



University of Lapland

This is a self-archived version of the original article. It may differ somewhat from the publisher's final version, as the self-archived version is typically the accepted author manuscript.

Evaluation practices and impact

Foss Hansen, Hanne; Aarrevaara, Timo; Geschwind, Lars; Stensaker, Bjørn

Published in:

Reforms, Organizational Change and Performance in Higher Education

DOI:

[10.1007/978-3-030-11738-2_8](https://doi.org/10.1007/978-3-030-11738-2_8)

Published: 18.04.2019

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (APA):

Foss Hansen, H., Aarrevaara, T., Geschwind, L., & Stensaker, B. (2019). Evaluation practices and impact: Overload? In R. Pinheiro, L. Geschwind, H. Foss Hansen, & K. Pulkkinen (Eds.), *Reforms, Organizational Change and Performance in Higher Education: A Comparative Account from the Nordic Countries* (pp. 235-266). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-11738-2_8

Document License
CC BY



Evaluation Practices and Impact: Overload?

*Hanne Foss Hansen, Timo Aarrevaara, Lars Geschwind,
and Bjørn Stensaker*

INTRODUCTION

We live in an era when all policy fields and organisations are expected to evaluate their activities (Dahler-Larsen 2012). Not least in universities, multiple evaluation practices have become integrated parts of everyday life. Academic peer review aimed at assessing the quality of publications and the competencies of scholars has been supplemented by other forms of evaluation practices related to accreditation systems and

H. F. Hansen (✉)

Department of Political Science, University of Copenhagen,
Copenhagen, Denmark
e-mail: hfh@ifs.ku.dk

T. Aarrevaara

Faculty of Social Sciences, University of Lapland, Rovaniemi, Finland
e-mail: timo.aarrevaara@ulapland.fi

L. Geschwind

School of Industrial Engineering and Management, KTH Royal Institute of
Technology, Stockholm, Sweden
e-mail: larsges@kth.se

B. Stensaker

Department of Education, University of Oslo, Oslo, Norway
e-mail: bjorn.stensaker@iped.uio.no

© The Author(s) 2019

R. Pinheiro et al. (eds.), *Reforms, Organizational Change and
Performance in Higher Education*,
https://doi.org/10.1007/978-3-030-11738-2_8

performance-based funding systems, as well as rankings imposed on universities from the universities' external environments (Stensaker and Maassen 2015). Actors at the European level, for example, related to the Bologna process, and actors at national levels are both drivers in developing these practices. Other types of evaluation practices, such as student assessments of courses, peer-review evaluation of departments and individual performance assessments, are initiated by the universities themselves (Karlsson et al. 2014).

In this chapter, we analyse the evaluation practices in and around the universities in Denmark, Finland, Norway and Sweden as an approach to better understanding the ongoing changes in the governance of the higher education sector in the region. As in many other regions around the world, the public sector in the Nordic countries has been exposed to a range of reforms in which the state has changed its governance approach, allowing for more institutional autonomy. At the same time, the reforms have introduced and changed other policy instruments in the sector, exposing the sector to strengthened demands for accountability on aspects such as quality, relevance, impact, effectiveness and efficiency (Verhoest et al. 2004; Stensaker and Harvey 2011).

Evaluation is a procedure for assessing how public organisations perform on these aspects (Vedung 2010). It can have different purposes and roles associated with changed governance, ranging from being an instrument of control to being a measure for stimulating formative improvement (Hansen 2005). Disclosing the configuration of the evaluative design present in a given country can accordingly inform us about underlying rationales and logic in the emergent governance of higher education in the Nordic countries. In this chapter, two research questions are addressed: (1) What are the major similarities and differences of evaluation practices across the Nordic countries? (2) What are the experiences of these practices from the points of view of academics and managers? The latter issue is of interest as input to our understanding of the meaningfulness and impact of the evaluative practices implemented. Our focus is on institutionalised evaluation routines. Ad hoc evaluations, for example, following up on implementation of reforms, are not included in the analysis.

The analysis takes a comprehensive approach to evaluation practices. By doing so, it adds to the analyses in the other chapters in Part II of this book. Whereas those chapters delve thoroughly into funding dynamics, managerialism and strategy work, this chapter is an attempt to link these aspects together.

The rest of this chapter is structured in four sections. In the section ‘[Conceptual Framework and Methodology](#)’, a conceptual framework of different types of evaluation models is presented. The framework is used for analysing the types of evaluation practices implemented. Further, this section briefly presents the methodology for analysis. In the section ‘[Mapping Evaluation Practices](#)’, evaluation practices in the four countries are mapped and compared. In the section ‘[Experiences of Evaluation Practices](#)’, academics’ and managers’ views on evaluation are presented and discussed comparatively. The section ‘[Discussion and Conclusion](#)’ holds the conclusion as well as a discussion on further perspectives.

CONCEPTUAL FRAMEWORK AND METHODOLOGY

Within public sector management, there is an increasing interest in how public governance can and should be constructed in more complex and internationally dependent societies (Treib et al. [2007](#)). In general, the concept of governance has implied a change in public management in which the state may allow for private sector actors to have or take a role; a range of instruments, including rules and regulations, voluntary agreements, standardisation and information are applied; and coordination rather than regulation characterises the operating mode (Levi-Faur [2014](#)).

The changing forms of governance often include the following three elements: (1) increased emphasis on institutional autonomy, which is meant to stimulate a stronger organisational actor-hood and improved management (Verhoest et al. [2004](#); Seeber et al. [2015](#)); (2) more emphasis on institutional accountability in terms of quality, relevance and overall performance (Stensaker and Harvey [2011](#)); and (3) the introduction of various evaluative measures to inform, control or stimulate both autonomy and accountability (Levi-Faur [2014](#)).

As such, it is possible to argue that evaluation is a key measure in the new emergent governance patterns in higher education. As discussed in Chap. 2, the literature on evaluation is rich in discussions on how to define the concept. As mentioned earlier, we define *evaluation* as procedures for assessing aspects such as the effectiveness and quality of public organisations’ activities, among others (Vedung [2010](#)). Evaluation can be performed by both public and private actors, and it can have both ‘hard’ and ‘soft’ consequences. Consequences are hard if organisations are sanctioned if they do not meet evaluation criteria; consequences are soft if evaluation routines are implemented as support for learning and quality development

(Weiss 1998). Further, evaluation can be policy driven, managerial or academic in its design (Hansen 2005). Performance-based national funding schemes are examples of policy-driven evaluation. University-driven systems assessing student satisfaction are examples of managerial-driven evaluation, and peer review-based appointment routines are examples of academically driven evaluations. As suggested, evaluations can be conducted at different levels of the higher education system (Stensaker and Harvey 2011), ranging from national systems of quality assurance to evaluation processes that concern universities and their performance or programmes. With *programme*, we refer to an ‘organized, planned, and usually ongoing effort designed to ameliorate a social problem or improve social conditions’ (Rossi et al. 2004: 29), in our case, educational programmes. However, the levels at which evaluations are conducted are indications of where autonomy is found within the system, what this autonomy is used for and the accountability demands associated with it.

