Struggle for Acceptance - Maintaining External School Evaluation as an Institution in Germany
Piezunka, Anne

Struggle for Acceptance – Maintaining External School Evaluation as an Institution in Germany

Anne Piezunka [*]

Abstract: »*Kampf um Anerkennung - Die Stabilisierung externer Schulinspektionen als Institution in Deutschland*«. In the educational field, evaluations based on standardised indicators play a major role in the determination of evidence-based regulations. To have an effect within a policy field, evaluations based on standardised indicators have to become institutionalised. Nevertheless, very little is known about the strategies that actors apply to institutionalise their procedure and whether these strategies are successful. This paper explores the strategies that German school inspectorates use to maintain external evaluations based on standardised indicators as an institution. The paper draws on educational governance research and institutional work. By conducting expert interviews with staff members of school inspectorates, I show which strategies are used to increase the acceptance by school representatives. Two strategies can be identified: school inspectors try to make less judgements and focus on descriptions of what they have observed (strategy I), and school inspectors give school representatives a greater say when formulating expectations for schools (strategy II). From a governance perspective, these strategies mean that school representatives are perceived as equal partners. They also imply that school representatives do not have to give up any of their autonomy, which is also a necessary condition for the institutionalisation of school inspections.

Keywords: Educational assessment, school inspection, institutional work, indicators, educational governance.

1.   Introduction

University rankings, test scores, and teacher evaluations – there are several ways to assess quality within the educational field. These types of evaluation have become highly institutionalised in recent decades: for example, Espeland and Sauer studied how universities react to rankings. According to their study,

some universities changed their admission policies to improve their position within the rankings (Espeland and Sauder 2007). Another possible indicator for the institutionalisation of a mode of evaluation is whether its existence is taken for granted (Scott 2008). This also applies to university rankings: "We are so accustomed to rankings that they have become a naturalised way of making sense of the world" (Espeland and Sauder 2016, 5; see also Kehm 2014, 102; Salais 2008). Nevertheless, many criticisms have arisen in recent years: researchers and practitioners have criticised several methodological aspects. Rauhvargers (2014) questions the selection of indicators that are used to compare universities. For example,

> University rankings continue to focus on the research function of the university and are still not able to do justice to research carried out in the arts, humanities and the social sciences. (Rauhvargers 2014, 40)

Nevertheless, while external evaluations such as university rankings in the US or the school inspection model in England (OFSTED) can be understood as highly institutionalised, certain types of evaluation within the educational field are only partly institutionalised.

In the case of the federal states of Germany, a new type of external evaluation was introduced in 2004/05 in every federal state. The states' ministries of education established external evaluations based on standardised indicators to estimate the quality of public high schools. When evaluating the quality of schools, school inspectors analyse documents, conduct personal interviews, and carry out classroom observations. Thereby, inspectors apply a set of indicators: frequently used indicators in classroom observation sheets are "The class starts on time" or "The children are treated with appreciation".

After the first cycle of school inspections, which took around five years, many criticisms arose. Several researchers questioned whether school inspections had any impact on school improvement measures (Gärtner, Wurster, and Pant 2013; Gärtner 2011). Furthermore, several teacher associations and unions argued that school inspections meant a large workload for schools but were not very effective (BLLV 2017; GEW Berlin 2008). In response to these criticisms and for financial reasons, some federal states abolished their school inspections (Hessen in 2016, Rhineland-Palatinate in 2015) or put them on hold (Saxony since 2015, Thuringia since 2015, Baden-Wuerttemberg since 2017, Bavaria since 2018, Schleswig-Holstein between 2009 - 2016). Within the remaining school inspectorates, some staff members articulated the need to undertake changes to ensure their continued existence (in the case of Lower Saxony, see Sowada and Dedering 2016, 183).

Against this backdrop, this paper explores the strategies used by staff members of school inspectorates to maintain external evaluation based on standardised indicators as an institution and discusses whether they were successful. Due to the fact that many decisions in the educational field are justified with

reference to external evaluations (Heinrich 2010), it is important to analyse which circumstances define the development of indicators:

> Decisions about how to make elements of experience commensurable are particularly influential during the development of metrics, and this can generate path dependence for subsequent interpretation of data and its effect on action. (Sellar 2014, 133)

To identify the strategies used by staff members of school inspectorates to maintain external evaluation as an institution, this paper draws on research on institutional work. In this regard, Lawrence and Suddaby describe the practices of individual and collective actors aimed at creating, maintaining, and disrupting institutions (Lawrence and Suddaby 2006). Furthermore, to analyse the role of school inspectorates within the German education system, the paper builds on the literature on school inspectorates (Heinrich 2015; Dietrich and Lambrecht 2012; Sowada and Dedering 2016) and relies on the educational governance approach (Altrichter, Brüsemeister, and Wissinger 2007) as a research perspective. The educational governance approach identifies dependencies and analyses expectations and interests formulated by different actors. In methodological terms, I conducted interviews with staff members of school inspectorates in several federal states and applied a qualitative content analysis (Gläser and Laudel 2009). I argue that staff members of school inspectorates in some federal states applied similar strategies to increase their acceptance by school representatives and thus seek to maintain external evaluation as an institution. In general, this study situates itself in a research field that understands indicators as socially constructed and not as a neutral instrument to gather data (Diaz-Bone 2017; Lamont 2012; Kehm 2014, 103).

