



Fairness and Flexibility in Oral Examination

**- A Qualitative Study of the Russian
Teacher Education**

Eva Andreasson

Undergraduate thesis, 10 credit points

Spring 2005

Abstract

This is a descriptive ethnographical study with the purpose of examining teachers' and students' experiences of oral examination at a State Pedagogical University in western Russia. The study also focused on finding the characteristics of oral examination and the contextual factors influencing its implementation. The research was done using participatory observations and interviews. The results show that interviewees experience oral assessment in general as positive. Their descriptions are summarised and analysed using a number of key concepts, of which flexibility, subjectivity, individualisation, and fairness are the most important. The study also shows that contextual factors such as culture, traditions, and organisational framework have large impact on how the examination is done.

The conclusion is that oral examination has both gains and losses, since the teacher's active participation creates possibilities for individualisation and deep probing of the students' knowledge, but is also a source of bias because of its subjectivity.

Key concepts: alternative assessment, individualisation, oral assessment, Russian higher education, test validity, test reliability.

Acknowledgements

I would like to express my gratitude to all the people who have aided me in writing this thesis. First of all, Dan, who have given me invaluable support when things got rough and did not protest when I went to Russia for two months.

I would also like to thank all my interviewees and other friends in Russia who always had time for me when I needed them. Without your co-operation, this would have been impossible. Last but not least, I thank my supervisor, Oleg Popov for valuable help and support, and Tim Honn who helped me with the English.

Eva Andreasson,
Umeå.

Table of Contents

| | |
|--|------------|
| ABSTRACT | I |
| ACKNOWLEDGEMENTS | II |
| TABLE OF CONTENTS..... | III |
| | |
| INTRODUCTION | 1 |
| <i>Purpose and Research Questions</i> | <i>2</i> |
| <i>Theoretical Framework.....</i> | <i>2</i> |
| METHOD..... | 5 |
| <i>Sample</i> | <i>5</i> |
| <i>Data Collection.....</i> | <i>6</i> |
| <i>Method of Analysis.....</i> | <i>7</i> |
| <i>Ethical Considerations.....</i> | <i>7</i> |
| RESULTS | 8 |
| <i>Context and Traditions of Russian Higher Education</i> | <i>8</i> |
| The Russian System of Higher Education..... | 8 |
| Continuous Assessment during the Courses in Mathematics and Physics..... | 9 |
| Oral Course Exams in Mathematics and Physics..... | 10 |
| Written Course Exams..... | 11 |
| State Exams..... | 11 |
| Cheating..... | 11 |
| <i>Teachers' and Students' Experiences</i> | <i>12</i> |
| Perceptions of what is Measured, and on what Scale | 12 |
| Perceptions of Strengths and Weaknesses of Oral Examination | 14 |
| Opinions on Cheating | 18 |
| ANALYSIS AND DISCUSSION | 21 |
| <i>Analysis and Implications of the Results.....</i> | <i>21</i> |
| Characteristics of Oral Examination | 21 |
| Influential Contextual Factors..... | 22 |
| Implications of the Results | 23 |
| <i>Discussion of Method.....</i> | <i>26</i> |
| <i>Questions for Further Research.....</i> | <i>27</i> |
| <i>Closing Words.....</i> | <i>27</i> |
| APPENDIX 1. INTERVIEW QUESTIONS | 29 |
| <i>Background (About Assessment and Testing in General)</i> | <i>29</i> |
| Teaching plans (both university and lower stages) | 29 |
| Assessment in compulsory and upper grade school..... | 29 |
| Assessment at exams | 29 |
| <i>Interviews with Teachers.....</i> | <i>29</i> |
| <i>Interviews with Students.....</i> | <i>30</i> |
| APPENDIX 2. INFORMATION TO THE INTERVIEWEES..... | 32 |

Introduction

Assessment plays a crucial role in any curriculum. It is especially important in teacher education curriculum, since prospective teachers' experiences of assessment will inevitably influence their future work. The purposes of assessment and the ways of conducting it depend on socio-cultural context and teachers' views on knowledge and learning (Black 1997, Roos & Hamilton 2005, Brown & Knight 1994). I became particularly interested in these issues when I was a student at the teacher education program in Umeå, where I had the possibility to experience and reflect about a broad variety of assessments forms and methods that was used in our courses.

