

**The Differing Purposes and Effects of State-Wide Exit Examinations**  
**A Comparison of Finland, Ireland and the Netherlands From a Governance Perspective**

E. Dominique Klein

Institute of Education, University of Duisburg-Essen, Germany

Abstract

Most OECD-states have state-wide exit examinations at the end of upper secondary education. By regulating work in schools and classrooms, the examinations are supposed to raise school effectiveness. However, due to specific governance structures, the examinations probably serve particular intentions and therefore have multifarious designs. According to *Educational Governance* concepts, these diverse designs may cause differing effects on school actors. This project compares the designs and purposes of the state-wide exit exams at the end of upper secondary schooling in Finland, Ireland and the Netherlands. The project intends to detect patterns in the examination structures that foster the ability to regulate school actors. A document analysis of official documents is used to elaborate the exam designs as well as the overarching governance structure and strategies to improve the school system. Expert interviews with representatives of the education authorities are used to reconstruct the intentions of the government that are connected with the conduct of state-wide exit examinations. The results show that within the three countries, exams have very different functions which are very much connected with the governance structure and the overall system of quality assurance in the country. The functions postulated in the literature can only partially be retrieved.

Keywords:

state-wide exit examinations, education policy, governance, school development, expert interviews

In a lot of countries all over the world, students at the end of upper secondary education sit state-wide exit exams. This type of exams is so popular because it is considered to be a “powerful instrument for moving schooling in a desired direction” (Noah and Eckstein 1992, ¶2). The exams are taken by students in order to graduate from a certain course of studies, and are controlled by an agency external to the schools, administered by a national or state government, and based on prescribed syllabi in curriculum (Kellaghan and Madaus 2003). It is assumed that, as state-wide exams are controlled and set by an external agency, people in schools have only little, if any, influence on the exams, and therefore are forced to align their work with the standards and goals necessitated in the curriculum (cf. e.g. Bishop 1998; Eckstein and Noah 1993; Keeves 1994, Wößmann 2003). The exams are very clear about the competencies students must have acquired by the end of upper secondary school and what therefore should be addressed in the classroom. They also show where teachers and schools have to improve. Additionally, they hold school members accountable for their work, and link the graduation of the students to the exams, and therefore are presumed to motivate both educators to elaborate strategies for better instruction and school improvement, and students to raise their efforts. Thus, controlling the *output* is supposed to be a good means for regulating and improving *processes* at both the school- and classroom level.

These are mainly normative assumptions about (1) how the exams can affect schooling, and (2) how a state or government might in the ideal want their exam system to affect schooling. However, there is no clear empirical evidence to support these assumptions. The purpose of the study described therefore is to identify purposes and intentions linked to the conduct of state-wide exit exams in three countries with very distinct governance procedures. It is the first part of a larger study set up to identify exam mechanisms and effects at school and classroom level. The next step of the study described will therefore be aimed at people in schools and inquire the effects of the exams to elaborate conditions under which state-wide exit exams can actually serve to improve schooling. This, however, has not been done yet and therefore will not be further described in this paper.

### **Background: State-wide Exit Exams in the OECD-States**

Education systems have established state-wide exit exams, preferably at the end of upper secondary schooling (ISCED level 3a), for differing reasons, within distinct time-frames, and have given their exams different “faces”, so that the range of different exam procedures –that can all be labeled “state-wide exit exam”– is enormous (cf. Klein et al. 2009). Some states have conducted state-wide exit exams for more than a century; these exams usually emanated from university entrance exams, or were introduced to replace such exams. Their primary function to provide an objective way of selecting students into tertiary education institutions has stayed the main purpose for conducting state-wide exams in most states, although it has been supplemented with additional functions during the years. In some other states, state-wide exit exams have only recently been introduced or revised or will be in the near future (e.g. Czech Republic, Germany, some states of the USA) with the explicit goal of improving effectiveness through the control of the outcome (cf. e.g. National Institute for Public Education 2003). In those states the driving force for change was the belief that state-wide exit exams are an instrument with which the quality of schools can be monitored and school effectiveness can be controlled (see above). Thus, today probably all countries use the exams to assure and raise quality by setting standards every student has to achieve in order to graduate. However, the different designs of the exams (cf. Klein et al. 2009) imply that the countries probably have diverging opinions about how this can best be done, and pursue different strategies with regard to the amount of control and regulation exercised through the exams.

