

Governance and Evaluation

- Consequences for teaching culture and teacher profession

Ane Qvortrup

Department for the Study of Culture, University of Southern Denmark

Campusvej 55, DK-5230 Odense M, Denmark

Tel: +45 6550 2224 E-mail: anq@sdu.dk

Dion Rüsselbæk Hansen (Corresponding author)

Department for the Study of Culture, University of Southern Denmark

Campusvej 55, DK-5230 Odense M, Denmark

Tel: +45 6550 4538 E-mail: dion@sdu.dk

Received: July 05, 2013 Accepted: July 12, 2013 DOI: 10.5296/jpag.v3i3.3962

Abstract

This article examines, whether contemporary teaching culture and the teaching profession are being transformed as a result of demands for evaluation and visibility, and what the consequences of this may be. We identify dominant education policy discourses that impact teachers and teaching, following which we show how the concept of ‘quality assurance’ has shaped contemporary evaluation demands. The paper then argues that we must challenge existing worries that evaluation initiatives have instrumental and mechanical consequences for teaching practices. Inspired by existing research and building upon empirical studies of Danish educational practices (upper secondary schools and universities), we show that indirect, so-called performance paradoxes may occur. This is because the effort to make one thing visible often renders other aspects difficult to access. From these considerations, we discuss which impact evaluation and visibility demands may have on the teacher profession and teaching culture. The article only points out tendencies, wherefore, we argue the need for more knowledge about how various evaluation demands impact teaching practices.

Keywords: Governance, Evaluation, Self-management, Visibility, Teacher Profession, Teaching Culture, Discourse Analysis, Pseudo Realities

1. Introduction

Ends: We want the best education system in the world. We want everybody to get an education. We want an inclusive education culture. We want the students learn more, faster, and longer.

Means: Upskill teachers. The teaching must be evidence-based. The teachers must evaluate their teaching and student learning. The teachers must document and visibilize what they do, how they do it, and why they do it.

Precondition: The right evaluation culture!

In Denmark, there is a constant, demand for ends and means as described above (Hjort 2012; Hjort & Raae 2011). At present, and in extension, focus is therefore upon improving our education system and upskilling our teachers. Almost every head master, teacher, and educational researcher is engaged in educational development and quality assurance. This is because it has been suggested that the right evaluation culture will automatically lead to educational improvement as well as better teachers. The emphasis upon quality assurance may however be explained by a widespread sense that the nations competitiveness is at stake and depends on the skills, competencies, and knowledge bestowed upon the younger generations, (The Danish Government 2006). There is a huge focus on teachers in this regard and in extension a consensus has appeared “that teacher quality is one of the most, if not the most, significant factor in students’ achievement and educational improvement” (Cochran-Smith 2004 p. 3). That is why many efforts in Denmark, as in the rest of the world, are taken to improve the quality of teaching (Apple 2001; Ball 2008; Biesta 2011; Lindblad & Popkewitz 2004).

Perhaps emphasis on improving the quality of teaching is not new. What is a novelty, however, is that institutions and teachers are ascribed new subjectivities and areas of responsibility in the process of quality assurance. It is up to the individual schools and teachers to enable observer *access* and be able to *judge* the quality of teaching. Schools and teachers are not only expected to provide quality teaching, they are also expected to ‘visibilize’ and document how they contribute to the required level of teaching. In this regard, evaluation seems to be the primary technology (Rose 1999) in improving teaching quality! Nowadays, it is not an easy task to refuse to evaluate or to be evaluated. High expectations are attached to evaluation.

In spite of wide spread assumptions about the ease of evaluation, it is in fact a very difficult task. In this article we will focus on *how* and *why* evaluation is less than a straightforward issue. In extension, we claim that this issue of better education is not merely a question of identifying the ‘right’ evaluation culture as some researchers argue (Ekholm, Mortimore, David-Evans, Laukkanen & Valijarvi 2004).

The present article applies a discourse analytical approach. Within this epistemological paradigm we are not interested in what the right evaluation culture is or could be. Rather, we are interested in the discourses that seem to be attached to a contemporary understanding of successful evaluation. Based on data gathered in Danish upper-secondary schools and

universities¹ we show how these discourses impact practices. At the same time we point out what consequences evaluation and visibility demands have for teachers and teaching culture. The question we seek to answer is: *What characterizes the new evaluation demands and what consequences do they have for contemporary teaching culture and the teaching profession?*

First we introduce the discourse analytical approach and explain how it leads to a particular analytical frame. We then focus on the transnational lines, tendencies, and discourses that seem to construct a growing need for educational evaluation. In extension of this we demonstrate how the concept of ‘quality assurance’ is central to understanding the contemporary desire for evaluation. Thirdly, we illustrate how this desire leads to the rise of a particularly form of evaluation and how different meanings, assumptions and rationalities are attached to it at different times, in different contexts, and in different situations. Fourthly, we point out how such demands seem to transform the teaching culture and profession. Fifthly, we argue that evaluation can blur the actual state of affairs and create ‘pseudo realities’, because evaluation is handled strategically. Finally, the article discusses the consequences of the above-mentioned transformation.

