Research Evaluation **22** (2013) pp. 105–117
Advance Access published on 7 January 2013
doi:10.1093/reseval/rvs041

Accountability, performance assessment, and evaluation: Policy pressures and responses from research councils

Creso M. Sá*, Andrew Kretz and Kristjan Sigurdson

Department of Leadership, Higher, & Adult Education, University of Toronto, Toronto, ON M4Y 1R6, Canada
**Corresponding author. Email: c.sa@utoronto.ca*

This study identifies contemporary government accountability requirements impacting research councils in North America and Europe and investigates how councils deal with such demands. This investigation is set against the background of rising policy frameworks stressing public sector accountability that have led many national governments to enact legislation requiring public agencies to collect more performance information and tie it to decision-making. Through documentary analysis and interviews with informants at several research councils we clarify how broader policy trends are reflected in the operation of public institutions that provide critical support for academic science. In addition to legislation cast broadly to regulate the activities of all government agencies, numerous regulations and guidelines have been targeted specifically at science and technology (S&T) activities. Regulations on S&T expenditures in general and on research councils more specifically include efforts to develop new metrics specific to science-based or innovation-based outcomes, to enhance the use of indicators in decision-making, to focus on tracing the broad impacts of programs, to increase the frequency of reporting, and to make agencies more responsive to business and public interests.

Keywords: accountability; performance assessment; research evaluation; research councils; research funding agencies; science policy.

1. Introduction

Many governments worldwide have become increasingly evaluative of public bureaucracies and service providers (Power 1997). Undergirding policymakers' scrutiny of public organizations is concern for assuring quality, typically entailing an emphasis on outputs and outcomes, productivity and effectiveness, responsiveness, and performance management (Lane 2005). Such notions are key to public policies steered in accordance with New Public Management (NPM), a global reform movement consisting of common notions for public governance that generally support the view that publicly supported organizations can be managed and evaluated in the same way as private organizations, namely through demands for accountability, transparency, and efficiency and responsiveness (Christensen and Lægreid 2007). Accordingly, performance indicators and outcome measures figure prominently in efforts to assure quality and the social and economic relevancy or impact of public policies, bureaucracies, and programs (Behn 2001; Bird et al. 2005).

Reflective of this general trend of ensuring quality in public sector activities are building pressures to better assess and demonstrate the impacts of public investments in scientific research (Himanen et al. 2009; Stensaker and Harvey 2011). Many countries have developed new indicators for the evaluation of academic research activity (Grant et al. 2009; Dill and van Vught 2010). The recent impetus to map the social and economic returns of investments in science was reflected in former US presidential science advisor and Director of the Office of Science and Technology Policy (OSTP), Dr John H. Marburger III's 2005 speech to the American Association for the Advancement of Science (Norrie 2012). During his

address Marburger advocated for the development of a 'science of science policy', a term which he later elaborated on as 'the data sets, tools, and methodologies needed to assist science policy decision-makers as they invest in Federal research and development and make science policy decisions' (Interagency Task Group 2008: 1). The push towards the rationalization of public investments in science implies greater attention to measuring, assessing, and evaluating research activity.

Research councils are at the nexus of contemporary demands of the NPM and growing expectations over the social and economic benefits of scientific research. How have widely discussed government accountability requirements influenced research councils? How do research councils deal with demands to demonstrate the value and impact of their investments? These questions guided this investigation of major research agencies in Canada, the USA, and Europe. Through the identification and analysis of agency responses to this environment, this article clarifies how broader policy trends are reflected in the operation of public institutions that provide critical support for academic science.

2. Methods

This study draws on documentary and interview data. Documentary data include publicly available annual reports, strategic plans, and evaluative reports of research councils, in addition to relevant legislative texts. We also conducted interviews with key informants at the following research councils: The National Science Foundation (NSF) and National Institutes of Health (NIH) in the USA, the Canadian Institutes of Health Research (CIHR) and Natural Sciences and Engineering Research Council (NSERC) in Canada, the Medical Research Council (MRC) and Economic and Social Research Council (ESRC) in the UK, and the European Union (EU)'s European Research Council (ERC).

For interviews, we contacted potential informants holding relevant positions in each of the research councils. Such positions include members of advisory bodies, senior administrators, programs directors, and support staff. In most cases our requests were mediated by agency administrators, who provided the research team with access to key personnel. The final number of informants interviewed included 19 participants from 7 different research councils. Informants participated in semi-structured telephone interviews, all of which were ~30 min in length. Interviews were audio recorded and transcribed for analysis. Themes concerning the influences of government accountability requirements on research councils and the approaches of these agencies in dealing with these demands were identified through content analysis (Budd, Thorp and Donohew 1967; Krippendorf 1980; Weber 1996). The authors independently coded interview transcripts to enhance inter-rater reliability, and develop a robust coding scheme.

Two major themes emerged from the analysis. The first theme concerns the policy landscapes within which research councils are embedded, as viewed from the agencies' perspectives. The second theme regards the organizational responses of research councils to government policies.

3. Policy landscape

The policy landscape in which research councils operate is characterized by accountability legislation, regulations on research councils, and agency reporting requirements. Each of these poses distinctive constraints on agency officials and influences policies at their institutions.

3.1. Accountability legislation

Research councils are subject to much of the overarching accountability legislation enacted by governments and attract particular scrutiny, as they not only receive government funding, but distribute it as well.

In the USA, the *Government Performance and Results Act (GPRA)*, which was signed into law in 1993 (Government Performance and Results Act 1993), and came into force in 1997, aims at enhancing the efficiency and effectiveness of federally funded programs by compelling agencies to focus on program results, the quality of services, and customer satisfaction (Radin 1998). In 2010, the GPRA was updated through the *GPRA Modernization Act (GPRA-MA)*, which required the development of strategic plans that specify measurable goals (Government Performance and Results Modernization Act 2010). The updated Act shifts the focus of accountability from the production of plans and reports to the use of performance data to measure outcomes.

Similarly, in Canada, a number of federal accountability measures have created a complex regulatory environment in which government agencies, including research councils, are pressured to set performance indicators and measure outcomes. One element of this environment is the *Federal Accountability Act* (2006), introduced by the new Conservative government in 2006 as a means to restore faith in the institutions and practices of government after a spending scandal plagued the previous government. John Baird, the Treasury Board President at the time, noted that '[t]he Federal Accountability Act is a significant and substantive step by Canada's New Government to help restore Canadians' trust in government and the democratic process' (Prime Minister's Office 2006). The wide-reaching legislation includes provisions mandating that all grant and contribution programs be reviewed at least every 5 years as a means of ensuring their relevance and effectiveness.