2. Case Study: School Inspections in Germany

All federal states introduced school inspectorates around the same time (Füssel 2008, 153). However, there are many differences between state-specific inspectorates. Each inspectorate developed its own indicators and procedures. For example, there are differences from state to state regarding who receives the report and whether former teachers are the only ones permitted to conduct school inspections. There are also many similarities between the school inspectorates. Most of them use standardised indicators (except Mecklenburg-Vorpommern and Bremen[1]) to assess school quality. They apply similar methods to gather data, such as personal interviews, classroom observation, and online questionnaires. Further, most of them have similar expectations regard-

---

[1] As they don't apply standardised indicators, the federal states Bremen and Mecklenburg-Vorpommern are not part of the sample.

ing how schools should deal with the inspection report: after the inspection, schools are tasked with negotiating target agreements with school authorities. Furthermore, in some federal states, if a school performs poorly, the inspection interval is shortened, and the next inspection takes place after two years instead of after four to five years.

Whereas in other countries there exists publicly accessible information on school quality, there are only three federal states in Germany that publish a shortened version of the inspection report online (Berlin, Hamburg, and Brandenburg), and no federal state publishes rankings. The fact that school rankings are not made public might be explained by the existence of strong teacher associations that are opposed to it (Piezunka 2018).

Furthermore, in the case of Germany, there is no federal state in which schools face sanctions such as school closure or financial cuts, as would be the case for failing schools in England and the Netherlands. Therefore, compared with other countries, the German inspection model is often described as a "low stakes system" (Altrichter and Kemethofer 2015). Furthermore, when assessing the quality of schools, German school inspectorates focus on data at the process level: they want to learn more about day-to-day practices. In other states, data at the output level, particularly the results of test scores, play a major role, which is not the case in Germany (Piezunka 2018).

The introduction of school inspections can be understood as a political measure implemented after the so-called "PISA-shock" in Germany (Böttcher and Keune 2011; Altrichter and Kemethofer 2016). The states' ministries of education wanted to know what was going on in schools after they had obtained bad results in the PISA study. The PISA study is an international programme of the OECD which evaluates national education systems by measuring 15-year-old students' performance on mathematics, science and reading. Hence, they introduced school inspectorates as a supplement to school authorities, which were and still are responsible for monitoring schools in Germany (Maritzen 2008; see also Heinrich 2015). The introduction of school inspections was a response to the so-called "crisis of school authorities" (Maritzen 2008): the rise of evidence-based governance in the educational field had identified a need for the introduction of school inspections in order to collect data systematically, as the school authorities didn't do this. Before this, school monitoring primarily consisted of peer feedback, examination of future teachers, and personal exchange with school authorities.

Furthermore, the understanding of school governance has changed (Heinrich 2007; Dedering and Sowada 2016). In the nineties, schools were given more autonomy regarding finances, pedagogy, and curricula (Heinrich 2007). The result was that educational ministries had to establish measures to monitor how schools dealt with their new degree of autonomy. This was another reason for the introduction of school inspections. Avenarius describes this development as a "contradictio in adiecto" (Avenarius 2006, 6-7), as the monitoring of schools

by school inspectorates would also imply that school representatives again lose some degree of autonomy.

The monitoring of schools is just one of several functions that school inspectorates should theoretically fulfil. According to Landwehr (2011, 39f; see also Maritzen 2008; Heinrich et al. 2014) further functions are ascribed to school inspectorates within the education system: second, school inspectorates are required to generate knowledge about schools that can be used for a variety of purposes (Landwehr 2011). To do so, they use standardised indicators to gain insights into the current state of the school (function of knowledge production). Therefore, school inspectorates generate empirical data that should theoretically provide an objective knowledge base for policymakers (Dietrich and Lambrecht 2012). Third, based on the knowledge they have gained, school inspectors are tasked with initiating school improvement measures (Landwehr 2011): the results of school inspections give schools an idea of the issues that should be addressed in the future (initiating function). Fourth, school inspections are intended to contribute to implementing standards of school quality within the federal state. The indicators thus represent normative understandings of what a good school is (function of implementing standards) (Landwehr 2011). Fifth, by defining school quality and interpreting policy guidelines, school inspectorates set expectations. Therefore, they act as policymaker (Piezunka 2018).

In this context, it should be considered that some of these functions might have contradictory implications for schools. If school inspectors want to initiate school improvement measures, they need to know what challenges schools are currently facing. For this to occur, it is necessary for schools to speak openly about their problems. However, given the monitoring function, schools want to create a good impression ("window dressing"). They therefore tend to hide the challenges and problems they are facing. Furthermore, the monitoring function and the initiating function represent different ideas of school governance. The monitoring function implies a hierarchical relationship between school inspectorates and schools, whereas the initiating function aims at "school development through reason" (Dietrich and Lambrecht 2012). "Inspectors and actors at school shall meet each other 'at eye level'" (Dietrich and Lambrecht 2012, 57). While the monitoring function means that schools must meet school inspectorates' expectations, the initiating function allows for the perspectives and interests of school representatives to be heard and considered during the inspection. The initiating function might even imply that school representatives have a different understanding of the schools' performance and therefore do not initiate any changes after the inspection has taken place (Dietrich and Lambrecht 2012).