2. Discourse analysis – an epistemological frame

In this article we are interested in the impact of dominant discourses upon the contemporary teaching profession and culture. We are not only interested in what the nature and character of these discourses are, but also how they affect something or somebody. Let us start by pointing out what discourses do in general.

The concept discourse has numerous meanings and a host of applications as it is employed within different theories and subject areas. In the following, we draw on Michel Foucault, but also Nikolas Rose’s (1999) and Mitchell Dean’s (2010) Foucault-inspired governmentality studies. Discourses influence what we can speak about, how we may speak about it, and what is excluded (Foucault 1990). Within different discourse regimes we find various assumptions, logics, and rationalities, which organize individual and collective conduct, attitudes, meanings, and practices (Foucault 1997b). Dominant discourses allow certain ways of thinking and acting to be considered correct, good, and acceptable, while others are viewed as incorrect, bad, or unacceptable (Foucault 1997a). Here, we must not forget the close relation between knowledge and power that influences how individuals are able to position themselves and, in return are positioned by others (Foucault 1980; Foucault 1991).

Using a discourse analytical perspective, we are not able to grasp the world as it is, the ‘pure essence’ of the world. We abandon the idea, that our experience or language might mirror the way reality actually is (Rorty 1980). We can only talk about constructions of the world. This does not mean, however, that we cannot grade these constructions. Some constructions are better or worse than others. Inspired by Foucault, we claim that better or worse constructions cannot be settled once for all, but depend on contemporary knowledge regimes and the historical a priori, which influences how we think, act, and judge.

We are always regulated by the society we live in (Foucault 1990) and are not – with a

¹ Classroom observations and interviews in upper secondary schools plus questionnaire survey and interviews with university teachers (see Rüsselbæk Hansen 2012; Keiding & Hansen 2012).

Freudian term – ‘masters of our own house’ (Freud 1916). This regulation ‘creates’ blind spots. There are things and premises that are so obvious that we rarely question and problematize them (Rose 1999). However, as the Danish professor of philosophy Lars-Henrik Schmidt says: “There is no such thing as free thought – but you can relate more or less freely to your premises (Schmidt 2011 p.10). Based on these considerations, we ask, which form of common sense knowledge and premises shape contemporary evaluation demands and how these establish a special kind of ‘social order’ with particularly consequences.

3. Contemporary educational policy trends

In this section we focus on contemporary trends, which dominate Danish as well as international educational contexts. Inspired by neoliberal thoughts and ideas, we are currently witnessing new need in educational policies. A need to *assure* and *secure* quality in the education system. In Denmark as well in the rest of the world, many new educational agendas have appeared since the late 20th century (Hood 1991). Inspired by New Public Management, a New Public Governance discourse has also come about (Greve & Hodge 2010). Within this ‘economic’ administration discourse, emphasis is more on cultural changes and less on structural changes within the educational system (Ejler, Seiding, Bojsen, Nielsen & Ludvigsen 2008).

The supremacy of economics in educational policy is closely related to the evaluation of evidence based best practice examples (Berkel, Aa, & Gestel 2010, Ball 2008). Here, evaluation is: “a central component of a resultant new regime of scrutiny, which defrays political responsibility and replaces it with the ‘virtuous circle of evaluation, evidence, performance’. This ‘turn to evidence’ is a new mode of political regulation, and evaluation is part of a wider logic of decentralized control. Evaluation, in this context, relies on an instrumental rationality based on technical evidence” (Taylor 2005 p.5). This form for ‘rational’ evaluation allows politicians to ask: what is worth spending more money on, and what is not? Politicians assume that this form of evaluation grants them access to the educational system and the teacher profession as a field of knowledge. Hereby, they can acquire knowledge about inputs (decisions, the use of technologies and financial aspects), processes (what is going on in teaching), outputs (the number of lessons the students are offered), outcomes (performativity, number of students who complete the education, and number of students who drop out on the way), and effects (the ‘work-readiness’ of students and the student employability) (Henrich 2007; Vedung 2009).

