

Policy Analysis Research in Higher Education: *Negotiating Dilemmas*

Análisis de políticas de investigación en la educación superior: negociar dilemas
Analyse de politiques de la recherche en éducation supérieure: négociant les dilemmes
Análise de políticas de pesquisa na educação superior: negociando dilemas

Received: JUNE 23, 2014 / Accepted: OCTOBER 4, 2014 / Available on line: DECEMBER 30, 2014

Find this article in <http://magisinvestigacioneducacion.javeriana.edu.co/>

doi: 10.11144/Javeriana.M7-14.PARH

Written by RONALD BARNETT
UNIVERSITY COLLEGE LONDON
LONDON, UNITED KINGDOM
r.barnett@ioe.ac.uk

GARETH PARRY
UNIVERSITY OF SHEFFIELD
SOUTH YORKSHIRE, UNITED KINGDOM
g.w.parry@sheffield.ac.uk

Abstract

Increasingly, qualitative research is funded by agencies — whether government or national agencies or in the private sector— that have a direct interest in the research that they are funding. Especially in qualitative research in such situations, methodological issues arise but so too do dilemmas, both in the actual conduct of the study and in the writing and the positioning of ensuing texts (for example, in relation to value neutrality and value commitment, in doing justice to the multiple and conflicting interests of various constituencies and in steering among competing ideologies). Here, in this paper, such dilemmas are brought out in an account of a study conducted in the UK in the 1990s, to review and to evaluate the UK's then quality assurance system.

Key words plus

Quality control, methodology, policy making, interest groups.

Transference to practice

In analysing and reviewing the UK's quality assessment system in the 1990s, the research also offered a set of (44) recommendations as to the ways in which the national quality system could be taken forward. The report arising from the study also had to gain the support of the higher education national funding body so as to secure its publication. That latter goal was secured, the report being published and so contributing to the national debate on quality assessment, with a national conference being held focusing upon the report itself. Many of the report's key recommendations were adopted. However, its central recommendation that there be established a "quasi-profile" of the quality of work in institutions' departments was not, thereby leading to the adoption of a numerical system, a national system that was ultimately to be discredited and which in turn gave way to yet further stages in the evolution of quality assessment in England.

To cite this article / Para citar este artículo / Pour citer cet article / Para citar este artigo

Barnett, R., & Parry, G. (2014). Policy Analysis Research in Higher Education: Negotiating Dilemmas. *magis, Revista Internacional de Investigación en Educación*, 7(14), 69-84.

Palabras clave descriptor

Control de calidad, metodología, formulación de políticas, grupos de interés.

Resumen

La investigación cualitativa está siendo cada vez más financiada por agencias —ya sean gubernamentales o del sector privado— que tienen un interés directo en la investigación que están financiando. En este tipo de situaciones, sobre todo en la investigación cualitativa, se presentan problemas metodológicos, al igual que dilemas tanto en la conducción del estudio, de la escritura y del posicionamiento de los textos posteriores (por ejemplo, en relación con la neutralidad y el compromiso de valor, al hacer justicia a los múltiples y conflictivos intereses de las diversas partes interesadas y en la dirección entre ideologías rivales). En este trabajo, estos dilemas se presentan a partir de un recuento de un estudio realizado en el Reino Unido en la década del noventa, para evaluar el sistema de calidad de aquel entonces.

Transferencia a la práctica

Tras el análisis y la revisión del sistema de evaluación de la calidad del Reino Unido en la década del noventa, la investigación también ofrece un conjunto de 44 recomendaciones sobre los modos como el sistema nacional de calidad podría seguir avanzando. El informe derivado del estudio tuvo que conseguir el apoyo del organismo nacional de financiación de educación superior, con el fin de asegurar su publicación y contribuir al debate nacional sobre la evaluación de la calidad; fue presentado en una conferencia nacional centrada en el mismo documento. Muchas de las principales recomendaciones del informe fueron adoptadas; sin embargo, su recomendación principal en cuanto a establecer un “cuasi-perfil” de la calidad del trabajo en los departamentos de las instituciones no estaba, lo que condujo a la adopción de un sistema numérico nacional que, en última instancia, iba a ser desacreditado. Esto, a su vez, dio lugar a otras etapas en la evolución de la evaluación de la calidad en Inglaterra.

Mots clés descripteur

Control de la calidad, méthodologie, formulation de politiques, groupes d'intérêt.

Résumé

A chaque fois la recherche qualitative est en train d'être financée par les agences —qu'elles soient les agences du gouvernement, nationales ou du secteur privé— qui ont un intérêt direct dans la recherche qui sont en train de financer. Dans ce type de situations, notamment dans la recherche qualitative, on rencontre des problèmes méthodologiques, de la même manière que quelques dilemmes, tantôt dans la conduction de l'étude, l'écriture et la situation des textes postérieurs (par exemple, par rapport à la neutralité évaluative et l'engagement d'estimation, lorsque l'on fait justice aux multiples et problématiques intérêts des divers parties qui sont intéressées que dans la direction entre les idéologies rivales). Dans ce travail ces dilemmes sont présentés à partir d'un recensement d'une étude réalisée dans le Royaume-Uni de ce temps.

Transfert à la pratique

Dans l'analyse et la révision du système d'évaluation de la qualité du Royaume-Uni dans la décennie de 1990, la recherche offre aussi un ensemble de 44 recommandations par rapport aux modes dont le système national de qualité pourra continuer en avançant. Le bilan qui provient de l'étude a eu le besoin de trouver le soutien de l'organisme national de financement d'éducation supérieure à fin d'assurer sa publication. Ce dernier objectif a été acquis, le bilan fut publié pour contribuer au débat national sur l'évaluation de qualité, et il fut présenté dans une conférence nationale centrée dans le même bilan. Beaucoup des recommandations du bilan ont été adoptées. Néanmoins, la recommandation principale par rapport à établir un quasi-profil de la qualité du travail dans les départements des institutions n'en était pas, cela a conduit à l'adoption d'un système national qu'en dernier lieu pourrait être discrédité et à la fois il a donné d'autres étapes à l'évolution de l'évaluation de la qualité en Angleterre.