The scant existing research confirms that state-wide exams affect prior instruction. However, these effects are very inconsistent and are influenced by more than just whether the exams are state-wide or school-based. Positive or neutral effects, with regard to teacher cooperation and student support (e.g. Maag Merki 2008), for instance, or the use of student-centered instructional methods (e.g. Vogler 2006; Vogler and Carnes 2009), are reported alongside negative effects like reallocating resources

concerning contents and aid (e.g. Gillborn and Youdell 2000). Also, it becomes apparent that even when looking at the same features, the effects can differ depending on schools (Maag Merki et al. 2008), individual persons (Vogler and Carnes 2009), subjects (Baumert and Watermann 2000) or different administrative areas (Kühn 2010). The differing effects support the assumption that the intentions linked to the exams, and the according design of the exams, collaterally also affect how the exams are seen and adapted at school and classroom level, and that at the same time country-, school- and subject-specific parameters influence the interpretation of the exams. The adaptation of requirements and goals necessitated in the exams therefore is informed in both a top-down process (purposes and design of the exams) and a bottom-up process (school and subject culture; individual dispositions).

### **Theoretical Framework: Educational Governance**

These different observations might be traced back to the fact that people at school and classroom level react differently to different exam structures. Such an understanding is rooted in (new) governance concepts transferred to education policy, also referred to as *Educational Governance* (cf. Altrichter et al. 2007). These concepts assume that there are discrepancies between what has been intended by central authorities and how their directions are perceived and carried out by local school actors. Instead of adopting directions, they *adapt* them to their specific individual and their school's situation, so that measurements taken by the authorities may have differing effects at school and classroom level. From this perspective, what happens in schools is regulated through formal directions, but also through informal arrangements, negotiations and the coordination of actions on and across the multiple levels of the school system (Hanf and O'Toole 1992; Kussau 2008; Kussau and Brüsemeister 2007).

How directions are carried out is partially dependent on the individual actors' dispositions, and the school's organizational culture, that is, its specific system of shared values, experiences, and attitudes towards quality, instruction, and innovations (cf. e.g. Holtappels 2004; Saunders and Rudd 1999; Swanson and Stevenson 2002). This may dilute or even contradict the interests of the administration. Whether state-wide exams are perceived as helpful instruments at school level and thus are used efficiently, or whether they are instead seen as a threat that incites actors to play the system depends to a large extent on the organizational culture of individual schools (cf. Maag Merki et al. 2008). The degree to which schools have freedom to deal with the exams and can deviate from the requirements necessitated in them is partially dependent from the standardization of procedures in the exams, or the degree to which actors at school and classroom level are involved in the exam process. A higher standardization reduces the influence individual and organizational factors have on the implementation of standards (as in requirements) set in the exams. For instance, if papers are marked by the class teacher (lower standardization), even with marking schemes, the teacher may adapt his or her marking to the way the content has been taught. If, on the other hand, the marking is done by an external marker (higher standardization), the teacher will have to adjust instruction more firmly to the standards, because he or she has no influence on the marking. In this respect, the exam procedures in the OECD-states prove to have very distinct degrees of standardization.

The aim of this study therefore is to investigate and analyze the design and functions of state-wide exit exams in an international comparison to provide groundwork for the further analysis of the effects the exams can have on school and classroom practice, and especially the conditions under which these effects can occur.

### **Methods and Conduct**

The project aims to investigate the differing intentions and functions that are linked to the conduct of state-wide exit exams in an international comparative case study. On grounds of the assumption that differing governance structures may result in differing purposes linked to the exams, three countries with rather distinctive governance structures were chosen. The choice was further informed by the overall structure of the exams, and especially by the standardization level of their procedures. Based on the assumption that the countries (a) must have differing governance structures and (b) must conduct highly standardized exit exams, the exams in Finland, Ireland and the Netherlands were

chosen for this comparative study on grounds of a prior comparison of exam structures (cf. Klein et al. 2009).