Regarding the functions described above, school inspectors' descriptions of their role within the education system and the schools' perception of school inspectorates vary (Heinrich et al. 2014). For example, it is possible that one

function is perceived to be more important than another or that one function isn't relevant at all. In addition, there are not only variations between actors, but also differences at the state level regarding the weighting of the functions. Therefore, in actual practice, there may be differences in the relative importance of these functions and, as a result, some attributes of the inspection process might vary depending on the relevance of the functions.

3.   Theoretical Framework

To understand how staff members of school inspectorates act and which strategies they apply to maintain school inspection as an institution, this paper relies on the educational governance approach as a research perspective. Furthermore, defining school inspections as an institution with a low degree of institutionalisation, the study draws on research about institutional work.

3.1   Educational Governance Approach

The educational governance approach can be understood as a research perspective that is interested in questions of regulation and is based on certain theoretical assumptions. Changes in the educational field are "not shaped by a single dominant actor" (Altrichter 2010, 148). Therefore, actors' interests and dependencies between actors must be considered to understand certain practices (Kussau and Brüsemeister 2007). Furthermore, formal hierarchies might not explain dependencies between actors.

Within the education system, school inspectorates and school authorities are part of the public educational administration and can be described as an intermediary actor between the educational policy level and the school level (Brüsemeister 2007, 93; Sowada and Dedering 2016, 175). School inspectorates have an ambivalent relationship with high-level educational policymakers: on the one hand, policymakers introduced the school inspectorates to monitor schools. On the other hand, policymakers will be held accountable by the public if too many schools are rated poorly by the inspectorates. Therefore, policymakers want schools within their federal state to perform well on average when they are being evaluated by school inspectorates (Piezunka 2018). Furthermore, school inspectorates depend on high-level educational policymakers because they decide whether school inspectorates continue to exist.

The relationship between school inspectorates and schools depends on the functions that school inspectorates are supposed to fulfil. The initiating function implies that school inspectors recognise school representatives as professionals and perceive them as an equal partner during the inspection process. By contrast, the monitoring function implies a strict hierarchical relationship as well as the requirement for schools to meet the expectations of inspectorates.

Therefore, when school inspectorates try to fulfil both functions, schools receive ambiguous signals regarding their position of power.

Another important actor for school inspectorates are school authorities. School authorities are tasked with negotiating target agreements with the schools after the inspection has taken place, but in practice they often perceive the inspectorates' data as unhelpful (Heinrich 2015). Furthermore, and as mentioned above, there are certain overlaps between school inspectorates and school authorities regarding their areas of competence within the educational system. Both are supposed to monitor what is going on at schools but they use different monitoring procedures and they have a different scope of action towards schools. Nevertheless, these overlaps require school inspectorates to legitimise their procedure within the education system (Heinrich 2015). To sum up, school inspectorates face diverging interests. In addition, school inspectorates also send ambivalent signals regarding their own function.

3.2   Institutional Work

This paper understands external evaluations based on standardised indicators as an institution. Scott describes them as follows:

> Institutions are comprised of regulative, normative and cultural-cognitive elements that, together with associated activities and resources, provide stability and meaning to social life. (Scott 2008, 48)

The regulative pillar of institutions is reflected in certain rules and it is based on a coercive mechanism (Scott 2008, 51). The normative pillar represents "rules that introduce a prescriptive, evaluative, and obligatory dimension into social life" (Scott 2008, 54). The cultural-cognitive pillar describes the "shared conceptions that constitute the nature of social reality and the frames through which meaning is made" (Scott 2008, 57). Whether all three elements (regulative, normative, cultural-cognitive) are at work varies from one institutional form to another (Scott 2008, 62).

Regarding school inspection, these pillars can be used to describe the degree of institutionalisation. The description of the degree of institutionalisation varies depending on its function. Therefore, it makes sense to distinguish between the monitoring function and the initiating function of school inspectorates. In theory, the monitoring function implies the following: school inspections are obligatory for all schools (regulative pillar). Schools should fulfil the expectations formulated by school inspectorates (normative pillar). Furthermore, school representatives are aware that monitoring through external evaluators is necessary (cultural-cognitive pillar).

By contrast, the initiating function has the following implications regarding the institutionalisation of school inspections: after the inspection has taken place, school representatives make target agreements with school authorities about future measures (regulative pillar). School representatives should use the

feedback provided by school inspectorates when implementing school development measures (normative pillar). School representatives are aware that they need external feedback from school inspectorates to develop further (cultural-cognitive pillar).

