



<http://www.diva-portal.org>

This is the published version of a paper published in *Education Inquiry*.

Citation for the original published paper (version of record):

Hult, A., Edström, C. (2016)

Teacher ambivalence towards school evaluation: promoting and ruining teacher professionalism.

*Education Inquiry*, 7(3): 305-325

<https://doi.org/10.3402/edui.v7.30200>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-125181>

# Teacher ambivalence towards school evaluation: promoting and ruining teacher professionalism

*Agneta Hult<sup>†</sup>\* & Charlotta Edström\**

## Abstract

Today's evaluation society makes teachers participate in a stream of external evaluations. How teachers experience evaluation in school and how this affects their work and professionalism is the focus of this article. Teachers' views of external and internal evaluations and of the consequences for school practice are described and analysed. The interviewed teachers emphasised the importance of internal evaluations performed close to daily teaching practice and jointly with students and colleagues. These evaluations are generally overlooked in evaluation and school-policy research and seldom attended to or appreciated by school providers. Further, teachers were critical of and reported several negative consequences of accountability and external evaluations, but still generally complied by participating in them. The present results are discussed in relation to professional responsibility and accountability as well as to possible constitutive effects. By emphasising that their daily informal evaluations represent their efforts to improve teaching, teachers are describing parts of their professional responsibility. However, the negative consequences of external evaluations signal constitutive effects on teachers' work, described as making it less creative, discretionary and autonomous as well as increasing mistrust, meaning that more tests are required in order to legitimate student grades.

*Keywords:* evaluation society, informal evaluations, school practice, teacher work, teacher views

## Introduction

This article examines the intersection between teacher professionalism and responsibility on one hand, and evaluation and accountability on the other. What teachers are responsible for and to whom, and how this is influenced by increased external accountability through evaluations and other measures, has been increasingly discussed in research (e.g. Ball 2003; Day 2002; Green 2011; Hargreaves 1994; Mausethagen 2013a, b). Green (2011) elaborates on how accountability and the control of measurable details in teachers' work have had devastating effects on teachers' professional responsibility as they direct attention to these details. Discussions have also treated teachers' possibilities to pursue social and moral goals via education during a period marked by a great emphasis on achievement (Biesta 2009, 2010). While attention has been paid to these matters, further empirical research is needed

---

<sup>†</sup>Correspondence to: Agneta Hult, Department of Education, Umeå University, SE-90187 Umeå, Sweden. Email: agneta.hult@umu.se

\*Umeå University, SE-90187, Umeå, Sweden.

©Authors. ISSN 2000-4508, pp. 305–325

into teachers' perspectives on how evaluation influences their work and how it can promote teacher professionalism. In her research review, Mauserhagen (2013b, 23) concludes the fact that most articles were either theoretical studies or policy analyses "highlights the need to conduct more empirical research and to investigate responses from teachers as well as school management, students, and representatives from the municipality in the current policy climate".

Further, teachers are important actors in policy processes because they are the ones who are supposed to enact education policy (Ball et al. 2012). This enactment includes "negotiation, contestation or struggle" at various levels (Ozga 2000, 2). Teachers negotiate (the meanings of) evaluations and the extent to which these are to be seen as legitimate parts of their professional responsibilities. Their experiences and views of this struggle are important.

The increasing use of evaluations in society and school has been described as an 'evaluation monster' (Dahler-Larsen 2012; Lindgren 2014). In the Swedish case, deteriorating performance in the international Programme for International Student Assessment (PISA) tests has not reduced interest in evaluating schooling. Evaluations of schools and teachers' work are carried out at various levels: the international (e.g. PISA and TIMSS), national (e.g. the Swedish Schools Inspectorate and national tests in various subjects), school provider<sup>1</sup> (e.g. quality reports and questionnaires on staff and student well-being) and school (e.g. teachers' ongoing reporting of student outcomes and questionnaires on equal treatment) levels. In addition, the mass media is very interested in reporting the results of international, national and municipal evaluations, especially negative results (Rönnerberg et al. 2013), and often hold teachers responsible for poor student performance. However, recent years have seen debate in the media about how this mistrust of teachers' work, expressed by the huge number of evaluations and tests, affects teacher professionalism.

This paper investigates teachers' experiences of and approaches to school evaluation and its consequences, exploring the following research questions:

- How are teachers' responsibility and professionalism described in relation to evaluation?
- What are the consequences of accountability and external evaluations for teachers' practice?

'Evaluation' is here used as a wide term in line with Dahler-Larsen's (2012) the 'evaluation society' and Michael Power's (1997) the 'audit society'. In Sweden, schools and school providers are obliged to organise and perform local quality work, a form of evaluation, to ensure constant improvement and thereby pass the assessment by the Swedish Schools Inspectorate (SSI)(e.g. Bergh 2015; Lundström 2015). The various steps in this evaluation discourse are presented and promoted as part of quality assurance work in an enormous number of documents, reports and webpages

from the National Agency for Education (NAE) and the SSI (e.g. <http://www.skolverket.se/skolutveckling/kvalitetsarbete>). The process involves several steps, starting with monitoring current conditions, defining areas in need of improvement, deciding on and carrying out interventions and, finally, restarting the evaluation cycle. Both school providers and schools carry out such quality assurance work. In other words, teachers are today required to undertake ongoing evaluations and, as Mausethagen (2013b) states, there is a need for more empirical research into teacher responses to evaluation demands and into whether and how these demands have affected the teaching profession. This article addresses this need.

### **Professional responsibility and accountability**

To evaluate teachers and make them accountable for how they and their students perform, evaluations are distributed mainly by school providers, but also by national and international officials. As mentioned above, evaluation in this article is used in the broad sense that Dahler-Larsen (2012, 12) described as an “umbrella category covering a range of activities with varying forms” that includes inspection, tests, questionnaires, assessments and informal teacher evaluation etc. Henceforth, when we refer to *external* evaluations we understand evaluations initiated primarily from levels above the local school and with a control focus, while evaluations initiated by teachers themselves (mostly informal) at the local school are henceforth termed *internal* evaluations.

