Prometheus
Vol. 30, No. 4, December 2012, 395–409

Routledge
Taylor & Francis Group

RESEARCH PAPER

Epistemic capacity in research environments: a framework for process evaluation

Tomas Hellström*

School of Economics and Management, Lund University, Lund, Sweden

This paper is based on a study of 40 centers of excellence evaluations performed by the Swedish Research Council in 2008 and 2010. Building on an analysis of these evaluations, a new quality concept, epistemic capacity, is proposed, and it is shown how this may be utilized in the evaluation of research programs and centers. Epistemic capacity refers to a scientific research center's capacity, as embodied in activities and processes, to generate first class research outcomes. In this sense, the concept is best applied in intermediary and process evaluation to ascertain how a research milieu is progressing towards future, but yet unrealized, research contributions. The framework is elaborated in the light of previous evaluation research, and conditions for its application are discussed.

Introduction

Evaluations of publically funded research programs and centers often focus on the extent to which the environments hold promise for, manage, or have succeeded in developing a sound organizational infrastructure for research. The idea is that the existence of certain organizational factors should, on the one hand, facilitate success in a nascent research program or center and, on the other, should be expected as results of successful programs. From a process perspective, it may be argued that research environments that develop these activities and structural conditions in the course of program implementation are more likely to endure and have positive effects on scientific contributions and research careers. At least this is the assumption behind emerging trend in research evaluation, which is to consider organizational and even behavioral antecedents to a successful milieu in addition to traditional output indicators, such as publications and citations of research (Hemlin *et al*., 2004; Heinze *et al*., 2009; Hellström, 2011). However, while evaluation reports are rife with statements about such organizational dimensions, there is a lack of systematic assessment and synthesis of how evaluators describe and assign value. In order to contribute such a synthesis, it was thought expedient that evaluation reports should be studied directly and that normative dimensions of value should be abstracted and systematized from these texts. This paper builds on the two-year evaluations of the Linnaeus centers of excellence conducted on behalf of the Swedish Research Council by two evaluation teams in 2008 and 2010. These evaluations focused on process and organizational aspects of 40 research environments selected

*Email: tomas.hellstrom@fek.lu.se

ISSN 0810-9028 print/ISSN 1470-1030 online

http://dx.doi.org/10.1080/08109028.2012.745959
http://www.tandfonline.com

in national competition across the sciences. By applying qualitative content analysis to these 40 evaluation reports, the study set out to capture how standards for such structural and process qualities were formulated by research evaluators. Several key dependent structural variables (processes and activities) were derived from this material, which may be said to summarize evaluators' notions about a research center's 'epistemic capacity', or its structural preparedness to generate high quality research.

The notion of epistemic capacity was derived from the text analysis and refers to those activities performed and structures maintained that are expected to influence the quality of research directly rather than indirectly. There are good methodological and theoretical reasons for evaluators as well as for researchers in evaluation studies to develop an empirically and conceptually sound understanding of how such process evaluation criteria can be categorized and applied, and to assess their possible contribution to a research environment's epistemic capacity. The present study is a move in this direction. In what follows, some of the theoretical preliminaries for the study will be outlined. Secondly the method employed will be described. Thereafter the evidence will be presented, and finally a discussion and synthesis will be offered of the evaluation of epistemic capacity in research environments.

Theoretical preliminaries

One of the central reasons for organizational/process evaluation in science is that *ex ante* speculations about future contributions to the growth of knowledge are very uncertain. The peer review process, while perhaps the best method available for this kind of evaluation, leaves much to be desired in terms of predictive success or, indeed, simple inter-judgment reliability (Cole *et al*., 1978; Thorngate *et al*., 2009). For obvious reasons, these evaluations will always be part of science; however, the size and concentration of public investments makes it necessary also to engage in assessments of the institutional and organizational capacity, or organizational viability, of large research programs, some of which are expected to receive more than a decade of public funding. Organizational process evaluations of this kind are now quite common, regularly performed by public and private research funders, by state governing bodies, accreditation institutes and by universities. They take a functional view of research organizations, implying that good research will depend, more often than not, on certain forms of organization. One can assume that, to a large extent, research organizations are simply organizations, and will depend for their success on the same general factors that affect any other successful or innovative organization. This is the domain of organization and management research, and numerous studies have outlined criteria for creative organizations; for example, decentralization and short power distances, individual autonomy and 'slack', encouragement of risk-taking, and cross-departmental interaction (Woodman *et al*., 1993; Leonard-Barton, 1998).

However, there may also be organizational factors that are typically (however contextually) associated with the capacity to make significant research contributions. This is the domain of research sociology, research evaluation and social epistemology (Merton, 1979; Kitcher, 2001; Scriven, 2007). The present paper will draw mainly on these traditions, especially the field of research evaluation, to address the problem of capturing epistemic capacity in research milieus. We will begin by

locating the study in this research context before addressing the literature on organizational/process evaluation. Finally, we will attempt to capture some key components of epistemic capacity.