Evaluation processes can be anchored in a number of evaluation models which stipulate the question in focus and specify how to set up criteria for assessment. Table 8.1 presents a typology of evaluation models drawn

Table 8.1 A typology of evaluation models

<i>Evaluation models</i>	<i>Questions addressed</i>	<i>Evaluation criteria</i>
Result models		
(a) Goal-attainment	Have goals been realised?	Derived from goals
(b) Effect	Which effects can be uncovered?	Open. All types of effects
Process models		
(a) Implementation	Are there implementation problems?	Assess links from the idea about an intervention to decisions about design, implementation, addressee responses and effects
(b) Activity	Are activity levels increasing/decreasing?	Improvement
Actor models		
(a) Users	Are users satisfied?	Formulated by users
(b) Stakeholders	Are stakeholders satisfied?	Formulated by stakeholders
(c) Peers	How do peers assess quality?	Formulated by peers

from the literature on organisational effectiveness (Cameron 1986), the literature on programme evaluation (Scriven 2003) and the literature covering both types of evaluands (Vedung 1997). The typology is a slightly revised and simplified version of the typology discussed in Hansen (2005).

The evaluation models in the typology are ideal types falling into three categories. The result models are summative. In the classical goal-attainment model, results are assessed according to predetermined goals. In the effect model, the scope is broader, as all types of effects, intended/unintended as well as anticipated/unanticipated, are assessed in principle. The process models are formative and explanatory, and the actor models are anchored in the different actors' own evaluation criteria.

All models can be said to represent different modes of governance, where result models in general are associated with more traditional hierarchical governance, process models are often associated with more horizontal and community-oriented governance modes and actor models are more related to market and user-type governance approaches (Treib et al. 2007). Of course, in practice, evaluation designs may often be hybrid phenomena drawing on several models. However, analytically, the models are fruitful tools for uncovering the regulatory logic behind the evaluation routines (Levi-Faur 2014). As such, they are used later for a comparative analysis of the country practices.

The analysis of the evaluation practices is based on several types of data. The mapping of the practices is based on official documentary material such as governmental reports, as well as on available scholarly analyses. The analysis of the experiences of evaluation practices are based on survey as well as interview data collected as part of the FINNUT-PERFACAD project (see Chap. 1 for a more elaborated presentation).¹ In this chapter, we use survey data for the comparative analysis of academics' experiences with evaluation practices and then use interview data in a supplementary and illustrative way to shed light on the dynamics and experienced impacts of evaluation practices of both academics and managers.

MAPPING EVALUATION PRACTICES

In this section, evaluation practices are mapped by country in order to shed light on similarities and differences across countries. In the mapping of the practices, we use the distinctions discussed earlier: policy-driven,

¹The survey data included in this chapter has been used also in an analysis of accountability relationships (Hansen et al. 2019).

managerial-driven and academic-driven practices. Further, we map evaluation practices related to education activities, research activities and other types of evaluands. In the comparison across countries, we use the typology of evaluation models to address the discussion on similarities and differences.

Denmark

There are eight universities in Denmark. Although the universities are very different regarding age, size, profile and structure, the policy-driven external evaluation practices they are confronted with are very much alike. However, the universities have considerable leeway regarding how to implement external evaluation practices in organisational routines and how to initiate and implement internal evaluation practices. The overall pattern of evaluation practices at Danish universities is presented in Table 8.2.

Policy-driven evaluation practices in Denmark are first and foremost related to performance-based funding streams. Such systems evaluate performance using indicators. When indicators have been decided on, such systems are implemented rather mechanically. In relation to education, nearly all resources have been linked to performance, concrete to the number of students passing exams, since the 1980s. Since 2009, this measure has been combined with bonuses given if students accomplish their studies in timely fashion (DEA 2011). Bonuses constitute nearly 10% of the total amount of funding for educational purposes.

Recently, the funding formula has been further developed, as the Parliament has decided on a new formula for 2019 which, besides student throughput, includes employability, goal attainment (according to contracts negotiated between the government and the individual university) and quality aspects. The quality aspects in the formula have yet to be decided, with student and graduate satisfaction and maybe teacher assessment of quality as dimensions being considered. The performance-based funding formula is, thus, becoming increasingly complex, and it seems more tightly politically governed.

In relation to research, approximately 20% of the total amount of ordinary funding is distributed according to a performance-based formula that includes the number of graduates from master's and PhD programmes, the ability to attract external funding and the number of published research publications. The future plan of the government is to distribute resources

Table 8.2 Overall pattern of evaluation practices at Danish universities

<i>Evaluand</i>	<i>Evaluation 'owners'</i>		
	<i>Policy driven</i>	<i>Managerial driven</i>	<i>Academic driven</i>
Education	<ul style="list-style-type: none"> • Funding formula emphasising student throughput, employability, goal attainment according to contracts and quality • Pre-evaluation and approval of new programmes • Accreditation of quality assurance systems and programmes • National database including information about students' assessment of quality and graduates' assessment of relevance, among other things 	<ul style="list-style-type: none"> • Students' satisfaction evaluation systems • Stakeholder curriculum dialogue with focus on relevance 	<ul style="list-style-type: none"> • Classroom evaluation dialogue with students • Examination of students
Research	<ul style="list-style-type: none"> • Partly performance-based funding based on several criteria including bibliometrics 	<ul style="list-style-type: none"> • Appointment and promotion decisions • Peer-review exercises (department level) • Human resource initiatives (systems for evaluation of managers, routine manager-individual staff dialogue) 	<ul style="list-style-type: none"> • Classic peer review (especially publication and funding and, to some extent, appointments) • Other scholarly qualifications such as third-stream activities in relation to appointments
Other evaluands	<ul style="list-style-type: none"> • Development contracts laying down institutional goals and followed up by goal-attainment evaluation • Since 2017, an annual report on HE- and research performance 		

on the basis of a new formula emphasising quality over quantity, but how to do this has not yet been decided.

According to the principle stated in the university law that universities are self-governing institutions, resources are distributed to the universities as lump sums. This implies that the university boards are responsible for the internal distribution and the principles for use of the resources. This again means that the individual university has leeway in deciding on the budget model and, thus, is able to strongly influence whether and how the

incentives in the funding formulae are implemented onwards in the organisation.

Further, in relation to education, the universities are confronted with evaluation practices built into accreditation requirements. In 2007, an accreditation scheme aiming at approving all bachelor's and master's programmes—new ones as well as existing ones—was established. When implemented, the system was criticised for being very bureaucratic. In 2013, it was decided to turn the accreditation regime into a system for improving the quality assurance systems at the universities, emphasising both programme quality and relevance. This regime transformation is still ongoing. In Denmark, the evaluation of PhD programmes is not included in the accreditation system but is handled in a more ad hoc fashion.

Both performance-based funding formulae and the accreditation system are laid down in political agreements, accreditation being a national political response to the Bologna process. However, considering both funding formulae and accreditation as evaluation practices, it becomes obvious that these represent hybrid evaluation models. Whereas the most recent teaching funding formulae combine different types of result models (e.g. goal-attainment, related to contracts, and effect, related to employability, combined with a user model—probably student satisfaction), the research funding formula combines a result model (graduates) more or less indirectly with peer review (external funding and publications). Finally, the accreditation scheme combines a user model, including both students and labour market representatives, with a peer-review model.