During the spring of 2005, I visited a State Pedagogical University in western Russia to get my final practice as a teacher trainee and gather data for my undergraduate thesis. There, I got to observe a pedagogy that differs in many ways from the Swedish practise. One intriguing novelty was the use of oral examinations, for final exams in compulsory and upper grade school as well as on most courses within the university. For me as a mathematics teacher, this was a new concept. Sweden has introduced an oral part at the national tests in mathematics in the recent years to put more focus on communicational skills, but only as a smaller part of the over all assessment (Nyström & Näsström, forthcoming). According to my own experience, oral examination is rarely used in lower as well as higher education and almost exclusively as a special measurement for children with reading problems.

The Russian use of oral assessment methods seemed to offer some great gains compared to the written assessment that I am more familiar with, but also a number of new problems. As I did some research in the subject, it turned out to be difficult to find literature discussing oral assessment in a wider sense. A lot has been written on assessment of oral skills in language, and there is also information on oral assessment in medicine, law and architecture (Joughin, 1998) but studies on oral assessment in the traditional school subjects were difficult to find.

Cheser Jacobs and Chase (1992) write in their book "Developing and using tests effectively" about oral examinations as a method with many limitations. They categorise it as time-intensive, subjective and stressing for the student, but also admit a number of benefits if the method is used at a smaller seminar. For example, the teachers can get a deepened view on the students' knowledge by asking additional questions and observe the students' ability to "think on their feet", and bluffing and guessing is harder to get away with. However they conclude that oral exams can only be recommended "when a physical disability or injury prevents the student from writing" (p.126). Others as Joughin (in Brown and Glasner, 1999) argue that oral examination promotes "deep approaches to learning" and is a valuable tool for assessment.

Oral assessment is by no means a new invention, rather the opposite. Because of its use of the oral medium, it is probably the oldest. As we put more and more focus on our communicational skills and students different approaches to learning, oral assessment again becomes an interesting option. Since Russia has a long tradition of oral assessment, there is a lot of experience and knowledge available for others

interested in trying the method. Therefore, I decided to make these experiences the target of my thesis.

Purpose and Research Questions

The purpose of this thesis is to examine Russian teachers' and students' experiences from oral assessment in mathematics and physics to find the strengths and weaknesses of the method from the perspectives of its both groups of users. Since teaching and learning always are done in a context and are highly influenced by culture and history, I have chosen an ethnographic approach, where interviews with and participatory observations of the life of Russian teachers and students have been my main tools.

My primary goal with this study is not to give a general picture of how oral assessment is used in Russia or to provide any general results on oral assessment, but rather to investigate how its users experience the method. The purpose of this is to understand some of the benefits and problems of the method from a user's point of view.

The questions of the study are posed as follows:

- What characterises oral examination in terms of how it is conducted and which results it produces, according to teachers and students?
- Which contextual factors affect the examination and in what way?

Theoretical Framework

Joughin (1998, p.367) defines *oral assessment* as “assessment in which a student's response to the assessment task is verbal, in the sense of being ‘expressed or conveyed by speech instead of writing’ (*Oxford English Dictionary*)”. This is the definition that is used in this thesis. According to Joughin, one needs to separate between two different kinds of qualities measured in oral assessment, namely “the student's command of the oral medium itself”, i.e. communicative and language skills, and the “command of content demonstrated through the oral medium” (p.367). This thesis is mainly interested in the second type.

Assessment is often described in terms of having a *summative* or *formative* character, referring to the purpose of the assessment. Formative assessment is seen as something that is done continuously to map the development of the pupil in order to adjust and develop methods and curriculum, and thereby alter an ongoing activity (Hamilton & Roos, 2005; Scriven, 1967 in Airasian, 1994). Summative assessment on the other hand is done at the end of an activity, to examine its results (Scriven 1967, in Airasian, 1994). Final exams, chapter tests and papers are typical summative assessments. However, several researchers stress that these should not be seen as opposites but rather parts of a continuum (Brown 1999 in Brown and Glasner (editors) 1999; Hamilton & Roos, 2005).