External evaluations of teachers’ practice have to make explicit the competencies and criteria to be measured. This “lure of the explicit” is problematic, according to Green (2011, 139), because trying to codify and articulate in detail this kind of professional and tacit knowledge (Polanyi 1966) tends to reduce and distort its complexity. Green (2011) claims that the desire to make teachers’ knowledge explicit also distorts their sense of responsibility. Professional responsibility cannot be reduced and formulated in an instruction, as “the more someone is tied down by specific instructions / . . . / the less they can be held responsible to see to it that things go well generally within their sphere of responsibility” (Green 2011, 91). When teachers have to relate to and be accountable for detailed aspects of their very complex school practice and teaching profession, they risk losing sight of the “responsibleness” (Green 2011, 61–64) that has traditionally been connected to their professionalism (also see e.g. Solbrekke and Englund 2011; Svensson and Evetts 2010). To develop and improve their professionalism, teachers need opportunities and time – individually and with colleagues – to reflect on professional values in their day-to-day practice (e.g. Biesta 2009; Hodkinson 2009).

Drawing on Ball et al. (2012), a starting point when studying teachers’ reasoning regarding and experience of evaluation in school is that evaluation initiatives, although imposed by policymakers outside school, are seen as done both *to* and, within the limits of the evaluation discourse, *by* teachers. In line with an enactment perspective, teachers are considered to do more than simply implement various evaluation policy initiatives (e.g. Ozga 2000).

In this paper, we focus on teacher professionalism and leave aside the sociological question of whether teachers can be considered professionals, such as doctors or lawyers. Instead, we are interested in professionalism as a quality (Englund and Solbrekke 2015), in this case, in the ways teachers practise their work in relation to evaluations. According to Hargreaves (2000), teacher professionalism has passed through four historical ages: pre-professional, autonomous, collegial and post-professional. The last age could describe the current situation, with teachers struggling to counter testing regimes, external surveillance, marketisation and economic cutbacks. Teachers as professionals have been associated with having, among other things, “[a] service ethic (commitment to service clients’ needs) and professional autonomy (control over classroom practice)” (Day 2002, 679). However, external prescriptive policy interventions and changes in management have resulted in teachers’ having less autonomy and less control over classroom decision making (e.g. Green 2011; Ozga 1995), in line with the post-professional age.

Responsibility is considered central to teacher professionalism (e.g. Solbrekke and Englund 2011; Mausehagen 2013a), and what teachers are responsible for, as well as to whom, has been studied and described in various ways. To analyse the implications of different kinds of responsibility for the work of teachers and other “social trustee professionals in a field of conflicting forces”, Solbrekke and Englund (2011, 849) distinguish two types of logic: professional responsibility and professional accountability. Based on this distinction, they consider the implications for professional work. The logic of accountability is oriented towards control rather than trust in the professional, and obtaining good service is ensured by means of “measuring and ‘accounting’ instruments, rather than by relying on professional discretion” (Solbrekke and Englund 2011, 855). Predetermined indicators construct these measuring instruments. The logic of professional responsibility rests on a moral obligation to see and respect the needs of the student and of the collective student group. It is difficult to predict the actions taken and their ‘outcomes’, making them difficult to report and account for as well. The two concepts are summarised in the table 1 below.

Table 1. The types of logic and implications of professional responsibility and accountability

Professional responsibility	Professional accountability
based in professional mandate	defined by current governance
situated judgement	standardised by contract
trust	control
moral rationale	economic/legal rationale
internal evaluation	external auditing
negotiated standards	predetermined indicators
implicit language	transparent language
framed by professions	framed by political goals
relative autonomy and personally inescapable proactive	compliance with employers’/politicians’ decision reactive

Source: Solbrekke and Englund 2011, 855

Solbrekke and Englund (2011) point out that the two categories are constructed as polarised for analytical purposes, but that they should not be perceived as static; preferably, the categories should be understood as “fluid and fluctuating in different systems of logic that are reconfigured over time” (2011, 855). Although stressing many adverse consequences of the present accountability and evaluation society, they also state that public accountability is needed in order to maintain trust and ensure good-quality public services. Still, it is important to be aware of the possible malpractices that accountability mechanisms may produce. The two types of logic can be related to Evetts’ (2011, 2013) two ideal types of professionalism, occupational and organisational professionalism, and to what have been referred to as ‘old’ collective and ‘new’ individual forms of professionalism (e.g. Helgoy and Homme 2007) and internal and external accountability (Mausethagen 2013a).

By defining the central aims and performances of education, evaluation activities might change what is considered important for school, teachers and students. Dahler-Larsen (2011, 2014) discusses this in terms of the constitutive effects of evaluation as these effects can help change how we understand the phenomena being evaluated. “By constitutive effects I refer to how QAE (quality assurance and evaluation) redefines the meaning of education and the practices of education by means of installing new discursive and cultural markers defining standards, targets and criteria” (Dahler-Larsen 2011, 153). This article pays attention to how evaluation might influence the meaning of teacher professionalism and responsibility via what is understood as a constitutive effect.

## **Methodology**

The present study is part of a broad research project, focusing on the consequences of evaluation for school practice<sup>2</sup>. We interviewed 43 compulsory school teachers in 2013 and the resulting interview material was complemented with the results of a teacher questionnaire. The interviews were carried out in eight case schools in four municipalities with populations of 75,000–100,000 inhabitants. Two schools from each of the four municipalities with a variation in school provider and socio-economic composition of the student body were chosen.