In recent years, a national database for comparison across educational programmes, UddannelsesZOOM, has been developed. The database holds information about study elements and dropout rates, for example, but also includes data on students' assessments of quality and graduates' assessments of relevance.

Development contracts have, for many years, been a political steering instrument. The contracts do lay down both national goals and individual university goals, and goal fulfilment has been monitored. Funding has not previously been a part of the contract regime, but as mentioned earlier, this will be the case in the future.

The development contracts are also used as *managerial-driven evaluation practices*, as faculties and departments are asked to deliver into goal fulfilment. Besides this, managerial evaluation practices are first and foremost related to education. Student satisfaction evaluation is a routine exercise. Evaluations of educational programmes by graduates and stake-

holders in light of the labour market requirements are carried out in a more ad hoc fashion.

From time to time, some universities conduct research evaluations of departments by flying in international peers. The peers produce an assessment report which may provide input on, for example, strategy processes, but in some situations seems to be more symbolic, legitimising ongoing activities. At the individual level, managers are obliged to conduct what are called ‘staff development conversations’ on an annual basis. Some managers use this occasion to evaluate staff performance related to publication activities and activities aiming at attracting external funding. In recent years, recruitment and promotion procedures related to academic staff have been changed. Traditional peer review is still part of these processes, but, today, heads of departments and deans have much more influence on these evaluation processes, as well as on the subsequent decisions.

Thus, the importance of *academic-driven evaluation practices* in relation to recruitment and promotion has been challenged. Collegial processes as well as collegial bodies have become advisory and not, as before, decision-making bodies. In the case of evaluation activities related to key academic activities, teaching and research, academic-driven evaluation practices are, however, very important. Examination processes related to learning outcomes have become more structured. The same goes for PhD programmes and the monitoring and evaluation of PhD students’ activities. In research, peer review constitutes the core of evaluation related to publishing and funding decisions.

Finland

There are 15 universities in Finland, 14 of which are under the Ministry of Education and Culture, and the direct government core funding they receive is about half of their total funding. For these 14 universities, the evaluation practices are generally based on legislation, but the universities also influence the evaluation practices themselves. The overall pattern of evaluation practices at the Finnish universities are presented in Table 8.3.

Common *policy-driven evaluation practices* have been developing since the 1990s in which the idea of Finnish universities accepting responsibility for the quality of their research and education has gained ground. A common Universities Act in 1997 replaced separate University Acts that had regulated each university in earlier years. This was the starting point for a

Table 8.3 Overall pattern of evaluation practices at Finnish universities

<i>Evaluand</i>	<i>Evaluation 'owners'</i>		
	<i>Policy driven</i>	<i>Managerial driven</i>	<i>Academic driven</i>
Education	<ul style="list-style-type: none"> • Funding formula emphasising focus areas, employability and national duties • Field-specific funding and external funding for specialisation studies and non-degree programmes • Quality assurance 	<ul style="list-style-type: none"> • Contracting, performance agreements, demand for relevance, student credits, student mobility 	<ul style="list-style-type: none"> • Traditional quality assurance, graduates' employment, student feedback after bachelor's degree
Research	<ul style="list-style-type: none"> • Policy initiatives as the government's key projects and reforms, including direct government funding for strategic development for national education and science policy aims 	<ul style="list-style-type: none"> • Performance agreements (internal and external), corporate funding 	<ul style="list-style-type: none"> • Scholarly publications (JuFo), international and domestic competitive research funding
Other evaluands	<ul style="list-style-type: none"> • State of Scientific Research in Finland (report) 	<ul style="list-style-type: none"> • Human resource management 	<ul style="list-style-type: none"> • Regulation-based university evaluation of research

university system in which evaluation could be a system-guiding and economic factor. Since 2005, the Finnish universities have been implementing the European principles of quality assurance at institutions and at the national level. The Finnish universities apply quality as a key element in evaluating higher education. The Universities Act (2009) obliges universities to participate in the evaluation of their operations and quality systems, but the universities also decide on their quality systems. The Finnish system is based on the idea of developing quality systems (enhancement-led evaluations) that correspond to the European principles of quality assurance (FINEEC 2016).

The role of the Finnish Education Evaluation Centre (FINEEC) is crucial in these institutional audits. FINEEC is a semi-independent institution funded by the Ministry of Education and Culture, and it produces the information for knowledge-based decision-making in the development of education. The process is based on enhancement-led principles to reach

universities' strategic goals. The final results of the university audits are decided by the Higher Education Evaluation Committee of FINEEC.

Further, the Ministry of Education and Culture conducts performance negotiations with each university annually, and indicators defined in the funding formula since 1998 play a key role in these negotiations (Hicks 2012). Even as recently as the early 2000s, these indicators were mostly quantitative, but the quality factor has been emphasised more in recent years. The significance of 'quality' in the Ministry's 2017 funding formula has a strong correlation with key performance indicators: education is 39%, research is 33% and other education and science policy considerations are 28% of the total government core funding.

The field-specific funding emphasises art, engineering, natural sciences, medicine, dentistry and veterinary medicine. Ten percent of the funding depends on the number of students who complete 55 study points a year. This is to improve performance by reducing the time spent in formal study. Two percent of funding is based on the number of employed graduates, and 3% is based on student feedback. All in all, the Finnish development reflects the international story of university quality assurance in the 2000s being about impact, quality and internationalisation and governments accordingly changing funding formulae (Jongbloed and Vossensteyn 2016).

In research, quality determines 9% of the total core funding. The model includes indicators for competitive research funding and corporate funding. The funding formula also includes other education and science policy matters covering strategy development, field-specific funding and national duties. These indicators in institutional negotiations are partly quality and partly impact and internationalisation.

Based on regulation, the Finnish Academy is responsible for the evaluation of research in Finland. A report, 'The State of Scientific Research in Finland', is launched every second year, aiming to strengthen knowledge-based policymaking in science policy. The report contains analyses of research personnel, funding, scientific impact, bibliometric analysis and co-publications. There are also comparisons with the most important reference countries.

The policy basis described earlier applies to the internal funding allocation of universities. However, *managerial-driven evaluation practices* are varied. Some universities follow the indicator of external financing as closely as possible. For some other universities, however, the internal allocation model is based primarily on historical allocations. The role of quality in performance management is strong for the universities that follow

external funding indicators in their internal allocations (Aarrevaara et al. 2018). University tools for the implementation of internal allocation in education include performance-based negotiation or the determination of block grants on a historical basis through a performance contract or a combination of these models (Pruvot et al. 2015).

Performance management and quality assurance started to be accepted almost simultaneously at Finnish universities in the 2000s. These two factors have strengthened the role of universities as autonomous institutions. However, *academic-driven evaluation practices*, academic traditions and collegial practices are still strong factors in elements of quality management. Some of the funding formula indicators relate to a publication forum, JuFo, classifying publications at three qualitative levels. The publication forum is a system maintained by scholarly communities. The importance of degree qualifications, academic evaluation and, in particular, peer-review practices are central to the university system. The Universities Act (2009) obliges universities to evaluate these practices regularly every year in negotiations between the universities and the Ministry of Education and Culture.

Norway

There are currently ten universities in Norway. These universities are very different according to age, size, profile and structure, but the policy-driven external evaluation practices they are confronted with are very similar. The universities still have considerable autonomy concerning their own internal evaluation practices, although some types of evaluations are required by law (Stensaker 2014).