In Denmark, politicians are very interested in summative teaching results, but also formative teaching aspects. Why is this the case? The answer seems to be that we all: leaders, politicians, researchers, and practitioners, must learn from summative and formative evaluations, in order to have a starting point for making informed decisions on what produces quality. It is hard to be against this wish for new and improved quality. However, quality to whom and when? Who has the power and ability to determine what is quality and what is not? How does evaluation informed decisions increase quality and/or produce quality? What evaluation information will be produced and how? Such questions will be examined in the next section.

4. Contemporary evaluation and visibility discourses

Evaluation is the mantra today. We are currently witnessing an ongoing and growing prevalence and exchange of new evaluation demands, new evaluation models, and new evaluation practices in schools. In Denmark, evaluation has become a management technology of vital importance. The purpose of evaluation is to identify, whether a given context is under development, to explain why this is the case or not, and on this basis to make informed decisions about opportunities for improvement. Evaluations are considered a major factor in formulating concise policies and political programs. By measuring performance, evaluation is supposed to help governments and public administration enforce decisions in an efficient manner. Evaluation is seen as a ‘correcting device of society’ (Stame 2006).

Evaluation takes many forms. In the New Public Governance-sphere, the dominant form may be characterised according to its purpose, which is to *highlight* the processes that move in the right direction and not to document cause-reason-relations. Systematic data gathering is meant to give indicators of progress, development, and/or goal achievement in proportion to allocated resources. The *visibilization* aspect is of crucial importance (Vedung 2010; Krogstrup 2011). In the data collection process, all members of the public sector are involved, i.e. each leader, teacher, and student in the educational system. A comprehensive set of data must be collected. The collected data must include self-observation (i.e. teaching intentions and planning) besides observation of that which is dealt with, so to speak (i.e. marks, drop-out rate, student satisfaction).

At least three closely-knit discourses permeate these evaluation tendencies:

First, we find a comparison and assessment discourse focusing on the performance of students, teachers, and institutions: “The performances of individual subjects or organisations serve as measures of productivity or output, or displays of ‘quality, or ‘moments’ of promotion or inspection. These performances stand for, encapsulate or represent the worth, quality or value of an individual or organisation within a field of judgement” (Ball 2008 p.49). Here, it is interesting to ask oneself, which aspects of a teacher’s performance can be measured. This is interesting in view of the risk, that: “what can be measured will come to determinate the concept of quality” (Schmidt 1999 p.245).

Second, we find a decentralisation and institution accountability discourse. Taking the New Public Management or Governance focus on market demands (i.e. free hold) as an outset, each Danish upper-secondary school and university is assigned increased responsibility. Here a new kind of ‘regulation-from-a-distance’ appears (Qvortrup & Rüsselbæk Hansen 2012a) - strategies of indirect regulation. This has led to the Danish Universities in 2003 and the Danish upper-secondary schools in 2007 being turned into self-governing institutions. Budgets now depend on the number of enrolled students, the number of graduating students, and on the ‘right’ evaluations. As self-governing institutions, universities and upper secondary schools are assigned increased responsibility for their self-management, but at the same time the state controls whether the institutions live up to their responsibilities and if they show accountability. The educational institutions are responsible for their activities and results (Qvortrup & Rüsselbæk Hansen 2012a). Hence, we have to do with a control system

placing itself in the field between the political system and the educational system. We see a special kind of individualization tendency in relation to the institutions self-management and self-accountability (Biesta 2011). A tendency that may create a blind spot facing the social matters and conditions which influence the institutions and teachers in question.

In the *third* discourse – a discourse of transparency – we find a conception that, by bringing about transparency and by enhancing the free choice of the individual, is able to solve: “the political problem of ‘public doubt’: Governments need (politically) to demonstrate that they are taking public services seriously (since publics continue to want services) and that they can ‘deliver’ services, standards, quality or improvement (promises may vary), while trying not to be directly responsible” (Clarke 2003 p.7). The discourse appears as both taking and disclaiming responsibility - centralized decentralization. The assumption seems to be attached to a logic that stems from rational choice theory. Citizens are provided knowledge about institutional matters and are made into consumers, who are able to decide which teaching institution is best (Taylor 2005).

