An International Comparison of Biology Tasks in State-Wide Exit Examinations

Mirko Krüger

Research Group & Graduate School Teaching and Learning of Science,

Institute of Education, University of Duisburg-Essen, Germany

mirko.krueger@uni-due.de

Abstract:

Most OECD-countries have installed state-wide exit examinations at the end of upper secondary education (ISCED 3A). These examinations are assumed to set standards for learning outputs, and thus secure comparable results and educational quality. However, only little is known about the effects of state-wide examinations, for instance how uniform standards affect different dimensions of examination questions. Against this background, the project aims to analyze differences and similarities in how requirements and biology tasks used in both state-wide exit examinations and the preceding classroom are related in six high- and low-performing OECD-countries with different political and cultural backgrounds. Focusing structure- and content-related dimensions by using a concept-oriented rating manual, the project shall disclose the general examination practice as well as trends concerning structure and content of the tasks and identify deviations from the examination standards. The results shall show whether the tasks used in the examinations are “tasks worth teaching to”, and whether they do have the assumed backwash effect on the tasks in classroom assessment. By identifying good practice and elements that are especially well suited to affect classroom practice, the results of the study can also practically contribute to the development of biology tasks designed to improve biology instruction.

Key words:

State-wide Exit Examinations, Biology Tasks, Analysis, Governance, European OECD-Countries

Currently, in most OECD-countries, students have to pass state-wide exit examinations to graduate from upper secondary education. These examinations are assumed to set standards for learning outputs, and thus secure comparable results and educational quality. However, only little is known about the effects of state-wide examinations, for instance how uniform standards affect different dimensions of examination questions. Therefore, this project aims to analyze content and structure of biology tasks in state-wide exit examinations and in this to produce information about how national standards can influence the format of examination tasks, and collaterally also the preceding classroom work.

In a first step, to describe the background of the project, the expectations towards state-wide exit examinations, normative discourses and empirical findings on their effects, as well as the significance of tasks related to this kind of examinations and the current state of research shall be outlined in an overview. Secondly, the project aims and methodology are described to give an idea of the research interests and methods used to answer the leading research questions. Finally, the expected outcomes are mentioned.

State-wide Exit Exams – Expectations, Discourses and Empirical Findings

State-wide exit examinations have recently been introduced in a number of countries because it is assumed that they create a more standardized framework for student graduation from general upper secondary schooling than school-based exit examinations do. In this context, that students deal with the same range of examination questions at the same time and date and within the same duration shall secure the comparability of students' achievements across an administrative area. Furthermore, state-wide exit examinations shall maintain that qualifications expressed in the certificates are comparable, as every student has to achieve external examination standards for each area to graduate from upper secondary education and to gain access to universities. Recently, state-wide exit examinations are also used to positively affect learning and teaching processes and outputs on school and classroom level, for instance quick and extensive implementation of new syllabi and innovative instruction methods (e.g. Bishop 1998; Maag Merki 2010; Klein et al. submitted for publication).

However, only little is known about the effects of this kind of exit examination. Up to now, there is a more normatively than empirically founded discourse concerning their (un-) intended effects (e.g. van Ackeren 2007; Maag Merki 2010). Additionally, existing empirical results show inconsistent (positive, negative or no) effects concerning different aspects of schooling; and these effects also depend on schools, individuals, subjects and states (e.g. Baumert and Watermann 2000; Vogler and Carnes 2009; Kühn 2010; Maag Merki 2010; Block et al. in press). Beyond that, a systematic international comparison of state-wide exit examinations in several OECD-states reveals different standardization levels of organizational conditions (e.g. Klein et al. 2009). Currently, the supposedly uniform label "state-wide exit exams" covers examinations of rather heterogeneous designs. Therefore, it must be questioned whether all of these examinations are able to fulfill the expectations linked to state-wide exit examinations, or whether different designs also cause different effects.