The study is guided by qualitative methods. The investigation consists of two succeeding steps. First, the procedures of the exams and their embedding in the overall system of governance and quality assurance are analyzed, and second, the explicit functions of the exams are being investigated.

The first step is accomplished with qualitative content analyses of already existing material, which mostly consists of official documents issued by the relevant authorities. This comprises for instance legal documents, newsletters, guidelines, websites etc. The analysis is based on a system of categories set up beforehand and amended during the process. The category system has two sections. Section A describes the procedures of the exams and consists of the five categories *Structure of General Upper Secondary* (e.g. organization of school time, compulsory subjects, 22 items), *Historical Context* (e.g. traditions, political philosophies, 2 items), *Organizational Framework* (e.g. number of examinations, task development 59 items), *Marking* (e.g. standards, students' anonymity, 21 items) and *Handling of Results* (e.g. meaning of results for final grades, feedback, stakes, 5 items). Section B describes the governance structures of the three countries and consists of the five categories *Decision Making* (e.g. levels, forms, participation, 24 items), *External Evaluation* (e.g. inspection, use of results, 23 items), *Forms of Accountability* (e.g. stakeholder, goals, 8 items), *School Development at school level* (e.g. methods, rules and guidelines, 21 items) and *School Autonomy and Leadership* (type of autonomy, participation, qualifications, 46 items).

For the second step, the analysis of the specific functions of the exams, the method was changed. The purposes for which the exams are being employed can only hardly be found in official documents, especially as some of them might be conceptualized as “common knowledge” in the respective country and therefore need not be written down. Therefore, instead of conducting secondary analyses of already existing material, we carried out expert interviews (cf. Gläser and Laudel 2009; Meuser and Nagel 2010) with experts of the authorities relevant for the exams. In each of the three countries, three experts were chosen so that altogether nine interviews were conducted. The interviews were based on an interview manual which mainly served as a mnemonic to make sure that all interesting contents were properly addressed. The interviews were transcribed completely and analyzed with methods of qualitative content analysis as well. The objective of the interview was not to elaborate personal feelings about the exams and reconstruct their origins, but rather to find a common notion of the functions of the exams in the respective country. Therefore, only those parts of the interview were analyzed that reported factual knowledge; expressions of opinion and assumptions were considered only if these were mentioned repeatedly by all three interview partners from the respective country, which was then interpreted as “common sense” and included in the analysis. The interviews were analyzed with a category system as well. The category system was set up deductively as well as inductively to meet the explorative character of the research objectives. The coding was carried out with the software MAXQDA®.

In a third step, the effects of state-wide exit exams on school and classroom level will be investigated with quantitative methods, using a standardized questionnaire for school heads, teachers and students, which will then be analyzed with multi-level analysis with HLM. With this multi method approach, the study shall serve the explorative character of the study and at the same time be able to produce results that at least within cases can be seen as generalizable.

## Results

In the following, for each country the exam in its embedding within the overall system of education policy and quality development will be described in short, and the main findings concerning the purpose of the exams will be displayed.

### Finland: Trust

The Finnish education system is much decentralized. The quality development is mainly the task of the Finnish National Board of Education and the municipalities. The national core curriculum is rather

vague and describes contents in very global terms. On grounds of this core, municipalities and schools establish individual curricula.

The Finnish exam – *Ylioppilastutkinto* – is a “descendant” of the entrance exam of the University of Helsinki. Students have to sit written exams in four semi-compulsory core subjects in a maximum of three succeeding exam periods. The exam papers are set by a central examination board, and are marked by the same after a preliminary (but non-decisive) marking of the class teacher. Today, the exam still has the purpose to select students for admission to tertiary education as fairly as possible. It is, however, complemented with possible additional university entrance exams which may contain practical tests, interviews and other alternative methods.