Situating this study

Scriven (2007) discusses research evaluation from the point of view of 'intradisciplinary' quality assessment; that is, those judgments of merit that are directed at experimental designs, hypotheses, data etc., and further – and perhaps more generally – directed at the worth of project proposals, academic manuscripts and job applicants. These assessments are an integral part of research; the intellectual tools applied here are the nuts and bolts of any science, and they require field expertise. In addition, there are at least three ways in which evaluation applies to research quality indirectly, via proxies. The first of these is meta-evaluation, or the assessment of evaluation and other assessment processes applied to research (Stufflebeam, 1981). Meta-evaluation focuses one level up from actual research practice. As a result, general methodological competence rather than field competence is normally required here. Secondly, evaluation can focus on the institutional and organizational framework for knowledge production, on the factors that operate outside the core epistemic activity of actually doing research. Such organizational and process evaluations focus on the operative qualities of the research organization (Hemlin and Barlebo Rasmussen, 2006). This typically requires competence in research management, research sociology, organization studies etc., but not necessarily core field competence.

Finally, there is bibliometric assessment, or scientometrics, pioneered by Eugene Garfield (Garfield, 1964), which bases quality on citation counts and associated composite indicators (e.g. journal impact factors). Disregarding the serious criticism advanced against bibliometrics as a quality assessment tool (e.g. Kostoff, 1998; Feller, 2002; Gläser and Laudel, 2007), one may note how this form of evaluation comes full circle back to science's own internal assessments, by attempting to generalize many such assessments of quality into one indicator saying something about epistemic worth (or 'impact', the term preferred in this tradition). Epistemic capacity is not an indicator of this kind, but refers to those factors that can be assumed to 'cause' productive epistemic processes because they are enacted in order to bring about certain effects in research [see Davidson (1980) for a discussion about motives, actions and causes]. In this sense, they can be referred to as 'near epistemic', 'proto-epistemic', or even 'epistemic conditionals'. For the sake of convenience, we will refer to them simply as 'epistemic capacity' to indicate their enabling relationship to scientific processes.

Research process evaluations

The evaluation of research by means of assessing processes and intermediary outcomes is usually referred to as formative, interim or real-time evaluations (Scriven, 1991; Georghiou and Meyer-Krahmer, 1992). However, it can also be conducted as intermediate summative evaluation to ascertain whether a program has reached the expected levels of maturity or capacity at the end of a build-up period. The interim evaluation of the EU's seventh framework program represents a large-scale example of such an evaluation which, while conducted on a family of funding activities

rather than centers, points towards a number of relevant process dimensions. Specific process or implementation indicators include contribution to regional infrastructure, attraction of appropriate research capacity, cross-sector collaboration, stimulation of participation of young and female researchers, and support of cross-disciplinary research (EC, 2010). Coordinators of collaborative projects, networks of excellence etc., are expected to report on management activities, such as consortium management tasks and achievements, project meetings, dates and venues, impact of possible deviations from the planned milestones and deliverables, and coordination activities (such as communication between beneficiaries, and possible cooperation with other projects/programs) (EC, 2012, p.11). These management process reports are expected to provide a picture of the development of research capacity and organizational maturity. However, such notions as organizational maturity or capacity are multidimensional and highly context-dependent. Abramo *et al*. (2009) note that the identification and mapping of centers of excellence (CoEs) require a unique methodology compared with simple efficiency/productivity measures, and that methodological challenges abound, such as the tendency of traditional output measures to favor pure size and consequently penalize smaller environments.

Moving one step away from the one size fits all framework, Pounder (2000) highlighted the imprecise and non-generalizable nature of output quality indicators for research and education organizations. However, he also showed that across disciplines the most robust quality indicators were organizational ones, such as cohesion, adaptability and goal setting; fairly broad qualitative dimensions that could easily be agreed upon among a diversity of actors, but that also had to be operationalized depending on context.

The more process evaluation is adapted to context, the more specific the indicators need to be to offer valid assessments. As a result, such evaluations, it has been argued, should rely less on proxies (such as publication counts) and more on what Klein (2008) refers to as 'intrinsic indicators' or epistemic indicators of 'good work'. These are exemplified by relevant changes in research frameworks over time, aesthetic quality and fruitfulness in pursuing new problems. Such indicators are likely to be more appropriate for young, interdisciplinary projects, and are sensitive to the fact that the goals of front-line research are variable (Klein, 2008).

Process evaluations taking these aspects into consideration have also been argued to benefit from a bottom-up and interactive approach, where value indicators are derived from project participants and (as far as possible) assessed interactively between evaluator and evaluated (Klein, 2008). One way to ensure the capture and validity of such indicators is to derive a program logic from the bottom-up. Program and outcome logic models work by explicating the intermediary causal factors thought to produce certain valued effects (e.g. Weiss, 1997). They may also be applied in a 'goal-free' mode – that is, without assuming any specific intended effects when studying the program – to elicit unexpected valued activities and outputs (Scriven, 1991). Trochim *et al*. (2008) apply this methodology to generate an outcome model for research which they divide into short-term, intermediate-term and long-term 'markers', where the first two include typical process variables (e.g. collaboration, training, internal/external recognition and support, and transdisciplinary integration).