I have chosen to focus not on oral assessment in general, but on oral *examination* and *exams*, meaning assessment done at the end of an activity with a mainly summative purpose, to establish a picture of the student's knowledge in order to set a

mark or in other ways assess the student's level of knowledge. The term *assessment* will be used for a broader range of evaluation activities.

Gordon Joughin (1998) has done a literature study where he identifies six dimensions of oral assessment. I will use these in the way that Joughin defines them, to describe oral examination. The first dimension is **primary content type**, which describes "what one is looking for, or remarking upon, in the people one is assessing" (Rowntree, 1987 in Joughin, 1998, p.369). It is split into four categories, *knowledge and understanding*, *applied problem solving ability*, *interpersonal competence* and *intrapersonal qualities*.

Interaction is the second dimension which "refers to reciprocity between examiner and candidate, with each acting on, responding to and being influenced by the other" (p. 369). It is ranging between two poles. At the *presentation pole* where the process is much alike the one in written assessment, where a task is set, the student delivers an answer and the answer is assessed without any additional questions given or any discussion with the teacher. At the other end of the spectrum, we have the *dialogue pole* that is "characterised by a high level of interaction between the examiner and student so that assessment takes the form of a conversation" (p. 371). There are also intermediate points, where for example the student starts with a presentation and then get questioned.

Joughin states that interaction is a key dimension of oral assessment. It is the possibility of a more complex interaction between the teacher and student that gives capacity to "probe a candidate's reasoning, ethics and knowledge (Luntz & Stahl, 1993, p.174)" (Joughin, p. 370). Therefore, interaction is seen as a principal advantage of oral assessment according to Joughin. However, he also address the risk of "bias that is introduced in the interaction" (Abrahamsson, 1983, p.34 in Joughin, 1998, p.370), both concerning the students performance and the teacher's assessment.

The third dimension is **authenticity**, referring to "the extent to which assessment replicates the context of professional practice or 'real life'." (p.371). It stretches between a *contextualized* and a *decontextualized pole*.

The fourth dimension, **structure**, is seen as a fundamental dimension in many studies according to Joughin. It refers to "the extent to which oral assessment is based on a pre-determined, organised body of questions or sequence of events" (p. 372). It varies between an *open* and a *closed structure*, where an assessment with a closed structure follows a strict set of protocols telling the teacher what to ask and in what order, regardless of the student's responses. Respectively, an open structure is characterised by a large amount of freedom for both student and teacher in how to present answers and ask questions, and can take the form of a dialogue. Therefore, this dimension is in that perspective tightly connected to the interaction dimension.

The **examiners** constitute Joughin's fifth dimension. He divides it into *self-assessment*, *peer assessment* and *authority based assessment*, where the last type is the most common. However, oral assessment differs from other kinds in the frequent use of panels, often with external members.

The sixth and last dimension is **orality**. It refers to the extent to which the assessment is conducted orally, ranging from the *purely oral* where only the oral

medium is used, to the *orality as secondary*, with oral explanations or defending of another product, for example a written paper or a manufactured item.

When discussing any kind of assessment, the questions of *reliability* and *validity* are of great importance. Reliability refers to the consistency of the measurement (Berg, 2004; Davis, 1998): will all students be compared in an equitable way and would the result be the same if the test were to be remade? The validity of the test states whether the test provides information appropriate for making assumptions; does the test measure what it was intended to measure? The concepts of validity and reliability are directly linked to what I have chosen to call the *fairness* of assessment, referring to a subjective feeling of whether the assessment is just or not.

In the analysis of my results, I have also used the concept of *flexibility*. The term refers to the teacher's possibilities to vary the manner in which the examination is made. This includes Joughin's dimensions of structure, interaction and primary content type, but it also stretches beyond the assessment situation per se, to include factors as rules for the examination, possibilities of making exceptions from these rules, and taking individual factors into account when asking questions or assessing answers. Flexibility should not be seen as a continuum with given poles or categories as for Joughin's dimensions, but rather as a dimension of variation. It is a measurement of how free the users are to choose and change the manner of the assessment. Consequently, a highly flexible assessment can for example still have a closed structure or low interaction, if there is a possibility to deviate from this if the participants so choose.