The overall pattern of evaluation practices at Norwegian universities is presented in Table 8.4.

Policy-driven evaluation practices in Norway are first and foremost related to performance-based funding streams. The number of credit points taken determines 25% of the total budget. This system was introduced as part of a major reform in 2003 intended to strengthen the quality, relevance and efficiency of Norwegian higher education (Stensaker 2014). Due to increased criticism of credit points as the key indicator for educational performance, recent changes include the introduction of a new indicator related to study programme completion. However, this indicator has so far only been linked to a small amount of the performance-based funding. A further expansion of possible indicators included in the

Table 8.4 Overall pattern of evaluation practices at Norwegian universities

<i>Evaluand</i>	<i>Evaluation 'owners'</i>		
	<i>Policy driven</i>	<i>Managerial driven</i>	<i>Academic driven</i>
Education	<ul style="list-style-type: none"> Partly performance-based funding (25% of total budget) emphasising student credits taken Institutional accreditation of all universities Accreditation of study programmes in university colleges at MSc and PhD level National student satisfaction survey (mandatory for institutions to participate in) 	<ul style="list-style-type: none"> Institutional QA systems, including routine reviews of study programmes Stakeholder forum where institutions and representatives of employers discuss relevance of educational offerings (mandatory) 	<ul style="list-style-type: none"> Student evaluation of teaching (mandatory to conduct, but autonomy is given regarding design) External examination system administered decentrally by the individual institution (mandatory)
Research	<ul style="list-style-type: none"> Partly performance-based funding (15% of the total budget) A national system for bibliometric registration of research output (used as input to performance-based funding) National evaluations of disciplines (rotating system) 	<ul style="list-style-type: none"> The development of performance reports on research output 	<ul style="list-style-type: none"> Classic peer review regarding external funding of projects, publications and academic appointments (mandatory)
Other evaluands	<ul style="list-style-type: none"> Development contracts currently piloted in some institutions A national database containing key performance indicators on a range of dimensions enable the possibility for institutional benchmarking An annual report on R&D, innovation and HE performance, indicators etc. 	<ul style="list-style-type: none"> Human resource initiatives (systems for evaluation of managers, routine manager-individual staff dialogues) 	

performance-based funding system is likely, though it might be in relation to recent experiments with developmental contracts between the Ministry of Education and individual institutions. In relation to research, approximately 15% of the total amount of ordinary funding is distributed according to a performance-based formula for research output regarding the number of publications, external funding from the EU and so on. The research-output funding of research is based on input from a national database for academic publishing called Cristin.

A national system of institutional accreditation is also an important element in the policy-driven education practices in Norway. This system accredits all public and private higher education institutions (HEIs), and accreditations determine the degree of institutional autonomy provided. For example, being given the status of a university implies full autonomy concerning the establishment of new study programmes at all levels. In Norway, PhD programmes are included in the accreditation system—with respect to both the institutional accreditation system and when university colleges without the independent right to establish PhD programmes apply to the national quality assurance agency for specific recognition.

It should be noted that a considerable portion of these resources is distributed to universities in the form of lump-sum funding (60%), allowing the institutions significant autonomy concerning their strategic development. Formally, university boards are responsible for the internal distribution and the principles for using economic resources, and the boards also have full autonomy regarding how the university should be internally organised. This again means that the individual university has considerable leeway in deciding on the budget model and, thus, is able to influence whether and how the incentives in the funding formulae are implemented onwards within the organisation.

Additional policy-driven forms of evaluation play an important role in the Norwegian higher education system. First, the Research Council conducts—on a rotating basis—their independent national assessments of specific disciplines, a practice that tends to have implications for how later external research funding schemes and programmes directed at these disciplines are designed. Partly linked to these evaluation processes, the Council also produces an annual report on R&D outputs, staff, performance, innovation and so on for the entire Norwegian higher education system.

More recent policy-initiated evaluations include a national student satisfaction survey conducted by the national QA agency, in addition to

experiments with national exams in particular disciplines aimed at establishing and upholding national academic standards.

Compared to the relatively high number of policy-driven evaluations, there are relatively few *managerial-driven evaluation practices* in Norway. Perhaps the most influential of these is the institutional QA system, mandatory for every university and college to establish, in which issues such as management, formal responsibilities and evaluation are central (Stensaker 2014). These QA systems have been a central element of the national accreditation system in place since 2004 and are tightly linked to the external checks conducted by the national QA agency at an institutional level every sixth year. In some institutions, this QA system has been expanded to the area of research and has become more integrated into the regular steering of the institutions.

Every higher education institution in Norway is also required to establish a formal forum where institutional representatives and representatives from employers evaluate and advise the institutions about the relevance of the educational offerings. Finally, several formal evaluations related to an expanding and more professional administration and the perceived need for more data and knowledge in the area of human resources management have been developed administratively within universities.

Academic-driven evaluation practices have become a more common phenomenon in Norway since World War II. However, an external examiner system had already been institutionalised decades earlier. Also, student evaluation of teaching has a relatively long history in Norway, although these activities were later incorporated into the institutional QA systems. While it is mandatory for institutions to conduct student evaluations of teaching according to national regulations, there is significant room for individual autonomy concerning how these systems are designed and implemented.

As in many other Nordic countries, a considerable number of evaluation activities are conducted as part of the daily running of the universities, including academics' involvement in evaluating the possibility of being granted externally funded research projects, evaluations in relation to academic appointments and so on. While universities have started to change the ways in which academic staff is recruited, there is still significant focus put on the traditional peer-review procedure in Norwegian higher education.

Sweden

There are 15 public and 2 private universities in Sweden. In addition, there are more than 30 university colleges providing higher education. There are many differences across the sector when it comes to size, scientific scope and relation between teaching and research. All universities report to the government, and their operations are regulated by the laws and statutes that apply to the higher education sector. They are nationally evaluated by the Higher Education Authority (UKÄ). In addition, they have the responsibility of initiating and undertaking their own evaluations (Table 8.5).

The *policy-driven evaluation practices* in Sweden are partially based on metrics. In education, a funding system based on the input and output of students was introduced in 1993. Approximately half of the money is allocated upon student admission, the other half when students complete their studies. This system does not include any evaluation component but,

Table 8.5 Overall pattern of evaluation practices at Swedish universities

<i>Evaluand</i>	<i>Evaluation 'owners'</i>		
	<i>Policy driven</i>	<i>Managerial driven</i>	<i>Academic driven</i>
Education	<ul style="list-style-type: none"> • Higher Education Authority (UKÄ): accreditation, institutional audits, reviews of subjects and programmes 	<ul style="list-style-type: none"> • HEI-initiated educational assessment exercises, institutional standardised systems for course evaluation 	<ul style="list-style-type: none"> • Collegial bodies' decisions on new courses and programmes, teacher-driven evaluation of courses, assessment and examination of students
Research	<ul style="list-style-type: none"> • Bibliometrics: publications, citations and external funding; national evaluations of subjects and programmes (like centres of excellence) 	<ul style="list-style-type: none"> • HEI-initiated research assessment exercises, performance indicators; ad hoc benchmarking within networks 	<ul style="list-style-type: none"> • Classic peer review (funding, publication, appointments)
Other evaluands		<ul style="list-style-type: none"> • Human resources initiatives (systems for evaluation of managers, routine manager-individual staff dialogue) 	

rather, is mechanical in character. However, the assessment of students' performances indirectly affects the outcome. This system has, at times, been criticised and is currently once again under review. The critical remarks are related to quality; the argument put forward has been that academic staff become pressured to lower standards in order to get students through educational programmes. This has been particularly prevalent in areas and institutions where student admissions have been less competitive.