Conceptual or logic models are useful, but not necessary for research process evaluations. Program monitoring (of organizational processes and delivery mechanisms) is a long-standing form of evaluation (e.g. Rossi and Freeman, 1993), and

checklists of organizational features to be covered by evaluators (such as size, structure, climate, resources and leadership) have been in existence for a long time (Love, 1983), and applied in research process evaluation (Hong and Boden, 2003). The challenge is how to adapt the existing frameworks to the specific challenges of the research organization and its essentially valued outcome; that is, high quality research. Pfirman and Martin (2010) suggest using Boyer's (1990) division of academic core activities – discovery, integration, application and teaching – as a framework for academic evaluation, so as to broaden the traditional output measures and introduce intermediate and process aspects. Scriven and Coryn (2008) propose a couple of novel approaches to research process evaluation; for example, the support net analysis (which elicits aspects of program/center structure, infrastructure and composition that significantly assist researchers in their work and reduces frustration) and the progression discontinuity design (which highlights the effect of a program on the individual's career). The second of these is essentially an outcome measure which focuses on enabling factors in the organization. As always, checklists face the challenge of reifying aspects of an organizational ideal that are then expected to hold for every research organization and academic field, regardless of epistemic or organizational conditions (Hellström and Jacob, 2003). As we will see, some of these problems are evident in the empirical material presented below.

Aspects of epistemic capacity

Tjissen (2003) makes a useful distinction between research excellence and scientific excellence, where the former denotes the activities and outputs specific to research and the latter a number of related cognitive activities, such as teaching, network building and infrastructure creation and maintenance, including research. The expansion of the relevant set of activities from research as such to its social and cognitive context is also illustrated by the concept of a creative knowledge environment (CKE), coined by Hemlin *et al.* (2004), which suggests that an organization's capacity to creatively generate new knowledge is established on three levels: the physical (infrastructure, facilities, location), the social (organizational processes and structures, communication and openness), and the cognitive (thought styles and problem solving approaches).

Upon reviewing the literature, these dimensions are found without any great difficulty in several other sources. However, conclusions are not always straightforward. Tjissen (2003), for example, concludes that organizational infrastructure is important, but it should also be able to produce creative tension between researchers. Availability of resources, economic and physical as well as human capital or competence, is central, but must be adapted and relevant to the problems faced (Tjissen, 2003). In the context of CKEs, according to Martin *et al.* (2004), collaboration with other groups is usually a robust predictor of creativity; yet such involvement also takes time away from research. Likewise, these authors argue, autonomy at the individual level may stimulate productivity, but lack of coordination at the organizational level reduces creativity overall. And while time pressures and high expectations can generate creative and productive solutions, this tends to be functional only up to a point. These authors suggest, in addition, that such qualities of research environments are typically phase dependent, meaning that their causal relevance to productivity will be contingent on where in the research process an individual or group happens to be (Martin *et al*., 2004).

On the organizational level, to utilize the distinction suggested above, factors such as program coherence, strategic approach to publishing, and distribution of publishing effort across the organization are indicators of excellence in a research milieu (Tjissen, 2003). Program coherence is an issue often encountered in the research quality literature, and here, too, it is important to adopt a balanced approach. So, for example, the importance of a research condition known as 'carefully constrained autonomy' is considered by a few authors. Researchers exercise freedom to pursue their own line of inquiry, but within a broad set of problems or a broadly defined framework (Hollingsworth and Hollingsworth, 2000; Heinze *et al*., 2009). Other organizational aspects mentioned include the presence of mechanisms which facilitate interdisciplinary contacts; for example, sharing of laboratory or office space, maintenance of large collaborative networks, and timely access to relevant and complementary skills for solving problems relating to research (Heinze *et al*., 2009). The last of these relates to what Hemlin *et al*. (2004) label the 'cognitive level'. Here we locate the common observation that organizations that are highly prone to significant discoveries and research contributions typically score high on visionary, nurturing and integrative leadership, scientific diversity, interdisciplinarity and integration of activities (Hollingsworth and Hollingsworth, 2000; Hemlin *et al*., 2004; Heinze *et al*., 2009). They also tend to be low on hierarchical and bureaucratic coordination, a quality which has both social and cognitive relevance (Hollingsworth and Hollingsworth, 2000). Huutoniemi (2010) suggests that an important independent variable is that of the integrative research environment; that is, how various specialisms are actually integrated into research. She suggests that the communication and collaboration process itself, including the development of integrative skills, should be assessed alongside the research process and traditional *ex ante* and *ex post* evaluations (Huutoniemi, 2010).