To sum up, school inspection in Germany would be highly institutionalised if these descriptors applied in practice. For some aspects, this might be the case. For example, school inspection is obligatory for schools in every federal state (except Schleswig-Holstein). Hence, schools cannot choose whether they want to be evaluated by school inspectorates. Furthermore, it is probable that schools try to meet the expectations formulated by school inspectorates to avoid having to justify poor results. This might be the case especially in those states where a shortened report is published online and is therefore accessible for potential future students and their parents. Regarding the normative pillar, an empirical study in lower Saxony shows that school inspection increases the awareness of principals and teachers of the strengths and weaknesses of their schools (Sommer 2011, 144; see also Senatsverwaltung für Bildung and Jugend und Wissenschaft Berlin 2014). Nevertheless, there are also many studies that prove that school inspection has a low impact on school development measures (for an overview, see Altrichter and Kemethofer 2016). Furthermore, the cultural-cognitive pillar does not play a major role in Germany. Whereas OFSTED, the school inspection model in England, is taken for granted, many actors in Germany question whether there is a need for data from school inspections to assess the quality of schools (BLLV 2017; GEW Berlin 2008). In addition, school inspectorates in some federal states were abolished and others are afraid that something similar is going to happen to them.

To sum up, school inspections are institutionalised to a certain degree as most of them have been in operation since 2004/05, they are obligatory for schools, and school representatives try to meet the requirements formulated by the indicators. Nevertheless, school inspections are only partly institutionalised because many school representatives oppose the procedure and still do not take school inspections for granted. Therefore, there is a need for action to maintain school inspection as an institution.

While a lot of research treats institutions as given (Scott 2008, 93), there is a lack of research analysing the measures that are used to maintain institutions (Scott 2008, 122; Lawrence and Suddaby 2006, 23). It is important to research the maintenance of institutions as there are only a few institutions that do not require intervention to ensure their maintenance (Lawrence and Suddaby 2006, 229). Lawrence and Suddaby argue that strategies used to maintain institutions should

> be distinguished from simple stability or the absence of change; institutional work that maintains institutions involves considerable effort, and often occurs as a consequence of change in the organization or its environment. (Lawrence and Suddaby 2006, 234)

Based on a review of the empirical literature, Lawrence and Suddaby (2006) describe six forms of institutional work that are applied to maintain institutions (Lawrence and Suddaby 2006, 230): the first activity, "enabling", refers to strategies which "facilitate, supplement and support institutions" (Lawrence and Suddaby 2006, 230). This might imply the introduction of a new actor who is tasked with ensuring the adherence to the norms of an institution. Furthermore, enabling work might also imply that existing conflicts between institutions are overcome (Lawrence and Suddaby 2006, 231). The next activity, "policing", suggests that strategies with a high degree of coercion are applied to ensure that actors comply with the rules of an institution (Lawrence and Suddaby 2006, 231). For example, there would be sanctions in the case of non-compliance. Furthermore, auditing and monitoring can also be understood as a form of "policing" (Lawrence and Suddaby 2006, 231). A third activity Lawrence and Suddaby (2006, 232) have identified is "deterring." It refers to strategies which prevent the development of new institutions or institutional change. In the context of school evaluations, obligating practitioners and researchers to use the data collected by school inspectorates to evaluate schools constitutes an example of deterring behaviour.

Furthermore, the fourth activity, "valourizing and demonizing", tries to link existing norms to the institution (Lawrence and Suddaby 2006, 232). School inspectorates could use narratives to show that their work is necessary to ensuring that children are able to learn and to feel safe. Another form of institutional work is "mythologizing" (Lawrence and Suddaby 2006, 233). In practice, this might imply the development and communication of founding myths which represent the norms of the institution. The sixth form of institutional work is "embedding and routinizing." It refers to strategies which make sure that the institution becomes part of the practitioners' daily practices (Lawrence and Suddaby 2006, 233). In the case of school inspection, one strategy of routinizing has been observable since the beginning, as the school inspection usually takes place every four to six years.

If German school inspectorates take into account the situation of school inspectorates in other countries, it is probable that they apply strategies which are already in use in those countries. This would be the case if school inspection introduced rankings or sanctions for failing schools which can be linked to "policing". Regarding the existing dependencies within the educational system, school inspectorates might apply strategies which can be linked to another activity: "enabling work." This form of institutional work might lead to overcoming existing areas of tension with schools because, as described above, school inspections threaten their autonomy.

Furthermore, the following has to be taken into account: Although actors apply different strategies to create and maintain institutions, they are not always successful in this endeavour (Scott 2008, 96).

4.   Methods

In order to capture strategies, I conducted 18 semi-structured interviews with the staff members of school inspectorates in nine federal states carrying out external evaluations based on standardised indicators[2]. Most of them worked as teachers or school principals before they started to work for the school inspectorate. Some of them continued working as teachers part time while they conducted external evaluations. Furthermore, some federal states employed so-called "scientific instructors" who were responsible for data analysis and further empirical questions. Some of the persons interviewed mainly conducted external evaluations at schools on a regular basis, while others were responsible for scientific issues. All of them were involved in the development of indicators.

The interviews lasted between thirty and ninety minutes. All interviews were audiotaped and transcribed. I used MAXQDA software and a qualitative content analysis approach as described by Gläser and Laudel (2009) to analyse the interviews. Following the educational governance approach, the analysis was based on the following categories: "relevant agents and their interests regarding school inspections," "changes regarding the indicators," and "functions of school inspections." In this process, I noticed certain patterns regarding the acceptance of school inspections by school representatives and the strategies applied by staff members to increase this acceptance. During the analysis, Lawrence and Suddaby's (2006) heuristic model allowed me to identify actions which are related to maintaining external evaluations based on standardised indicators as an institution.