The national evaluation of higher education is under the authority of the UKÄ, which undertakes accreditations of new programmes, subject and programme reviews, thematic evaluations and institutional audits. It evaluates education at all three levels. The focus of the evaluation system has varied over time but has always comprised control, enhancement and information as the three main aims, currently with more emphasis on enhancement-driven institutional audits as compared with the previous, more control-oriented subject and programme reviews.

In research, a performance-based funding system was introduced in 2009 based on indicators: external funding, citations and publications. This was a major shift in Swedish research policy and part and parcel of a larger reform agenda. Over time, 10–20% of the total funding has been allocated on the basis of these indicators. Until then, the direct state funding was allocated entirely based on the size and number of academic staff, sometimes referred to as 'historical principles'.

Managerial-driven evaluation practices are also in place. In addition to the national systems, there have been a number of initiatives at the institutional level, both in education and research. The national policy-driven systems described earlier have, to varying degrees, 'trickled down' to the universities (Hammarfelt et al. 2016). The first research assessment exercise at the institutional level was launched at Uppsala University in 2007 (called Quality and Renewal), which was subsequently followed by a number of similar exercises at other institutions. They have all used peer-review panels and bibliometrics as standard procedure. Some of the Swedish universities have repeated the exercises more than once, usually with a slightly altered methodology (Bomark 2016). They have, for instance, put more focus on research environments (Quality and Renewal, Uppsala University 2017) or societal impact (RAE, KTH Royal Institute of Technology). The outcomes of the evaluations have been used in various ways: for funding allocations (e.g. in the case of bonuses), for further support (for less-than-excellent environments) or simply for recognition (Karlsson 2017).

Similar initiatives from the HEI management have been taken on the educational side. This was particularly the case during a period when some leading universities, in light of more formal autonomy and a criticised national evaluation system, decided to instigate their own subject and programme reviews (Karlsson et al. 2014). The variation in aims and methodology was even larger than in the case of research. These reviews were also affected by the research assessments, contributing to a general feeling of evaluation fatigue in Swedish higher education. The current national system puts much emphasis on the universities' own responsibility for assuring and enhancing quality. This might indicate more management-led evaluations in the near future, both in the form of large-scale, comprehensive exercises like the ones mentioned earlier and by introducing more formalised quality assurance cycles and processes on an annual basis.

Although both policy- and managerial-driven evaluation practices have been intensified, *academic-driven evaluation practices* do not seem to have decreased, either in number or in importance. Many of them are intertwined with managerial practices, such as the hiring of academic staff and promotion decisions. The relation between the collegial bodies at Swedish universities and the line management has been discussed quite frequently in the last decade. As a consequence of the autonomy reform in 2011, collegial bodies have, like in Denmark, become more advisory than decision-making. However, all major decisions regarding academic core activities shall still be made by academically qualified staff, according to the Higher Education ordinance.

Evaluation practices regarding academic activities start with the assessment of students. This is an area in which academic judgement is key, and traditionally, this has been under the discretion of academic staff themselves with few guidelines. Increasingly, the assessment of students has become more structured and more related to intended learning outcomes, a development fuelled by the Bologna process. This development is intended to increase the transparency of the assessment process. Furthermore, doctoral education has also become rationalised and structured, moving away from the traditional master-apprentice model to becoming an education based on a curriculum, structured supervision, occasional mentoring and individual study plans to be followed up at least once annually.

In the research realm, peer-review activities seem to grow continuously due to an increasing number of publications, journals, book publishers and conferences. Official reports from the Swedish Research Council show that Swedish researchers are indeed very productive in terms of publica-

tions but also that their papers are not as cited as the other top nations (Vetenskapsrådet 2017).

Closely related to both policy-driven and managerial practices are the evaluations performed in research councils and private foundations that inform decisions on funding. Typically, leading professors form peer-review panels who grade research applications. Since a large portion (55%) of Swedish public research funding is external and competitive, these activities are important (Geschwind 2017).

Comparison

The mapping of evaluation practices in the Nordic countries has revealed both similarities and differences across countries. In this respect, our findings are in line with former studies focussing more narrowly on national evaluation systems related to education (Hansen 2014; Schmidt 2017). In all four countries, policy- and managerial-driven evaluation practices are widespread. And these types of evaluations have become more important compared to academic-driven practices. Further, policy-driven evaluation practices seem to include more performance elements across time as still more indicators and measures are developed and included in evaluation practices.

The process can be observed in relation to both education and research evaluation. In evaluation practices related to education, student activity measures have become supplemented with measures focussing on timely student throughput, student satisfaction and employability. And in evaluation practices related to research, classical academic peer-review routines have been supplemented with bibliometric measures in the form of publication and citation counting, as well as with an attention to impact. This development has raised discussions on whether the performance-based funding systems have promoted quantity more than quality. In recent years, there seems to have been a turn towards giving more weight to quality.

There are, however, differences across countries regarding which new measures to adopt and when to adopt them. The Nordic countries seem to monitor each other's evaluation practices as well as those of other northern European countries, thereby seeking inspiration for further developing their own practices. Differences across countries are also seen in how they have responded to the evaluation demands implemented according to the Bologna process. In Denmark, the accreditation system,

although now in transformation, has constituted a hard regulation system stating that every educational programme has to be accredited. In Norway and Sweden, universities have more authority to establish new programmes. In Finland, a softer, enhancement-led quality assurance approach, rather than an accreditation approach, has been adopted, and in Sweden, institutional audits have now become the main feature of the national system, evaluating the quality assurance systems rather than the quality itself.

Further, performance-based research funding schemes are anchored in different approaches. Denmark, Finland and Norway have schemes focusing on counting publications, while Sweden's scheme centres on assessing the number of citations. Although the publication counting schemes look similar at first sight, emphasising scholarly publications and competitive funding, there are important differences. For example, the transparency in the Norwegian system is much stronger than in the Danish system. This fact probably makes it much easier to use the performance information in the system for other purposes than the official purpose related to the redistribution of resources at the national level.

In Norway and Finland, national research evaluation systems driven by the Norwegian Research Council and the Finnish Academy have been developed, whereas the approaches in Sweden and Denmark have been university led or faculty anchored.

Table 8.6 summarises our findings on evaluation models; result models, process models and actor models are in use.

Overall, similar evaluation models have been implemented across the Nordic countries. We find result, process and actor models in use in all four countries. At the same time, however, there are many differences in the details. For example, activity-based educational funding formulae are used in one way or another in all four Nordic countries. In Denmark and Norway, the number of credit points taken are important; in Sweden, the number of completed degrees is prioritised. In Denmark and Norway, the emphasis on timely student throughputs is more intense than in Finland and Sweden.