The Canadian government's 2006 strategic plan, *Advantage Canada: Building a Strong Economy for*

Canadians (Department of Finance Canada 2006), also includes initiatives aimed at ensuring the effectiveness of government programs and spending, seeking to hold all government offices, departments, and agencies to greater account for their spending. Further accountability requirements are set out by the Treasury Board of Canada (2008a) *Policy on Transfer Payments*, which requires fund granting departments to establish policies and procedures to ensure a results-based management and accountability framework for measuring and reporting of results of use of funds. Government agencies are also subject to accountability requirements set out in the Treasury Board's Policy on Internal Audit (Treasury Board of Canada 2012) and Policy on Evaluation (Treasury Board of Canada 2009), as well as to government expectations for the implementation of Results-based Performance Frameworks facilitated through the Treasury Board's Centre of Excellence for Evaluation to monitor and report on the performance of programs (Treasury Board of Canada 2010). In addition, federal departments and agencies are required to report to parliament on priorities through a Report on Plans and Priorities (RPP) (Treasury Board of Canada 2008b) and on performance through a Departmental Performance Report (DPR) (Treasury Board of Canada 2011).

In the UK, budgetary legislation plays a major role in setting accountability requirements for government departments. Through the 2010 *Spending Review* (HM Treasury 2010), resource allocations across all government departments are made in accordance with government priorities. Moreover, in 2010 the Review required all government departments to undertake reforms towards greater accountability, such as by publishing business plans, and by establishing the Transparency Framework that mandates each government department provide performance and spending indicators. Following the requirements of the Spending Review, government departments, including the Research Councils UK, as well as individual research councils, publish Delivery Plans that set out their key deliverables over a 3 year time frame. The plans form part of a comprehensive Performance Management Framework intended to facilitate long-term planning for science built on an evidence-based foundation (Research Councils UK 2005). Agencies are required to develop scorecards based on the strategic deliverables specified in the multi-year delivery plans and report outcomes to the Department of Business, Innovation and Skills (BIS).

Accountability measures embodied in the NPM approach appear to have been pursued with less vigor in the programming of the supranational governance context of the EU. The European Commission is the body responsible for ensuring the application of EU law; the Commission proposes and implements legislation, executes the EUs budget, and manages the Union's programs (Cipriani 2010). In 1998, the *Annual Report* published by the European Court of Auditors chastised the Commission for its lack of use of 'performance targets in terms of outputs, outcomes and the costs involved in achieving these' (p. 05). According to Harlow (2002), The European Court of Auditors and European Commission have been at odds, as the former has advocated extending accountability through the introduction of NPM techniques, whereas the latter has preferred to retain the substantial discretion to which it is accustomed as the motor of the EU. The nature of EU spending is fundamentally compliance-oriented and not linked to sets of concrete objectives and the achievement of results. Factors such as these have led to what some have referred to as an 'accountability gap' (Cipriani 2010).

In the three countries, the rise of accountability frameworks has led governments to enact legislation that requires public agencies and departments to collect more performance information and tie it to decision-making. Each country's governance history and mix of institutions and structure makes for a different variety of accountability legislation, which may be attached to budget setting, such as in the UK, or to omnibus accountability legislation, such as in Canada. The EU's unique status and the associated complexities of supranational governance have led to debates about accountability. A clear trend is apparent towards stronger overarching laws requiring government agencies to adhere to particular decision-making frameworks focused on the use of performance measurement and results evaluation as key inputs to decision-making.

3.2. Regulations on research councils

In addition to legislation cast broadly to regulate the activities of all government agencies, numerous regulations and guidelines have been targeted specifically at science and technology (S&T) activities. Regulations on S&T expenditures in general and on research councils more specifically include efforts to develop new metrics specific to science-based or innovation-based outcomes, to enhance the use of indicators in decision-making, to focus on tracing the broad impacts of programs, to increase the frequency of reporting, and to make agencies more responsive to business and/or public interests.

Comprehensive, multi-year S&T strategies, such as Canada's 2007 *Mobilizing Science and Technology to Canada's Advantage*, are framed as extensions of overall economic strategy, but highlight government S&T decision-making. Canada's 2007 strategy laid out a framework to guide investments of public funds in this area including new directives for S&T spending and requirements for research councils. One of the four core principles that guide the Strategy and its policy commitments is 'Enhancing Accountability', by which the Government of Canada promises to 'implement stronger governance and reporting practices to deliver and demonstrate results'

(Industry Canada 2007: 6). The document's policy commitments include such measures as increasing the representation of business and community interests on research funding agency governing councils as well as improving the 'ability to measure and report on the impact of S&T expenditures' (Industry Canada 2007: 88). These changes are intended to ensure that science funding in the country is guided by appropriate evidence and authorities that will lead the country towards a more efficient use of resources for increased economic competitiveness.

In the USA, similarly targeted efforts to enhance accountability and increase the use of information in S&T decision-making have taken place through the issuance of joint memos by the Office of Management and Budget (OMB), and the OSTP. The 2005 joint memo, which outlined federal R&D priorities for the 2007 financial year, stated that an interagency process was needed 'to develop a new framework for understanding the impact of R&D investments, define appropriate data elements for monitoring and assessing this impact . . . and improve the basis for national science policy decisions' (Marburger and Bolten 2005: 6). In 2006, the National Science and Technology Council's Subcommittee on Social, Behavioral and Economic Sciences (SBE) established an Interagency Task Group (ITG) to develop a coordinated Federal approach to the science of science policy and in 2008, the ITG published *The Science of Science Policy: A Federal Research Roadmap*. This document identified shortcomings of US science policy decision-making, which included heterogeneous and unsystematic data collection about the science supported by federal agencies. In addition, agencies used different models, data, and tools to understand their investments in S&T, and the data infrastructure was inadequate for decision-making. The report found that '[n]ew tools and data sets could be developed and used to quantify the impact that the scientific enterprise has had on innovation and competitiveness' (Interagency Task Group 2008: 1). The ITG also outlined the Federal efforts necessary for the long-term development of a science of science policy intended to lead to a maximization of the country's investments.

More recently, a 2009 OMB OSTP joint memo outlining federal priorities for the 2011 financial year, took a step further by asking federal agencies to

> 'develop outcome-oriented goals for their science and technology activities, establish procedures and timelines for evaluating the performance of these activities, and target investments toward high-performing programs. Agencies should develop 'science of science policy' tools that can improve management of their research and development portfolios and better assess the impact of their science and technology investments. Sound science should inform policy decisions, and agencies should invest in relevant science and technology as appropriate' (Orszag and Holdren 2009: 2).

These directives, informed by the findings of the ITG, create new requirements for coordinated efforts by USA research councils to work towards models, tools, and metrics intended to provide policy-makers with better evidence to inform decisions about investments in science.