5.   Results

In the following, I describe the strategies that staff members use to maintain school inspections as an institution. These strategies aim at increasing the acceptance of these inspections by school representatives. As Scott argues, social acceptance from certain actors has important implications for the degree of institutionalisation (Scott 2008, 59). School inspections often lack acceptance from many school representatives because of the negative positions of some lobby associations described above. In this regard, staff members of school inspectorates assume that the lack of acceptance may be one reason why school inspections have a limited impact on school development measures (initiating function):

---

[2] This paper is based on interview material that is part of the author's PhD-project on the development of indicators regarding inclusive education (Piezunka 2018).

> If you want to initiate measures at schools, you need a procedure that is accepted by schools. School representatives should have the impression that they can accept what school inspectors observe and report. (18_19; also; 8_36; 15_12; 18_15; 5_38; 11_17)

Therefore, staff members assume that a higher acceptance of inspections by school representatives would allow school inspectorates to fulfil their initiating function to a higher degree (see also Sowada and Dedering 2016, 192). Furthermore, school representatives' increased acceptance might have a positive impact on policymakers' opinions of school inspection, and this might be the case irrespective of whether or not school inspectorates fulfil the initiating function.

Some interview partners even link the school inspectorates' right to exist to its acceptance by school representatives: one interviewee argued that the inspectorate she works for is not at risk because "we treat schools very carefully and we are very open and appreciative from the beginning" (2_9). In addition, one interviewee said that the school inspectorate he was working for faced closure because it was strongly rejected by lobby associations (18_17). In sum, staff members argue that the acceptance of school inspections by school representatives is necessary if they want to fulfil their initiating function towards schools. In turn, they assume that fulfilling the initiating function to a higher degree could stabilise school inspection as an institution. In the following, I will describe strategies used by staff members of school inspectorates to increase their acceptance by school representatives and to maintain school inspection as an institution: 1) by having more descriptive and less evaluative components in the inspection report and 2) by involving school representatives when formulating expectations.

5.1 More Description, less Evaluation

Even if their database allows rankings, no German school inspection publishes any rankings. In addition, after the first cycle, some school inspectorates began dividing their inspection report into a descriptive and an evaluative part (for Lower Saxony, see Sowada and Dedering 2016). Some features of schools are described but not evaluated. For example, they write that "40% of the observed classes started on time." They do not link the description to any benchmarks or make evaluative statements.

> In the classical sense of 'good', there is no school that scores well or badly. [...] We don't want to evaluate. This is why we don't say 'regarding our standards of school quality, you are a good school or you are a bad school.' (9_49)

Interviewees include descriptive parts in their inspection reports in order to leave the task of evaluation to school representatives: "I don't know if it is good or bad. Only the school knows – the school knows if this is what they

want" (4_18; also 17_34). Others claim that school representatives can take into account the circumstances of the specific school:

> We don't evaluate whether the school is good. Maybe they have tried really hard for five years or for three years to reach a certain level. Maybe they had 3% and now they have 15%. (5_17)

To sum up, in some states, school inspectors only describe and evaluate certain features of schools in their reports. They want to leave the task of evaluation to school representatives. Nevertheless, they report that many schools perceive the descriptive part as evaluative: "It's their self-perception: 'We are bad. [...] We have been evaluated as bad.' School inspectors say: 'No, we didn't evaluate anything. We just observed and described. You are evaluating'" (11_16). The reason might be that school representatives don't have the initiating function of school inspectorates in mind but instead perceive them as a controlling body (monitoring function).

Furthermore, it has to be taken into account that "each choice of indicators and weighting reflect value judgements" (Kehm 2014, 103). For example, the indicator "the class starts on time" represents a commonly held norm, namely that it is preferable for all classes to start on time. Hence, even if they try just to describe what they see, it already implies a certain degree of evaluation.

Irrespective of whether the school inspectorates' strategy might not be understood as intended, it can be seen as an attempt to increase the acceptance of inspections among school representatives and therefore to have a higher impact on school development measures (initiating function).

5.2   Formulating Expectations for Schools

Various actors are involved in the development of indicators: educational researchers, school authorities, and further representatives from the ministries of education. To increase acceptance of inspections by school representatives, school inspectors also want to consider the views of school representatives when they formulate expectations towards schools. "It's like that… we try, no … we know that we have to and we want to cooperate with others if we want to have an impact" (16_15). The issue of inclusive education illustrates the involvement of school representatives regarding the development of indicators: in 2009, Germany signed the UN Convention on the Rights of Persons with Disabilities. As a result, Germany has to implement an inclusive education system. However, there is no shared understanding of how to define inclusive education in Germany (Moser 2012; Piezunka, Schaffus, and Grosche 2017). School inspectors considered the views of school representatives when they developed indicators regarding inclusive education: "Schools should implement inclusive education, but we must not expect too much of schools. We have to take into account what schools are able to do in light of their circumstances" (8_16; also 7_28; 12_31). Furthermore, many staff members consid-

ered their experience as school inspectors. When evaluating the schools within their federal state, they got an idea of what school representatives think is feasible regarding inclusive education in their specific circumstances. When formulating expectations about inclusive education, school inspectors weighed up their own ideas of how inclusive education should be implemented in schools against what they knew about the concrete implementation process in their federal state (Piezunka 2018).