Research approach

The empirical material

The Swedish Linneus program is an initiative of the Swedish Research Council and aims at 'a concerted, long-term and strategic investment in Sweden's leading research environments' (Vetenskapsradet, 2006). The aim of the Linneus grant is further to 'enhance support for research of the highest quality … to encourage universities and colleges to prioritize research fields and to allocate funding for them' (Vetenskapsradet, 2006). The mechanism for this program was to establish well-funded centers of excellence, and to support these for a maximum of 10 years. In 2006, grants were awarded in 20 environments, and in 2008 another 20 were funded. The case that will be elaborated below consists of two separate interim (two-year) organizational/process evaluations of the Linneus centers of excellence conducted in 2008 and 2010 by two evaluation teams on behalf of the Swedish Research Council. Altogether, 40 center evaluations are covered in these reports. The terms of reference stipulated a focus on issues of organization, cooperation and leadership in reference to the plans presented in the program applications. In addition to these topics, each of the evaluation reports also covers opportunities created by the Linneus grant, as well as the strategic and international implications of the centers. The composition of the teams spanned the disciplinary spectrum, ranging from higher education policy, economics, molecular medicine and sociology to physics.

Procedure and analysis

The text analysis on which this case builds takes these reports as its empirical material to elicit how evaluators phrase and posit qualities of research environments likely to lead to successful research, or what is referred to here as 'epistemic capacity'. These characteristics refer mainly to the organizational level; that is, what can be expected to be observed in successful research organizations. Evaluations apply normative categories or proxies to elicit such characteristics, and in this process evaluative dimensions or standards are abstracted from organizational life so as to be applicable to a variety of cases. In doing this, the evaluator will enact or instantiate a value framework, which may be more or less tacit. Some of this framework comes as part of the terms of reference for the evaluation, and some from the value universe of that evaluator, including social, epistemic and other values. The latter is always needed to make sense of and apply the former. Typically, one could say three things about the nature of values as they are expressed in evaluations. They are *social* in the sense that they assume validity beyond the individual expressing them. They are *locally enacted* in the sense that they are brought out in relation to a specific example. And finally they are *performative* in the sense that they are articulated and applied with an audience in mind, and to achieve an effect of some kind. This is to say that values as derived from evaluation texts represent social acts that have implications for the academic community, and that, as such, they can be studied by means of text analysis.

The approach to text analysis used here is based on what Thomas (2006) refers to as the 'general inductive approach', and interpretative phenomenological analysis (IPA) (e.g. Giorgi, 1997). The procedure follows a standard approach by first going through the evaluation reports in detail, identifying evaluative statements and motivations which denote activities explicitly or implicitly assumed to affect the research process directly. These statements may be referred to as 'evaluation units', and they are taken to represent the evaluators' ascriptions of how factors in the research environment impinge on research quality. Evaluation units were captured by assigning codes in the form of short descriptive labels or simple keyword summaries to positive and negative evaluative statements. Codes were then clustered into broader themes based on commonalities identified by the researcher. These themes were re-interpreted and broken down into lower level categories according to the same method of identifying similarities and co-extensive qualities in the evaluation units. The dimensions identified in this way were *production*, which denotes factors to do with the actual generation of research; *coordination*, which involves factors relating to interaction and exchange internal to the research center; and *extension*, which refers to activities that facilitate boundary spanning into new fields and contexts. In what follows, these dimensions will be elaborated on in more detail, starting with epistemic production.

Dimensions of epistemic capacity

Epistemic production

This category comprises activities that are closely related to the cognitive aspects of research, specifically knowledge creation and learning (scholarly development) and the technical infrastructure that is directly used in generating research results (as opposed to simply enabling or facilitating research). Scholarly development is a

central component of epistemic production in that it represents the reproductive and competence-enhancing efforts present in the academic milieu. Mentoring programs were especially prominent among the activities selected for positive mention by evaluators, particularly senior–junior mentoring partnerships that were more formalized than usual; for instance, mentoring programs which were subject to some project management. Such 'projectified' mentoring activities were seen by evaluators to guarantee continuity and accountability in a type of activity, which is often very difficult to assess in terms of quality. Mentoring of this more formalized type also took the form of systematic integration of junior scholars into the research networks of senior scholars. Sometimes formalized mentoring was accomplished by using a research school (for example, governed by a group of Ph.D. students) to integrate junior–senior mentoring activities. In addition to this type of development support, mentoring was also directed at junior scholars to ensure sustainable junior to senior transitions, including structures enabling researchers to transit to leadership roles. This is a more cognitive-institutional aspect of academic leadership, which is sometimes forgotten, but which was brought up by evaluators and highly valued where it occurred. In this general category, one would perhaps have expected to find more activities relating to the scholarly production of senior researchers. Such activities are absent from the evaluation accounts. One interpretation is that senior scholars do not view their own scholarly development as a distinct activity (in the sense that it should leave organizational traces). There is intellectual coordination/communication between scholars of course (and this will be considered below), but in terms of learning it seems that the junior–senior relationship is the most visible and distinct form of scholarly development, at least in the evaluation accounts analyzed here.

A second category concerns the research infrastructure directly involved in conducting experiments and in other ways generating results (e.g. from databases). Evidence that a research milieu was investing in building up databases for supporting research was put forward as a sign of strength and appropriate priority setting for research, and as a way of focusing research. Direct access to critical and advanced experimental infrastructure, either through local access or through strong ties with national resources, was treated as a strong evaluative point. Easy access to crucial research infrastructure was seen as facilitating the testing and uptake of new methods, and as attracting research talent to the center, while the absence of accounts of such access was considered a failure (at least in a few cases). It is interesting that physical manifestations of research capacity are given such prominence in evaluators' accounts, while the informal aspects of research culture (e.g. casual research alliances and interactions) are almost completely absent. We will return to this phenomenon later, but for now it may suffice to note that research infrastructure and projectified, formalized knowledge-building activities provide a strong signal value to research evaluators.