In the UK, the government's *Science and Innovation Investment Framework 2004–14* set out a strategy for improving the UK's innovation performance based on greater responsiveness of the publicly funded research base to the needs of the economy. The report was underpinned by the requirement that progress towards scientific and innovation system development targets be tracked by a range of new indicators (HM Treasury, Department of Trade, & Industry and Department of Education and Skills 2004). These indicators were largely built from the 2006 report*: Increasing the Economic impact of the Research Councils*, a report of the Research Council Economic Impact Group, chaired by Peter Warry. Known as the 'Warry Report', the report included recommendations on how the UK's research councils could better demonstrate the economic impact of their investments. The Warry Report called on the Research Councils to make 'strenuous efforts to demonstrate more clearly the impact they already achieve from their investments', while simultaneously noting the difficulties in doing so given the time-delayed or indirect nature of these outcomes (Research Council Economic Impact Group 2006: 5). The report, and the government directives that followed, recommended that the Research Councils proactively seek out information required to evaluate impact and that, once such data have been obtained, full impact analysis of all Research Council knowledge transfer schemes is conducted (Research Council Economic Impact Group 2006: 20). The findings and recommendations of the Warry Report were largely built into the current performance management system of the Department for Business, Innovation, and Skills, which funds the Councils and reports to parliament for the Councils' activities.

The EU's 7th Framework Programme for Research and Technological Development (FP7) is the main instrument for funding research in selected priority areas in order 'to respond to Europe's needs in terms of jobs and competitiveness, and to maintain leadership in the global knowledge economy' (European Commission 2007: 06). With a total budget of over €50 billion for 2007–13, the framework includes a variety of funding schemes and is implemented by the European Commission (Cipriani 2010). The propensity towards input-based and compliance-oriented accountability discussed above with regard to the European Commission's general management approach also applies to its implementation of the FP7. However, recent efforts indicate that changes in this area may be on the way. A 2010 communication from the European Commission to the European Parliament, the European Council, the Economic and Social Committee, and the Committee of the Regions suggests ways that the

implementation of research framework programs can be simplified. Notably, it included the proposal of 'far-reaching changes towards a results-based funding using lump sums' as one of three major recommended actions for the future (European Commission 2010: 04). Thus, while the input-based control mechanisms characterizing the FP7 appear fixed until the program's conclusion in 2013, a more business-like, results-based approach to accountability for S&T funding appears to be on the horizon.

Mounting political initiative has aimed at increasing the quantity and quality of information available to policy-makers tasked with overseeing research councils. This has led to new directives for the exploration, production, and use of new metrics to guide science policy. As major destinations for and sources of government funding, research councils in particular are increasingly subject to close scrutiny from the government departments that oversee them to be accountable for the money spent on their activities. As expectations for the socio-economic impact of science and innovation mount, particular focus has been cast towards research councils that are asked to find better ways to measure their activities and outcomes and then act on these measurements.

3.3. Agency reporting requirements

Heightened accountability has translated into increasing reporting requirements on research councils. Agencies are required to report with more frequency and cover a wider variety of their activities and the outcomes of these activities. Reporting requirements range from quarterly or annual, to multi-year or one-off reports, depending on the country, the agency and the type of information to be reported.

In the USA, under the GPRA, the National Science Foundation must submit an Annual Performance Report (National Science Foundation 2011). The National Institutes of Health likewise reports to the government under the GPRA, primarily through an Annual Performance Plan and Report (National Institutes of Health 2012), which describes the agency's mission, resources, programs and strategies, and provides an assessment of the progress towards achieving its performance goals and targets.[1]

In Canada, the three federal research councils are required to report to parliament on an annual basis. The 2007 Science and Technology strategy praised NSERC for having strengthened its reporting to government by initiating quarterly reporting on activities and results, demonstrating the political weight behind increasing the frequency of research funding agency reporting in Canada.

In the UK, research councils are required to issue an annual performance report to Parliament as stipulated in the Science and Technology Act 1965. The councils are also required to compile Annual Economic Impact Reports, which are part of the BIS performance management framework. In 2008/09 new metrics were added to expand reporting on the social and economic dimensions of the output of the research councils with the aim to make reporting more robust and consistent across councils (Medical Research Council 2012d). Some councils also compile additional reports, such as the MRC's Annual Review, which is aimed at the general public and explores aspects of the Council's work on a selected theme chosen each year (Medical Research Council 2012b).

One of the largest recipients of UK research funds, the Medical Research Council, communicates research outcomes in a variety of ways and to a number of different audiences. According to the Medical Research Council's *2012/2013 Delivery Plan Scorecard*, the Council currently reports to the government on the outputs of its activities every 6 months (Medical Research Council 2012c). The agency also produces an *Annual Report and Accounts* intended for parliament and other government audiences, which details the agency's activities and achievements in various areas. Although the delivery plan scorecard and Annual Report and Accounts provide the most comprehensive picture to government and parliament of annual operations, the MRC Economic Impact Report contains information requested by the government on selected aspects of the organization's performance specifically related to economic and social dimensions of research output. The Annual Review—aimed primarily at the general public—has since 2010 been published as an interactive website.

The European Research Council, as a component of the EU's FP7, operates in a different type of governance environment than its national counterparts. In line with the other FP7 funding schemes, and with the management dynamics of the European Commission more generally, the ERC employs a fundamentally cost-based funding system wherein accountability is sought primarily through cost-reporting and financial auditing (European Commission 2010). However, some performance information is reported through annual Monitoring Reports, which cover the implementation of the Framework Program as a whole. Additionally, high-level, independent evaluation panels have been established by the European Commission to evaluate the performance of the ERC under the FP7.[2]

The array of information issued by research councils in the UK, USA, Canada, and the European Commission with regard to their activities is diverse and complex. Some reports may be required directly by federal legislation and others by executive bodies or parental departments overseeing the functions of the agencies. Still other reporting may be undertaken outside of what is formally required as agencies seek to preemptively increase the robustness of measurement and to ensure that no relevant audience is left uninformed. Evidence points to the existence of a changing environment with increasing pressures

on research councils to measure and communicate in myriad ways and to deliver information to diverse audiences about how money is being spent and what returns are being gained. Although there appear to be vast differences in the avenues by which agencies are required or choose to evaluate and publicize results, there is commonality in the overarching trend towards additional pressure for measurement and communication. This, in turn, has led to a number of observable organizational responses.

4. Organizational responses

The trends reported above have led to the on-going search for effective evaluation systems, comprising methodologies, indicators, and standards for the measurement and reporting of research outcomes. This has created a need among agencies to deal with the inherent uncertainties in the evaluation of research. Research councils have attempted to overcome these difficulties internally, as well as externally, by better communicating the outcomes of their investments, sharing practices and benchmarking internationally, and refining their evaluation systems.

4.1. Enhancing evaluation systems

The evidence agencies collect when monitoring funded projects concerns both program administration (e.g. grant application data) and research performance. Agency staff members consider the first step in program evaluation as occurring during the panel review of submitted grant applications. Agencies tend to collect information from applications for internal circulation to inform the administration of their own programs. Application data such as application rates, applicants' demographics, proposed topics, and grant success rate are collected. Such information is used to inform policy decisions about the direction and scope of current funding programs. As one informant describes, '[a] very good example in this respect is the low participation of women researchers in the call.... When we see these numbers we go and look at the population, the potential population that could apply.... we try to see whether this low percentage of participation is representative of the presence of women in senior positions.' Although research councils traditionally use panel review and application data to review projects and programs, recent policy and initiatives target post-project impact assessment.