Tasks in the Context of State-wide Exit Exams – Significance and Status of Research

(Written) tasks play an important role in the context of state-wide exit examinations. The use of tests with uniform tasks serves as a comparable assessment and classification tool of students' performances at the end of general upper secondary education. They show how far students possess knowledge and competencies in general and as scaffolding for the various target vocational qualifications they will acquire after graduation, and match their competencies with a national standard. From an "Educational Governance" point of view (cf. Altrichter et al. 2007), tasks in state-wide exit examinations are also an essential element for influencing instruction in schools, implementing innovations in classes and governing school systems as they make the covering of certain contents and instructional methods binding which is something that school-based examinations cannot provide because they are based on tasks set by the respective class teachers, and are therefore be aligned to the preceding instruction.

Regarding the high hopes connected to the introduction of these examinations, the actual findings concerning the effects of tasks in external and state-wide exit examinations are remarkably sparse. In Germany, only little is known about how uniform standards can affect different dimensions of examination questions (e.g. Bolle-Bovier 1994; Brockhage and Weghöft 1994; Paul 2002; Kirsch 2003; Giar 2007). Kühn (2010) shows that science tasks in the state-wide examinations of three selected German *Bundesländer* partially differ from the national standards and contain only few experimental and context-oriented tasks. Summarizing, the state of research concerning examination tasks in state-wide exit examinations in Germany can only offer shattered findings. The same is true for the research on examination tasks in other countries. Currently, most task analyses focus on specific configurative aspects of tasks, especially in connection with Bloom´s Revised Taxonomy of Cognitive Objectives as theoretical framework (e.g. Eberle et al. 2008; Karamustafaoglu 2003; Azar 2005; Tikkannen 2009). Up to now, there is a lack of an international comparison of examination tasks that will also consider the importance of their embedding in the respective national standards and the intended and unintended guiding function that they are carrying.

Therefore, the aim of this project is to analyze and compare national standards, biology examination tasks and biology tasks used in the preceding classroom in six OECD-states which have state-wide exit examinations at the end of upper secondary education.

Study Design

Research Questions

This study aims to investigate and analyze the national ’status quo’ of the abovementioned elements for each of the selected countries, and to explicate transnational similarities and differences. Particularly interesting is in how far biology tasks are corresponding to the examination standards. In this context, the leading project questions are:

1. Which national standards do exist for the development of examination tasks and what are the main similarities and differences between the countries?
2. Which are the distinctive features of biology examination tasks in the selected states as well as in an international comparison? Of what quality are these tasks, also against the background of national and international discussions about ’good’ biology examination tasks?
3. Are there any congruities or incongruities between national standards and the contents and structures of biology examination questions? How far can differences in the implementation of standards in the tasks give evidence about the possibility to influence schooling with policy measures in the different countries?
4. How far do biology tasks in state-wide exit examinations have (un-) intended effects (f.e. the backwash-effect) on tasks, contents (and processes) in the preceding classroom?

Focusing the content and structure of biology examination tasks on one hand and the examination requirements on the other hand, it is expected that there are (a) differences between the countries in all aspects and (b) common trends regarding consideration and design of quality features of tasks. Against the backdrop of existing research results, it can be expected that biology examination tasks do not entirely match the examination standards. Beyond that, it can be assumed that they have a backwash effect on the tasks in classroom assessment.

Choice of Countries

The research focuses examination tasks in science subjects, especially in biology. Research in Germany points out that the subject biology is considered as the leading discipline in science from the perspective of teachers and students (e.g. Baumert and Köller 2000; Block et al. in press). A scan of examination statistics shows, that in a lot of other countries, biology is the most popular science (examination) subject as well. For that reason, this study focuses those states in which biology is semi-

compulsorily or optionally examined in a state-wide exit examination. To avoid too great heterogeneity in the standards within one country, only non-federal states are considered. Moreover, the study only includes European OECD-countries which have a comparable wealth and economic situation, so that possible effects will not be distorted by demographic or economic extremes. Apart from that, the chosen states all showed a good or considerably improved performance in PISA and TIMSS, which is interpreted as an indicator for best practice with regard to science education. Upon these criteria, the following six European OECD-countries were chosen:

- England
- Finland
- France
- Netherlands
- Scotland
- Poland

As an additional example, the results shall also be compared to already existent results of examination task analyses in Germany.