The increasing use of evaluation is filled with tension however. This influences the possibility of succeeding (Schmit 1995; Perrin 1998, Perrin 1999; Van Thiel & Leeuw 2002; Pidd 2005; Qvortrup & Rüsselbæk Hansen 2012b). Often, tension originates from the fact that evaluations are *not* neutral and objective technical mechanisms. They contribute to a *certain* evaluation practice. The tensions and resulting consequences are, however, not always immediately perceptible; they do not necessarily show up directly in the context of evaluation, but on other levels. For instance when evaluation results or products gain the status of exclusive *knowledge* or *documentation* and determine the onward course of events (Qvortrup & Rüsselbæk Hansen 2012a), subsequently repressing other equally valid perspectives.

It varies according to *three* dimensions: the fact dimension, the temporal dimension, and the social dimension. In each single event, the three dimensions appear in combination, but can be deconstructed analytically (Luhmann, 1995 p.86). This can be illustrated by the utterances, ‘What did she say?’, ‘When did she say this?’, and, ‘Who said this?’, which emphasizes the fact, the temporal, and social dimensions respectively. An event offers certain positions, i.e. potential social relations and identities. These positions are offered in a given situation. There – in continuation of the above-mentioned factors – you cannot grasp the substance of a given evaluation practice out of context.

5. An evaluation culture

As mentioned above, the OECD report from May 2005 concludes that an evaluation culture is missing in the Danish schools (Ekholm, Mortimore, David-Evans, Laukkanen & Valijarvi 2004). This is an insuperable barrier for quality. This announcement was followed by great political interest in building up such a culture (Local Government Denmark² 2005; Nielsen 2009; Dahler-Larsen 2006). However, the concept ‘evaluation culture’ is intangible. Dahler-Larsen (2006) refers to it by saying that it must standardise and level out differences

² Denmark is divided into five regions and 98 municipalities. Even it is voluntary to be a member of Local Government Denmark (LGDK), all 98 municipalities are members (for more information see <http://www.kl.dk/English/>).

between evaluation practices. We see similar claims in the OECD report, where evaluation culture is linked to clear national objectives and test systems (Ekholm, Mortimore, David-Evans, Laukkanen & Valijarvi 2004 p.71). This tendency can also be found in the CEPOS report from 2006 (CEPOS 2006).³

Clear structures and objectives seem to be a prerequisite for a sufficient evaluation culture. What these structures and objectives consist in and of is less clear. We are here dealing with vague and indefinite formulations: informal norms, values, habits, and practices (Dahler-Larsen 2006 p.91). These cannot be ‘set up’, but must occur progressively as a negotiation in a cultural cross-roads and between various cultures: school culture, teacher culture, teaching culture, teacher college culture, etc. (Dahler-Larsen 2006 p.86).

Furthermore, the evaluation culture is closely related to a concept of evaluation capacity, which in itself contains several dimensions such as human competencies and attitudes, organizational structures and processes, management support and resources (Dahler-Larsen 2006). These are also difficult to define. The concept of evaluation culture not only thus seems complex, but is *hyper complex*: The concept contains several meaning components, which affect, disturb, and change each other. However, they remain connected and exhibit an irrefutable link (Fink 1988 p. 22). When looking for instructions, an evaluation culture is characterised by having a preference for data, which means, that we are informed and guided from these data (Dahler-Larsen 2006 p.75). We do not make decisions substantiated only in ideology, tradition, and belief, but learn from experience of experiments. The approach is open in seeking to try out different suggestions (ibid. p.75-76). Here, the concept of evaluation culture takes form as a necessary reduction of contingency from contingent evaluation situations that will never become non-contingent. The *right* evaluation strategy and method can never be found, but a strategy and method have to be *chosen*, with different – positive and negative consequences as a result. This can explain the on-going and vague attempts to define and explain the concept. No single and univocal definition exists.

The succeeding sections look at empirical examples that serve as a topicalization of the way evaluation and evaluation culture is implemented. In extension, we focus on ‘the reality’ that the evaluation initiatives construct.

6. Transformations of teaching culture and teacher profession

On-going research – including our own – has focused on how evaluation initiatives affect teaching culture and profession. What happens when teachers are forced not only a) to do their work, but must also b) make their efforts visible to other, such that it can c) become the subject of assessment and learning? The number of variables in a teaching situation makes it impossible or at best very troublesome to solve the two later tasks (Berliner 2002). Katrin Hjort points out that it doubles the teacher’s work (Hjort 2012). Elsewhere we have argued, that you cannot talk about doubling as Hjort does, since a) will not remain a), when it is coupled with b) and c) (Qvortrup & Rüsselbæk Hansen 2012b). The work a) is to a certain

³ CEPOS (Center for Political Studies) was founded in 2004 and is an independent Danish think tank promoting a society based on freedom, responsibility, private initiative and limited government (for more information see <http://www.cepos.dk/english>).

degree changed or *transformed* during the process of making visible and assessing.