Output indicators (e.g. publications, citations, patents) are typically used to assess program and project impact. Agencies also perform bibliographic analysis of papers resulting from sponsored programs. In addition to publications, agencies monitor awards and honors connected to funded research, the contribution of the funded project to the career of investigators (in the case of career-development grants), and evidence of international collaboration. One agency officer explains, 'we do our own monitoring [through] international databases... we are monitoring also the award of various scientific prizes and we see whether [our] grantees are being included in [the] list of awardees.' As documented above, there are growing pressures on agencies to demonstrate program success in terms of the broad impacts of the results of funded research. Research output indicators provide agency evaluators with evidence of the outcomes of agency investments.

The need to document, track, and analyse research processes and outcomes has generated a large data gathering effort at research councils. As one informant argues, 'internationally, agencies are really striving to develop the metrics that allow you to say these are the socioeconomic benefits that came out of a particular program of research.' NSERC, for instance, collects data through surveys, file reviews, and administrative reporting on research outputs, knowledge transfer (self-reported data); training (exit survey, career survey), commercialization (post-award studies), as well as through qualitative approaches such as interviews, case studies, focus groups, and expert opinion. The overall context for evaluation is preoccupied with 'value for money', relevance, impacts, and relatedly, an 'increased focus on measures of effectiveness, efficiency and economy' (Morris and Townsend 2008). Methodological approaches for evaluating grant programs therefore include program need, 'alignment with government priorities, consistency with federal roles and responsibilities, achievements of expected outcomes, and demonstration of cost-efficiency and economy' (Morris and Townsend 2008). In Europe, the ERC has strengthened its commitment towards monitoring systems for collecting information through indicators on an ongoing and systematic basis since the end of Sixth Framework Program in 2006. Methods used to evaluate the impacts of research include interviews, surveys of program participants, and expert panels. Bibliometric information is also analysed (Olsen and Merrill 2011).

Monitoring and evaluation is embedded in the operation of research councils. At the NIH, for example, the Evaluation Set-Aside Program supports such activities. The Office of Program Evaluation and Performance coordinates NIH-level reporting, including the analyses required for the GPRA. The goal is to produce authoritative, qualitative, and quantitative analyses of return on investments by studying selected and measurable outputs and outcomes of NIH investments. At the ESRC, through its Research Evaluation Committee, impact evaluations attempt to demonstrate broader impacts of work supported by the agency.[3] Organizational structures and processes in each agency are dedicated to data gathering, analysis, and reporting activities.

There have also been broader efforts to develop inter-agency frameworks and systems for measurement and evaluation. These have been particularly strong in

the UK. In 2002, the Research Councils UK formed as a partnership among seven research councils in the UK to coordinate activities to enhance their overall impact and effectiveness. Four of these Research Councils (AHRC, BBSRC, ESRC, and EPSRC) have harmonized the collection of research outcomes through the Research Outcomes System (ROS), a web-based system that allows users to provide research outcomes to the four participating councils[4]. To further integrate activities, RCUK produces Delivery Plans that set aims and indicators for enhancing excellence, impact, and efficiency. These plans enable participating councils to demonstrate the contribution funded research is making towards achieving government priorities and targets.[5] Moreover, the MRC's Economic Impact Reporting Framework (EIRF) was implemented across all research councils in 2005 and forms part of the performance framework managed by the Department for Business, Innovation, and Skills.[6] In the USA and Canada, more preliminary steps have been taken towards integrated approaches. A recent inter-agency initiative in the USA aims at developing data and tools to assess the impact of R&D investments. Dubbed 'Star Metrics', it seeks to develop measures of the impact of federal science investment on scientific knowledge, social outcomes, workforce outcomes, and economic growth. In Canada, a 'Granting Councils Indicator Working Group' is intent on identifying indicators to capture impacts of funded R&D and assess the feasibility of adopting a common approach to measure impact (Morris and Townsend 2008).

4.2. Dealing with uncertainty in the evaluation of research

Demands for better evidence of returns on investments in science underscore a continuing search for appropriate evaluation methodologies and indicators that can satisfy demands for objective metrics. Agency officials need to manage the uncertainties in evaluating research: beyond the unpredictability of scientific inquiry, the manner in which projects are funded can be highly complex and variegated, cobbling together a mix of short- and long-term objectives, complementary research programs, and research sites. While there is perhaps little concern that substantial benefits indeed accrue from science, the political pressure to measure precisely what those benefits are and report them clearly remains a major challenge for agency officials.

A number of complications arise for research councils when measuring the impact of funded research. First, the complexity of scientific research and policy making make it challenging to attribute a particular funded project to a particular development in policy or practice, or to new insights and common understandings. The indirect and hard to measure influence of research on policy has been long recognized (Weiss and Bucuvalas 1980). Research findings are typically incremental, building on previous studies and knowledge. Moreover, research alone is rarely the sole basis for policy decision-making, and when research does have an impact, it is likely to be influential at different times throughout policy decision-making (Nutley, Walter and Davies 2007). Therefore, the influence of research may not be visible. Furthermore, in cases where research projects have various funding sources, it can be difficult to attribute findings directly to a specific funding agency. Second, there is also a time dimension to capturing and assessing research impacts. NSF Director Subra Suresh has noted, '[t]he time frame in which you assess the impacts is much longer than the duration of the grant itself. That makes it much harder to say, at the end of 3 years, whether you've have an impact or not' (Mervis 2011: 1493). One informant from another agency echoed how time frames for evaluation were a limiting factor in evaluating grant programs for broad impacts and noted that the longer you wait, the more difficult attribution becomes.

A third challenge in measuring the impacts of funded research is the lack of standardization: the variety of initiatives and sectors/fields supported by research councils complicates the development of measures with which to evaluate research impacts. This issue is particularly challenging for those research councils that support numerous research centres and organizations, as different operational, technical, and cultural factors can complicate evaluation, even more so when there exists multiple program responsibility centres (Economic and Social Research Council 2007). A related issue is the difficulty in operationalizing and obtaining data for evaluation. One of the main problems besetting attempts at economic impact measurement relates to the establishment of reliable indicators of relevant variables for outputs and outcomes (Economic and Social Research Council 2007). As agencies grapple with quantifiable, verifiable, and accessible metrics to demonstrate their impact, they also recognize the threat of over relying on these types of imperfect or partial measurements. Agencies understand that there is no obvious solution to the lack of indicators related to more intangible benefits of research, such as the satisfaction of social curiosity or other benefits to social welfare. One informant asserted, '[to] assess at a dollar level an impact of a particular program of research is extraordinarily hard for everyone. This . . . socioeconomic impact is exceedingly difficult [to measure]. Thirty years later you can say, look at the impact that it had. Look at the dollars that it just saved us. Just look at the lives it has saved. That's often hard to assess in the short term.'