Most of the 43 interviewed teachers were women, work in lower secondary school and have experience teaching various subjects. The interviews were carried out as focus group interviews at the schools during the day, with two interviewers present. Since we wanted to investigate possible differences between newer and more experienced teachers in relation to their approaches to evaluation, teachers were distributed into two groups at each school based on their years of teaching experience. Teachers with ‘less experience’ (i.e. 1–5 years) were interviewed in one group and ‘more experienced’ teachers, mostly with 10 or more years of experience, in another group. The groups consisted of three to four teachers each. The sample of teachers was not chosen systematically; rather, we selected teachers who happened to have spare

time to talk to us on the day we visited the school and the principal helped us distribute them to the ‘right’ groups. To ensure confidentiality, the names of both municipalities and schools have been anonymised in the presentation of the findings. School provider and made-up names for communities and schools are presented in table 2.

Table 2. The selected anonymised municipalities and schools

Municipality	School	Provider
East	Salvia	Municipal
East	Coriander	Independent
West	Nettle	Municipal
West	Parsley	Independent
North	Basil	Municipal
North	Carnation	Independent
South	Garlic	Municipal
South	Lavender	Municipal

In the focus group interviews, we made it clear that ‘evaluations’ was being used in the broad sense to cover a spectrum extending from less formal oral follow-ups to formal written evaluations, assessments and tests. The interview questions concerned teachers’ experiences of and approaches to international, national, school provider, school and their own evaluations. An interview guide with open-ended questions was used as a checklist to ensure that all of our questions had been answered. The interviews lasted 60–120 minutes, and were recorded and transcribed verbatim.

With the help of the school principals, we distributed a questionnaire to teachers in all municipal schools in the four municipalities. In total, 145 teachers completed the questionnaire, representing a low response rate. However, the exact response rate could not be calculated as the principals were responsible for distributing the questionnaire and we were not informed of the number distributed. We interpret the low response rate as a sign of teachers’ heavy workload, including increasing streams of evaluations and questionnaires. While the overall results of the questionnaire study are very consistent with those of the interview study, due to the low response rate we only use the questionnaire results to back up some of the main results of the interview study.

In analysing the interview transcripts, the first step was to carefully read and analyse the transcripts without any preconceived interpretive schema in mind, and with an emphasis on identifying and describing the teachers’ approaches to and views on the various consequences of evaluations for teacher work. This first step resulted in the discovery of important themes in the teachers’ views and appraisal of the external and internal evaluations. The second step of the analysis concentrated on more thoroughly analysing the teachers’ elaborations of the consequences of external and internal evaluations, respectively. This step also revealed a third theme and possible synthesis between the two kinds of evaluations. Drawing on the theoretical framework

described above, these findings were mainly analysed along the lines of the two types of logic, i.e. those of professional responsibility and accountability (Solbrekke and Englund 2011).

To illustrate the teachers' views and appraisal of internal and external evaluations, the findings include quoted comments from both the more and less experienced individual teachers from the participating schools. However, the interviews were carried out in groups so the presentation emphasises the group level, i.e. experienced or less experienced teachers in aggregate.

### **The importance of teacher-initiated internal evaluations**

The analyses identified a gap between teacher experience and approaches to their own evaluations on one hand, and externally initiated (i.e. by school providers, national and international bodies) evaluations on the other. The first theme concerns teachers' elaboration on the importance of their own evaluations closely connected to their teaching practice. The evaluations performed close to teachers' daily practice were those they claimed enabled them to improve their teaching practice. When performing these evaluations, the teachers were motivated: "I analyse the results, I get insights and change things, often" (Salvia School, less experienced teacher). The teachers stressed that, to enhance quality in school, change must be bottom-up and not top-down.

The 'smaller' the evaluation is, the more important it is, I feel. I mean, if I have an evaluation with a small group of students, then it's close, you can feel it and it's possible to make changes that are fair and clear. The national ones or PISA, well, indeed, they don't concern me at all. But the ones that are close, small and hands-on – it's something that you can influence and change (Carnation School, less experienced teacher).

When stressing the importance of "evaluations on one's own initiative, those you know why you are doing and where you are aiming" (Salvia School, less experienced teacher), the teachers often described experiencing external evaluations that had unclear objectives.

Virtually all the interviewed teachers talked about their teaching as incorporating continuous assessments and evaluations based on professional judgements. Reflection and dialogue were repeatedly emphasised as a means to improve teaching, and a perceived need for more time for doing this was mentioned. There were divided opinions on the need to have these evaluations 'on paper'. Whereas writing things down on paper was sometimes mentioned as a way to increase clarity, the teachers uniformly expressed fatigue with the increased documentation workload, a pattern found in other studies as well (e.g. Helgoy and Homme 2007; Power 2013).

When describing their own daily or weekly evaluations of teaching, the teachers generally did not distinguish between evaluations conducted jointly with students in the classroom, ones they carried out themselves during lessons, directly after

lessons, or later on, and those carried out together with the work team or other collegial groupings. The teachers often appeared to conduct their own evaluations informally during or after regular working hours, a way of exercising the discretion and situated judgement that characterises professional responsibility (Solbrekke and Englund 2011; Green 2011). Carrying out these informal ongoing evaluations was described by one of the experienced teachers as something that was “ingrained in the bone” (Basil School). The less experienced teachers also expressed similar opinions: “You have to do it in your own head, like in the shower or on your way home” (Parsley School).

### ***What information is needed?***

Regarding necessary information and feedback from students, the teachers highlighted student views of teaching and social aspects as well as information related to results. According to the teachers, valuable student opinions on teaching included feedback on what students believed did and did not work well in current teaching and, consequently, how the situation might be improved:

I used to evaluate [lessons] together with my students, this informal evaluation, and it was reasonable, for example, after a chapter or something. Then I'd ask, 'How does it feel? What worked and did not work?' ... And I'd get feedback from them, and I'd develop as a pedagogue (Garlic School, experienced teacher).