Apart from that, the exam is the only nationwide test in the Finnish education system and in that is meant to serve schools as a benchmark for estimating their students’ performances. In this, there is hope that schools use the results of their own school to elaborate strategies for improvement. This anticipated further processing of exam results, however, is not in any way prescribed. The municipalities and schools are free to react on the results in whatever way they feel is best. Moreover, the results are not linked to accountability purposes at all and used for system monitoring only in an aggregated state. The examination commission keeps, administers and processes the results, but cannot issue any directives for schools, and does not for instance transfer school-based results to the ministry or the National Board of Education. Thus, on a national level, there are no stakes attached to the exams for anyone except the students. They can be used for accountability purposes on the local and school level though – and they usually are.

The functions of the exams and the handling of the results mirror the fundamental principle of trust in each other and in the expertise and professionalism of teachers that shapes the Finnish education system.

### **Ireland: Objectivity**

In the Irish education system, the state has traditionally played a mere administrative role, although from the 1960s onwards, it has struggled to strengthen its impact on development issues. Still, the Irish education system is characterized in that schools enjoy a huge amount of freedom with regard to administrative and also pedagogical matters, although contents are prescribed in the national curricula.

Students sit the *Leaving Certificate Examination* in at least five subjects of choice; many subjects have an additional component such as practical work, which is examined externally as well. The exam papers are set by the State Examinations Commission, and marked anonymously by teachers recruited by the same commission.

In a country with only 3.5 million inhabitants, it is traditionally not the task of the teacher to carry out assessment that will lead to awarding marks. Thus, the state-wide exams have the purpose of an assessment that is objective and free from bias. The results of the exams are the only criteria used for allocating students in tertiary education in a central application system. In the authorities, there is awareness of the pressure that students and schools are under due to the high stakes attached for students, and momentary there are cautious considerations as to how the system can be changed without compromising the objectivity of the exams (e.g. integrating coursework into the final grade). This, however, goes with a huge logistic effort, as even those practical components are marked by an external agency.

The accountability function of the exams is rather weak, and schools are not held accountable for the exam results directly. Instead, the –rather unsystematic– inspection of schools, called Whole School Evaluation, may take the exam results into consideration as one aspect among many. There are no consequences for schools based on exam results only.

### **Netherlands: Accountability**

The Dutch system is mainly characterized by a high autonomy that is accompanied by a pyramid system of accountability and inspection, in which the results of the central exams have a prominent, but not singular position.

In the Netherlands, the final grade consists equally of grades of the *centraal examen* and of grades from additional school-based exams. The system was introduced in the 1960s. In 2007, all formal prescriptions on the school exams were abolished so that today, schools are free to choose on number, form and – for the most part – contents of the school exam. Thus, students sit school exams of any form in all subjects (about twelve), and state-wide written exams in about three quarters of their subjects. The latter are produced by the semi-private testing institute Cito on grounds of a construction assignment issued by the state-run *College voor Examens*. The papers of the state-wide exams are marked by the class teacher and an additional teacher from another school. The contents of the state-wide exams only cover a part of the curriculum; the rest must be assessed in the school exam.

The state-wide exam in the Netherlands primarily has the function of benchmarking results for schools and the state: the results are used by the inspection to classify other data, such as class repetitions, drop outs and other material provided by schools. Only if these show insufficient results does the inspection take a closer look at the processes at school, and only if schools do not manage to improve their offering for some years will there be real consequences. Inspection reports and the exam results, integrated in subject groups, can be accessed by everyone on the website of the inspection. In this, the exams also form part of a competitive type of school development that is characterized with “pressure and support” (cf. van Ackeren 2003). Thus, the exams, in contrast to Finland and Ireland, do explicitly serve an accountability and controlling function, but are at the same time not meant to urge schools to do exactly as wished by the government, as can be seen in the school exam, which gives the schools a lot of freedom.

The school exam, in this context, cushions the pressurizing effect the exams may have on students and teachers. It also has the purpose to raise the quality of the judgment about a student’s abilities. In the state-wide exams, which serve to benchmark students’ performances and keep standards at a comparable level, only those aspects are examined that can be examined in a standardized test and are easily comparable. The school exam, on the other hand, is supposed to cover those other aspects that can hardly be tested in a standardized test, and thus take care that contents are broadly covered and that there is place for individual interests.