Palavras-chave descritor

Controle de qualidade, metodologia, formulação de políticas, grupos de interesse.

Resumo

Cada vez mais a pesquisa qualitativa está sendo financiada por agências —sejam agências governamentais, nacionais ou do setor privado— que têm um interesse direto na pesquisa que estão financiando. Neste tipo de situações, especialmente na pesquisa qualitativa, apresentam-se problemas metodológicos e dilemas, tanto na condução do estudo, como na escrita e no posicionamento dos textos que daí resultam (por exemplo, com relação à neutralidade avaliativa e ao compromisso de valor, ao fazer justiça aos múltiplos e conflitivos interesses das diversas partes interessadas e na direção entre ideologias rivais). Neste trabalho estes dilemas são apresentados a partir de um estudo realizado no Reino Unido na década de 90, para revisar e avaliar seu sistema de qualidade daquela ocasião.

Transferência à prática

Na análise e revisão do sistema de avaliação da qualidade do Reino Unido na década de 90, a pesquisa também oferece um conjunto de (44) recomendações quanto aos modos em que o sistema nacional de qualidade poderia continuar avançando. O relatório derivado do estudo também teve que conseguir o apoio do organismo nacional de financiamento de educação superior a fim de assegurar sua publicação. Este último objetivo se conseguiu, o relatório foi publicado para contribuir ao debate nacional sobre a avaliação da qualidade, e foi apresentado em uma conferência nacional centrada no mesmo relatório. Muitas das principais recomendações do relatório foram adotadas. No entanto, sua recomendação principal quanto a que se estabelecesse um “cuasi-perfil” da qualidade do trabalho nos departamentos das instituições não estava, o que conduziu à adoção de um sistema numérico, um sistema nacional que em última instância ia ser desacreditado e que por sua vez deu lugar a outras etapas na evolução da avaliação da qualidade na Inglaterra.

Introduction

Essentially, there are two dominant strands in the literature on what might crudely be termed policy analysis methodology: technical expositions of favoured methodologies (Richie & Spencer, 1994; Sadoulet & De Janvry, 1995; Yanow, 2000; Srivastava & Thomson, 2009) and guides to writing reports arising from policy analyses that treat such writing as a set of technical accomplishments. There appear to us to be two gaps, one being that of explorations of challenges that may arise in conducting fieldwork in this domain (policy analysis) and the other being that of challenges and even dilemmas that may arise in the writing of reports that ensue from such research, especially where such research has a controversial character.

Such challenges are likely to be especially apparent in situations in which the policy (and possibly associated practices) in question are freighted with politics, ideologies, large interests and so forth. In such circumstances, those with responsibilities for the policies and/or practices being explored may be reluctant to be fully open with researchers. Indeed, they may try to thwart the researchers' best endeavours at various points in the research process. It is challenges such as these that we want to explore in this paper, and in doing so, we shall visit a study in which the two of us were involved some two decades ago.

The study was that of an evaluation of the national quality assessment process in higher education in England and Wales conducted in the 1990s. The environment within which the study was undertaken was explicitly ideological, with different parties having heightened and conflicting expectations of our research. Accordingly, this study gave rise to some profound dilemmas, especially in the conduct of the research and in the writing of the report. This paper describes and analyses those dilemmas, as well as setting the research in the context of the then existing national quality assessment environment and it is the latter with which we begin our paper.

The study in question

The study in question was that of a review and evaluation of the national system of quality evaluation of higher education in force in England and Wales in the early 1990s. At that time, quality evaluation of teaching in England and Wales was the responsibility of the two national funding councils, respectively the Higher Education Funding Council for England (HEFCE) and the Higher Education Funding Council for Wales (HEFCW). The two national funding bodies had been established in 1993, following a 1992 Act of Parliament in the UK, and had been specifically given responsibility for assessing the quality of higher education (in England and Wales) by the then Government. It had been expected that the Funding Councils for England and Wales would carry out a review of their quality system and this review, therefore, was commissioned by the two Councils¹.

The HEFCE began its work in assessing the quality of higher education quite quickly, instituting in October and November 1992 "test assessments" in business and management, and law. Following those test assessments, in 1993, the HEFCE published a document setting out the essence of its intended quality evaluation method. The HEFCE and the

¹ An entirely separate set of arrangements existed in Scotland, under the purview of the Scottish Funding Council for Higher Education. Those arrangements form no part of the evaluation being explored here.

Article description | Descripción del artículo | Description de l'article | Artigo descrição

This paper of reflection is based on research: "Assessment of the Quality of Higher Education: a review and an Evaluation". The paper reflects on a piece of national policy-oriented research, identifying qualitative dilemmas that arose therein, both in the conduct of the research and in the writing of the project report, commissioned by the national funding body and attracting much public controversy. The research was funded by the Higher Education Funding Councils for England and for Wales.

HEFCW adopted somewhat different methodologies, a not unimportant difference. However, for the most part, the differences will not concern us in this paper unless they are germane to the matters that we wish to discuss.