Another concept central in my description of oral examination is *individualisation*. With this, I mean the adjustment of procedures and/or tasks to improve the assessment of an individual student. This concept is connected to the concept of flexibility in the way that high flexibility in some sense is necessary for individualisation.

Method

This is a descriptive qualitative study done with an ethnographic approach. Semi-structured interviews and participatory observation have been the main methods of data collection.

Sample

I have studied the use of oral examination at the Faculty of Physics and Mathematics at a State Pedagogical University in western Russia. According to the dean of the faculty, this is a rather typical Russian university, with regards to its size, age and quality.

Five students and ten teachers from the faculty were interviewed.

Five of the teachers were teaching physics, of whom one also taught physics didactics, and the other five taught mathematics, of which one mostly worked with didactics. The experience of the teachers reached from two to forty-six years in the profession. Three of the teachers who were young also answered the questions from a student's perspective.

The students were on different levels in their education (one third year and four fifth year students) having different combinations of the subjects mathematics, physics and informatics as major and minor subjects.

The teacher sample consisted of four men and six women, while of the students two were men and three were women.

Additional interviews were also made with people outside this faculty to get a broader view on the subject. These were a student studying computer science at the State University in the same city, a student that studied Swedish at the State University who also worked as a teacher in Swedish as a foreign language at a school in the city, a mathematics teacher and three ninth grade pupils at another school and a physics teacher from the "technologicum"¹ tied to the State Pedagogical University. All the additional interviewees were women.

The selection of the interviewees were made through convenience sampling, with a touch of purposive sampling (Berg, 2004). The university was chosen because I was an exchange student there during the period of my investigation. I held a series of seminars at the university and worked closely with several teachers at the faculty. That and the fact that I was staying with a student from the university gave me good opportunities to form contacts with the students and staff there. The interviewees were chosen to get a diverse sample regarding subjects and experience form assessment, but also largely from language considerations since the interviews were mainly done in English – a language that rather few people spoke at an acceptable level (see data collection).

¹ A professional education program.

Data Collection

All data was gathered during the spring of 2005. To get a picture of the context, formal areas of use and procedures of oral assessment, I gathered information through active participation in the life of the university and at two schools where I was doing my practice work. I shared an apartment with a fifth-year student at the university through whom I came in contact with other students, and spent a lot of time with teachers at the university who acted as participants and supervisors at seminars I held and teachers that I shared office with. Some of these people also participated in the interviews. My observations were documented daily in the form of diary notes.

Semi-standardised interviews were made with the interviewees (Berg, 2004). A battery of open questions based on the results of the participatory observations were used as a foundation (see appendix 1), but the questions were asked in different order during different interviews, not all questions were asked every time. The questions were sometimes reformulated and other questions were added to deepen the answers from the interviewee.

This method was used to minimise my influence on the responses of the interviewee; asking broad questions and then focusing on the aspects that the interviewee chose to address made it possible to examine areas that I had not thought of. It made the interviews differ quite a lot in how they were made, which is considered as a strength rather than a weakness in ethnographical studies (Kullberg, 1996) since the purpose is not to compare the interviewees but to gather as broad material as possible.

To ensure a formal consent, each interviewee was given a written text in the beginning of the interview, where the purpose of the study and conditions of the interview was stated (see appendix 2).

The interviews were made in different environments depending on practical matters. The majority was made at the university, in offices or empty classrooms. Some were also made in a private setting, in the interviewee's or my home. The time spent on each interview was between one and three hours, in the longer cases divided on two or three occasions.

Most of the interviews were made in English, in some cases with help from a teacher at the university as a translator. Two interviews were made entirely with translation from English to Russian and vice versa, since the interviewees did not speak English themselves. The interview with the Swedish teacher was made in Swedish. The English proficiency level amongst the interviewees varied a great deal, from almost fluent to rather basic, and this is of course a source of error in the interviews. In some cases, the interviewees even expressed that they felt that they were leaving things out because they lacked the ability to express it in English. However, the rather personal relations I developed with many of the interviewees made the interviews quite relaxed, and this made it possible to try to talk about things that were hard to formulate. Another consequence of the language issue was that a lot of time was spent on making sure that we understood each other, which limited the number of issues possible to address in the given time.