Epistemic coordination

This category consists of specific social processes considered by evaluators to be central to knowledge creation; specifically activities of centers that facilitate interaction, coordination and integration between individuals and research groups. The first set of activities can be labeled 'research group interaction' and includes items referring to evidence indicating active research cooperation, specifically among groups.

These can take the form of regular meetings to facilitate interaction among groups, the regular movement of researchers among groups, and other types of cross-fertilization. Concrete outcomes of such collaboration are highly valued, as is a culture of collaboration, expressed through the presence of many interrelated activities. One type of recurring topic is academic theme integration, for example through joint supervision of Ph.D. students. This can take the form of cross-group mentoring programs (when mentors come from different groups).

The previous type of process comes close to interdisciplinary facilitation, another epistemic coordinating mechanism. Here the emphasis is on the extent of interplay between different disciplines/sub-fields, expressed through center processes and organizing principles. Concrete activities were especially valued, such as frequent researcher meetings across themes, interdisciplinary theme seminars, internal training in the participating research traditions through journal clubs, Ph.D. courses to facilitate cross-field integration, and research themes that integrate disciplines and not just research projects. Also important was the way center strategy integrated interdisciplinary integration among the participating fields and an associated formal interdisciplinary governance structure involving all contributing specialisms.

Interdisciplinary or sub-specialty interaction in the above sense is a principle for coordination among groups that may otherwise operate independently within an interdisciplinary center, thereby perhaps increasing short-term efficiencies for the group, but foregoing larger epistemic collective benefits in the longer term. The category of capturing intellectual synergies attempts to summarize what evaluators saw as efficacious in seizing the epistemic effects of such interaction. Valued activities ranged from the normal seminar retreats to more innovative mechanisms, such as the creation of a center core curriculum or the use of action groups for time-limited missions, which was considered a way of devolving action capacity and fostering a culture of collaboration. The presence of a strategy for intra-center research dissemination was highly valued, as was the existence of regular research and idea exchange seminars, where, for instance, leaders and groups had the opportunity to identify common topics and bottlenecks to research progress.

This set of mechanisms and processes is close to, but not the same as, critical mass facilitation. This sub-category concerns activities and mechanisms that create synergies among existing groups to achieve critical mass. Critical mass in this instance is a way of describing an outcome where center resources are being assembled to make possible the pursuit of a particular epistemic goal. It may require that some projects cover all the groups of the center, or the coordination among clusters of researchers to enable a certain line of inquiry. Multi-group projects and other activities that bring together several research groups to achieve a particular goal are expressions of critical mass mechanisms.

An interesting aspect of the evaluative statements in this category was the recurrence of another aspect of coordination, which turned out to be prominent enough to receive its own sub-category here, namely voluntary scholarly formation. There were several instances, highly appreciated by evaluators, indicative of organic/bottom-up research group formation within the program topics of the centers. This was expressed in terms of post-doctoral students self-selecting research groups, and movement between research groups stimulated by curiosity or the need to pursue a new line of inquiry. But it was also promoted by mixed models; for example, a procedure of forced circulation of new Ph.D. students among research groups, followed by voluntary affiliation by students to groups; or the presence of an academic com-

munication function (for example, a dedicated cross-area and inter-project coordinator).

Epistemic extension

This category refers to center activities aimed at, on the one hand, extending current research in ways that secure future creative combinations and, on the other, structuring future research and facilitating future research opportunities through funding and other management mechanisms. The praxis of establishing or enabling creative knowledge alliances refers to the first of these types of mechanisms. When the center manages a diversity of funders and fosters a number of cooperative activities, this was viewed by evaluators as a form of stimulation of future diversity and avoidance of lock-ins to existing paths of inquiry. There was similar positive recognition when new and unexpected activities resulted from such funding and alliances.

Another way to view this is as the value of researchers having access to outside knowledge when it is needed (i.e. where potentially useful knowledge alliances are maintained across the boundaries of the center, and this is part of a conscious center policy). These initiatives typically consisted of high-end international cooperation, or strong international relationships based on real cooperation (rather than professed or intended, long-term ties with leading research environments), and documented network ties with other centers of excellence. On the more epistemic side, creative knowledge alliances can also be used to refer to cross-subject/interdisciplinary or transdisciplinary integration of academic inquiry and other sector interests, as well as to integration of the social sciences and natural/engineering sciences. Evaluators typically valued cross-university and cross-faculty collaboration involving the center, and other signs of diversity among cooperative partners. One example of how such effects can be built into the governance structure was internally funded projects being run by researchers from at least two faculties.