Methods

The study is conducted in an international comparative approach and has an explorative and descriptive character since there is only little known about the design of tasks in the context of state-wide exit examinations.

First, secondary analyses of official documents and guidelines concerning the construction of biology examination tasks are used to elaborate the characteristics of these procedures for every state. Based on this, differences and similarities between the states can be identified and systematically compared, and options for the task designers to deviate from national standards can be detected.

Second, examination tasks and tasks used in the preceding classroom are analyzed using a concept-oriented rating manual. The analysis of both types of tasks will cover examination papers from the years 2000 through 2011. An existing concept-oriented rating manual which was developed and used for a task analysis in Germany will be adapted for the international comparison and aligned to the specific circumstances within the different states. In this, the national standards of the respective countries will serve as a basis for the adaption with regard to both contents and structure of the examination tasks. The development of categories for the manual takes place in a deductive and inductive way. Thus, biology-related instructional aspects, especially with regard to supposedly ’good’ examination tasks, can be considered as well as empirical findings gained by looking into the examination tasks. Both perspectives will be the guidelines for generating the rating manual. This will cover the areas Format, Content and Cognitive Operations, which will be further defined in several sub-categories (see figure 1).

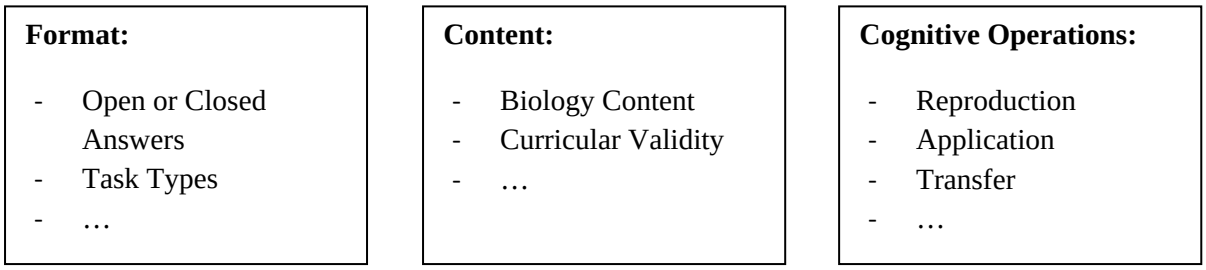

Figure 1: Categories and examples of sub-categories of the concept-oriented rating manual

With this concept-oriented rating manual both the examination tasks and the tasks used in the preceding classroom are analyzed. Among other things, the findings are supposed to answer in how far the tasks used in the examinations are "tasks worth teaching to", and whether they do have the assumed backwash effect on the tasks in classroom assessment.

Third, standardized questionnaires for biology teachers and students are used to produce information about whether the tasks in state-wide exit examinations have (un-) intended effects on the preceding classroom that go beyond the tasks used in classroom assessment.

Expected Outcomes

The study is supposed to generate information with both theoretical and practical significance. From a theoretical point of view, the empirical findings can help to consolidate and distinguish the discussion about the effectiveness of state-wide exit examinations in that it produces statements about the current practice of biology examination tasks and their structure and content in European countries. Furthermore, and the study will provide evidence about their capability to direct work at school and classroom level and thus to contribute to the improvement of classroom practice. The study will deliver a screen on which the German exams as well as those of other European countries can be positioned with regard to both their design and their utilization as policy tools.

From a practical perspective, the study allows for identifying international trends concerning biology examination tasks at the end of general upper secondary education. Also, the project aims to contribute to the optimization of this kind of tasks and to recognize congruities and incongruities between biology examination tasks and tasks used in Large Scale Assessments like PISA or TIMSS.

References

Ackeren, I. van (2007). Zentrale Abschlussprüfungen. Entstehung, Struktur und Steuerungsperspektiven. *Pädagogik, 59*(3), 12-15.

Altrichter, H., Brüsemeister, T. and Wissinger, J. (Eds.) (2007). *Educational Governance: Handlungskoordination und Steuerung im Bildungssystem.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Azar, A. (2005). Analysis of Turkish High-School Physics-Examination Questions and University Entrance Exams Questions According to Blooms Taxonomy. *Journal of Turkish Science Education, 2*(2), 144-150.