In addition to specific feedback on how to improve current teaching, teacher knowledge of students' varied interests and preferences, information partly obtained from evaluations, was described as relevant to improving teaching. For the teacher quoted below, it also involved an ambition to make the students reflect on and evaluate their own learning:

Today I told the students, 'Now there are three weeks left in the semester. I have a plan for what to do, but I want you to confirm it or not. So what do you want to work on for the rest of the semester, what do you need help with?'. Because they have to reflect on their own learning as well (Coriander School, experienced teacher).

Many teachers also noted the need to stay informed about the students' safety and well-being. Further, teacher information on the students' knowledge level was described as a prerequisite for adapting everyday teaching to varied individual needs. In the following, a teacher contrasted this sort of information to the more peripheral information provided in the evaluations initiated by the school providers:

What I do in the classroom is quite separate from that [i.e. the school provider's evaluations]. Because in any case I need to adapt, I need to individualise everything. I have a Grade 9 class now in which the students vary from Grade 5 level to having passed

Grade 9. The [overall] level is very heterogeneous. It does not work to do things in one way, then. I maybe need to do it in five different ways. I sometimes bring in five lessons so that everyone should get his or hers somehow. Then one does not have the time to worry about what the local authority statistics say (Nettle School, less experienced teacher).

The teachers also argued that they, as illustrated in the following quotation, needed information to evaluate how they had succeeded in raising student awareness of their own performance in relation to the new grade criteria.

Have they [i.e. the students] understood? Can they read the knowledge criteria? Can one translate the knowledge criteria into basic Swedish so that the students understand what it is we're doing? And, based on that, then practise knowledge and skills (Parsley School, experienced teacher).

The preferred evaluations with the students were ones that were short, clearly delimited and carried out shortly after the lesson. In discussions, the teachers primarily reasoned about oral follow-ups. Though the formality and frequency of these evaluations varied, the importance of direct communication and continuous dialogue with students was repeatedly highlighted, as was subsequent teacher reflection. As mentioned, carrying out these kinds of evaluations was described as being sensitive, “capturing their views *in situ* and in time” and “having a discussion all the time” (Salvia School, less experienced teacher). In these discussions, evaluation was described as something that was being done all the time *by* teachers and, as in the following example, *of* teachers:

Then one is being evaluated every day, every lesson one is being evaluated by the students. It's 'bang on' if a lesson has been bad. I will find that out immediately. This was not good, so one really gets feedback. And then one has to be modest in one's teacher role and really take it in, not get into a defensive position. But what was good one should of course take credit for, too, and what one needs to redo (Nettle School, less experienced teacher).

When the teachers cited examples of different aspects of the kind of information they needed when evaluating, they also illustrated the diversity of the situated judgements that characterise professional responsibility. Solbrekke and Englund (2011, 855) emphasise this space for professional discretion “in deciding what is ‘best’ for their client and for society” – a space that these teachers often complained was too small.

### ***Collegial evaluations***

The importance of reflecting on and evaluating issues jointly with one's closest colleagues, and a perceived need for more time for doing this, as also noted as valuable in other studies (e.g. Biesta 2009; Hodkinson 2009), was often emphasised in the

interviews. These collegial discussions appeared to be more valued the closer they were to the individual's own teaching practice. The work team was a highly valued context for mutual reflection on, for example, the new grading system in compulsory school, which was repeatedly mentioned. Reflection together with colleagues working on the same subject, in the same or other schools in the municipality, was also noted as especially relevant. The aim of joint teacher reflection and evaluation included that of increasing professional consensus on pedagogical matters. A recurrent pattern noted in these discussions was the need, in the wake of recent educational reforms that included the introduction of a new grading system, to increase equivalence among students with different socio-economic backgrounds.

The importance of joint reflection and evaluation with teacher colleagues to create a common interpretation and understanding of, for example, the new grading system, is in line with what Solbrekke and Englund (2011) refer to as an implicit language framed by professions to characterise their professional responsibility. This also seemed to be the case when the teachers highlighted the importance of feedback from colleagues' observations of their teaching. Some interviewed teachers had experienced being observed during lessons by teacher colleagues from the same or other schools. These collegial observations, which included both oral and written feedback, were unanimously praised. One of the less experienced teachers described it as a "privilege" to be observed by a more experienced teacher, after which the teaching efforts were jointly evaluated (Nettle School). In Salvia School, where teachers from other schools in the same municipality had observed lessons and then provided oral and written feedback, collegial observation was described as "an ideal type of evaluation for developing teaching" (Salvia School, experienced teacher). One important feature here was being observed and receiving feedback from someone knowledgeable and whom "you trust".

In line with the above findings, the questionnaire results also indicated the importance of evaluations conducted close to the teachers' practice. Such evaluations carried out jointly with teacher colleagues, on one's own, or with students were believed to positively influence teaching to a "high" or "certain degree".

### **The (un)importance of external evaluations**

Despite the significance for teacher professionalism of teachers' own evaluations, the teachers did not seem to trust school providers to understand this. The teachers believed that the external evaluations were considered the important ones by school providers and the government:

They [i.e. external evaluations] get a certain weight. Our own reflexive evaluations that we do after every lesson or teaching segment, they are hard to show [in external evaluations]. That knowledge stays with us, as our experience instead (Lavender School, experienced teacher).

The above quotation exemplifies the difficulty of ‘translating’ teachers’ implicit language, which characterises professional responsibility, into the transparent language that is a feature of professional accountability (Solbrekke and Englund 2011).