Even where indicators are agreed upon, research councils report difficulty in collecting objective and valid data. Evidence is often reported by the awardee of research funds or from the use community, making the collection of reliable and unbiased information difficult. Brian Sloan, Directorate-General for Research and Innovation for the

European Commission, has been noted as pointing out that '[s]urveys can be a burden to participants, especially when long and detailed answers are required. This can influence the quality of their response. Response bias and partial responses are also a concern' (Olsen and Merrill 2011).

Lastly, impact is difficult to assess because the ways in which research findings are delivered, and the particular environment into which they are introduced, shape their influence on policy and practice. The research message itself is only one part of the picture, and both the '[p]rocesses (user engagement, dissemination, networks, consultancies etc.) and contexts (for example, policy relevance, user receptiveness, timing, financial or political issues) will all have a bearing on the uptake of research findings' (Economic and Social Research Council 2009: 4).

As regards the early stages of research, widely accepted methods have been developed and refined to measure research performance such as the peer review of proposals, manuscripts, and projects that are in progress, as well as bibliometric analysis of publications (Wagner and Jordan 2009). However, effective methods for measuring the long-term impact generated by public expenditure on research funded by such mechanisms as federal granting agencies remains elusive (Orians et al. 2009).

4.3. International benchmarking and collaboration

Alongside the efforts to manage uncertainty in the evaluation of research impact, research councils have engaged in national and international dialogue and collaboration around developing data infrastructure for monitoring and analysis. The challenges of developing indicators for the outcomes of research funding has led agencies to cooperate in seeking out agreed upon metrics that might be used to measure the effectiveness of research funding activities.

Since 2005, the ESRC hosted a number of international events on assessing research impacts. These forums have provided an opportunity for representatives from various research councils to work together in building the early foundations of impact evaluation for funded research.[7] In Canada, NSERC also formed an international panel of experts in 2011 to conduct an assessment of approaches used nationally and internationally to evaluate research performance. Using this information, NSERC expects to develop the methodology in consultation with key stakeholders that allows for objective budgetary allocations based on performance assessments. In the USA, the NIH organized a meeting about the Science of Science Management in 2008, focused on the management of biomedical research to improve decision-making.[8] These various initiatives suggest active steps to exchange information, in search of new and improved practices to monitor and evaluate research programs.

In addition to symposia and workshops, several interviewees indicated that they regularly communicate with staff of other national and international agencies as a means of fine-tuning approaches to evaluation and in seeking new methodologies. Regarding the difficulty of measuring the longer-term socioeconomic impact of research funding discussed above, one respondent noted that 'to really drill into our particular study or a body of studies and say this is a socioeconomic impact that we can attribute to that [study]... is an exceedingly difficult area of research'. Circumventing this problem is a shared concern among research councils in different countries, the informant argues: 'a lot of agencies are sharing with each other the ways that they try to do this.'

The paucity of widely accepted methods for evaluating the impact of funded research has led agency staff to seek to work with other agencies formally and informally to share information (Sá, Tamtik and Kretz in press). This practice appears to be relatively new, and as pressure to account for and report on outcomes such as socioeconomic impacts of research increase, it is foreseeable that agencies will continue to work together to build new metrics and establish consensus on practices and methodologies that can link research funding to impact.

4.4. Communicating outcomes

One way in which research councils deal with the challenge of appeasing the demands for greater accountability when the measurement of outcomes are difficult if not impossible to effectively undertake and comprehensively report is by stepping up efforts to communicate those outcomes that can be reported. Although they are required by law to produce reports, research councils often go beyond the mandated requirements and produce additional types of information for various audiences as part of a strategy supporting accountability efforts.

In Canada, the CIHR issue a wide range of reports as a means of communicating outcomes of funding in a variety of formats to different audiences. A number of reports are issued to Parliament on an annual basis.[9] CIHR also develops a variety of other publications showcasing strategies, outcomes, and success stories[10] and periodically published documents such as *Show me the Evidence,* a pamphlet published twice yearly with stories about the use of health research evidence produced by the council (Canadian Institutes of Health Research 2011d).

Overall, efforts to communicate outcomes from research funding use a variety of approaches to presenting information and are tailored to different audiences such as government representatives, the general public, and the research community. Research councils appear to go well beyond the minimum reporting requirements and are increasingly using more appealing electronic and print publications as a means to inform their stakeholders of their activities and the results thereof.

5. Conclusion

This study extends a growing body of research describing and analysing how accountability regimes influence public institutions (Lane 2000; Schiavo-Campo and McFerson 2008; Pidd 2012), and more specifically the role of performance assessment in the governance of science (Cozzens and Turpin 2000; Georghiou and Larédo 2006; Martin 2011). In this article we shift attention away from the evaluation of academic research to the workings of research councils and their policy environment.

Performance measures and outcome reporting are central to increased monitoring and evaluating, as they help to report, plan and budget; manage and evaluate; learn and improve; and compare performance and results (Behn 2001; Bird et al. 2005). Research councils thus face considerable pressure to formalize, elaborate, and increase their reporting and evaluation structures. In the research councils investigated, performance metrics and increased emphasis on outcomes provide information necessary for organizational governance, and also address external pressures for transparency and accountability. Apart from general demands to demonstrate efficiency and effectiveness, accountability legislation, regulations on research councils, and agency reporting requirements work to systematize and direct research towards social and economic priorities and impacts.

Demands for accountability, transparency, efficiency, and responsiveness have encouraged the widespread and systematic monitoring and evaluating of public sector activities (Whitley and Gläser 2007). The emergence of an 'indicator culture' in many academic systems parallels broader public sector trends under the influence of the assumptions of NPM. However, concerns over the negative and unintended side effects of introducing NPM-style managerialism to academic systems continue to be voiced. Power (1997) suggests that NPM has the potential to adversely affect research culture and may end up being a 'fatal remedy' when used to improve the performance of higher education systems. He stresses that the activities for which performance measures have been, or can be, devised are not necessarily constitutive of the entirety of activities needed in a healthy scientific community.

In the research funding agencies that formed part of this study evidence points to the emergence of an 'indicator culture'. New organizational structures such as the ESRC's Research Evaluation Committee or the Office of Program Evaluation and Performance of the NIH add powerful managerialist elements to operations that have traditionally relied on the professional expertise of the researchers themselves. The launch of novel evaluation tools and information repositories such as Star Metrics in the USA or the UK's ROS suggest a growing entrenchment of indicators as instruments for guiding the operations of funding agencies and the strategies of the governments that support them.