In fieldwork, we have observed that transformations cannot be understood as direct cause-relationships owing to evaluation initiatives. This observation stems from an experience at the conference: *Culture and Learning* at University of Southern Denmark, 30th November 2011. We gave a talk about contemporary evaluation demands and the way these are met. We then discussed how these demands are sometimes counter-productive for schools, teachers, and students. When something is made visible, other things are often elided or receive less attention. This caused a discussion between an upper-secondary school teacher and a local government representative. The teacher told us that these consequences were well-known and that the teachers acted strategically, because they are aware of how important it is to present evaluation results that look good so the institution and the teachers not will get in 'bad standing': 'We know that what we document and evaluate has nothing to do with what is actually going on in reality. But we make it look as if it has'. This suggests that we must challenge existing ideas about the outcome of evaluation, because it can make good sense to evaluate in a special way.

Evaluation initiatives do not automatically make teachers substitute reflections on teaching practice and professional judgment with methods and unreflected 'teaching to the test' (Andersen 2010; Rømer 2012). The teacher profession cannot renounce allegiance to the professional reason and judgment, which is a key aspect of the profession (Schmidt 1999; Luhmann 2006). As illustrated, we do not think that the teacher profession will become a rational, mechanical and instrumental affair due to the evaluation demands. However, we see indicators, that the knowledge base from which teachers justify their choices do not remain unaffected. Teachers continue to practice their professional judgment, but this judgment does not itself seem unaffected. Let us provide an example. Based on 116 university teachers' choice of teaching content, Keiding and Hansen (2012) show that university teachers do not take their students (prerequisites and future values) into consideration when selecting content. They suggest that new teachers at the university in question read Biggs & Tang (2011) at the postgraduate teacher training course. Here didactical reflections and recommendations are disconnected from the participants (Keiding & Hansen 2012). We suggest another possible reason, namely that a new initiative of the university's quality assurance strategy, requires detailed teaching plans that must be handed in for approval by the Academic study board 5 months before the subsequent semester for evaluation. This means that the teachers cannot take student (prerequisites and future value) into consideration when choosing content. The knowledge base from which teachers make their choices has been transformed because of quality assurance initiatives.

7. Construction of 'pseudo realities'

Regarding the epistemological issues articulated earlier, the 'real practice' will never be the 'evaluated practice' and vice versa (Foucault 1990). In continuation of the above-mentioned, however, we can fear that increasing pressure on teachers and schools to produce 'right' evaluations, results in the construction of 'pseudo realities', which do not necessarily have anything to do with what is going on in practice. Therefore, we should not be blind to the fact that schools and teachers may have a particular interest in giving us 'pseudo information' of

what is going on.

Still, we must be aware that not all teachers evaluate and change their teaching even if they must (Rüsselbæk Hansen 2011) and it is not an easy task getting teachers to evaluate and change their teaching practice: “ they must first be dissatisfied with their exiting beliefs in some way; second, they must find the alternatives both intelligible and useful in extending their understanding to new situations; third, they must figure out some way to connect the new beliefs with their earlier conceptions” (Prawat 1992 p. 357).

Our fieldwork shows, however, that if and when changes and transformations take place, it may have several consequences. Let us give an example from classroom observations in a Danish upper-secondary school.⁴ In this class, there exists a general discourse among the students that the teacher knows what is right and what is wrong. Such a discourse sets the ground for the teacher’s evaluation practice and places constraints on student answers considered relevant. What the teacher can do better was not a theme in the evaluation. Instead, much of the evaluation concerned what the students could do better. At no point during the evaluation process were the teacher’s actions problematized by the students. This made sense because the students assumed that the teacher knew what was right/wrong. What does this observation tell us? It tells that even if something is going wrong in the class, the students will not blame the teacher, but only themselves. Obvious, this is problematic. Not to question the teacher may contribute to the status quo. This example shows that evaluation will not always provide the ‘right’, helpful or even relevant information. This means that we must handle the information we get very carefully. We need to take into consideration, how information is generated and on what basis.

8. Concluding remarks

In this article we have pointed out educational discourses and tendencies, which produce an increasing need for evaluation of educational matters. It has been argued that the right evaluation culture in the Danish educational system will not necessarily produce and lead to increased quality of teaching, instead it can do the opposite.