Turning to the teachers’ views of external evaluations that focus on the whole school, all schools in the municipality or even all municipal staff, many critical views were expressed in almost all schools. Despite this criticism, it seemed difficult for the teachers to defend themselves from the evaluation discourse and the tempting logic that they had embraced when referring to their own evaluations and presumably also to the local quality work prescribed by the Education Act (SFS 2010:800) and promoted by the NAE and the SSI. Somewhat paradoxically in relation to their critical stance, some interviewed teachers expressed a hope for an evaluation instrument that would solve many of their problems. However, they had not yet found such an instrument:

If we could find a good evaluation instrument, and get a ‘virtuous circle’ going, where you could evaluate and then rearrange and change things, then it would help many of the things that stand in our way all the time. . . . all these thousand acute matters that have to come before and that interfere with one’s sight could be adjusted efficiently then (Lavender School, experienced teacher).

Although the teachers often expressed agreement with the logic of the evaluation discourse, most of them (with some exceptions primarily found in two schools, to which we will return later) did not describe much experience of successful external evaluations beyond their own or the ones conducted jointly with their immediate colleagues. Their delineation of external evaluation experiences was a list of examples of ‘no use – nothing happened’, ‘it takes far too much of our time’, and even has a ‘negative influence’ on school practice.

A common experience of the teachers, especially the more experienced ones, usually concerning external evaluations from the municipal council, was that even if they seriously strove to complete the evaluation questionnaires they were of no use as they rarely resulted in any changes. One teacher described how when teachers complained about things not working, their complaints simply bounced back to the teachers and the school:

So we do report it, and they [i.e. the school provider] take part of it and come back and say: ‘All right, we can’t do that now, but you still have to live with it every day, so what are you going to do about it?’. It makes you really mad! (Garlic School, experienced teacher).

The teachers experienced a lack of relevant feedback from the external evaluations and thought that repeatedly making the same complaints to the school provider was frustrating and made them feel the evaluations were being conducted for their own sake. In other words, from their perspective, these evaluations seemed to have

a 'ritual' (Vedung 1998) or 'symbolic use' (Mark and Henry 2004), making them a 'misuse' (Mark and Henry 2004) of their time. In our questionnaire, most teachers also agreed with the statement: "It is frustrating to complain about things that are not then improved". The results indicate that few of the teachers found that external evaluations had any favourable effects on their working conditions.

### ***Deprofessionalisation and uncertainty***

Altogether, evaluations from levels above the teacher's own impose a burden: "From the state, from the school provider, from the principal – and all of it ends up with us, in our laps and must be done" (Garlic School, experienced teacher). All these evaluations take a lot of time that the teachers argued could be better spent on planning upcoming lessons:

It's preparation and planning time you have to give up when you are going to complete all the questionnaires. *One* would not be a big deal but, together with all the others, it's just 'NO, NOT ANOTHER ONE!', and then I switch my computer off so I can use that time for something substantial that I can apply in my teaching ... that feels more meaningful (Salvia School, experienced teacher).

There were discussions among the teachers that more or less explicitly concerned teacher deprofessionalisation. Some teachers, as in the example below, explicitly argued there was a link between increased external evaluation and what they found to be an impoverishment of the teaching profession. Not only was there less time for planning, but some teachers also mentioned that all the extra administrative tasks had an effect on teacher creativity: "The free planning – that you are a bit creative in your thinking – I think that gets inhibited. I feel blocked sometimes because there's so much else" (Garlic School, experienced teacher). The teachers described an unwanted 'colonisation' of their time (Hargreaves 1994). In the questionnaire study, most teachers agreed with the above teacher's experience. Increased stress, less time for creative planning of teaching, and less time for reflection were cited as negative consequences of evaluations. That increased teachers' workload, in this case through evaluations, may inhibit creativity has also been noted in other studies (e.g. Day 2002; Hultqvist 2011; Mausethagen 2013b).

We noticed one difference between the experienced and less experienced teachers in that negative views of external evaluations were more common among the experienced teachers, probably due to their experience of times when much less evaluation and documentation was required. Some of the less experienced teachers instead stressed indifference: "These big evaluations of all school activities – I really don't feel part of them, so I can say that they are rather unimportant to me" (Salvia School, less experienced teacher). A similar difference was also observed in that teachers working in public schools expressed more negative views of external

evaluations than did teachers from independent schools. However, our sample of schools is very small with only three independent and five public schools.

The current external evaluations also had emotional consequences for teachers as they felt suspected of not carrying out their work satisfactorily: “They [the education department] really don’t trust us teachers to do our jobs without forcing us, controlling us – are we really present at work? They don’t think we are doing our jobs” (Garlic School, experienced teacher). It was fairly common for the teachers, especially the more experienced ones, to say they no longer felt fully trusted. They increasingly perceived teachers as being surveilled, mostly by the school providers but also by their ‘clients’ due to increased parental pressure. Greater parental demands were repeatedly mentioned in the interviews:

Now, it is more as if we are not trusted [by the parents] to grade their kids, so we have to have documentation to support the grades [in contacts with parents] (Carnation School, experienced teacher).

This sense of mistrust the teachers said they perceived on the part of both school providers and parents can be interpreted as indicating an erosion of trust in teachers’ professional responsibility and as a reaction to the external auditing and control focus characterising the concept of professional accountability (Solbrekke and Englund 2011). Earlier studies also have found teacher feelings of mistrust related to audit mechanisms and performance management (e.g. Evetts 2013; Fitzgerald 2008; Troman 2000). In a working paper, Michael Power (2013) elaborates on the routine production of documents as *audit trails* shaped by institutional demands for accountability and manifesting larger performance-monitoring regimes. Audit trails reflect how documentation and external evaluations have evolved in the Swedish education system. In recent years, Swedish school documentation, particularly of equal treatment and anti-discrimination work, has grown as a way to avoid parental pressure and prevent complaints to the SSI, which are still increasing in number (Skolinspektionen 2016).