A second category of epistemic extension consists of activities where processes and activities directly relating to future research are undertaken or are built into center governance. One may refer to these activities as research futures management. This is where the center actively works with identifiable subjects of consolidation and ‘forward thrust subjects’ for research, based on more or less formal principles and processes. One such mechanism is to use an active and eminent scientific advisory board that identifies future research options, evaluates output in cycles in terms of these, and has a real connection to the research management of the center rather than simply a symbolic presence. The role of leadership in maintaining and supporting research futures is emphasized throughout the evaluations. Typical examples include leadership’s ability to capitalize quickly on unpredicted results and allocate resources accordingly; that is, to exploit unforeseen opportunities by reinforcing certain research projects. It also includes an ability, in making these choices, to maintain reasonable trade-offs between long- and short-term focus in the center’s portfolio of research activities, and to recruit academic personnel to open new fields and support existing ones.

A tangible way of demonstrating that leadership is supporting future research options is to have the right financial mechanisms for such research in place. The notion of prospective research funding implies a clear internal funding strategy that targets renewal in the research portfolio. Examples of this include a certain

percentage of the research budget (say, 10%) being set aside for grant development, or new research projects being funded competitively with originality and novelty being the evaluative criteria. Evaluators gave resource allocation made on the basis of internal proposals a high rating, but only in so far as it supported scientific quality. In this case, this includes involvement of other researchers in the center and long-term sustainability. As a result, this last category spans epistemic production, coordination and extension to some extent.

Discussion and conclusions

The main ambition of this study is to outline the concept of epistemic capacity as a quality of research environments and, using the case of the Swedish Linneus evaluations, to suggest a number of evaluative dimensions. The attributes of this quality,

Table 1. Dimensions of epistemic capacity in research environments

	Supporting epistemic processes and mechanisms	Summary
Production	• Scholarly development	• Research competence development. Mentoring activities supporting learning. Junior–senior integration
	• Research infrastructure	• Infrastructure creation as a way of focusing activities. Infrastructure stimulates uptake of new methods and attracts talent
Coordination	• Research group interaction	• Culture of collaboration across groups. Many concrete interrelated interaction activities
	• Interdisciplinary facilitation	• Coordination between specialisms/disciplines through learning mechanisms. Strategic integration of disciplines
	• Capturing intellectual synergies	• Capturing the effects of collaboration and integration via common topics, core curricula and elimination of bottlenecks
	• Critical mass facilitation	• Mobilization of groups or researchers to work on a particular problem requiring the input of many. Assembling epistemic resources to achieve a particular goal
	• Voluntary scholarly formation	• Freedom of movement and affiliation between groups. Self-selection to pursue topics
Extension	• Creative knowledge alliances	• Managing and incentivizing research partnerships to explore new fields and avoid lock-ins
	• Research futures management	• Organizing 'forward thrust' projects. Evaluating projects in terms of future trends. Recruiting personnel to open new fields.
	• Prospective research funding	• Allocation of funds according to future research prospects and pathways of renewal of the portfolio

as they have been outlined, taken together with the antecedent literature covered, are not exhaustive. However, the evaluative statements synthesized above clearly suggest an outline of the concept as well as several key instantiations of epistemic capacity. Table 1 provides an overview and summary of the concept as developed so far.

Epistemic production is akin to building new knowledge, and it is the most difficult dimension to capture using process indicators since it is usually easiest observed in terms of outcomes. Here the focus is on two separate but closely interrelated aspects of knowledge production: the competence development and knowledge build involved in mentoring and learning from others, and the research infrastructure developed and drawn on to generate new results (i.e. a human and a technological dimension of knowledge production). This dimension can be stretched to include many more aspects, and each sub-category can be deepened and made more specific. The most interesting observation in the present account relates to the relative invisibility of primary knowledge production, such as effort spent on actual experiments and publication, and resorting to social–physical proxies to represent such processes. These resonate with, for example, Scriven and Coryn's (2008) notion of a support net for research. It may also be that this dimension should be sensitized to some of the tensions mentioned previously, namely that between infrastructure/routine and the ease of developing new perspectives, and between the availability of resources and their use in bringing about significant results (Tjissen, 2003). These aspects relate epistemic production to the second main dimension of the framework, epistemic coordination.

Epistemic coordination captures the duality of seeking synergies between human resources in order to pull the mission together (realize strategy, mobilize around key problems and eliminate duplication), and differentiating the mission creatively by encouraging academic freedom under minimal constraint. These ambitions reflect tensions mentioned earlier between academic diversity and integrative activities (Hemlin *et al*., 2004; Heinze *et al*., 2009), and among collaboration, individual autonomy and organizational coordination (Martin *et al*., 2004). The question is how to balance these qualities and when. One way of interpreting the material here is that the centers have been quite successful in fostering intellectual coherence, but less so in fostering the complementary quality of divergence. The encouragement of voluntary movement among groups, self-selection and curiosity-driven affiliation within a larger topical framework ought to be beneficial for research progress in that it taps into the intrinsic motivation of researchers, and to some extent reproduces the ideal of the research community to select and pursue topics freely. Following this line of argument a bit further highlights another critical point for front-line milieus, namely that in some cases the only way to reap the fruits of a collection of field experts is for the coordinator to abstain from direct epistemic leadership in terms of subject matter or even methodological choice. This means yielding a large part of his/her authority to the collective, while at the same time recognizing that part of the center of excellence idea is that researchers in such milieus have somehow signed an epistemic contract, or research agreement, with the center at the expense of their own academic freedom. This is perhaps the essential balancing act for front-line research leaders, and it may be related to Hollingsworth and Hollingsworth's (2000) notion of nurturing integrative leadership.