Whitley and Gläser (2007) have similarly argued that the import of NPM measures into systems of public science changes the conditions under which scientific knowledge is produced, potentially for the worse. While the scale and significance of this change depends on the national context, particularly for those countries where higher education is more centrally controlled, shifts can have major, and not necessarily positive, effects on performance. The replacement of direct state control with the monitoring and evaluation of indicators intended to reflect broad objectives, which is characteristic of the shift to NPM, can have powerful effects on intellectual competition, innovation, and coordination (Whitley 2011). Increasing reliance on monitoring and evaluation through bibliometric data or scientific awards, as reported by managers interviewed for this study, may reward conformity rather than risk-taking or unorthodox approaches to research. This threatens to reduce research diversity and the variety of approaches being taken to solve problems, as failure ceases to be an option for the researcher who wants to continue to be awarded funding. If the indicators available cannot measure the output of the research, its value is reduced. Another possibility is that institutions focus ever-more exclusively on hiring faculty with established researcher careers, posing difficulties for young researchers starting their academic careers (Whitley and Gläser 2007). With funding agencies looking at quantitative measures of publications, citations, awards, and patents to determine increasingly large sections of university budgets, it might be expected that institutions will be forced to rely more heavily on the same types on information in deciding who gets hired.

Although it has not been the purpose of this study to investigate in detail the effects of greater accountability pressures on the public systems of science, it is important to acknowledge that these effects may be taking place. For their part, research funding agencies, as intermediaries between government and science face a dual challenge as they are pressured towards greater adoption of the 'indicator culture' and the managerialism embodied in NPM approaches, while also ensuring the wellbeing of research cultures and academic systems. As they grapple with imperfect indicators and metrics of scientific performance and outcomes, agencies may strive to avoid being seen as unaccountable to protect their budgets, while also guarding against over-reliance on such measures since this could lead to the accountability cure being worse than the disease.

Evaluators of research, such as the research councils studied here, continue to grapple with the limitations of prevailing methodologies to provide straightforward indicators of research impact desired by policy makers. Despite the extensive methodological literature on the topic (e.g. Moed, Glänzel and Schmoch 2005; Garbarino and Holland 2009; Donovan 2011), this remains a problematic activity from the viewpoint of research councils.

The demand for measuring and evaluating research performance notwithstanding, the actual capacity of agency staff to do so remains underdeveloped, particularly with regard to longer-term impact and outcomes. Previous efforts at inter-council cooperation and coordination at the national level (Garrett-Jones and Aylward 2000; Holbrook 2000) have become solidified and have since expanded to attempts at international cooperation in developing new evaluation mechanisms. Research councils continue to rely heavily on peer review and output indicators. However, as councils seek more sophisticated and systematic means of evaluating research, the peer review process is being integrated with other methods for evaluating research outcomes and guiding decisions about council policy and investments.

On this policy landscape, political expectations around public investments in science place considerable pressure on research councils to formalize, elaborate, and increase their reporting and evaluation structures. More than 10 years ago, in a special edition of Research Evaluation, Cozzens and Turpin (2000) highlighted the world-wide emergence of these reporting and evaluation structures and noted, '[t]he days of the occasional, large-scale, program evaluation seem to be gone. The current evaluation environment is tightening through the introduction of regular monitoring and assessment systems, fed by more systematic reporting on the activities of academic researchers through new information systems.' (p. 3). The findings in this study confirm Cozzens and Turpin's insight. It also shows that the further development of research evaluation systems has prompted multiple changes in how research councils track, assess, and communicate the activities they support. Even in the USA, where evaluation by peer review has traditionally been the dominant system (Cozzens 2007), research councils are finding funding for research conditionally tied, or contractually connected with, performance assessment.

Notes

1. For more on NSF reporting requirements see the Performance Information section of the FY 2012 Budget Request to Congress <http://www.nsf.gov/about/budget/fy2012/pdf/fy2012_rollup.pdf>. For NIH see the FY 2013 Annual Performance Report and Performance Plan <http://www.hhs.gov/budget/performance-appendix-fy2013.pdf>.
2. These reports include the *Interim Evaluation* of FP7 which is intended to assess implementation and compile information to be used in the next framework program and the one-off 2009 report *Towards a World Class Frontier Research Organization,* aimed at reviewing the structures and mechanisms of the ERC and which proposed both immediate changes as well as longer term changes to be made prior to the start of FP8. This report also recommended that a follow-up report be commissioned to ensure that recommendations have been followed (European Commission 2009).
3. To date, two summary reports have been produced. In 2009, The Council's Research Evaluation Committee reported on ESRC's work to evaluate the impact of research on policy and practice in *Taking Stock*. This report focuses primarily on the ESRC's series of Policy and Practice Impact Case Studies and the lessons learned from this work (Economic and Social Research Council 2009). The second, 'Branching-Out: New Directions in Impact Evaluation from the ESRC's Evaluation Committee' reviews the latest findings from the Council's Impact Evaluation Program and proposes further activity to be taken forward (Economic and Social Research Council 2011).
4. Outcomes are categorized into nine areas: Publications, Other Research Outputs, Collaboration/Partnership, Further Funding, Staff Development, Dissemination/Communication, IP and Exploitation, Award/Recognition, and Impact.
5. In addition, the research councils, in coordination with other funders of research, have developed a Strategic Framework for Health Informatics in Support of Research, identifying five key areas deemed significant in strengthening the UK's capacity and capability in health informatics research.
6. The framework was revised in 2009 to expand reporting on the economic and social dimensions of funded research inputs, performance, and outputs. Another avenue for regular reporting of outcomes is the MRC Researchfish system. The system, launched in 2009 under the title MRC e-Val, replaced end-of-grant reporting and requires grant recipients to make online reports each year rather than a long report at the end of the grant summarizing years of progress (Medical Research Council 2012a).
7. These include an international symposium entitled 'New Approaches to Assessing the Non-Academic Impact of Social Science Research', which aimed to provide an opportunity to share approaches and discuss current international developments in demonstrating and assessing the impact of social science research. Another workshop was held in 2007 to discuss the development of methods for assessing the impact of social science research. The third workshop to host academics, consultants, and representatives from across the Research Councils and government was in 2010 with the aims of disseminating the latest findings from ESRC Impact Evaluation Program and to discuss possible new activity to drive the area forward.
8. Discrete program evaluations are also performed within an international frame of reference.

The NSERC's 2007 *Report of the International Review Committee on the Discovery Grants Program* compares the funding agency's programs to those in other countries in order to benchmark effectiveness and impact (Natural Sciences and Engineering Research Council 2007).

9. These include the Annual Report, the DPR, and the RPP. The Annual Report provides a synthesized overview of agency activities including a variety of informational graphics, charts, and statistics, intended to communicate what the CIHR is, what it has done in the past and where it is headed (Canadian Institutes of Health Research 2011b). In contrast, the RPP provides information on how each funded activity area is linked to strategic outcomes and how these activities benefit Canadians (Canadian Institutes of Health Research 2011c), whereas the DPR detail the progress that has been made in these activity areas through a variety of indicators (Canadian Institutes of Health Research 2011a).
10. See the CIHR's *Research, Knowledge, Impact* at <http://www.cihr-irsc.gc.ca/e/documents/cihr_research_knowledge_impact_e.pdf> and *Impacts of CIHR-funded research: a compendium of results* at <http://www.cihr-irsc.gc.ca/e/documents/compilation_success_stories_e.pdf>.