A recurring comment by the teachers was that a change of school principal – which was fairly frequent – meant new directions in what was considered important to evaluate. The evaluations and quality work the teachers had started to grasp and even appreciate would suddenly be replaced by a new evaluation regime:

During the ten years I’ve worked here, the school evaluations have changed their names, not ten, but at least three, four, five times. So we have evaluated, and then a new principal, a new organisation, someone new [appears], and then they change the instrument . . . and it makes it much harder to compare the results (Lavender School, experienced teacher).

The school provider’s education committee might also change and, with this change, the committee’s policies and evaluation focus would also be altered. One teacher

illustrated this by citing the municipality's aim of fostering democracy, which had earlier meant that schools were evaluated according to whether their students took part in school planning: "Then it also meant a certain focus on this for us, but not so much attention has been paid to that since it disappeared [from the municipal aims]" (Lavender School, experienced teacher). These continuous changes in teachers' working conditions create uncertainty for them in accomplishing their tasks.

Although much discontent was expressed about the various external evaluations, the teachers nevertheless complied with and carried out most of these evaluations. Some resistance or response strategies (e.g. Hanberger and Gisselberg 2008), illustrated, for example, by the teacher who turned their computer off and instead did something that mattered more for the teaching, were described by interviewees. In another case, the teachers at one school, despite reminders, did not complete a compulsory questionnaire from the school provider. However, such resistance was rare and the overall impression from the interviews was that the teachers had completed most of the required evaluations.

### **A possible synthesis**

In two schools, the teachers were quite in favour of some evaluations from the principal and the school provider. The principals at these schools had succeeded in positioning these evaluations close to the teachers' actual practice. Coriander School started with an annual evaluation questionnaire from the school provider to students, parents and teachers. The principal organised thorough processing of the results and much effort, according to the teachers, went into identifying and analysing focus areas to be developed the next year:

It's a yearly questionnaire in February–March that teachers, students and parents complete. From that, we analyse at different levels how things are working and this will identify our working areas for the next year. In this evaluation, we note things that maybe we didn't achieve and if we got a very good result, we discuss how we achieved that result.

In our work teams, we analyse what the students in our grade say. Then, of course, you check at the school level. And I check as a tutor how I succeeded or how the students perceived what I did in my base group [Note: One teacher is a tutor for 20 students]. (Coriander School, experienced teachers).

In this school, the principal and teachers also created an operational plan based on the questionnaire. According to the teachers, continuous emphasis was placed on students' individual goal-setting and helping students to set goals every week followed by evaluations of those goals at the end of the week. The teachers also had individual weekly mentoring conversations with the students as well as evaluations of those conversations. Overall, the focus was on evaluations close to the students, on their individual results, but they also paid attention to areas to be developed during the year.

Garlic School had developed weekly evaluations of both teaching and student performance. These evaluations included frequent teacher self-evaluations of their own lessons and well-being as well as evaluations of student performance in all subjects. The evaluations were very elementary in form, simply intended to identify possible problems. They were carried out by the teachers and, when problems were identified, the teachers discussed them with colleagues and often also with the principal – depending on the kind and seriousness of the problem. A main purpose of these evaluations was to provide principals and teachers with an up-to-date overview of the current situation at the school, for example, regarding different teachers' or work teams' workloads and how different classes and students work, to be able to speedily direct effort to where it is most needed. If a teacher indicated or identified a problem, the colleagues or principal could offer to help:

We signal how we feel to the principals, and they usually take it seriously and ask, 'Why aren't you feeling all right?'. And then you have to explain yourself, and this usually results in some measures being taken – that's really good! (Garlic School, experienced teacher).

According to the teachers, these evaluations worked to the extent that some actions, within the limits of the local authority budget for the school, were taken:

So therefore those measurements – instead of just hearing things in the corridors, we have them clear and distinct on paper . . . and that means that we must take action there, do something about it. And so those [evaluations] have been rather good, positive (Garlic School, experienced teacher).

The teachers at this school also seemed to appreciate that the evaluations gave the principals quick information about problems with teaching, workload and so on. The principals could therefore take an active stance while listening to teachers' opinions, "and the principals had to see their part in the total situation" (Garlic School, experienced teacher).

Student performance in each subject was monitored by each work team on a board using simple markers showing the status of each student – everything ok, some worries, or alarming – to the teachers so they could take action if needed. This evaluation gave the teachers an overview of the current performance of the students in all subjects. Having these kinds of overviews and thereby being able to make further assessments of various matters was repeatedly noted as valuable. For example, it gave the teachers knowledge of whether a student was experiencing difficulties in one subject or was close to not passing in a number of subjects.

It [i.e. the board] is very good and explicit. You get a very clear picture of how the students are doing. It's a form of evaluation of the student, for us (Garlic School, experienced teacher).

To summarise, these two schools seemed to succeed in different ways in positioning their evaluations close to the teachers' own practice. This could be described as the 'transformation' of external evaluations into internal ones, which created possibilities to make situated judgements directed by the professionals themselves, the type of judgements characterising professional responsibility (Solbrekke and Englund 2011). In the first school, the evaluations identified focus areas that the teachers agreed on and then addressed with their students in the classroom. In the second school, the teachers seemed to appreciate that the evaluations were crude and simple to perform but still gave a snapshot of the teaching and classroom situation on which basis they and the principals could act.