### **Discussion**

The closer comparison of exam functions observed in Finland, Ireland and the Netherlands suggests that in countries that have had state-wide exams for a long time, and where the schools can act quite autonomously at the same time, the exams do not seem to be the “last straw” of the government to keep control of schooling. None of the interviewees said it was a function of the exams to “push” schools into a certain direction – which does not mean that they are not aware of the power the exams can have with regard to implementing innovations in the curriculum. However, it seems that the exams are rather seen as an offer to schools, from which they can –but do not have to– draw information (Finland) or even something that the schools should not draw too many conclusions from at all (Ireland). It is only in the Netherlands that the exam results are systematically used to hold schools accountable for their work, so that the state really plays an active role in development processes and tries to regulate work at school. This probably can be traced back to the fact that the exam in the Netherlands is embedded in a partially competitive school system. At the same time, there are no market elements in either the Irish or the Finnish school system, so that for example the publication of exam results is being rejected, and the influence of the state as a whole is rather limited.

### **Conclusion**

The main function of the exams in Finland and Ireland remains a fair and comparable student certification and selection. A connection between the standardization of the exam and this function can

at least partially be detected in Ireland, where there seems to be a connection between the very standardized process and the exam's task to provide an objective assessment of student abilities.

There is no hint that the exams are meant to raise students' and teachers' motivation, and the control function seems to be prominent only in one of the three countries. The functions of state-wide exit exams, as they are postulated in the literature and at least theoretically taken as a basis for change in countries like Germany, can only partially be detected in the realities of these three countries.

It is likely, however, that these functions are weighted differently in countries that have just switched to state-wide exams, and in countries in which there is traditionally more control (input and output) of schools in general. In this, the study is facing certain limitations due to its case design. However, the comparison shows that there is no one-size-fits-all approach to the conduct of state-wide exit exams. Rather than that, the different practices display widely diverging views on the best solution for the organization of the exams. These views have emerged from specific historical, political, and cultural traditions and beliefs.

Given that, it must be discussed whether some procedures are, within their context, better suited than others to support quality development in schools. Answering these questions goes beyond the results provided in this contribution and shows a desideratum for research. Further research into the effects of state-wide exams on school quality will have to compare different types of exams as to their structures and purposes, and the actually observed effects on the actions at school and classroom level.

## References

- Ackeren, I. van (2003). *Evaluation, Rückmeldung und Schulentwicklung: Erfahrungen mit zentralen Tests, Prüfungen und Inspektionen in England, Frankreich und den Niederlanden*. Münster: Waxmann.
- Altrichter, H., Brüsemeister, T., and Wissinger, J. (Eds.) (2007). *Educational Governance: Handlungskoordination und Steuerung im Bildungssystem*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Baumert, J., and Watermann, R. (2000). Standardisierung durch die Abiturprüfung: Zentralabitur oder dezentrale Prüfungsorganisation? In J. Baumert, W. Bos, and R. Lehmann (Eds.), *TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie-mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 2* (pp. 341–351). Opladen: Leske + Budrich.
- Bishop, J. H. (1998). The effect of curriculum-based external exit exams on student achievement. *Journal of Economic Education*, 29, 172–182.
- Eckstein, M. A., and Noah, H. J. (1993). *Secondary School Examinations: International Perspectives on Policies and Practice*. New Haven: Yale University Press.
- Gillborn, D., and Youdell, D. (2000). *Rationing Education: Policy, practice, reform and equity*. Buckingham: Open University Press.
- Gläser, J., and Laudel, G. (2009). *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen* (3., überarbeitete Auflage). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hanf, K., and O'Toole, L. J. (1992). Revisiting Old Friends: Networks, Implementation Structures and the Management of Interorganizational Relationships. *European Journal of Political Research*, 21(1-2), 163–180.
- Holtappels, H. G. (2004). Schulprogramm und Organisationskultur – Ergebnisse aus niedersächsischen Schulen über Bedingungen und Wirkungen. In H. G. Holtappels (Ed.), *Schulprogramme – Instrumente der Schulentwicklung. Konzeptionen, Forschungsergebnisse, Praxisempfehlungen* (pp. 175–194). Weinheim, München: Juventa.
- Keeves, J. P. (1994). *National examinations: design, procedures and reporting*. Paris: UNESCO: International Institute for Educational Planning.