It is a common assumption that evaluation can provide politicians with valuable information about the educational system. The assumption is that politicians attain insight into practices and are then be able to hold schools and teachers accountable and responsible for good or bad performances (Ball 2008; Biesta 2011).

As we have argued that it is not without consequences when insisting upon an evaluation and visibility culture. Our point is that evaluation is not always reliable and may not have anything to do with the ‘real’ practice, but are often strategically ‘pseudo’ constructions of a ‘wished’ practice. There is no linearity and causality between evaluation and practice, but neither they are undisturbed by each other or completely arbitrary or random (Luhmann 2006).

Who does what, how, and why are questions that must be asked over and over again, just as

⁴ These data stem from a researching project about social inclusion and exclusion in a Danish upper-secondary school (see Rüsselbæk Hansen 2012).

questions as who does what, to whom, and with what consequences must. Despite the fact that evaluations always construct a sort of ‘pseudo reality’ it does not mean that some constructions cannot be more accurate than others. Nevertheless, it is not always strategically wise to present ‘accurate’ evaluations. Sometimes it can be unwise or even stupid to do so.

In this article it has not been our intention to reject evaluation as a fruitful and helpful practice technology (Dean 2010; Rose 1999). Research has shown that evaluation is crucial to good teaching (Hattie 2012, s. 41; 17-18). Instead, we have argued that the limits and risks, which are attached to evaluation, must be recognized, because much is at stake for the schools and the teachers. That is why we find it problematic if evaluation data is assumed to represent the real ‘practice’ and thereby attains its own ontological life (Popkewitz 2009). We also wanted to problematize sceptical assumptions based on the idea that evaluation seems to create an instrumental and mechanical teaching culture where the teachers act as if: “education is like a complex automobile engine: if only we make the right adjustments – in teaching, in learning, in assessment – it will hum, and transport us to our destination, the promised land of high test scores” (Pinar 2004 p. 1). We do not see such a tendency. Teachers always make decisions based on their judgements in the specific situation and context. From our perspective it set limits for *how* mechanical the teacher profession can be. It does not mean that evaluation demands and discourses do not influence teacher judgements and teaching practices (Foucault 1997b).

We need more knowledge about the possibilities and limitations of different evaluation models, and about the performance paradoxes we see in practice. We need knowledge about the complex situations that we see in practice. This knowledge can be used for a better definition of the situation and for choosing the aspects that is or can be handled with a given evaluation strategy and model (Luhmann 2006: 171). This knowledge will make it possible to make informed decisions about evaluation strategies and models and their consequences. Furthermore, we need new and updated research based, reflection programmes that teachers can use when planning, improving, and evaluating their teaching practice (Luhmann 2002; Qvortrup & Wiberg 2013). These programmes have to take into account the tension between politics and practice that today’s teaching is marked by.

References

- Andersen, H. L. (2010) “Constructive alignment” og risikoen for en forsimplende universitetspædagogik, [”Constructive alignment” and the risk of a simplifying university didactic], *Dansk Universitetspædagogisk Tidsskrift*, 9: 30-35.
- Apple, M. W. (2001) *Educating the “Right” Way: Markets, Standards, God and Inequality*. New York and London. RoutledgeFalmer.
- Ball, S. J. (2008) *The Education Debate*. The Policy Press.
- Berkel, R.; Aa, P. v. d.; Gestel, N. v (2010) Professionals Without a Profession? Redesigning Case Management in Dutch Local Welfare Agencies. *European Journal of Social Work* 13(4): 447-463.