There were a number of key elements to the national quality assessment methodology (in those early 1990s) and it would be helpful quickly to enumerate them. In England, every institution of higher education (including therefore some world-famous universities) was required to undertake a self-assessment, discipline-by-discipline. There were three categories of judgement —excellent, satisfactory and unsatisfactory— and in their self-assessments, English institutions were invited to make a claim for “excellence” (against which the self-assessment was then evaluated). All self-assessments were subject to a “template exercise” by officers at HEFCE, on the basis of which an assessment visit might or might not be undertaken. At that time, assessment visits were made just to 58% of English institutions, whereas all institutions in Wales were subject to an assessment visit. Visiting parties were made up of assessors who were drawn quite widely (beyond university disciplines) and were led by “contract researchers” who were typically former national inspectors of technical colleges and schools. No attention was paid to the self-assessments as part of the national evaluation process.

A judgement of “unsatisfactory” did not lead to any overt sanction. However, the judgements were public and could find their way into the press, and an unsatisfactory judgement would lead to a repeat visit from the Funding Council. In the background lay an unspoken possibility that funding could be withdrawn, even if in practice such a course of action would have been difficult for the Funding Council to take.

In summary, universities that had been more or less autonomous for all their histories (perhaps several hundreds of years) were having to comply with a national system of evaluation imposed upon them by the state, a system that possessed a number of features that not surprisingly had given rise to misgivings publicly expressed by the institutions. There had previously been a university-run system of quality assurance but this was the first time that there had been a state-backed national system.

Perhaps enough has already been said (by way of bald description of the matter) to intimate some of the sensitivities present in this study. The national system of quality assessment had only recently been set up, it applied to every institution of higher education in the country (including, as stated, world-famous universities), and there were differences of operation in the separate jurisdictions (England, Wales, Scotland). There were some evident oddities in the English system (visiting teams were led by non-academics, decisions to visit were made only on the basis of an office exercise, and visits were made only to a proportion of institutions and self-evidently lacked a degree of systematicity).

Accordingly, this study took place against the background of considerable unease across the higher education sector. There were sensitivities present at the time and they made their presence felt at various points in the study. To put it simply, the Funding Council had been given (by the national government) responsibility for conducting a systematic assessment of the quality of teaching across the whole higher education sector and was intent on fulfilling that obligation. The academic community, on the other hand, had profound misgivings both about there being an external system of quality assessment in general and about several of the actual features of the assessment system as it stood at that point in time.

Consequently, there were opposing expectations of our study. The Funding Council presumably wished the study largely to endorse its general approach while understanding that the report that flowed from the study would make some recommendations for the development of the assessment methodology. The academic community, on the other hand, was looking to the study—and its ensuing report—to be highly critical of the national quality assessment methodology, to endorse the misgivings that were apparent in the public debate, and to make some trenchant proposals for wholesale change in the Funding Council's approach. The study, therefore, was conducted in the midst of a febrile atmosphere, with opposed interests placing high expectations on the study and the report flowing from it.

The team that undertook this study was a group of higher education researchers at the Institute of Education, University of London, who had responded to an open invitation to submit an application to the Funding Council to tender for the project. The competitive process had a number of stages, successively working up the detail of the proposed study and concluding with a (competitive) interview at the Funding Council's offices, the interviewing panel being led by the chair of the Council's Quality Committee. This is a significant matter since the work could, in principle, have been conducted by private sector consultants (who were able to bid for the project). As academic researchers, we were keen to establish our own academic credentials and to conduct the project according to academic norms, rigour and standards. This orientation on our part, on the one hand, imparted a heightened sensitivity to the deep ideological environment governing this project and, on the other hand, led us explicitly to articulate our own approach and value perspective (of which more in a moment).

Considerations and principles

Policy analysis poses researchers and analysts with a number of challenges. Quite apart from the matter of methodology—and we shall say something about our methodology in a moment—there are actually some even more fundamental challenges. These include:

- The definition of the problem: given that the policy in question will have been chosen by an external body, institution or entity—be it government, or a public service or a single major corporation—that responsible body will almost certainly have formed both its sense of the key problem (or problems) and be looking for a practical solution. There may, therefore, be a tension between the researcher's framing of the presenting situation and the key problem and that of the body in question, between—in other words—an intellectual problem and a practical problem to be solved pragmatically.
- The nature of the presenting situation: the presenting situation—for example, in health policy, in transport policy, in energy, or in education policy—may be ideologically infused. Large interests, backed by power, may be present such that a policy analysis may be itself perceived or framed by interested parties in particular ways. There may be no neutral space available to the analyst (Shore & Wright, 2005).
- A moving situation: As just intimated, the presenting situation may be saturated by power or, at least, comprise a space inhabited

by rival discourses and rival value systems, all of which may be influencing the situation even while the research and analytical process is underway. The research object may be changing significantly during the course of the research.

In a situation that contains some or indeed all of these elements—as was the situation in question here—challenges befall the researcher and analyst straightaway. How, ideally, do the researchers want to position their work? What stance do they wish to adopt towards the situation under review? To what degree are researchers intent only on describing a situation and making technical judgements and to what extent might researchers attempt to design and conduct their research to raise critical questions about the policy framework itself and/or so as to have a normative effect on a situation? Talk, for instance, of “value neutrality” is of little help in answer to questions such as these. For the situation in question is super-saturated with values and explicitly contested sets of values at that. The whole of the research process and the analysis of the data—quite apart from the formulation of any recommendations for policy development (if that be appropriate)—will find itself embroiled in values and, as stated, it cannot be presumed that there is available any position of value neutrality, however rigorously and scrupulously the research and the analysis is conducted.

Just such a thicket of values, expectations, discourses and ideologies formed the situation in which this study was placed. We were, indeed, faced here with our first dilemma, even before the fieldwork was underway: we wished for our own study to be seen by the interested parties as possessing its own integrity, but those interested parties were bringing to bear different and conflicting values in their perceptions of our study (as to what count as “integrity”). A position of value neutrality was not available to us empirically, even if it had been theoretically available. Even more, it was not obvious that there was room for the research universally to be seen as possessing integrity, for it was not obvious that it would or could be seen as fulfilling the conflicting agendas of the parties concerned.