References

Behn, R. D. (2001) *Rethinking Democratic Accountability*. Washington, D.C.: Brookings Institute Press.

Bird, S. M. *et al*. (2005) 'Performance Indicators: Good, Bad, and Ugly', *Journal of the Royal Statistical Society A*, 168/1: 1–27.

Budd, R. W., Thorp, R. K. and Donohew, L. (1967) *Content Analysis of Communications*. New York, NY: The Maximillian Company.

Canadian Institutes of Health Research (CIHR). (2011a) *Departmental Performance Report 2010-11* <http://www.tbs-sct.gc.ca/dpr-rmr/2010-2011/inst/cri/cri00-eng.asp> accessed 29 Aug 2012.

——. (2011b) *Moving Forward: Canadian Institutes of Health Research Annual Report 2010–11* <http://www.cihr-irsc.gc.ca/e/44139.html> accessed 29 Aug 2012.

——. (2011c) *Report on Plans and Priorities 2010-11* <http://www.tbs-sct.gc.ca/rpp/2010-2011/inst/cri/cri-eng.pdf> accessed 29 Aug 2012.

——. (2011d) *Show Me the Evidence* <http://www.cihr-irsc.gc.ca/e/documents/show_me_evidence_v1_e.pdf> accessed 25 Sept 2012.

Christensen, T. and Lægreid, P., eds (2007) *Transcending New Public Management: The Transformation of Public Sector Reforms*. Aldershot, England: Ashgate.

Cirpriani, G. (2010) *The EU Budget: Responsibility without Accountability?* Brussels, Belgium: Centre for European Policy Studies.

Cozzens, S. E. (2007) 'Death by Peer Review? The Impact of Results-Oriented Management in US Research'. In: Whitley, R. and Gläser, J. (eds) *The Changing Governance of the Sciences*, pp. 225–42. Dordrecht: Springer Science.

Cozzens, S. E. and Turpin, T. (2000) 'Processes and Mechanisms for Evaluating and Monitoring Research Outcomes from Higher Education: International Comparisons', *Research Evaluation*, 8/1: 3–4.

Department of Finance Canada. (2006) *Advantage Canada: Building a Strong Economy for Canadians*. Ottawa, ON: Finance Canada.

Dill, D. and van Vught, F. (2010) *National Innovation and the Academic Research Enterprise*. Baltimore, MD: John Hopkins University Press.

Donovan, C. (2011) 'State of the Art in Assessing Research Impact: Introduction to a Special Issue', *Research Evaluation*, 20/3: 175–79.

Economic and Social Research Council. (2007) *Report of the ESRC Impact Evaluation Methods Workshop 20th March 2007* <http://www.esrc.ac.uk/_images/ESRC_Impact_Evaluation_Methods_Workshop_tcm8-3814.pdf> accessed 20 Aug 2012.

Economic and Social Research Council (ESRC). (2009) *Taking Stock: A Summary of ESRC's Work to Evaluate the Impact of Research on Policy and Practice* <http://ebookbrowse.com/gdoc.php?id=51187813&url=f11e2edc4e340114719ca40be4e694ca> accessed 20 Aug 2012.

——. (2011) *Branching Out: New Directions in Impact Evaluation from the ESRC's Evaluation Committee* <http://www.esrc.ac.uk/_images/Branching%20Out_tcm8-14881.pdf> accessed 25 Sept 2012.

European Commission. (2007) *FP7 in Brief: How to Get Involved in the EU 7th Framework Programme for Research* <http://ec.europa.eu/research/fp7/pdf/fp7-inbrief_en.pdf> accessed 1 Nov 2012.

——. (2009) *Towards a World Class Frontier Research Organization: Review of the European Research Council's Structures Mechanisms* <http://ec.europa.eu/research/evaluations/pdf/archive/other_reports_studies_and_documents/erc_review_panel_report_-_23_july_2009.pdf> accessed 25 Sept 2012.

——. (2010) *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Simplifying the Implementation of the Research Framework Programmes* <http://ec.europa.eu/research/fp7/pdf/communication_on_simplification_2010_en.pdf> accessed 1 Nov 2012.

Federal Accountability Act. (2006) *Retrieved from the Department of Justice Canada Website* <http://laws-lois.justice.gc.ca/PDF/F-5.5.pdf> accessed 25 Sept 2012.

Garbarino, S. and Holland, J. (2009) *Quantitative and Qualitative Methods in Impact Evaluation and Measuring Results* (Governance and Social Development Resource Centre) <http://www.gsdrc.org/docs/open/EIRS4.pdf> accessed 25 Sept 2012.

Garrett-Jones, S. and Aylward, D. (2000) 'Some Recent Developments in the Evaluation of University Research Outcomes in the United Kingdom', *Research Evaluation*, 8/1: 69–75.

Georghiou, M. and Larédo, P. (2006) *OECD Science, Technology and Industry Outlook 2006*. Paris: Organisation for Economic Co-operation and Development.

Government Performance and Results Act. (1993) *Retrieved from U.S. Office of Management and Budget Website* <http://www.whitehouse.gov/omb/mgmt-gpra/gplaw2m> accessed 26 Sept 2012.

Government Performance and Results Modernization Act. (2010) *Retrieved from the U.S. Office of Management and Budget Website* <http://www.whitehouse.gov/omb/performance/gprm-act> accessed 26 Sept 2012.