- Berliner, D. C. (2002) Educational Research: The Hardest Science of All. *Educational Researcher*, 31(8): 18-20.
- Biesta, G. J. J. (2011) Good Education in an Age of Measurement: Ethics, Politics, Democracy. Paradigm Publishers.
- Biggs, J. & Tang, C. (2011). Teaching for Quality Learning at University. UK: Open University Press/McGraw-Hill Education.
- CEPOS (2006) Folkeskolen med nye øjne. 22 forslag til udvikling af en evalueringskultur i folkeskolen [New eyes on the preschool. 22 suggestions for developing an evaluation culture in preschool]: https://www.cepos.dk/uploads/media/folkeskolen_med_nye_oejne.pdf [11.07.12].
- Clarke, J. (2003) Performing for the Public: evaluation, evidence and evangelism in social policy. Paper for the inaugural ESPAnet conference “Changing European Societies - The Role for Social Policy” Organised by the Danish National Institute of Social Research Copenhagen, November 2003,13-15.
- Cochran-Smith, M. (2004) Taking Stock in 2004: Teacher Education in Dangerous Times. *Journal of Teacher Education*, 55(1), January/February 2004, 3-7.
- Dahler-Larsen, P. (2006): Evalueringskultur. Et begreb bliver til [Evaluation culture. A concept emerges]. Odense: Syddansk Universitetsforlag.
- Dean, M. (2010) Governmentality. Power and Rule in Modern Society. London. SAGE Publications.
- Ejler, N.; Seiding, Rosenberg, H.; Bojsen, Schwarts, D.; Nielsen, Bohni, S.; Ludvigsen, F. (red.) (2008) Når måling giver mening. Resultatbaseret styring og dansk velfærdspolitik i forvandling. [When measurement makes sense. Result-based management and Danish welfare politics in transformation] København: Jurist- og Økonomforbundet.
- Ekholm, M.; Mortimore, P.; David-Evans, M.; Laukkanen, R. & Valijarvi, J. (2004). OECD-rapport om grundskolen i Danmark - 2004. [OECD-report about Danish schools] København.
- Fink, H. (1988) Et hyperkomplekst begreb – kultur, kulturbegreb og kulturrelativisme. [A hyper complex concept – culture, notions of culture and cultural relativism]. H. Hauge & H. Horstbøll (ed.) *Kulturbegrebets Kulturhistorie*. Århus: Århus Universitetsforlag: 9-23.
- Foucault, M. (1980) Truth/Power: In: Collin Gordon (ed.) Selected Interviews & Other Writings 1972-1977. New York. Vintage.
- Foucault, M. (1990) The History of Sexuality, Vol. 1: An Introduction. New York. Pantheon Books.
- Foucault, M. (1991) Governmentality. G. Burchell, C. Gordon & P. Miller (ed.) The Foucault Effect. Studies in Governmentality. Chicago: The University of Chicago Press.

- Foucault, M. (1997a) *The Archaeology of Knowledge*. London. Routledge.
- Foucault, M. (1997b) The ethics of the concern for self as a practice of freedom. P. Rainbow (ed.) Michel Foucault. *Ethics, Subjectivity and Truth: Essential works of Foucault 1954-1984*, vol 1. New York. The New Press.
- Freud, S. (1916) *Introductory Lectures on Psychoanalysis*. James Strachey (ed.) *The Standard Edition of the Complete Psychological Works of Sigmund Freud (1963)*, Vol. 16: 284-5.
- Greve, C; Hodge G. (2010) *Public-Private Partnerships and Public Governance Challenges*. S. Osborne (ed). *The New Public Governance?: Emerging Perspectives on the Theory and Practice of Public Governance*. Routledge: 149-162.
- Hattie, J. (2012) *Visible Learning for Teachers*. New York & London: Routledge
- Heinrich, C. J. (2007) Evidence-Based Policy and Performance Management: Challenges and Prospects in Two Parallel Movements. *The American Review of Public Administration* 37: 255-277.
- Hjort, K (2010) NON SENSE. A Discourse Analysis of the Danish Upper Secondary School Reform 2005: *International Journal of Education*. 2, No. 1
- Hjort, K. (2012) Det affektive arbejde. [The affective work] Frederiksberg C. Samfundslitteratur.
- Hjort, K. & Raae, P.H. (2011) Strategic self-management. Danish gymnasium management between playing solo and showing solidarity. *Journal of Public Administration and Governance*. 1, No. 2.
- Hood, C. (1991) A public management for all seasons? *Public Administration*, 69, Spring 1991: 3-19.
- Keiding, T. & Hansen, J. D. (2012) Undervisningens indhold: universitetsdidaktikkens stedbarn [Teaching Content: the Stepchild of University Didactics]. *Dansk Universitetspædagogisk Tidsskrift* 13: 106-119.
- Krogstrup, H. K. (2011). *Kampen om evidens. Resultatmåling, effektevaluering og evidens*. [The struggle about evidence. Result-based measurement, effect evaluation and evidence] København: Hans Reitzels Forlag.
- Lindblad, S., & Popkewitz, T. S. (Eds.). (2004). *Educational restructuring: International perspectives on traveling policies*. New York: Information Age Publishers.
- Local Government Denmark (2005): *Evalueringskultur – en ny dansk skoletradition* [Evaluation culture – a new Danish school tradition]. <http://www.kl.dk/English/>
- Luhmann, N. (1995) *Social Systems*. Standford University Press.
- Luhmann, N. (2006). *Samfundets uddannelsessystem* [Society's Educational System] København: Hans Reitzels Forlag.
- Luhmann, N. (2002) *Das Erziehungssystem der Gesellschaft*. Frankfurt: Suhrkamp.
- Nielsen, M. (2009). *Intelligent brug af evaluering er vejen frem* [Intelligent use of evaluation

is the road ahead]:
<http://www.eva.dk/e-magasinet-evaluering/evaluering-jubilaeumsnummer-2009/intelligent-brug-af-evaluering-er-vejen-frem>: 10.07.12.