Accordingly, we decided not on attempting a position of value-neutrality but rather the reverse, that of stating a set of principles that, for us, guided our study, including the formulation of recommendations with which we concluded our report. And we did so in a somewhat oblique way, by identifying a set of principles that we felt were “appropriate to the administration of a public service such as higher education” (as we put it in our report). At the outset of our research, we identified five such principles, namely:

1. Partnership: that the responsible body seeks to carry its associated institutions with it in the formulation of its policies.
2. Transparency: that its operations are transparent, for example, in the clarity of the assessment criteria in practice and the link between the evidence obtained and the judgements reached.
3. Justice: that institutions are enabled to present themselves adequately, that judgements are made on all the available evidence and that like cases receive similar treatment.
4. Appropriateness: that the operations and approach [of the body being analysed] are appropriate to the character of the service in question (here, higher education).
5. Quality and rigour of process itself: that the processes adopted are likely to achieve their purposes and that their operation is relatively robust.

These propositions, determined by the research team, served as a set of value horizons against which the team conducted its work. Worth noting, perhaps, is that these value propositions were formulated in the context of the research project itself. They reflected the value orientation of the research team but they were intended also to have a normative aspect, both guiding our work and also, thereby, serving as horizons to which any impact that our research might have would strain towards. We did not believe that value neutrality was available to us in this situation —our research was going in any case to be positioned by opposing sets of values— and so we took a deliberative value-stance, as it might be termed, in explicitly determining and setting out our own value position (Fischer & Forester, 1987).

Methodology and research design

At one level, the key problem that our research was intended to address was quite straightforward: to what extent were the national arrangements for quality assessment in higher education (in England and in Wales) meeting their objectives? On the basis of that analysis, the research team was also expected to make recommendations for the further development of the quality assessment system. However, as intimated, the study was conducted within an ideological environment, in which different interest groups had their own expectations of the study, even before the study commenced its work.

The envelope, therefore, within which this study was conducted, contained a number of domains to which the study had to be sensitive, and here, we follow Greene (2006):

- i) A methodological domain: this domain took on a heightened intensity, even beyond that of policy analyses more generally, for the study was focused on the academic community, across all the disciplines, and the academic community across the nation had a particular interest in this study. The “rigour” of the study was likely to be scrutinised very carefully, therefore, and across all disciplines each with its own methodological proclivities, and with both academic and practical interests in play;
- ii) A philosophical domain: this domain was necessarily even if subtly in play. Just what was the entity on which the study was focused? Was it the overt entity, that of a national quality assessment system or was it more the relationships between the state and the academic community, at a time when —in

higher education and more broadly still— “the evaluative state” was emerging? In other words, there was an ontological matter that lay at the heart of this study.

- iii) A practical domain: the practical domain was present from the inception of the project, and in several ways. The commissioning agent, the national funding body for higher education, imposed a tight deadline within which the study was to take place; key actors were distributed right across the nation and needed to be contacted; interactions between the commissioning agent and the research team had to be negotiated with care throughout the project (of which more later); and a research team needed to be assembled and find a way of working collaboratively together within an intense setting.
- iv) A socio-political domain: as indicated, this domain was also vividly present throughout the project, with heightened expectations being placed on the study from different and even antagonistic camps. Here, in fact, the term “political” could be understood as an upper case “Political” , for the study attracted the attention of the national political sphere (again, we shall come back to this point later in this paper).

Within such a complex environment, the adoption of a “mixed methods” methodology on our part was a natural course of action. There were, in any event, a number of subsidiary (and technical) questions that the research was expected to address (concerning the details of the national assessment methodology) that would have probably led us in this direction but our point here is that the presence of multiple national political and academic interests (and antagonistic interests at that) was such that it was politic on our part to adopt a set of mixed methods so as to minimise the difficulties that the study was bound to face in its exposure (right from the start) to the external world. A mixed methods approach had political value —quite apart from its methodological value— in heading off possible critiques to the effect that this research could be disregarded, with an alleged deficiency in its research methods.

The data collection techniques that we adopted were as follows:

- Documentary evidence: this was voluminous, and included sample assessment reports, committee papers from the Quality Committee of the Funding Council, formal documents issued by the Funding Council,

key documents in the national domain, and internal working documents;

- Written “evidence” submitted voluntarily from over twenty institutions and from some individuals;
- Questionnaires, designed for and administered separately to institutions and to the national assessors;
- Interviews with members of the national quality committee of the Funding Council (N=7) and officials at the English and Welsh Funding Councils (N=8) (including the then Chairs of the Funding Council itself and of its Quality Committee), national assessors (N=17) and key individuals in institutions (N=9);
- Observations of assessment visits to two institutions (in History at a Welsh university and in Chemistry at an English university) conducted by the Funding Council as part of its quality assessment process and of a number of other events (each with its own ideological colouring).
- Participative events, namely regional conferences organized by the review team and a de-briefing meeting for assessors.

It will be noticed that this array of research methods includes participative events in which the research team sought to interact with key groups — especially the then existing national assessors. This strand of the methodology precisely reflects the ontological question posed above (just what was lying at the heart of this study?) and the socio-political element of the research envelope with which the team was presented: there were considerable sensitivities towards this research among the national assessors, who were concerned that the research findings and recommendations might have considerable implications for their role (again, we shall return to this matter shortly).

It will be evident that here lies a second dilemma for such a research team. In conducting a form of policy analysis in a heightened ideological setting, to what extent should a research team engage directly with relevant constituencies, so as to prepare the way for the outcomes that might ensue? Here, we judged it politic somewhat to step outside (beyond, even) our roles as researchers and, anticipatorily, engage with a particular constituency that we felt could be hostile to our findings and were in a position to influence the reception of our report.