In some cases, staff members admitted that their expectations regarding the implementation of inclusive education were very low:

> We require very, very low standards – this might change over the years when there are more schools that implement inclusive education. Then, we could increase the standards. (17_47)

They formulated expectations that are achievable for the majority of schools in order not to overburden schools.

Nevertheless, school inspectors are aware that the implementation of inclusive education is prescribed by law and that this legal requirement exists independently of the school-specific circumstances and opinions of school representatives. Hence, they are acting in a context riddled with conflict: on the one hand, they want to control the implementation of inclusive education as a legal entitlement. As a result of this, some inspectors argue that certain minimal standards should be fulfilled to make sure that students with disabilities receive the support they need. On the other hand, many inspectors do not want to expect too much of schools because they assume this would decrease the acceptance of school inspections in general. To sum up, the feasibility of expectations plays a major role in the formulation of expectations of schools by school inspectorates.

6.   Discussion

Staff members of school inspectorates used different strategies to increase their acceptance by school representatives. They changed the presentation of their data (more descriptive, less evaluative) and gave school representatives a greater say when formulating expectations. Therefore, these strategies have an impact on which indicators are applied to gather data and how the results of the data collection are presented to the public. By increasing acceptance, staff members of school inspectorates hope to have a higher impact on school development measures (initiating function). Furthermore, school representatives' acceptance is not only important regarding the initiating function but also because high-level educational policymakers consider what school representatives think when they decide whether school inspections should continue to exist.

Comparing the strategies of school inspectorates with the six practices of institutional work described by Lawrence and Suddaby (2006, 230), we can see evidence of "enabling" as one form of institutional work. Increasing school representatives' acceptance can be understood as a strategy which is applied to overcome existing areas of tension with school representatives ("enabling work"). By contrast, the practice of "policing" has decreased in relevance between the first and second cycle of school inspection. There are two reasons which might explain the situation in Germany: first, it seems as if the school representatives' associations have a greater say in Germany than in other countries. This is because associations representing the interests of school representatives are very powerful in Germany. Hence, when making decisions regarding school inspectorates, policymakers consider their interests (Piezunka 2018). Second, to ensure school representatives' autonomy is a very dominant norm in Germany. Therefore, it is not surprising that school inspectorates apply strategies which can be linked to "enabling work" as they are not able to use strategies with a higher degree of coercion.

Given the mutual dependencies and the functions that school inspection should theoretically fulfil within the educational system, the situation is paradoxical: on the one hand, school inspectorates need the acceptance of school representatives and policymakers to ensure their existence. On the other hand, to be able to fulfil the initiating function and, in particular, the monitoring function, it is also necessary that school representatives give up a certain degree of autonomy. However, the strategies described above mean that school representatives are treated as equal partners, which implies that they don't have to give up a certain degree of their autonomy. Therefore, striving for more acceptance from school representatives might imply that school inspectorates' efforts to reach a higher degree of institutionalisation might not be successful because school representatives don't have to give up a certain degree of their autonomy, which is also necessary in regard to school inspectorates' monitoring and initiating function. Considering the initiating function, schools have to give up a certain degree of autonomy because they have to accept that external evaluators are a necessary prerequisite for learning more about their school's strengths and weaknesses and to conduct school improvement measures. Compared to the initiating function, the monitoring function implies even more strongly that schools lose a certain degree of autonomy because they are asked to comply with the standards formulated in school inspectorates' indicators and their own understanding of school quality is ignored.

These observations raise the question of whether these strategies to maintain inspection as an institution have been successful. Acceptance of school inspection by school representatives is seen as a necessary condition for initiating school improvement measures (initiating function). However, it is not clear whether these strategies do in fact lead to higher acceptance or whether higher

acceptance leads to a stronger impact regarding school development measures. Regarding the latter, one person interviewed admits,

> To be honest: the most successful function is the monitoring function. This might not be our desired function. I would prefer if we were able to initiate school development measures. (15)

Therefore, a higher acceptance of school inspection by school representatives does not necessarily imply that the initiating function is fulfilled to a higher degree. However, the question of whether these strategies are successful cannot be resolved due to a lack of data.

Considering the inspection model in England and in the Netherlands, it is possible that an increased focus on the monitoring function and the application of strategies that ensure compliance through enforcement ("policing") or that establish coercive barriers to institutional change ("deterring") would be more successful at institutionalising school inspections than the strategies described above. Potential strategies might include the introduction of higher sanctions for failing schools or the publication of school rankings. These strategies imply that the initiating function would lose its relevance. Furthermore, focussing on the monitoring function would mean that it is not necessary for school inspectors to be accepted by schools. Irrespective of the desirability of the implications of these strategies, their implementation is not possible in Germany because the scope of action of school inspectorates in the education system is different from other inspectorates such as OFSTED, which describes itself as "independent and impartial" towards the ministry of education (Rutter 2013). Therefore, other forms of institutional work, such as "valourizing and demonizing" (Lawrence and Suddaby 2006, 230), which are within the school inspectorates' scope of action, could be more successful in maintaining school inspection as an institution. However, it is unknown whether these strategies have ever been applied by German school inspectorates.