Epistemic extension is perhaps one of the least expected dimensions in the sense that it relates to activities that are about managing the future rather than focusing

on existing expectations. Capacities relating to an environment's ability to extend current practices and competences in unexpected ways by developing new lines of research and using funding, personnel and alliances to such ends must be central to epistemic capacity. Yet, this is very seldom mentioned in the research evaluation literature. The notions of constrained autonomy (Hollingsworth and Hollingsworth, 2000; Heinze *et al*., 2009) discussed above, as well as the reflection by Huutoniemi (2010) that communication and integration of specialisms should be assessed as they develop (as should integrative skills), are both somewhat future-oriented, but far from as explicitly prospective as the evaluative categories of research futures management and prospective research funding elaborated above.

The evaluation concept of epistemic capacity and the framework proposed here satisfy some of the normative suggestions made in the literature. In particular, two dimensions mentioned previously seem important to such a framework. One is its sensitivity to context and the other is temporal sensitivity. In the case of the former, we may recall Tjissen's (2003) notion that research excellence is framed by the broader context of scientific excellence. In other words, a robust assessment instrument for the core qualities of research must be sensitive to the context of research, which is science and its institutional conditions. The current framework abstracts from research activities, such as experimenting and publishing (often represented as infrastructure availability and publication record in more outcome-oriented evaluations), to other social and strategic activities necessary to accomplish these outcomes (e.g. use of infrastructure for competence enhancement, mentoring and strategic disciplinary integration). This reinforces the idea that creative knowledge environments are established on the physical, cognitive and social level (Hemlin *et al*., 2004).

The latter requirement relates to temporal sensitivity in research evaluations. The framework elaborated here suggests that the present and the future are both valid objects for quality management, and proposes which aspects of these can be evaluated. Klein (2008) correctly suggests that frameworks for research change over time, and likewise the present study highlights how fruitfulness of inquiry may be dependent on time-sensitive evaluation. The above categories of epistemic capacity could be integrated into an outcome model of a center in the short-, intermediate- and long-term, like the one suggested by Trochim *et al*. (2008), where training, collaboration and interdisciplinary integration form sections on a center's knowledge creation cycle. Further elaboration along those lines is beyond the ambitions of this paper.

Finally, a brief but important methodological note: the framework derived above can be taken, on the one hand, to designate preferred ways of communicating about academic quality (evaluators' espoused theory), and on the other to reveal the preferences applied in evaluations (evaluators' theory in use). Either way, these notions of quality end up guiding actual decisions by the funder and they affect the actions of the researcher, as indeed they are meant to. They can be assumed to play a central role in how academic values are promoted, and generally give some indication about how such values are considered by evaluators. As utilized here, they provide content to a potential new evaluative dimension for research centers and programs.

References

Abramo, G., D'Angelo, C. and Di Costa, F. (2009) 'Mapping excellence in national research systems: the case of Italy', *Evaluation Review,* 33, 2, pp.159–88.

Boyer, E.L. (1990) *Scholarship Reconsidered: Priorities of the Professoriate*, Princeton University Press, Princeton, NJ.