### **Concluding discussion**

The aim of this paper was to investigate teachers' experiences of and approaches to school evaluations, with a special focus on how teachers' responsibility and professionalism are described in relation to evaluation. Our overall main finding is the great emphasis and importance attributed to internal teacher-initiated evaluations, which we have interpreted as signs of their professional responsibility (Solbrekke and Englund 2011). These evaluations, indicating teachers' ambitions to improve teaching, are generally overlooked in evaluation and school-policy research and seldom paid attention to or appreciated by school providers. Such evaluations are performed by the teachers themselves, together with their students and guided by collegial feedback and reflection on their teaching (principals who were also interviewed in this study appeared to agree with the teachers in highly valuing these evaluations, see Hult, Lundström and Edström 2016). When talking about these evaluations, the teachers express most of the characteristics of the logic of professional responsibility, such as situated judgement, a basis in professional mandate, and a moral rationale (Solbrekke and Englund 2011, 855). In addition, teachers arguably take responsibility for and defend their professionalism by sometimes ignoring external evaluations, which they consider take valuable time away from their core task. Performing their own evaluations and sometimes ignoring external evaluations can, with reference to Green (2011, 61–64), be considered as protecting their "responsibleness". However, their professional responsibility appears to be threatened by the volume of external evaluations and demands for accountability and by the lack of trust in teachers as practitioners. This lack of trust characterises the logic of accountability imposed on them largely from levels above the school.

Our second research question concerned the consequences of accountability and external evaluations for teachers' practice. A main finding is that the external evaluations illustrating the logic of professional accountability, as elaborated on by Solbrekke and Englund (2011), have a number of negative consequences for teachers' work. The control and external auditing performed according to an "economic/legal rationale" and using "predetermined indicators" (Solbrekke and Englund 2011, 855)

takes time that teachers claim could be better spent on planning and designing new lessons. Similar patterns have also been emphasised by several previous scholars (Ball 2003; Biesta 2009; Day 2002; Green 2011; Hargreaves 1994; Mausethagen 2013a). Accountability and external evaluations seem to make teachers less independent and creative, thereby having constitutive effects on the teacher profession (Dahler-Larsen 2014). Also tied to the experience of control were the teachers' feelings of being mistrusted by both the school provider and parents, which likewise could have constitutive effects on the teaching profession. For example, parents' lack of trust is forcing teachers to test students more often to justify student grades to parents.

Can the negative consequences of accountability and the gap between external and internal evaluations be reduced? Our results imply that the principal may sometimes succeed in transforming external evaluations into internal ones so close to the teachers' practice that the teachers consider them relevant to improving their work and thereby strengthen their professional responsibility. This is the case in two of the eight schools. These two schools, where the teachers were quite in favour of external evaluations coming from the principal or school provider, differ in several respects. First, as illustrated in the results, the types of evaluations carried out in these two schools differ. Second, the contextual characteristics of the two schools also differ. For example, one is an independent school with high average grades in an area with well-educated parents and mainly academically oriented students. The other is a municipal school in an area with a high proportion of immigrants and refugees and a high proportion of students whose mother tongue is not Swedish, meaning that some of them even have difficulty understanding the language their teachers speak. These results indicate that, instead of one 'ultimate' evaluation instrument that works for all, different schools with different contextual characteristics may find different types of evaluations valuable. For example, the overview of how the students were doing served as an early warning system in the municipal school, while most of the students in the independent school were not considered to need that kind of monitoring by their teachers. Following the logic of professional responsibility, situated judgements framed by professions can adapt some external evaluations to local school conditions.

The overall emphasis in these more 'relevant' (to the teachers) external evaluations as well as in the more distant ones is on student performance. Despite the strong emphasis on fostering democratic citizens in the steering documents governing schooling in Sweden, there appears to be a narrowing of goals when it comes to evaluating school quality. The heavy weighting on performance measures in school helps define the desirable qualities and goals of school for teachers, students and parents – a constitutive effect of the way school evaluation is carried out today (Dahler-Larsen 2014). Another indication regarding the constitutive effects of the evaluation society and the dominant school evaluation discourse on what it means to be a teacher and how teaching work is to be executed was the unquestioning way in which all

evaluations, according to the teachers' answers, were performed, despite many critical views. For most teachers, it seemed to be a given that a stream of external evaluations was now part of the teacher profession. A sign that this marked a change in teachers' perceptions of the profession was that the more experienced teachers were more negative towards external evaluations than were the less experienced ones. For the less experienced teachers, these evaluations seemed a more obvious and inherent part of the profession.

Still, the primary finding of this study – which deserves to be emphasised in relation to evaluation and school-policy research and to be disseminated among school providers, policymakers, politicians and teacher educators – is the significance attributed to internal evaluations performed by teachers on a daily basis. If the importance of these evaluations becomes understood and respected by those governing and influencing school policy, it might be possible to reclaim some lost parts of teachers' professional responsibility.

---

*Agneta Hult* is Associate Professor at the Department of Education, Umeå University. Her research interests are mainly focused on the implications and consequences of evaluations and assessment in the field of education. Email: [agneta.hult@umu.se](mailto:agneta.hult@umu.se)

*Charlotta Edström* has a PhD in educational work. Her primary research interests are within the areas of evaluation policy and practice and gender equality issues. Email: [charlotta.edstrom@umu.se](mailto:charlotta.edstrom@umu.se)

## **Notes**

- 1 The school providers are either municipal school boards, independent school boards or companies (often major concerns and sometimes organisations based on a specific pedagogical philosophy, as in Montessori schools).
- 2 The project “Consequences of evaluation for school practice: governance, accountability and school development” is funded by the Swedish Research Council.