To sum up, given the existing dependencies, school inspectorates aim at increasing their own acceptance and focus on the initiating function to ensure their continued existence. Nevertheless, these strategies might not be successful at institutionalising school inspections because these strategies also imply that school representative don't have to give up any autonomy. The latter is a necessary condition for school inspectorates to be able to fulfil certain functions within the education system.

References

Altrichter, Herbert. 2010. Theory and Evidence on Governance: Conceptual and Empirical Strategies of Research on Governance in Education. *European Educational Research Journal* 9 (2): 147-58.

Altrichter, Herbert, Thomas Brüsemeister, and Jochen Wissinger, eds. 2007. *Educational Governance: Handlungskoordination und Steuerung im Bildungssystem*.

Altrichter, Herbert, and David Kemethofer. 2015. Does Accountability Pressure through School Inspections Promote School Improvement? *School Effectiveness and School Improvement* 26 (1): 32-56.

Altrichter, Herbert, and David Kemethofer. 2016. Stichwort: Schulinspektion. *Zeitschrift für Erziehungswissenschaft* 19 (3): 487-508.

Avenarius, Hermann. 2006. Regelungen zur Rolle der Schulaufsicht und zum Umfang ihrer Aufgaben und Befugnisse. Visselhövede-Jeddingen. <www.gew-nds.de/Aktuell/archiv_dez_06/erweiterte_selbstaendigkeit.pdf> (Accessed May 30, 2017).

BLLV. 2017. *Schulleitungen klagen an: Mehr geht nicht: BLLV schreibt Brandbrief an Ministerpräsident.* <https://www.bllv.de/Brandbrief.12743.0.html> (Accessed February 1, 2018).

Böttcher, Wolfgang, and Miriam Keune. 2011. Wirkungsanalyse der Schulinspektion: Empirische Ergebnisse aus Deutschland. In *Wirkungen und Wirksamkeit der externen Schulevaluation*, ed. Carsten Quesel, Norbert Landwehr, Peter Steiner and Vera Husfeldt, 123-44. Bern: hep.

Brüggemann, Nele, Susan Jäkel, and Tabea Riemer. 2011. Schulinspektion – ein historischer Vergleich. In *Schulinspektion: Fluch und Segen externer Evaluation*, ed. Matthias von Saldern, 32-63. Schule in Deutschland 5. Norderstedt: Books on Demand.

Brüsemeister, Thomas. 2007. Steuerungsakteure und ihre Handlungslogiken im Mehrebenensystem der Schule. In *Governance, Schule und Politik*, ed. Jürgen Kussau and Thomas Brüsemeister, 63-95. Wiesbaden: VS Verlag für Sozialwissenschaften. doi: 10.1007/978-3-531-90497-9_3.

Dedering, Kathrin, and Moritz G. Sowada. 2016. Reaching a Conclusion – Procedures and Processes of Judgement Formation in School Inspection Tams. *Educational Assessment, Evaluation and Accountability*: 1-18.

Dietrich, Fabian, and Maike Lambrecht. 2012. Menschen arbeiten mit Menschen: Schulinspektion und die Hoffnung auf den zwanglosen Zwang der „besseren Evidenz". *Die deutsche Schule* 104 (3): 57-70.

Diaz-Bone, Rainer. 2017. Classifications, Quantifications and Quality Conventions in Markets – Perspectives of the Economics of Convention. *Historical Social Research* 42 (1): 238-62. doi: 10.12759/hsr.42.2017.1.238-262.

Espeland, Wendy N., and Michael Sauder. 2007. Rankings and Reactivity: How Public Measures Recreate Social Worlds. *American Journal of Sociology* 113 (1): 1-40.

Espeland, Wendy N., and Michael Sauder. 2016. *Engines of Anxiety: Academic Rankings, Reputation, and Accountability*. New York: Russell Sage Foundation.

Füssel, Hans-Peter. 2008. Schulinspektion und Schulaufsicht, Schulinspektion oder Schulaufsicht, Schulinspektion versus Schulaufsicht, Schulinspektion als Schulaufsicht. In *Externe Evaluation von Schulen: Historische, rechtliche und vergleichende Aspekte*, ed. Hans Döbert and Kathrin Dedering, 153-64. Münster, New York, München, Berlin: Waxmann.

Gärtner, Holger, Sebastian Wurster, and Hans A. Pant. 2013. The Effect of School Inspections on School Improvement. *School Effectiveness and School Improvement*. 25 (4): 489-508.

Gärtner, Holger. 2011. Die Wirkung von Schulinspektion auf Schulentwicklung. In *Wirkungen und Wirksamkeit der externen Schulevaluation,* ed. Carsten Quesel, Norbert Landwehr, Peter Steiner and Vera Husfeldt, 145-61. Bern: hep.

GEW Berlin. 2008. *Selbst Schuld und allein gelassen: Die Schulinspektionen müssen anders konzipiert werden.* <https://www.gew-berlin.de/1298_1790.php> (Accessed February 6, 2018).

Gläser, Jochen, and Grit Laudel. 2009. *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Heinrich, Martin. 2007. *Governance in der Schulentwicklung: Von der Autonomie zur evaluationsbasierten Steuerung.* Educational Governance, Wiesbaden: VS Verlag für Sozialwissenschaften.