Activity- and effect-based funding formulae are also in use in the allocation of resources to research. Research funding formulae, however, also differ across countries. For example, Sweden is the only country that includes effect indicators in the form of citation counts.

In all the countries, we also find examples of policy-driven indicators reflected in managerial practices at both the organisational and individual

Table 8.6 Evaluation models in use

<i>Evaluation models</i>	<i>Used in evaluation procedures</i>	<i>Country</i>
Result models		
• Goal attainment	• Contract steering	• DK, FI, NO
• Effect	• Employability focus	• DK, FI
	• Ability to attract external funding, influence funding for research	• DK, FI, NO, SWE
	• Citation patterns influence funding for research	• SWE
Process models		
• Activity	• Funding formula education	• DK, FI, NO, SWE
	• Formula including student throughput	• DK, NO
	• Funding formula research (publications)	• DK, FI, NO, SWE
Actor models		
• Users	• Student surveys	• DK, FI, NO. Gaining importance at the national level
• Stakeholders	• Relevance in education	• DK, NO
• Peers	• Research evaluation	• All countries. Classic use in relation to publishing and funding decisions. Challenged in relation to appointment and promotion decisions

employers' level, as some universities implement national indicators in internal resource allocation, hiring and firing procedures and decisions concerning individuals' reward-based salaries. However, universities apply these practices in very different ways.

EXPERIENCES OF EVALUATION PRACTICES

The previous section mapped and compared evaluation practices across the Nordic countries. In this section, the focus is on how academics experience these practices. Key questions on this aspect include: Do they find it legitimate and meaningful? How do they experience its impacts?

Academics' Views on Evaluation: Meaningful?

The survey data (Table 8.7) show that Nordic academics perceive evaluation as a fairly legitimate task. However, there seems to be some misalignment between the personal opinions regarding academic performance and measured academic performance. Evaluation is experienced most negatively in Denmark. In particular, there is a striking difference between Denmark and the other Nordic countries in perceiving measurement as a sign of mistrust.

Many interviewees reflected upon the meaning of the growing number of evaluations. An interviewee from Sweden described how policy-driven evaluations fuel managerial-driven evaluations, as some universities undertake their own evaluations to prepare for the national reviews carried out by the UKÄ:

There are, like, so many evaluations. Before UKÄ, there are also a couple of internal evaluations, etc. (Flagship, Social science)

A department head from Denmark (flagship, natural science) gave voice to an experience of evaluation overload. This person felt that too much evaluation is conducted, with some evaluation procedures being principles without specific purposes. Further, the same individual had experienced evaluation procedures taking a lot of time but seldom producing new knowledge.

Various evaluations on research, teaching, university activities and the innovation system in Finland have also caused a lot of work for the interviewees. Although various evaluations produce legitimacy according to

Table 8.7 Respondents' views on the legitimacy of evaluation and measurement (percentage of those who answered 4 (*agree*) and 5 (*strongly agree*))

	<i>Denmark</i> (%)	<i>Finland</i> (%)	<i>Norway</i> (%)	<i>Sweden</i> (%)
Control and evaluation of my work is a legitimate task	42	46	49	51
Internal procedures for measuring academic performance are in accordance with my understanding of academic performance	18	25	22	23
In my opinion, performance measurements are signs of mistrust	47	23	36	28

the survey results, the knowledge base of these evaluations is subject to much criticism. The main argument of the criticism is that the information used does not sufficiently support the academic tasks. One of the leaders of an academic unit stated:

I feel that our own observations, development work, monitoring and student feedback, locally, and our general evaluation are more useful compared to the management system, even at the faculty level. (Regional, academic leader, sciences)

One of the Swedish interviewees belonging to the social sciences reflected upon the implications of changes in evaluation practices and experienced them as mistrust:

I think it used to be part of the profession to be able to ... just like teachers, and there has been a deprofessionalisation, that's for sure. [...] Well, they (academic assessments) have been replaced by these quantitative evaluation systems rather than showing confidence in those who are educated to the assessments themselves. (Flagship, social sciences)

When asked about the degree of changes regarding evaluation and accountability issues in the last decade, all the groups reported an increased focus on this. In Norway, an administrator stated:

[...] as management, we are required to have a little more accountability by our owner, KD [Ministry of Education], than we were 10 years ago [...] But we are working hard to fulfil what we think are the orders we have received. Also, in other areas as well, I feel that we must be more responsible for the good management of human resources [...] We have had financial problems. There is more focus [now] on having proper management and control processes. It must be quality assured ... We got a quality assurance system from Bologna. So, yes, I really feel that it has become more [focus on evaluation]. (Flagship, administrator)

Still, there were large variations in how meaningful this development was. Some found evaluation meaningful:

I am stimulated by the demand for higher performance. I want to do more. I get motivated to do more when people around me appreciate what I am doing and give feedback. (Regional, manager, sciences)

Others were more detached and did not pay attention to this regime, while some were more critical, questioning the role of universities as independent institutions if they were met by such indicators and questioning the impact from New Public Management.

Academics' Views on Impacts of Evaluation

Table 8.8 shows that academics from all the countries are quite pessimistic about the positive impacts of measurement and evaluation, with respect to both performance and the atmosphere at work, regardless of the fact that they consider evaluation a rather legitimate activity. This observation holds true for both research and teaching tasks. Denmark differs most from the other three countries, particularly concerning research rather than teaching. In the perceptions of impacts of research performance measurement on work atmosphere specifically, the Danish figures are considerably lower than the figures for Sweden, especially, but also for Norway.

The interview data illustrate a range of different impacts of *educational evaluation*, from non-positive to positive. A Danish academic (flagship, natural science) characterised the teaching team to which he belonged as anarchistic. They routinely conduct student evaluations using surveys.

Table 8.8 Respondents' views on the impact of evaluation and measurement (percentage of those who answered 4 (*agree*) and 5 (*strongly agree*))

	<i>Denmark</i> (%)	<i>Finland</i> (%)	<i>Norway</i> (%)	<i>Sweden</i> (%)
Measurements increase my performance in teaching	13	16	18	18
Measurements increase my performance in research	16	26	26	28
Teaching performance measurements have a positive impact on the atmosphere surrounding academic work	11	12	16	19
Research performance measurements have a positive impact on the atmosphere surrounding academic work	11	15	19	26
Control and evaluation of my work has a positive impact on my teaching performance	16	20	23	29
Control and evaluation of my work has a positive impact on my research performance	14	23	22	26

These results are read, and sometimes colleagues have fun doing that, but in his experience, the evaluations do not influence practice.

Others, however, characterised the routine student evaluations as a kind of fire alarm. If an evaluation uncovered problems, action was necessary. Sometimes, not-so-good evaluations also had the consequence of the responsible programme leader having to explain upwards in the hierarchy: 'Up to father and over the knee', as one (flagship, academic, social science) phrased it. In this way, evaluation routines can be seen as enhancing accountability.