Different means of documentation was used during the interviews. Some were recorded on tape and later transcribed into text. In other cases, where the interviewees declined recording, their main opinions were documented through written notes that were transformed into text directly after the interview. The transcriptions also varied in form. Parts of the taped interviews were transcribed word by word, leaving out hesitations and such since they occurred frequently due to language problems and would have made the text unreadable if included. The rest of the taped interviews and the ones made with written documentation was transcribed as describing opinions expressed by the interviewee in a descriptive way and structured around a number of central themes. This was made to make the analysis of the material easier.

Method of Analysis

The transcripts of the interviews and the notes from my participatory observations were read and analysed to find essential constructs that characterised the interviewees' experiences. As a first means of analysis, the dimensions of oral assessment that were proposed by Joughin (1998) were used to give a picture of the character of the investigated methods. For some matters falling outside Joughin's dimensions, new constructs were made.

To make sure that data are correct concerning the context of the examination and avoid too wide generalisations, the results-chapter was checked independently by a teacher and the dean of the faculty where I conducted the study.

Ethical Considerations

The questions posed in the study concerns areas where certain information could be sensitive for the institution as well as the interviewees, not only in relation to the public but also within the group of interviewees, since they have internal relations as colleagues or teacher-students. Therefore, it is important to keep all identities confidential.

In the cases where an interpreter was needed, great effort was made to make sure that it was a person that the interviewee could trust to tell also sensitive things, and that I could trust in giving me the entire information. On some occasions it was hard to tell if the gain from using an interpreter was great enough to compensate the risk of loosing information due to lacking trust, but my estimation is that the interpretation worked well, much thanks to the fact that the interpreter and interviewees were already acquainted. Interpretation was offered as help for the interviewee rather than demanded by me, and several claimed that it made them more comfortable.

Results

The results are divided into two parts, the first giving a picture of the context of oral examination, describing the system of education and the procedures of assessment. It is based on my participating observations and interviews, and is to a large extent valid for Russian universities at large since they are parts of the same system and tradition. Conditions and relationships that are likely to be more specific for this particular university or a particular department or teacher are described as such. Since there was no opportunity to observe an examination during the period of the study, all information about that area is taken from interviews.

The second part is devoted to more subjective experiences and opinions of teachers and students. Quotes from interviewees are taken from the transcriptions of the interviews. In some cases, the grammar has been corrected in order to make the quotes more clear.

Context and Traditions of Russian Higher Education

The Russian System of Higher Education

All students at Russian universities study for five years, in a program following a special teaching plan endorsed by the Russian Ministry of Education. The student takes ten to twelve parallel courses each semester, and only a few courses are personally chosen. At the teacher programs, the students pick one major and one minor subject, in which they are entitled to teach after the education.

There are two kinds of courses with different procedures for marking and examination. Smaller courses are only marked “pass” or “fail”, while bigger courses that are considered central to the program have exams at the end where a grade is decided. Most courses last only one semester, but if they last longer, a “pass” or “fail” is given at each of the semesters before the final exam. There are rather rigorous syllabuses for each course, stating the number of hours and content of the courses.

The grading system has four levels, “two”, “three”, “four” and “five” where “two” means “fail”. There are no formal criteria for the grades given by the government. However, there is a possibility for local initiatives.

The manner of grading at this faculty is basically criteria-based, in the sense that the marks are tied to certain levels of skills or results. There is also an implicit normative thinking, in the sense that some teachers adjust the level of demands to the average level of the class. The teachers’ and students’ perceptions of marking are further discussed below.

There are often no formal tests on the non-graded courses. Instead the teacher makes the decision based on factors as whether the student have been present or not, level of activity on lessons and on smaller tasks during the course. If the student fails, he or she has the right to three retrials. The character of this retrial varies between teachers and courses.

The students have about four exams per semester. To be allowed to take the exams, they need to pass all the non-graded courses taken during the semester. At the

end of the program, the students take state exams in their major and minor subjects that cover all the courses taken during the program.