Cole, S., Rubin, L. and Cole, J. (1978) *Peer Review in the National Science Foundation*, National Academy of Sciences, Washington, DC.
Davidson, D. (1980) *Essays on Actions and Events*, Clarendon Press, Oxford.
European Commission (EC) (2010) *Interim Evaluation of the Seventh Framework Programme*, European Commission, Brussels.
European Commission (EC) (2012) *Guidance Notes on Project Reporting*, European Commission, Brussels.
Feller, I. (2002) 'Performance management redux', *American Journal of Evaluation,* 23, 4, pp.435–52.
Garfield, E. (1964) 'Science Citation Index – a new dimension in indexing', *Science,* 144, 3619, pp.649–54.
Georghiou, L. and Meyer-Krahmer, F. (1992) 'Evaluation of socio-economic effects of European Community R&D programmes in the SPEAR network', *Research Evaluation,* 2, 1, pp.5–15.
Giorgi, A. (1997) 'The theory, practice and evaluation of the phenomenological method as a qualitative research procedure', *Journal of Phenomenological Psychology,* 28, pp.235–60.
Gläser, J. and Laudel, G. (2007) 'Evaluation without evaluators: the impact of funding formulae on Australian university research' in Whitley, R., Barker, K. and Gläser, J. (eds) *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*, Springer, Dordrecht, pp.101–23.
Heinze, T., Shapira, P., Rogers, J. and Senker, M. (2009) 'Organizational and institutional influences on creativity in scientific research', *Research Policy,* 38, pp.610–23.
Hellström, T. (2011) 'Homing in on excellence. dimension of appraisal in Center of Excellence program evaluations', *Evaluation,* 17, 2, pp.117–31.
Hellström, T. and Jacob, M. (2003) 'Knowledge without goals: Evaluation of knowledge management programs', *Evaluation,* 9, 1, pp.55–72.
Hemlin, S., Allwood, C. and Martin, B. (2004) 'What is a creative knowledge environment?' in Hemlin, S., Allwood, C. and Martin, B. (eds) *Creative Knowledge Environments: The Influences on Creativity in Research and Innovation*, Edward Elgar, Cheltenham, pp.1–28.
Hemlin, S. and Barlebo Rasmussen, S. (2006) 'The shift in academic quality control', *Science, Technology and Human Values,* 31, 2, pp.173–98.
Hollingsworth, R. and Hollingsworth, E. (2000) 'Major discoveries in biomedical research organizations: perspectives on interdisciplinarities, nurturing leadership, and integrated structures and cultures' in Weingart, P. and Stehr, N. (eds) *Practicing Interdisciplinarity*, University of Toronto Press, Toronto, pp.215–44.
Hong, H. and Boden, M. (2003) *R&D Programme Evaluation – Theory and Practice*, Ashgate, Aldershot.
Huutoniemi, K. (2010) 'Evaluating interdisciplinary research' in Frodeman, R., Klein, J. and Mitcham, C. (eds) *Oxford Handbook of Interdisciplinarity*, Oxford University Press, Oxford, pp.309–20.
Kitcher, P. (2001) *Science, Truth, and Democracy*, Oxford University Press, Oxford.
Klein, J. (2008) 'Evaluation of interdisciplinary and transdisciplinary research: a literature review', *American Journal of Preventive Medicine*, 35, 2S, pp.116–23.
Kostoff, R. (1998) 'The use and misuse of citation analysis in research evaluation', *Scientometrics,* 43, 1, pp.27–43.
Leonard-Barton, D. (1998) *Wellsprings of Knowledge: Building and Sustaining the Sources of Innovation*, Harvard Business Review Press, Harvard, MA.
Love, A.J. ed (1983) Special Issue: Developing Effective Internal Evaluation. New Directions for Program Evaluation, 20, pp.1–103.
Martin, B., Allwood, C. and Hemlin, S. (2004) 'Conclusions: how to stimulate creative knowledge environments' in Hemlin, S., Allwood, C. and Martin, B. (eds) *Creative Knowledge Environments: The Influences on Creativity in Research and Innovation*, Edward Elgar, Cheltenham, pp.193–220.
Merton, R. (1979) *The Sociology of Science Theoretical and Empirical Investigations*, University of Chicago Press, Chicago, IL.
Pfirman, S. and Martin, P. (2010) 'Facilitating interdisciplinary scholars' in Frodeman, R., Klein, J. and Mitcham, C. (eds) *Oxford Handbook of Interdisciplinarity*, Oxford University Press, Oxford, pp.386–403.

Pounder, J. (2000) 'Evaluating the relevance of quality to institutional performance assessment in higher education', *Evaluation,* 6, 1, pp.66–78.
Rossi, P. and Freeman, H. (1993) *Evaluation: A Systematic Approach*, Sage, Newbury Park, CA.
Scriven, M. (1991) *Evaluation Thesaurus*, Sage, Newbury Park, CA.
Scriven, M. (2007) 'The logic of evaluation' in Hansen, H. (ed.) *Dissensus and the Search for Common Ground*, OSSA, Windsor, ON, pp.1–16.
Scriven, M. and Coryn, C. (2008) 'The logic of research evaluation' in Coryn, C. and Scriven, M. (eds) *Reforming the Evaluation of Research. New Directions for Evaluation*, 118, pp.89–105.
Stufflebeam, D. (1981) 'Meta-evaluation: concepts, standards and uses' in Berk, R. (ed.) *Educational Evaluation Methodology: The State of the Art*, Johns Hopkins University Press, Baltimore, MD, pp.146–63.
Thomas, D. (2006) 'A general inductive approach for analyzing qualitative evaluation data', *American Journal of Evaluation,* 27, 2, pp.237–46.
Thorngate, W., Dawes, R. and Foddy, M. (2009) *Judging Merit*, Psychology Press/Taylor and Francis, New York.
Tjissen, R. (2003) 'Scoreboards of research excellence', *Research Evaluation,* 12, 2, pp.91–103.
Trochim, W., Marcus, S., Masse, L., Moser, R. and Weld, P. (2008) 'The evaluation of large research initiatives: a participatory integrative mixed-methods approach', *American Journal of Evaluation,* 29, 1, pp.8–28.
Vetenskapsradet (2006) Press releases 2006-07-07, available from http://www.vr.se [accessed 24 July 2012].
Weiss, C. (1997) 'How can theory-based evaluation make greater headway?', *Evaluation Review,* 21, 4, pp.501–24.
Woodman, R., Sawyer, J. and Griffin, R. (1993) 'Toward a theory of organizational creativity', *Academy of Management Review,* 18, 2, pp.293–321.