Some of the Swedish interviewees also reflected upon the positive impacts of the external educational evaluations carried out by the UKÄ. They saw the evaluations as opportunities to work with quality development at the organisational level:

And actually then, there has been good quality work as a consequence. Although nobody thinks this last evaluation system has been particularly good, the result has been good ... You got an opportunity to reflect and go through the education holistically and scrutinise this with successive progression and coherence, etc. (Flagship, Social Sciences)

If a programme faces a negative outcome in an external evaluation, this seems to lead to extensive internal activity, as reputation could be threatened:

Well, of course, that we made it the first round was considered kind of good. I know my colleague at the XX programme, the programme director, when they failed the first round, they had a huge amount of work inwards in the organization. (Flagship, Social Sciences)

Also, if national evaluations are linked to funding, and good performance is rewarded with extra money, impact was thought to increase:

Well, when there was talk about linking funding to evaluation, then people really got busy. (Regional, Social Sciences)

Further, some interviewees also reflected upon the impact of educational evaluations at the system level. Here, the experience was that evaluations have led to a greater awareness of who is good and who is less good. This is particularly the case in evaluations where grading is used:

One can only see that we had this [name of evaluation]; it led to an increased awareness of: they are good, they are less good. Even if it's not exactly a competition, I think it leads in that direction. (Flagship, Regional)

As for the Norwegian academics, they are also performing student evaluations of the different subjects, but the impact of the evaluations varies:

What comes out of the evaluations depends—here, I am very arrogant—it depends on which students are showing up during the evaluation. (Regional, Academic, Sciences)

Most of them emphasised that the feedback from students was of high importance for them; however, the informal system through the daily contact with students was of higher importance than the institutionalised evaluation systems, as illustrated here:

Well, I have a strong focus on the student feedback. But not through the formal system. I organise it myself. Informal and self-organised. (Flagship, Academic, Social Sciences)

In relation to *research evaluation*, the Danish interview data show that managers who 'own' evaluations are positive regarding the impact on performance. For example, a former and a present dean at different universities (both social sciences) who had introduced performance-based research funding had both experienced a positive impact on research performance. However, they were also aware that undesirable side effects could occur, and, when such effects were recognised, the evaluation practices had been changed.

The interviews clearly indicate that evaluations also have an impact on the atmosphere of the academic units. The evaluations' consequences for the atmosphere of academic life are not necessarily constant, because the competitive situations are temporary in character. These situations, especially before an evaluation, highlight the contradiction between unit-based interests and collegiality. A leader of a Finnish academic unit stated:

Every time, a little depends on our situation, and if we are evaluated, we cannot truly be collegial. Frankly, the academic leaders of the large units somehow have their own interest. (Regional, academic leader, sciences)

From the Norwegian data, we can read that the impact of evaluations is more closely connected to research than teaching. This is seen through

the employment process but also through the allocation of funds. The metric system for publications is of particular importance for allocations at both the individual level (e.g. allocation of funds for daily operating activities, conferences and sabbaticals) and the institutional level (e.g. getting PhD students). There are critical voices that question whether this focus in the evaluations is of importance, but they also question this emphasis on research over teaching. A statement from a manager regarding the hiring procedures illustrates this:

Research is what comes first in the review of the competences of applicants ... The evaluation from the review committee contains 14 pages related to research, and then there might be a couple of sentences at the end, summing up: "The applicant has been teaching for ten years, so he must be competent." So, we are also focussed on highlighting development projects on the teaching side. (Regional, Manager, Social sciences)

Although the overall survey data indicate that academics are quite pessimistic about the direct impacts of evaluation and measurement, the interview data show that there may be, indirectly, more positive dynamics following these activities.

DISCUSSION AND CONCLUSION

The analysis in this chapter has drawn on a broad conceptualisation of the concept of evaluation. *Evaluation* has been defined as 'procedures for assessing the effectiveness and quality of public organisations'. This broad conceptualisation has made it possible to bridge the analyses in the other chapters in Part II of this book. While those chapters have gone thoroughly into the dynamics and influence of evaluation-based funding systems, managerialism and strategy work, this chapter has given the broader picture of how these themes are interrelated.

The analysis has shown that there are different evaluation practices within the Nordic region, though the ideas behind developing evaluation practices are similar; they aim at improving performance on a wide range of aspects, such as quality, effectiveness and, in some contexts, especially in recent years, internationalisation, impact and employability. But looking into the specific practices, evaluation systems are varied. In relation to education, the Nordic countries adopt slightly different compared to the evaluation requirements in the Bologna process, and they include slightly different indicators in their performance-based funding systems. In rela-

tion to research, Finland and Norway have developed national evaluation systems: in Finland, driven by the Finnish Academy, and in Norway by the research council. In Denmark and Sweden, there are no national systems as such. Here, the universities have more autonomy to organise evaluations themselves. Also, performance-based funding systems related to research include different indicators and are organised differently. Internationally popular governance and evaluation ideas are, thus, translated into national policy agendas and administrative cultures. We find convergence in policy talk but less convergence in practices (Pollitt 2002).

Even though quality enhancement has been an important topic on the national agendas, a discussion is ongoing on whether evaluation practices and indicators related to both research and education have caused the focus to move from quality to quantity. This has raised an agenda about how to develop systems and practices focussing more on 'real' quality. Future studies should look into how this agenda develops and how the initiatives implemented influence university performance.

A general pattern across the four countries is that policy-driven evaluation schemes have been institutionalised and expanded, and management-oriented schemes—sometimes mirroring the national systems—have gained importance. Last but not least, the academic-driven evaluations have proliferated in systems with fierce competition for recognition and rewards. Given the public nature and the long tradition of public sector steering in the Nordic region, the national policy-driven initiatives are still seen as quite legitimate by the academic staff. As seen in these case studies, the growth of policy- and managerial-driven schemes have not meant a reduction in academic forms of evaluations, resulting in an overall growth of evaluations in the system. There are, however, signs that academic forms of evaluation are changing, as indicators used in the policy-driven evaluation systems are finding their way into academic evaluation practices, just as academic evaluation is becoming a stepping stone for managerialism.

Although the policy-driven evaluation schemes in all the countries seem to carry some legitimacy, academic staff throughout the Nordic region do not consider these very effective as tools for improving performance, either in research or in education. It seems that evaluation criteria in policy- and managerial-driven evaluation schemes often do not match academic definitions of what constitutes and supports good performance. Danish academics, in particular, were found to be quite negative towards the potential performative impact of evaluations. Why Denmark stands

out is not easy to identify through our data, but one possible explanation is that evaluations have perhaps been perceived as more ‘intrusive’ when compared to those implemented in the other Nordic countries. A related explanation is that Denmark also seems to have more expansion in managerial-driven evaluations than the other countries, which may have contributed to the negative atmosphere. It is also possible that the negative perceptions identified among the Danish academics are an indicator of ‘evaluation fatigue’. Maybe evaluation has been overdone, and a proper balance between higher education policy initiatives, managerial initiatives and academic duties has not yet been found.

Returning to some of the concepts introduced in our analytical framework, it could be argued that evaluations have taken up a central role in the changed governance of higher education in all four countries. Given the historical forms of governance of higher education found in the Nordic region, it is striking that the growth of evaluation schemes, to a large extent, is policy driven and, as such, under the supervision of the national authorities. While the state has delegated a substantial number of evaluations to intermediate bodies and agencies, the governance of the sector is still very much a public affair in all the countries. This pattern reflects the intensified accountability demands due to the growth of the higher education sectors and the corresponding increases in tax-financed resources spent. It also shows that decentralisation in the form of increased institutional autonomy occurs in tandem with centralisation initiatives, a pattern also known from the hospital sectors in some of the countries (Torjesen et al. 2017).