The design of the study was chosen by the research team and presented to the interview panel commissioning the work as part of the application

to undertake the study. On our side, we set out a multi-method design and approach that defined the work as “research” rather than a technical evaluation. In so doing, we sought to emphasise our independence (as an academic team) and so distinguish ourselves from rival bidders (such as teams of consultants). On their side, the appointing panel recognised that a review and report by an academic team was likely to have some legitimacy, especially among higher education establishments. At the same time, the short timetable set for the work (just some months across 1992 and 1993) was not untypical of the conditions and pressures confronting policy research of this kind.

Conducting the study

Four sources of evidence were collected by the research team. First, interviews were undertaken with those responsible for leading and conducting quality assessment visits to institutions; and with those responsible for administering and overseeing the whole process. The purpose was to collect their views on the appropriateness and effectiveness of the assessment method, including its purposes, its analysis and use of self-assessments, the composition, performance and training of assessment teams, the assessment categories and the methods of reporting. The lead assessors were interviewed face-to-face. The majority of reporting assessors were interviewed by telephone.

Within the Funding Council, the conduct of subject assessment was the responsibility of a Quality Assessment Division (QAD). Its Director and officers were interviewed together with the Chief Executive and Secretary of the English Funding Council and the responsible officers at the Welsh Funding Council. Oversight of the scheme was exercised by the Quality Assessment Committee of the English Funding Council, with members drawn from higher education establishments as well as other bodies. Interviews with members and officers from each of these organisations were all undertaken face-to-face.

A second source of evidence was assembled from face-to-face interviews with senior university managers with experience of supporting subject departments in their self-assessments and team visits. The views of the representative body of heads of universities were also sought in this way. A similar set of questions was posed about the purposes, methods, measures, costs, benefits and impact of the scheme, as they were for a third type of evidence. This involved questionnaires sent to institutions whose programmes formed part of a subject assessment visit in History, Chemistry, Law and Mechanical Engineering, these subjects being chosen as representative of the range of programmes examined within the national

evaluation process. Questionnaires were also despatched to institutions which were not visited in the same four subjects. In both sets of institutions, the questionnaire was addressed to those responsible for the self-assessments in each subject. A questionnaire was sent as well to specialist assessors. With a deadline for return within 10-14 days, useable response rates ranged between 54% and 68%. The fourth and final source of data was generated through observations of subject assessment visits at two long-established universities: the assessment visit in History at the University of Wales, Aberystwyth; and the assessment visit in Chemistry at the University of Birmingham. They were the most contested, sensitive and difficult parts of the research design, requiring negotiated agreements and compromises between the Quality Assessment Division (QAD) and the research team on what elements should (and should not) be observed by the researcher.

The request to join the assessment visits at Aberystwyth and Birmingham was not exactly welcomed by QAD officers, with a period of stand-off between the research team and the QAD, followed by agreement at short notice for the researcher (conducting that fieldwork) to accompany the assessment team visiting the University of Wales, Aberystwyth. That QAD officers were uncomfortable with the review and its design was understandable. As a result of the earlier abolition of a national inspectorate for schools and colleges, some inspectors had moved into roles at the Higher Education Funding Council, which in turn exposed their methods and judgements to early scrutiny and criticism from some powerful quarters in higher education.

In the event, the terms of access for the researcher were not defined in advance of the visit but were negotiated with one of the Lead Assessors who was present for parts of the visit. The researcher was enabled to join the assessment team in its initial meeting with senior staff of the university and then at the end of the visit when an oral report on the judgement reached was given to senior members of the subject department and the institution. In between, the researcher was able to speak with team members about how they went about their tasks and with members of the department on how they experienced these activities. The researcher was not granted access to the teaching events observed and scored by individual assessors; nor to the meeting at which the overall judgement was determined. However, there was sufficient interaction with both parties to report on the tone and style of visits, including their public and private dimensions.

The tensions and dramas accompanying these encounters, including the demands placed on assessment teams, were well-illustrated at both institutions. At the University of Wales, Aberystwyth, the welcome extended to the assessment team by the Vice-Chancellor (a distinguished academic historian) was heavily qualified. While very effort would be made to accommodate their needs, it was made plain to the team—individually and collectively—that the University had little confidence in their capacity to judge the quality of teaching in history, including where this required an understanding of the Welsh language. The reception given to the assessment team at the University of Birmingham was different again but equally intimidating. On arrival at the University, and outside the planned programme, the team were ushered into a lecture room to meet with the full ranks of the staff in chemistry and to justify their credentials in the subject. These were tense situations for both parties and would be replayed again at the feedback session at the end of the visit. Ahead of this meeting, the visiting team at each university had agreed and rehearsed what would be

relayed, with the reporting assessor conveying the outcome and limiting any discussion to factual information and accuracy. Other members of the team were asked not to speak or respond to questions directed at them. The aim was to leave the premises in good order, if not in good standing.

Wicked issues

Policy analysis characteristically poses wicked issues, issues—in other words—that do not yield straightforward stances on the part of researchers (Halpin & Troyna 1994; Pawson, 2006). Several were present in this study, and while they in part emerge out of the particular context—that of quality assessment in higher education—the challenges surely have a wider resonance. In this section, we identify just a few such challenges by way of exemplification:

- **Weak evidential base:** Characteristically, in policy studies, a time horizon is present in a particular complex way. One might even say that the principle of ripe time (Cornford, 1908) is present, which is to say that the analysis of a policy has purpose at a point in time. In turn, there is a time horizon that presses on the conduct of the research and the production of the ensuing publication. And in turn, a close time horizon may lead to there being a limited period of time in which to conduct the fieldwork and assemble the relevant evidence and data. And, in turn, researchers are sometimes obliged to form their analyses and judgements on an evidential base that is weaker than they would wish. And yet, the discursive and ideological context may be looking for clear and incisive judgements. All of these elements were present in this project.
- **“Ought” and “Is”:** Policy analysis often includes an element of policy framing. A careful set of judgements on a policy framework—if only implicitly—may have a normative orientation, offering indications as to how a policy framework might be developed. In our project, that normative element was explicit and significant. Whether implicit or explicit, there arises the long-standing philosophical problem as to how to extract from an analysis of a situation—or build onto it—proposals (however subtly expressed) for ways in which a policy framework might be developed; how, in short, to derive an “ought” from an “is”. We commented earlier on the value orientation that we adopted that served to orient our proposals but there is a further point here. In the antagonistic ideological context in which this project was conducted, any prescriptive stance was likely to be seen as a matter of “taking sides” in one direction or another.
- **Gatekeepers:** Not infrequently in policy studies (Walford, 1994), especially in conducting the fieldwork, individuals or groups will be encountered who are significant players in the policy field. Such individuals and groups may act as gatekeepers in a double sense: they may have some measure of control over the researchers’ gathering of the evidence and/or they may attempt to control or influence the researchers’ developing narrative of the field (Williams, 2012). As indicated earlier, we encountered just such gatekeepers in the form of the “contract assessors” who were concerned that our research was in effect a device for critiquing their role. We were therefore faced with a critical, not to say

hostile and intransigent, “community of practice” even while we were conducting this research study.

- **Conceptual confusion:** A policy field is typically fraught with conceptual confusion, in the sense that key terms would be in widespread public use (say in health, or energy, or transport studies) and yet lacking in clarity. A research study could quickly become ensnared by such confusion. Here, for instance, difficulties were present in relation to the concepts of “quality”, “excellence” and “the student experience”, with the actors and constituencies implicated understanding and using these terms in their own ways. Detailed conceptual clarification of such terms would have been inappropriate here —the focus was elsewhere, on the efficacy and soundness of the national quality assessment methodology— but the research work had still to be conducted with considerable degree of conceptual sensitivity, care and rigour.
- **Underlying mechanisms:** In any large policy field, there are large forces at work such that there is a kind of ontological layering present. In the situation with which we were faced, there were close to hand interests of the state in securing a more accountable higher education system in which institutions were offering a “value-for-money” educational experience to students. There were also indications of a “performativity” at work, with the assessment process focusing on the teaching situation, and the teacher’s overt performances in that situation, the reports on institutions’ teaching as a result being characterised —as we judged it— by an excessive blandness. The way teachers fulfilled their wider professional and educational responsibilities towards their discipline or towards their students or the student’s fulfilment of their responsibilities were not matters for assessment. Accordingly, there would be no mention in a report on history teaching of “history” or of the intellectual rigour of a programme of study. One might even say, thereby, that this was a “postmodern” assessment, shorn of any location in any “grand narrative” (Lyotard, 1984).

Deeper still, at play here were the early signs of neoliberal moves on the part of the state, to orient higher education more towards the wants of the student-as-customer and to heighten competition between institutions (with the use of a graded system of assessment). In effect, our study was a nice example of critical realism in action, in which our own analyses and judgements were not content with a reading of the empirically evident features in front of us, but were also sensitive to a reading of the ontologically deep “generative structures” at work (Bhaskar, 2008). Accordingly, even if it was not made explicit, underlying our empirical findings and observations lay a “depth” ontology of the field confronting us.

- **Ideological options:** As stated, part of the policy field facing us was a determination on the part of the state to secure —through the quality assessment process— a greater measure, and a uniform, measure of accountability of the higher education sector. This concern was apparent in the judgemental process, in which institutions were graded against a common “template” of criteria (this resulted in some confusion, as it appeared to us, in that there was a rhetoric on the part of the Funding Council of being sensitive to each institution’s mission even while a

common set of criteria were in force in making quality judgements). As a result, the research team found itself in the middle of an ideological contest: was the quality assessment system intended to illuminate institutions' "fitness-for-purpose" or their "fitness-of-purpose"? That is to say, were institutions fully adequate in realizing the (common and uniform) purposes expected of them by the state as public institutions of higher education (and which therefore applied to all institutions, from world-famous institutions to more local institutions) or did institutions each have its own (and therefore particular) set of purposes against which the state's quality judgements could be made?

In this very tiny linguistic inflection ("fitness-for-purpose"/"fitness-of-purpose") we see then an ideological cleavage not merely between uniformity on the one hand, and sensitivity to difference on the other hand but between the emergence of the evaluative state on the one hand (with its insistence on uniformity and graded cross-institutional judgements) and the emergence of neoliberalism in higher education (with its encouragement of competition and difference) on the other hand. Such a reading poses a dilemma to researchers: to what extent might the researchers' analysis favour one of the ideologically-laden orientations held by the actors in the field (here, "fitness-for-purpose"/"fitness-of-purpose") or to what extent do the researchers attempt to circumvent the ideological field by inserting and even proposing an orientation of their own?

- Basis for recommendations: Within the quality evaluation methodology, the national assessors formed their judgements on a very limited basis. Institutions were required to compile a self-assessment but, where assessment visits to institutions took place, these self-assessments formed no part of the judgemental process. There was no attempt in those visits to examine the rigour and frankness of the self-assessment. Nor was there any dialogical meeting with a teaching team to assess the extent to which the teaching staff were working as a team and had a common set of aims and hopes for the programme in question. Rather, as intimated, the assessors formed their judgements on the overt features in front of them: the facilities and the teaching "event". The assessment moved on a superficial plane of empirical observations and summative judgements. There was no attempt, for example, to forge—in concert with the programme team—an agenda for quality enhancement.

