellectual structure to be a reification and that the inevitable biographical succession of leaders will with equal inevitability change it. Indeed, the actual 'center' of the discipline is the set of rules deployed in making these changes. These process rules define what makes work 'interesting', 'excellent', or whatever we wish to call it. They may include universal rules. In the sciences, for example, excellent work must subsume most earlier knowledge, as Whewell argued long ago. The process rules may also include "reflecting barriers". In literary studies, the cry against "excessive scientism" has emerged many times in the last century and a half, always pushing the discipline back towards literary appreciation (Graff 1987). But they also include quite local, particular desiderata. That is, there is not any one excellence for all times and all places, even in a single discipline. Excellence must always be judged dynamically and locally, in relation to the particular moment and particular intellectual setting from which it aims to move.

But this means that excellence has no permanent qualities beyond novelty. It cannot be given a particular name, any more than in universal religions God can be given a particular name without losing the limitless, infinite, transcendent quality that is God-ness. So it is little surprising that scholars commenting on the RAE would not mention, even in the most cursory way, the underlying projects of their disciplines, even though the comments were a place for bold claims and grand gestures. To give a name to a discipline's project is to set limits on it, to reduce its infinite possibility. It is also to concede to others the right to define the discipline's future direction, when in fact that direction is made in the performance of research. Put another way, "discipline as a cultural system" is what we get when a group of scholars loosely agree on a general way to move through the universe of things to know. They teach their recruits how to decide, from any given particular position in that universe, what are excellent next directions in which to proceed. With such rules, the recruits will have guidance even when the original leaders are gone.

In this argument, 'naming' arises because at the core of each discipline is a process rule, not a picture of some particular ideal. Such a picture would lead to dangerous reification. It would substitute for the boundless future of the discipline a fixed core of worship. As countless religious and revival movements show, this would be fatal to survival. At the same time, we know from the theory of Gibbs sampling that it is quite possible to construct algorithms that are perpetually iterative, but that tend to circle around a given area. That is, the fact that the disciplines invent their futures in part by locally-derived process rules does not in any way condemn those disciplines to purely random wandering. But it does mean that the rules for novelty and excellence at one point in disciplinary time and space are not likely to be the same (except in the most purely formal sense) as those of another. Hence there is no way to specify them once for all, either for the RAE or for anyone else.

This then is the heart of tacit knowledge (Polanyi 1958). It is the knowledge of how to move forward from where we are right now, of how to combine long-run aims and short-run paradigms to drive a judgment of local novelty.

My paper therefore must end on the brink of a new theory of knowledge institutions. We sought the different projects of the humanities and the sciences. Yet we found they thought about their projects quite similarly. On theoretical reflection, this

seems to be caused not by an actual similarity, but by the fact that the cultural unity of the lineages that are disciplines arises in a set of continuous recipes for generating novelty. These bring together long-run phenomena like canons and methods and short-run phenomena like research paradigms and trendy topics. If these recipes have some enduring qualities, they do not have any particular identity with one particular end point. It is no surprise then that the RAE did not find IT.

References

- Abbott, A. 2001. *Chaos of disciplines*. Chicago: University of Chicago Press.
- Abbott, A. 2004. *Methods of discovery: heuristics for the social sciences*. New York: W. W. Norton & Company.
- Abbott, A. 2008. The traditional future: a computational theory of library research. *College & Research Libraries* 69: 524-545.
- Abbott, A. 2011. Library research infrastructure for humanistic and social scientific scholarship in America in the twentieth century. In *Social knowledge in the making*, eds. C. Carnic, N. Gross, and M. Lamont. Chicago: University of Chicago Press.
- Burawoy, M. 1979. *Manufacturing consent: changes in the labor process under monopoly capitalism*. Chicago: University of Chicago Press.
- Gierzyn, T. F. 1983. Boundary-work and the demarcation of science from non-science: strains and interests in professional ideologies of scientists. *American Sociological Review* 48: 781-795.
- Graff, G. 1987. *Professing literature*. Chicago: University of Chicago Press.
- Polanyi, M. 1958. *Personal knowledge: towards a post-critical philosophy*. Chicago: University of Chicago Press.