Perrin, B. (1998) Effective Use and Misuse of Performance Measurement. *American Journal of Evaluation* 19(3): 367-379.

Perrin, B. (1999) Performance Measurement: Does the reality Match the Rhetoric? A Rejoinder to Bernstein and Winston. In: *American Journal of Evaluation* 20(1): 101-111.

Pidd, M. (2005) Perversity in public service performance measurement. *International Journal of Productivity and Performance Management* 54(5): 482-493.

Pinar, W. (2004) Curriculum and study, NOT curriculum and teaching. Unpublished manuscript, 1-14.

Popkewitz, T. S. (2009) Numbers in grids of intelligibility. Making sense of how educational truth is told (manuscript).

Prawat, Richard S. (1992) Teachers' Beliefs about Teaching and Learning: A Constructivist Perspective. *American Journal of Education*, 100(3), May: 354-395.

Qvortrup, A. & Rüsselbæk Hansen, D. (2012a) Læreralkymisten og voyeurisme. Om samtidens behov for synliggørelse [Teacher alchemist and voyeurism. About contemporary desire for visibility] I: Hjort, Katrin; Qvortrup, Ane & Raae, Peter Henrik (ed.) Der styres for vildt. Om paradokser i styring af pædagogik. Klim.

Qvortrup, A. & Rüsselbæk Hansen, D. (2012b) Samtidens spin: Konstruktionen af pseudovirkeligheder [Contemporary spin: Construction of pseudo realities]. *Turbulens.net*.

Qvortrup, A. & Wiberg, M. (2013) Kapitel 1. Viden, erkendelse, læring og undervisning [Chapter 1. Knowledge, realisation, learning and teaching]. A. Qvortrup & M. Wiberg (red.) Læringsteori og didaktik. København: Hans Reitzels Forlag

Rorty, R. (1980) *Philosophy and the Mirror of Nature*. Princeton University Press.

Rose, N. (1999) *Governing the soul. The Shaping of the private self. (Second Edition)*. London. Free Association Books.

Rüsselbæk Hansen, D. (2011) Én socialanalytisk tendensanalyse af samtidens italesættelser af den gymnasiale lærergerning [A social-analytical analysis of contemporary upper secondary teaching profession discourses] Ph.D. dissertation. Odense: SDU/Det humanistiske fakultet.

Rüsselbæk Hansen, D (2012) Social inklusion/eksklusion i tale og handling på HF & VUC [Social inclusion in action and speech on HF and VUC). *Gymnasiepædagogik*, nr. 90. University of Southern Denmark,

Rømer, T. A. (2012) Education and Evidence. *Pædagogisk Psykologisk Tidsskrift*, 49(3).

Schmidt, L.-H. (1999) *Diagnosis II - Socialanalytiske betragtninger [Diagnosis II – Social-analytical reflections]* København: Danmarks Pædagogiske Institut.

Schmidt, L.-H. (2011) *On Respect*. Aarhus University Press.

Smith, P. (1995) On the Unintended Consequences of Publishing Performance Data in The Public Sector. *International Journal of Public Administration* 18(2&3): 277-310.

Stame, N. (2006) Governance, Democracy and Evaluation. In: *Evaluation* 12(1): 7-16.

Taylor, D. (2005) Governing Through Evidence: Participation and Power in Policy Evaluation. *Jnl Soc. Pol.*, 34(4): 601-618.

The Danish Government (2006) Fremgang, fornyelse og tryghed. Strategi for Danmark i den globale økonomi [Progress, change and safety. A plan for Denmark in the global economy]. http://www.stm.dk/index/mainstart.asp/multimedia/55686_strat.pdf.

van Thiel, S. & Leeuw, F. L.(2002) The Performance Paradox in the Public Sector. *Public Performance and Management Review* 25(3): 267-281.

Vedung, E. (2009) Utvärdering i politik och förvaltning. [Evaluation in politics and management] Lund: Student-litteratur.

Vedung, E. (2010) Four Waves of Evaluation. *Evaluation* 16(3): 263-277.

Copyright Disclaimer

Copyright reserved by the author(s).

This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).