A dilemma that this raises is the following: to what extent might the research team base its own analysis deliberately different from the presenting rhetorical field? Here, for instance, might it be legitimate to propose a more dialogical engagement with teaching staff that addresses the intellectual landscape in question and its associated educational aims, and is oriented more towards quality enhancement)? Our own value framework (see earlier in this chapter) was helpful in providing a platform from which to build just such a normative set of proposals.

Writing dilemmas

The writing of the report posed a number of challenges; dilemmas indeed. There was one large challenge in particular. The study had been

commissioned, it will be recalled, by the Higher Education Funding Council for England. The principal document flowing from this study was that of a report to the Funding Council, setting out the research team's analysis and recommendations as to how the national quality assessment regime might be taken forward. However, as intimated, the academic community across England had an interest in the report; and both institutions of higher education and many individuals in the academic community were very interested in having sight of the report. Indeed, as noted, the academic community had placed high expectations upon the study—hoping to see it make some trenchant criticisms of the assessment regime and to come forward with recommendations for radical changes to it, and had thereby a very strong desire to see the report. Given our value orientation (cited above), which included an interest in transparency and partnership, we judged too that it was right and fitting that our report should be published by and placed into the public domain by the Funding Council.

This situation immediately posed a large challenge to us as the drafters of the report. After all, for the funders of the study—the Funding Council—the report was formally a report on the review written for and delivered to itself. The Council made no commitment to publishing the report, at least in advance of having sight of the report. Accordingly, the essential challenge before us was this: could we, on the one hand, draft a report that both made—where necessary—trenchant criticisms of the Funding Council's quality assessment approach and put forward recommendations for substantial changes in that approach and, on the other hand, be drafted in such a way that—despite any criticisms that the report might contain of such a major national body—the Funding Council would not be so discomforted by the report that it would feel able to release fully the report into the public domain? To return to an earlier point, could our research and its ensuing report command respect and be seen as having integrity by all the parties concerned?

The challenge was indicative of a dilemma, at once both professional and intellectual, that might have been felt to be facing us: either speak truth to power and provide an unvarnished analysis and incisive set of recommendations that flowed intellectually from the data (but then run the risk that the Funding Council might have been so embarrassed by the report as not then to feel able to place it in the public domain) or produce a softer reading of the data and proffer fewer and less trenchant set of recommendations such that the Funding Council would more readily publish the report (but then run the risk that the academic community would feel that the report had censored it-

self and had been economical with the truth). In short, the report was written within a context in which sensitivities were systematically present and which were bound to affect its pragmatic options.

Indicative of the sensitivities present at the time of the drafting and completion of the report were the following instances:

- An article appearing in the weekly national magazine, *The Times Higher Education Supplement*, entitled "Publish Barnett or be Damned".
- A question being posed in the upper house of the UK's national parliament (the "House of Lords"), asking as to why the Funding Council had not published the report on its quality assessment system.
- The chairman of the UK's Committee of University Vice-Chancellors and Principals paying a visit to the principal investigator, to "discuss" the study while the report was being drafted.
- More than one meeting with the senior officer at the Funding Council with whom we were liaising during the course of the study, in which comments were made on the precise wording of the draft report, the gist of which was "Does the research team really want to be judged on the basis of such a judgement in the report?", so placing moral pressure on the team in the drafting of the report.
- Mention in a satirical weekly column on the back page of the UK's *Times Higher Education* magazine (the item being implicitly critical of the Funding Council).

Such public exposure is indicative of pressures on researchers in policy studies, so presenting the drafting of the report with a severe discursive context, which in turn bore in upon the fine wording and crafting of the report. Every word, every phrase, was being scrutinised by external parties and read in multiple ways. The discursive challenges and dilemmas weighed upon the drafting of every sentence, and every clause in every sentence in attempting to secure the twin aim of perceived integrity and rigour across all of the contending parties and especially not discomforting the Funding Council to the extent that it might feel unwilling to release our report fully into the public domain. Accordingly, the drafting of the report took the form of constructing a multiply persuasive argument, an argument persuasive of multiple constituencies in positions of power (academic and political) and persuasive through multiple argumentative ploys (Fischer & Forester, 1993; Hoppe, 1999).

Outcomes

In the end, our report was published by the Funding Council and put into the public domain (Barnett, Parry, Cox, Loder & Williams, 1994). In the late stages of the drafting, it underwent, as intimated, some gentle interrogation from the Funding Council and some massaging of the text. At every moment of redrafting—of a word, of a phrase—the textual dilemmas (on which we have touched) were heightened. Could we find a word, a phrasing, that still retains the force of the point we are making and yet not discomfort the Funding Council such that it would shrink from publishing our report (and so also would gain legitimacy across the academic community)? Not only was our report published but it occasioned a certain amount of national debate (with a large day-conference in London, entitled “After Barnett, What?”) as well as being given attention in the national quality press.

With its forty-four recommendations, did it have any independent impact on the Funding Council’s evolution of its quality assessment methodology? We cannot say. We can say that, on the one hand, our recommendation that every quality assessment (of every subject in every discipline) should lead to a visit to the institutional department concerned was taken up. Also, ideas that we put forward for a wider array of educational matters to form part of the assessment canvas were adopted. On the other hand, our other central recommendations were not accepted, namely that the summative graded judgement element be abandoned and be replaced by a threshold judgement backed up by a “quasi-profile” in relation to different dimensions of educational provision. On the contrary, the Funding Council moved towards a more detailed and more intrusive assessment regime, in which institutions were given numerical scores on six categories of judgement. There was no attempt to move towards a more dialogical process, oriented towards quality enhancement.