One can, however, also argue that the new element found in the region is not so much the dominant position taken by the state but, rather, that many of the evaluation schemes introduced are summative and result-oriented with elements of user and stakeholder orientation. What we see, therefore, is that the nation states are strengthening competition in the sectors and developing more market-like governance structures while still holding on to the Nordic universal welfare model.

Acknowledgements The data presented in the current volume and individual chapters emanate from a comparative study funded by the Norwegian Research Council under its FINNUT flagship program, a long-term program for research and innovation in the educational sector program. The project number was 237782, and the project was titled ‘Does it matter? Assessing the performance effects of changes in leadership and management structures in Nordic Higher Education’.

REFERENCES

- Aarveaara, T., L. Wahlfors, and I.R. Dobson. 2018. The Higher Education Systems and Institutions, Finland. In *Encyclopedia of International Higher Education Systems and Institutions*, ed. J.C. Shin and P. Teixeira. Berlin: Springer.
- Bomark, N. 2016. *Drawing Lines in the Sand: Organizational Responses to Evaluations in a Swedish University*. PhD diss., Uppsala University.
- Cameron, K.S. 1986. Effectiveness as Paradox: Consensus and Conflict in Conceptions of Organizational Effectiveness. *Management Science* 32 (5): 539–553.
- Dahler-Larsen, P. 2012. *The Evaluation Society*. Stanford: Stanford University Press.
- DEA. 2011. *Taxametersystemet under lup (The Taxi-Meter System Investigated)*. Copenhagen: DEA.
- FINEEC. 2016. *National Plan for Education Evaluation 2016–2019*. Helsinki: The Finnish Education Evaluation Centre.
- Geschwind, L. 2017. Higher Education Systems and Institutions, Sweden. In *Encyclopedia of International Higher Education Systems and Institutions*, ed. Jung Cheol Shin and Pedro Nuno Teixeira. Dordrecht: Springer.
- Hammarfelt, B., G. Nelhans, P. Eklund, and F. Åström. 2016. The Heterogeneous Landscape of Bibliometric Indicators: Evaluating Models for Allocating Resources at Swedish Universities. *Research Evaluation* 25 (3): 292–305.
- Hansen, H.F. 2005. Choosing Evaluation Models. *Evaluation* 11: 447–462.
- . 2014. ‘Quality Agencies’: The Development of Regulating and Mediating Organizations in Scandinavian Higher Education. In *Building the Knowledge Economy in Europe: New Constellations in European Research and Higher Education Governance*, ed. C. Meng-Hsuan and Å. Gornitzka, 188–218. Cheltenham: Edward Elgar Publishing.
- Hansen, H.F., L. Geschwind, J. Kivisto, E. Pekkola, R. Pinheiro, and K. Pulkkinen. 2019. *Balancing Accountability and Trust: Higher Education Reforms in the Nordic Countries*. Higher Education. Published online 21 January. <https://doi.org/10.1007/s10734-019-0358-2>.
- Hicks, D. 2012. Performance-Based University Research Funding Systems. *Research Policy* 41: 251–261.
- Jongbloed, B., and H. Vossensteyn. 2016. University Funding and Student Funding: International Comparisons. *Oxford Review of Economic Policy* 32 (4): 576–595.
- Karlsson, S. 2017. Evaluation as a Travelling Idea: Assessing the Consequences of Research Assessment Exercises. *Research Evaluation* 26 (2): 55–65.
- Karlsson, S., K. Fogelberg, Å. Kettis, S. Lindgren, M. Sandoff, and L. Geschwind. 2014. Not Just Another Evaluation – A Comparative Study of Four Educational

- Quality Projects at Swedish Universities. *Tertiary Education and Management* 20 (3): 239–251.
- Levi-Faur, D. 2014. The Welfare State: A Regulatory Perspective. *Public Administration* 92 (3): 599–614.
- Pollitt, C. 2002. Clarifying Convergence: Striking Similarities and Durable Difference in Public Management Reform. *Public Management Review* 4 (1): 471–492.
- Pruvot, E.B., A.-L. Claeys-Kulik, and T. Estermann. 2015. Strategies for Efficient Funding of Universities in Europe. In *The European Higher Education Area – Between Critical Reflections and Future Policies*, ed. A. Curaj et al., 153–168. Heidelberg: Springer.
- Rossi, P.H., H.E. Freeman, and M.W. Lipsey. 2004. *Evaluation: A Systematic Approach*. 7th ed. Thousand Oaks, CA: Sage.
- Schmidt, E.K. 2017. Quality Assurance Policies and Practices in Scandinavian Higher Education Systems: Convergence or Different Paths? *Journal of Higher Education Policy and Management* 39 (3): 247–265.
- Scriven, M. 2003. Evaluation Theory and Metatheory. In *International Handbook of Educational Evaluation*, ed. T. Kellaghan, D.L. Stufflebeam, and L.A. Wingte, 15–30. Dordrecht: Kluwer Academic Publishers.
- Seeber, M., B. Lepori, M. Montauti, J. Enders, H. de Boer, E. Weyer, I. Bleiklie, K.L. Hope, S. Michelsen, G. Nyhagen Mathisen, et al. 2015. European Universities as Complete Organizations? Understanding Identity, Hierarchy and Rationality in Public Organizations. *Public Management Review* 17 (10): 1444–1474.
- Stensaker, B. 2014. Troublesome Institutional Autonomy: The Governance and Distribution of Authority in Norwegian Higher Education. In *International Trends in University Governance: Autonomy, Self-Government, and the Distribution of Authority*, ed. M. Shattock, 34–48. New York: Routledge.
- Stensaker, B., and L. Harvey, eds. 2011. *Accountability in Higher Education: Global Perspectives on Trust and Power*. New York: Routledge.
- Stensaker, B., and P. Maassen. 2015. A Conceptualization of Available Trust-Building Mechanisms for International Quality Assurance of Higher Education. *Journal of Higher Education Policy and Management* 37 (1): 30–40.
- Torjesen, D.O., H.F. Hansen, R. Pinheiro, and K. Vrangbæk. 2017. The Scandinavian Model in Healthcare and Higher Education: Recentralising, Decentralising or Both? *Scandinavian Journal of Public Administration* 21 (1): 57–80.
- Treib, O., H. Bähr, and G. Falkner. 2007. Modes of Governance: Towards a Conceptual Clarification. *Journal of European Public Policy* 14 (1): 1–20.
- Vedung, E. 1997. *Public Policy and Program Evaluation*. New Brunswick, NJ: Transaction Publishers.
- . 2010. Four Waves of Evaluation Diffusion. *Evaluation* 16 (3): 263–277.

- Verhoest, K., B.G. Peters, G. Bouckaert, and B. Verschure. 2004. The Study of Organizational Autonomy: A Conceptual Review. *Public Administration and Development* 24 (1): 101–118.
- Vetenskapsrådet. 2017. *The Swedish Research Barometer 2017*. Stockholm: Vetenskapsrådet.
- Weiss, C.H. 1998. Have We Learned Anything New About the Use of Evaluation? *American Journal of Evaluation* 19 (1): 21–33.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