- Kellaghan, T., and Madaus, G. (2003). External (Public) Examinations. In T. Kellaghan and D. L. Stufflebeam (Eds.), *Kluwer international handbooks of education: Vol. 9. International handbook of educational evaluation* (pp. 577–602). Dordrecht: Kluwer Acad. Publ.
- Klein, E. D., Kühn, S. M., Ackeren, I. van, and Block, R. (2009). Wie zentral sind zentrale Prüfungen?: Abschlussprüfungen am Ende der Sekundarstufe II im nationalen und internationalen Vergleich. *Zeitschrift für Pädagogik*, 55(4), 596–621.
- Kühn, S. M. (2010). *Steuerung und Innovation durch Abschlussprüfungen?* Wiesbaden: VS Verl. für Sozialwiss.
- Kussau, J. (2008). Governance der Schule im Kontext von Interdependenzen und sozialem Wissen. In T. Brüsemeister and K.-D. Eubel (Eds.), *Evaluation, Wissen und Nichtwissen* (pp. 203–232). Wiesbaden: VS Verl. für Sozialwiss.
- Kussau, J., and Brüsemeister, T. (2007). *Governance, Schule und Politik: Zwischen Antagonismus und Kooperation*. Wiesbaden: VS Verl. für Sozialwiss.
- Maag Merki, K. (2008). Die Einführung des Zentralabiturs in Bremen: Eine Fallanalyse. *Die Deutsche Schule*, 100(3), 357–368.
- Maag Merki, K., Klieme, E., and Holmeier, M. (2008). Unterrichtsgestaltung unter den Bedingungen zentraler Abiturprüfungen: Differenzielle Analysen auf Schulebene mittels Latent Class Analysis. *Zeitschrift für Pädagogik*, 54(6), 791–808.
- Meuser, M., and Nagel, U. (2010). Experteninterviews - wissenssoziologische Voraussetzungen und methodische Durchführung. In B. Frieberthäuser, A. Langer, A. Prengel, H. Boller, and S. Richter (Eds.), *Handbuch qualitative Forschungsmethoden in der Erziehungswissenschaft* (3rd ed., pp. 457–471). Weinheim: Juventa-Verl.
- National Institute for Public Education (2003). *The System of Content Regulation in Hungary: Public Policy Analysis*. Retrieved from [ftp://ftp.oki.hu/english/Content\\_Regulation.pdf](ftp://ftp.oki.hu/english/Content_Regulation.pdf).
- Noah, H. J., and Eckstein, M. A. (1992). The Two Faces of Examinations: A Comparative and International Perspective. In H. J. Noah and M. A. Eckstein (Eds.), *Examinations. Comparative and International Studies* (pp. 147–170). Oxford: Pergamon Press. Retrieved from <http://www.hku.hk/cerc/3e.html>; Reprinted by permission of Butterworth-Heinemann Ltd.
- Saunders, L. and Rudd, P. (1999). *Schools' use of ,value added' data: A Science in the service of an art?* NFER: Slough.
- Swanson, C. B., and Stevenson, D. L. (2002). Standards-Based Reform in Practice: Evidence on State Policy and Classroom Instruction from the NAEP State Assessments. *Educational Evaluation and Policy Analysis*, 24(1), 1–27.
- Vogler, K. E. (2006). Impact of an exit examination on English teachers' instructional practices. *Essays in Education*, 16(Spring). Retrieved from <http://www.usca.edu/essays/vol162006/vogler.pdf>.
- Vogler, K. E., and Carnes, G. N. (April 2009). *Comparing the Impact of a High School Exit Examination on Science Teachers' Instructional Practice*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, California.
- Wößmann, L. (2003). Central Exams as the "Currency" of School Systems: International Evidence on the Complementarity of School Autonomy and Central Exams. *CESifo DICE Report*, (4), 46–56.