As a result, the assessment regime became more punitive, and permitted the national press to produce “league tables” of quality assessments (since institutions came to be graded 1-4 on six criteria, so achieving a maximum score of 24). This was precisely the scenario that our proposal of a “quasi-profile” had been intended to head off. Here, manifestly, we were unsuccessful. Subsequently, the quality system that followed our report itself ran into difficulties and was itself followed by a succession of stages in the evolution of the national quality system. National quality assurance in the UK became an unstable matter (Brown, 2004). This is hardly surprising. After all, the evaluative state is likely to go on in ever refining its techniques of surveillance and regulation (Neave, 2012).

Conclusions

Policy research, especially in a situation freighted with ideological conflict (as here) is a site of challenges: is there available a position of value neutrality to the research team? If not, what value framework might a research team adopt? How might such research be steered among the ideological thickets? Is it possible for such research to gain legitimacy across the various interest groups? What kinds of reading might come into play: to what extent does the research remain within rather empirical readings (and the frames provided by the key actors in the policy field being examined) and to what extent does it reach down into the deep structures in order to expose the field with a more critical lens? Just what considerations might come into the very drafting of any report or publication arising from the research? If the research might have or is expected to have some kind of practical and normative impact, what kind of impact is it legitimate to aspire towards? How might a report or publication position itself as between being purely technicist and being transformative?

Such research work has a particular element of onerousness attaching to it. For the battery of questions above—which could easily be elongated—do not permit of straightforward answers. More conventional research housed within disciplinary boundaries—whether of an empirical or a scholarly variety—has typically a largely secure intellectual home, where the criteria, values and textual norms are more or less understood. Here, in the kind of policy research reported and analysed here, no such boundaries were present. Consequently, as we have seen, every step of the research process—the design of the research methods, the conduct of the fieldwork, the analysis of the data, and the writing of the report—posed its dilemmas. Nevertheless, when such dilemmas can be recognized and confronted, there is a possibility that those dilemmas can be resolved to some extent and so allowing policy research to play a part in helping directly to influence policy and practices in the wider world and so take on a form of “deliberative” policy analysis (Hajer, 2003).

There will be a limit, however, as to the extent such dilemmas can be resolved. After all, the ideological context cannot be dissolved and nor can the ontological and epistemological issues be entirely overcome. The separate parties will continue to hold their own (often conflicting) definitions of the situation and (opposed) expectations of the study, the object of the inquiry will be changing even while it is being investigated, and characteristically judgements will need to be formed by the researchers in an undue compress of time and on the basis of inadequate data. Policy

analysis research of this kind, accordingly, presents a research context with challenges and problems that simply cannot be entirely dissolved but rather will necessarily accompany the research.

All this raises issues over the academic identity of policy analysts and researchers: what are their values? To whom do they feel accountable? What effects are they striving for (across the intellectual, practical, professional and political domains)? These are abiding questions to be continually revisited.

About the authors

Ronald Barnett has been developing a social philosophy of the university and higher education for around 40 years. He has held senior management and leadership positions and also acts as a higher education consultant. His latest book is *Imagining the University*.

Gareth Parry has been scholar and researcher in higher education for around 30 years, with an interest in system change and reform.

References

- Barnett, R., Parry, G., Cox, R., Loder, C., & Williams, G. (1994). *Assessment of the Quality of Higher Education: A Review and an Evaluation*. Report for the Higher Education Funding Councils for England and Wales. London: Institute of Education University of London.
- Bhaskar, R. (2008). *A Realist Theory of Science*. London: Verso.
- Brown, R. (2004). *Quality Assurance in Higher Education: The UK Experience since 1992*. London: RoutledgeFalmer.
- Cornford, F. M. (1908). *Microcosmographia Academica. Being a Guide for the Young Academic Politician*. Cambridge: Bowes and Bowes.
- Fischer, F., & Forester, J. (1987). *Confronting Values in Policy Analysis: The Politics of Criteria*. Newbury Park: SAGE.
- Fischer, F., & Forester, J. (eds.). (1993). *The Argumentation Turn in Policy Analysis and Planning*. Durham: Duke University Press.
- Greene, J. C. (2006). Towards a Methodology of Mixed Methods Social Inquiry. *Research in the Schools*, 13(1), 93-98.
- Hajer, M. (2003). Policy without Polity? Policy Analysis and the Institutional Void. *Policy Sciences*, 36(2), 175-195.
- Halpin, D., & Troyna, B. (1994). *Researching Education Policy. Ethical and Methodological Issues*. London: Falmer Press.
- Hoppe, R. (1999). Policy Analysis, Science and Politics. *Science and Public Policy*, 26(3), 201-210.
- Lyotard, J. F. (1984). *The Postmodern Condition: A Report on Knowledge*. Manchester: University of Manchester Press.
- Neave, G. (2012). *The Evaluative State, Institutional Autonomy and Higher Education in Western Europe: The Prince and His Pleasure*. New York: Palgrave Macmillan.
- Pawson, R. (2006). *Evidence-based Policy. A Realistic Perspective*. London: SAGE.
- Richie, J., & Spencer, L. (1994). Qualitative data analysis for applied policy research. In Bryman, A., & Burgess, B. (eds.). *Analysing Qualitative Data*. London: Routledge.
- Sadoulet, E., & De Janvry, A. (1995). *Quantitative Development Policy Analysis*. Baltimore: Johns Hopkins University Press.

- Shore, C., & Wright, S. (eds). (2005). *Anthropology of Policy. Critical Perspectives on Governance and Power*. London: Routledge.
- Srivastava, A., & Thomson, S. B. (2009). Framework Analysis: A Qualitative Methodology for Applied Policy Research. *JOAAG*, 4(2), 2-79.
- Walford, G. (1994). *Researching the Powerful in Education*. London: UCL Press.
- Williams, C. (2012). *Researching Power, Elites and Leadership*. London: SAGE.
- Yanow, D. (2000). *Conducting Interpretive Policy Analysis*. Thousand Oaks: SAGE.