## References

- Ball, S. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228.
- Ball, S., Maguire, M., Braun, A., Hoskins, K. & Perryman, J. (2012). *How schools do policy: policy enactments in secondary schools*. London and New York: Routledge.
- Bergh, A. (2015). Local quality work in an age of accountability – between autonomy and control. *Journal of Education Policy* 30(4): 590–607, doi: <http://dx.doi.org/10.1080/02680939.2015.1017612>
- Biesta, G. (2009). Values and ideals in teachers' professional judgement. In *Changing Teacher Professionalism: International Trends, Challenges and Ways Forward*, S. Gewirtz, P. Mahony, I. Hextall and A. Cribb (eds.), 184–194. London: Routledge.
- Biesta, G. (2010). *Good education in an age of measurement: ethics, politics, democracy*. Boulder, CO: Paradigm Publishers.
- Dahler-Larsen, P. (2011). Afterword: Evaluation as a field and as a source of reflection: comments on how QAE restructures education now and in the future. In *Fabricating Quality in Education: Data and Governance in Europe*, J. Ozga, P. Dahler-Larsen, C. Segerholm and H. Simola (eds.), 150–159. London and New York: Routledge.
- Dahler-Larsen, P. (2012). *The evaluation society*. Stanford, CA: Stanford University Press.
- Dahler-Larsen, P. (2014). Constitutive effects of performance indicators: getting beyond unintended consequences. *Public Management Review*, 16(7), 969–986.
- Day, C. (2002). School reform and transitions in teacher professionalism and identity. *International Journal of Educational Research*, 37(8), 677–692.
- Englund, T. & Solbrekke, T. D. (2015). Om innebörder i lärarprofessionalism. [On meanings of teacher professionalism. In Swedish]. *Pedagogisk Forskning i Sverige. [Pedagogical Research in Sweden]*, 20(3–4), 168–194.
- Evetts, J. (2011). A new professionalism? Challenges and opportunities. *Current Sociology*, 59(4), 406–422.
- Evetts, J. (2013). Professionalism: value and ideology. *Current Sociology*, 61(5-6), 778–796.
- Fitzgerald, T. (2008). The continuing politics of mistrust: performance management and the erosion of professional work. *Journal of Educational Administration and History*, 40(2), 113–128.
- Green, J. (2011). *Education, professionalism and the quest for accountability: hitting the target but missing the point*. New York: Routledge.
- Hanberger, A. & Gisselberg, K. (2008). *IFAD's, EuropeAid's and Sida's management response system: a comparison of the systems' design and practice*. Evaluation reports No. 23. Umeå, Sweden: Umeå University, Umeå Centre for Evaluation Research.
- Hargreaves, A. (1994). *Changing teachers, changing times: teachers' work and culture in the postmodern age*. Trowbridge, UK: Redwood Books.
- Hargreaves, A. (2000). Four ages of professionalism and professional learning. *Teachers and Teaching: Theory and Practice*, 6(2), 151–182.
- Helgoy, I. & Homme, A. (2007). Towards a new professionalism in school? A comparative study of teacher autonomy in Norway and Sweden. *European Educational Research Journal*, 6(3), 132–149.
- Hodkinson, H. (2009). Improving schoolteachers' workplace learning. In *Changing Teacher Professionalism: International Trends, Challenges and Ways Forward*, S. Gewirtz, P. Mahony, I. Hextall and A. Cribb (eds.), 184–194. London: Routledge.
- Hult, A., Lundström, U. & Edström, C. (2016). Balancing managerial and professional demands: school principals as evaluation brokers. *Education Inquiry*, 7(3), 29960, doi: <http://dx.doi.org/10.3402/edui.v7.29960>

- Hultqvist, E. (2011). Om lärarnas förändrade yrkesvillkor [About teachers' changed working conditions]. *Pedagogisk forskning i Sverige*, 16(3), 203–213.
- Lindgren, L. (2014). *Nya utvärderingsmonstret: om kvalitetsmätning i den offentliga sektorn*. 2:a reviderade upplagan. [The new evaluation monster: on quality measurement in the public sector]. 2nd revised edition. Lund, Sweden: Studentlitteratur.
- Lundström, U. (2015). Systematic quality work in Swedish schools. Intensions and dilemmas. *Scandinavian Journal of Public Administration*, 19(1), 3–23.
- Mark, M. M. & Henry, T. G. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35–57.
- Mausethagen, S. (2013a). Accountable for what and to whom? Changing representations and new legitimization discourse among teachers under increased external control. *Journal of Educational Change*, 14(4), 423–444.
- Mausethagen, S. (2013b). A research review of the impact of accountability policies on teachers' workplace relations. *Educational Research Review*, 9, 16–33.
- Ozga, J. (1995). Deskillling a profession: professionalism, deprofessionalisation and the new managerialism. In *Managing Teachers as Professionals in Schools*, H. Busher and R. Saran (eds.), 21–37. London: Kogan Page.
- Ozga, J. (2000). *Policy research in educational settings: contested terrain*. Buckingham, UK: Open University Press.
- Power, M. (1997). *The audit society: rituals of verification*. Oxford: Oxford University Press.
- Power, M. (2013). Organizations and audit trails. Paper presented at the Swedish Evaluation Society (SVUF) conference in Stockholm, 13–15 November 2013.
- Polanyi, M. (1966). *The tacit dimension*. New York: Doubleday.
- Rönnerberg, L., Lindgren, J. & Segerholm, C. (2013). In the public eye: inspection and local newspapers: exploring the audit–media relationship. *Journal of Education Policy*, 28(2), 178–197.
- SFS 2010:800. *Skollag*. [Education Act. In Swedish.] Stockholm: Skolverket.
- Skolinspektionen. (2016). <http://skolinspektionen.se/sv/Statistik/Statistik-om-anmalningar/anmalningar-och-beslut-2015/> [Statistics of complaints and decisions 2015. In Swedish]. (Accessed 2016-02-19).
- Solbrekke, T. D. & Englund, T. (2011). Bringing professional responsibility back in. *Studies in Higher Education*, 36(7), 847–861.
- Svensson, L. G. & Evetts, J., eds. (2010). *Sociology of professions: continental and Anglo-Saxon traditions*. Gothenburg, Sweden: Daidalos.
- Troman, G. (2000). Teacher stress in the low-trust society. *British Journal of Sociology of Education*, 21(3), 321–353.
- Vedung, E. (1998). *Utvärdering i politik och förvaltning* [Evaluation in politics and administration]. Lund, Sweden: Studentlitteratur.