Grant, J. *et al*. (2009) *Capturing Research Impacts: A Review of International Practice*. Report prepared for the Higher Education Funding Council for England. Cambridge: RAND Europe.
Harlow, C. (2002) *Accountability in the European Union*. Oxford, England: Oxford University Press.
Himanen, L. *et al*. (2009) 'Influence of Research Funding and Science Policy on University Research Performance: A Comparison of Five Countries', *Science and Public Policy*, 36/6: 419–30.
HM Treasury. (2010) *Spending Review* <http://cdn.hm-treasury.gov.uk/sr2010_completereport.pdf> accessed 20 Aug 2012.
HM Treasury, Department of Trade, & Industry and Department of Education and Skills. (2004) *Science & Innovation Investment Framework 2004–2014*. London: HM Treasury.
Holbrook, A. (2000) 'Evaluation of Research Sponsored by Federal Granting Councils in Canada: The Social Contract', *Research Evaluation*, 8/1: 47–56.
Industry Canada. (2007) *Mobilizing Science and Technology to Canada's Advantage*. Ottawa, ON: Industry Canada.
Interagency Task Group. (2008) *The Science of Science Policy: A Federal Research Roadmap*. Washington, DC: National Science and Technology Council and the Office of Science and Technology Policy.
Krippendorf, K. (1980) *Content Analysis: An Introduction to Its Methodology*. Thousand Oakes, CA: Sage Publications, Inc.
Lane, J.-E. (2000) *New Public Management*. New York: Routledge.
——. (2005) *Public Administration and Public Management: The Principal-Agent Perspective*. London: Routledge.
Marburger, J. H. III and Bolten, J. B. (2005) 'Memorandum for the Heads of Executive Departments and Agencies: FY 2007 Administration Research and Development Budget Priorities'.
Martin, B. R. (2011) 'The Research Excellence Framework and the "Impact Agenda": Are We Creating a Frankenstein Monster?', *Research Evaluation*, 20/3: 247–54.
Medical Research Council. (2012a) *About MRC e-Val (Now Researchfish)* (pubd online 2012) <http://www.mrc.ac.uk/Achievementsimpact/Outputsoutcomes/e-Val/About/index.htm> accessed 25 Sept 2012.
——. (2012b) *Annual Review* (pubd online 2012) <http://www.mrc.ac.uk/Newspublications/Publications/AnnualReview/index.htm> accessed 25 Sept 2012.
——. (2012c) *Delivery Plan Reporting Framework 2012/13* <http://www.mrc.ac.uk/consumption/idcplg?IdcService=GET_FILE&dID=36187&dDocName=MRC008769&allowInterrupt=1> accessed 27 Aug 2012.
——. (2012d) *MRC Economic Impact Report* (pubd online 2012) <http://www.mrc.ac.uk/Newspublications/Publications/EIRF/index.htm> accessed 27 Aug 2012.
Mervis, J. (2011) 'Suresh Expects Low-Cost Ideas Will Mean a Big Payoff for NSF', *Science*, 334: 1492–3.
Moed, H. F., Glänzel, W. and Schmoch, U., eds (2005) *Handbook of Quantitative Science and Technology Research*. New York: Kluwer Academic Publishers.
Morris, S. and Townsend, S. (2008) 'Assessment of S&T Outcomes in Canada—The Current State in the Federal Government', Powerpoint slides from *The 11th Annual Meeting of the Heads of Research Councils Working Group, Workshop on the Science of Science Policy* <http://www.nsf.gov/about/performance/SciencePolicyWrkshp_Presentations/Morris&Townsend.pdf> accessed 12 Aug 2012.
National Institutes of Health. (2012) *NIH Performance Reporting (GPRA)* (pubd online February 15 2012) <http://dpcpsi.nih.gov/opep/performance_reporting.aspx> accessed 25 Sept 2012.
National Science Foundation. (2011) *FY 2012 Budget Request to Congress* <http://www.nsf.gov/about/budget/fy2012/pdf/fy2012_rollup.pdf> accessed 25 Sept 2012.
Natural Sciences and Engineering Research Council. (2007) *Report of the International Review Committee on the Discovery Grants Program* <http://www.nserc-crsng.gc.ca/_doc/Reports-Rapports/Consultations/DGinternational_review-rpt_e.pdf> accessed 25 Sept 2012.
Norrie, J. (2012) *Push to Quantify Social Impact of Science Goes Global* (The Conversation) <https://theconversation.edu.au/push-to-quantify-social-impact-of-science-goes-global-7068> accessed 18 May 2012.
Nutley, S. M., Walter, I. and Davies, H. T. O. (2007) *Using Evidence: How Research Can Inform Public Services*. Bristol, UK: The Policy Press.
Olsen, S. and Merrill, S. (2011) 'Measuring the Impacts of Federal Investments in Research: A Workshop Summary', *National Academies (US) Committee on Measuring Economic and Other Returns on Federal Research Investments*. Washington, DC: National Academies Press.
Orians, C. E. *et al*. (2009) 'Scientific and Public Health Impacts of the NIEHS Extramural Asthma Research Program: Insights from Primary Data', *Research Evaluation*, 18/5: 375–85.
Orszag, P. R. and Holdren, J. P. (2009) 'Memorandum for the Heads of Executive Departments and Agencies: Science and Technology Priorities for the FY 2011 Budget'.
Pidd, M. (2012) *Measuring the Performance of Public Services: Principles and Practice*. Cambridge: Cambridge University Press.
Power, M. (1997) *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press.
Prime Minister's Office. (2006) *Federal Accountability Act Becomes Law* <http://www.pm.gc.ca/eng/media.asp?category=1&id=1455> accessed 20 Aug 2012.
Radin, B. A. (1998) 'The Government Performance and Results Act (GPRA): Hydra-Headed Monster or Flexible Management Tool?', *Public Administration Review*, 58/4: 307–16.
Research Council Economic Impact Group. (2006) *Increasing the Economic impact of the Research Councils* ("The Warry Report"), <http://www.vitae.ac.uk/cms/files/DTI-Warry-Report-July-2006.pdf> accessed 29 Aug 2012.
Research Councils UK (RCUK). (2005) *Research Councils Publish Delivery Plans* <http://www.rcuk.ac.uk/media/news/2005news/Pages/050526.aspx> accessed 12 Aug 2012.
Sá, C., Tamtik, M. and Kretz, A. (2013) 'Evidence and Expertise in Policy Making for University Research'. In: Axelrod, Paul *et al*. (eds) *Policy Formation in Postsecondary Education: Issues and Prospects in Turbulent Times*. Montreal: McGill-Queen's University Press.
Schiavo-Campo, S. and McFerson, H. M. (2008) *Public Management in Global Perspective*. Armonk, NY: Sharpe.
Stensaker, B. and Harvey, L., eds (2011) *Accountability in Higher Education: Global Perspectives on Trust and Power*. London: Routledge.
Treasury Board of Canada. (2008a) *Policy on Transfer Payments* <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=13525&section=text> accessed 5 Aug 2012.
——. (2008b) *Report on Plans and Priorities* (pubd online December 8 2008) <http://www.tbs-sct.gc.ca/rpp/index-eng.asp> accessed 25 Sept 2012.
——. (2009) *Policy on Evaluation* <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=15024> accessed 25 Sept 2012.

——. (2010) *About the Centre of Excellence for Evaluation* (pubd online February 25 2010) <http://www.tbs-sct.gc.ca/cee/index-eng.asp> accessed 25 Sept 2012.

——. (2011) *Departmental Performance Reports* (pubd online November 2011) <http://www.tbs-sct.gc.ca/dpr-rmr/index-eng.asp> accessed 25 Sept 2012.

——. (2012) *Policy on Internal Audit* <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=16484&section=text> accessed 25 Sept 2012.

Wagner, R. M. and Jordan, G. B. (2009) 'Moving Towards Impact When Evaluating Research Programs: Introduction to a Special Section', *Research Evaluation*, 18/5: 339–42.

Weber, R. (1996) *Basic Content Analysis*, 6th edn. Newbury Park, CA: Sage Publications.

Weiss, C. H. and Bucuvalas, M. (1980) *Social Science Research and Decision-Making*. New York, NY: Columbia University Press.

Whitley, R. and Gläser, J., eds (2007) *The Changing Governance of the Sciences*. Dordrecht: Springer.

Whitley, R. (2011) 'Changing Governance and Authority Relations in the Public Sciences', *Minerva*, 49: 359–85.