Heinrich, Martin. 2010. Bildungsgerechtigkeit durch Evidence-based-Policy? Governanceanalysen zu einem bildungspolitischen Programm. In *Evaluation, Bildung und Gesellschaft: Steuerungsinstrumente zwischen Anspruch und Wirklichkeit*, ed. Wolfgang Böttcher, Jan N. Dicke and Nina Hogrebe, 47-68. Münster, New York, München, Berlin: Waxmann.

Heinrich, Martin, Maike Lambrecht, Oliver Böhm-Kasper, Thomas Brüsemeister, and Jochen Wissinger. 2014. Funktionen von Schulinspektionen? Zum Governance-Programm der Vergewisserung und Weiterentwicklung der Qualität schulischer Arbeit. In *Damit Unterricht gelingt: Von der Qualitätsanalyse zur Qualitätsentwicklung*, ed. Christian Fischer, 19-51. Münstersche Gespräche zur Pädagogik 30. Münster: Waxmann.

Heinrich, Martin. 2015. Zur Ambivalenz der Idee evidenzbasierter Schulentwicklung: Das Beispiel „Schulinspektion" – fortschrittlicher Rückschritt oder Innovation. *Zeitschrift für Pädagogik* 61 (6): 778-92.

Kehm, Barbara. 2014. Global University Rankings Impacts and Unintended Side Effects. *European Journal of Education* 49 (1): 102-12.

Kussau, Jürgen, and Thomas Brüsemeister. 2007. Educational Governance: Zur Analyse der Handlungskoordination im Mehrebenensystem der Schule. In *Educational Governance: Handlungskoordination und Steuerung im Bildungssystem*, ed. Herbert Altrichter, Thomas Brüsemeister and Jochen Wissinger, 15-54.

Lamont, Michèle. 2012. Toward a Comparative Sociology of Valuation and Evaluation. *Annual Review of Sociology* 38 (1): 201-21. doi: 10.1146/annurev-soc-070308-120022.

Landwehr, Norbert. 2011. Thesen zur Wirkung und Wirksamkeit der externen Schulevaluation. In *Wirkungen und Wirksamkeit der externen Schulevaluation,* ed. Carsten Quesel, Norbert Landwehr, Peter Steiner and Vera Husfeldt, 35-70. Bern: hep.

Lawrence, Thomas B., and Roy Suddaby. 2006. Institutions and Institutional Work. In *Sage Handbook of Organization Studies,* ed. Stewart R. Clegg, Cynthia Hard, Walter R. Nord and Thomas B. Lawrence, 215-54. London: Sage.

Maritzen, Norbert. 2008. Schulinspektionen. *Die deutsche Schule* 100: 85-96.

Moser, Vera. 2012. Inklusion: Standards und Perspektiven. *Sonderpädagogik in Berlin* 2: 23-7.

Piezunka, Anne, Tina Schaffus, and Michael Grosche. 2017. Vier Verständnisse von schulischer Inklusion und ihr gemeinsamer Kern: Ergebnisse von Experteninterviews mit Inklusionsforscherinnen und -forschern. *Unterrichtswissenschaft* 4: 207-22.

Piezunka, Anne. 2018. *Ist eine gute Schule eine inklusive Schule? Entwicklung von Messinstrumenten durch die Schulinspektionen.* Stiftung Universität Hildesheim.

Rauhvargers, Andrejs. 2014. Where Are the Global Rankings Leading Us? An Analysis of Recent Methodological Changes and New Developments. *European Journal of Education* 49 (1): 29-44.

Rutter, Jill. 2013. *The Strange Case of Non-Ministerial Departments*. <https://www.instituteforgovernment.org.uk/publications/strange-case-non-ministerial-departments> (Accessed August 20, 2018).

Senatsverwaltung für Bildung, and Jugend und Wissenschaft Berlin. 2014. *7 Jahre Schulinspektion in Berlin*.

Salais, Robert. 2008. Evaluation und Politik: Auf der Suche nach guten Indikatoren für die Forschung. In *Wissenschaft unter Beobachtung: Effekte und Defekte von Evaluationen*, ed. Hildegard Matthies and Dagmar Simon, 193-212. Wiesbaden: Springer.

Scott, Richard. 2008. *Institutions and Organizations: Ideas and Interests*. Thousand Oaks: Sage.

Sellar, Sam. 2014. A Feel for Numbers: Affect, Data and Education Policy. *Critical Studies in Education* 56 (1): 131-46.

Sommer, Norbert. 2011. Wie beurteilen schulische Gruppen die erlebte Schulinspektion? Ergebnisse einer Befragung. In *Schulinspektion in Deutschland: Eine Zwischenbilanz aus empirischer Sicht*, ed. Sabine Müller, Marcus Pietsch and Wilfried Bos, 137-64. Münster: Waxmann.

Sowada, Moritz G., and Kathrin Dedering. 2016. Die Reform der Reform. In *Schulinspektion als Steuerungsimpuls? Ergebnisse aus Forschungsprojekten*, ed. Arbeitsgruppe Schulinspektion, 169-99. Wiesbaden: Springer VS.

All articles published in HSR Special Issue 44 (2019) 2: Governing by Numbers.