Baumert, J. and Köller, O. (2000). Unterrichtsgestaltung, verständnisvolles Lernen und multiple Zielerreichung im Mathematik- und Physikunterricht der gymnasialen Oberstufe. In J. Baumert, W. Bos und R. Lehmann (Eds.), *TIMSS/III. Dritte internationale Mathematik- und Naturwissenschaftsstudie-mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 2* (pp. 271-316). Opladen: Leske + Budrich.

Baumert, J. and Watermann, R. (2000). Standardisierung durch die Abiturprüfung: Zentralabitur oder dezentrale Prüfungsorganisation? In J. Baumert, W. Bos und 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.

Block, R., Kein, E. D., Ackeren, I. van and Kühn, S. M. (in press). Leistungseffekte des Zentralabiturs? Eine kritische Auseinandersetzung mit bildungsökonomischen Interpretationen zu den Effekten der Prüfungsorganisation auf der Basis von PISA E 2003-Daten. *bildungsforschung.*

Bishop, J.H. (1998). The effect of curriculum-based external exit exams on student achievement. *Journal of Economic Education, 29*, 172-182.

Bolle-Bovier, H.-P. (1994). Schriftliche Abiturprüfung im Fach Deutsch – zentrale oder dezentrale Aufgabenstellung? Ein Vergleich zwischen Aufgabenstellungen in Niedersachsen und

Bayern. *Mitteilungen des Deutschen Germanisten-Verbandes 41*(4), 45-52.

Brockhage, R. and Weghöft, R. (1994). Zentralabitur und Naturwissenschaften. *Gymnasium in Niedersachsen, 26*(4-5), 80-82.

Eberle, F., Gehrer, K., Jaggi, B., Kottonau, J., Oepke, M. and Pflüger, M. (2008). *Evaluation der Maturitätsreform 1995 (EVAMAR). Schlussbericht zu Phase II.* Bern: Staatssekretariat für Bildung und Forschung SBF.

Giar, H. (2007). Schülerexperimente im Zentralabitur? *Praxis der Naturwissenschaften- Chemie in der Schule, 56*(1), 26-31.

Karamustafaoglu, O. (2003). Analysis of Turkish high school chemistry examination questions according to Bloom's taxonomy. *Chemistry Education: Research and Practice, 4*(1), 25-30.

Kirsch, W. (2003). Abituraufgaben auf dem Prüfstand – Chemie. Ergebnisse und Konsequenzen eines Erfahrungsaustauschs zwischen den Ländern – kontextorientierter Chemieunterricht als Basis für eine neue Aufgabenkultur. *Schulverwaltung, Ausgabe Hessen, Rheinland-Pfalz und Saarland, 7*(10), 267-268.

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.

Klein, E. D., Ackeren, I. van, Block, R. and Kühn, S. M. (submitted for publication). The Possible Impact of Differing State-Wide Exit Examination Procedures on School Improvement – A Comparison of Standardization Levels in 16 OECD-States.

Kühn, S. M. (2010). *Steuerung und Innovation durch Abschlussprüfungen?* Wiesbaden: VS Verl. für Sozialwiss.

Maag Merki, K., Holmeier, M., Jäger, D. and Oerke, B. (2010). Die Effekte der Einführung zentraler Abiturprüfungen auf die Unterrichtsgestaltung in Leistungskursen in der gymnasialen Oberstufe. *Unterrichtswissenschaft*, 2, 173-192.

Paul, R. (2002). Abituraufgaben auf dem Prüfstand. Ergebnisse und Konsequenzen eines Erfahrungsaustauschs zwischen den Ländern – Teil IV: Physik. *Schulverwaltung, Ausgabe Hessen, Rheinland-Pfalz und Saarland, 6*(6), 173-175

Vogler, K. E. and Carnes, G. N. (2009). *Comparing the Impact of a High School Exit Examination in Science Teachers´ Instructional Practice.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, California.

Tikkanen, G. (2010). *Chemistry matriculation examination as summative assessment tool.* Doctoral Dissertation. University of Helsinki.