If a student fails at an exam, there are three retrials. At the third, at least two other teachers are present except for the teacher responsible for the course, and they decide the grade together. This is to ensure the student of a fair judgement, so that the decision does not depend only on the opinion of one teacher. If the student fails at all the retrials, he or she must take a break in the studies, retake the failed course and pass the exam, and can then continue with the program. This means that an entire year of study time is lost. The student is not allowed to continue with any other subjects during this year. Sometimes, it happens that students get a fourth retrial, if they have valid reasons. This is a judgement made by the teacher and the faculty administration.

It happens that students get an “automatic exam”, which means that students who have worked hard and showed good understanding during the course get their grade (usually a five) without taking the exam.

The grades from the exams have important financial consequences for the students. If a student have only fours and fives, he or she gets a scholarship of 600 rubles per month (about 150 Swedish crowns) during the coming semester. Only fives renders a scholarship of 800 rubles. (This should be put in relation to the 600 rubles per month that students get if their parents’ incomes are low enough. In other cases, the scholarship is the only financial support the students get.) The scholarship is only given if the student passes the exam on the first trial.

If the student at the end of the studies have fives on both state exams and on the diploma work as well as in seventy-five percent or more of the courses, he or she gets a “red” diploma (the name comes from the colour of the document) instead of the ordinary “blue” one.

”Better to have a red face and a blue diploma, than a blue face and a red diploma²”

(Russian saying, according to Teacher A)

Continuous Assessment during the Courses in Mathematics and Physics

There are a large number of control activities during the semesters in the Russian educational system. It is of both formative and summative character, even if the grades are formally given at the exam. These activities consist of written homework assignments, written control tests (interim exams) and blackboard presentations, where students solve problems in front of the class. The presentations can be both voluntary and assigned by the teacher and occur at almost every practice lesson in both subjects. In physics, there are also written lab reports that are orally defended.

The physics teachers sometimes use so-called “colloquia”. The form and function of the colloquium differs between the teachers. Some teachers use it as a voluntary interim exam where the students can pass chosen parts of the curriculum and thereby

² The blue face refers to being very exhausted after studying too hard.

reduce the amount of material covered at the course exam. These colloquia are similar to the course exams in its performance.

One of the teachers interviewed uses the colloquium as an obligatory checkpoint that all students need to pass in order to take the exam. He also has a different structure on the presentation: the students are divided into groups of six to eight and were given one question each. They get about fifteen to thirty minutes to prepare and then the answers are presented to the rest of the students, who are expected to participate by asking questions or commenting on the answer. Questions can also be asked to the teacher. This is however not considered being a part of the grading process but rather an educational activity, and the material covered at the colloquium is repeated on the course exam.

Oral Course Exams in Mathematics and Physics

Traditionally, the course exams are oral, even if some teachers choose to have written tests.

At an oral course exam, the students are given thirty to fifty theory questions at least two weeks before the exam, sometimes earlier. A few days before the exam, there is a consultation where the students have the opportunity to ask the teacher for clarifications and explanations of the questions.

The structure of the exams is very open; it is up to the teacher to choose methods and set guidelines, and different students can get treated in different ways during the exam. However, there are a lot of traditions surrounding the exams, which make them quite similar to each other in their performance. The exams are always given during two days, and the students decide themselves what day they want to come. They are normally conducted in an ordinary classroom. Five or six students enter the room and draw a “ticket”³ each with three to four questions of which about half is theory questions and the rest are problems. They then have about one hour to prepare some notes for the answer. Usually, no aids like calculators or tables are allowed. The answers are presented orally to the teacher, in the same room as where the preparation takes place. The students normally choose for themselves when they feel ready to present their answer. Due to this, it is possible to spend up to three hours preparing, but if no one want to present after one hour, the teacher can ask someone to answer. When a student is done with the exam, another student enters the room and starts preparing.

The presentation of the answers is done differently depending on the teacher. The level of interaction between teacher and student varies between teachers and exams as well as between presentations within an exam. It is to large extent depending on the performance of the student.

The most common procedure is that the student sits down with the teacher at a table to look at the written notes together. The student starts telling about the answer of the first question. If the teacher finds something to be unclear, he or she might stop the student and ask for a clarification. If the answer on the paper looks correct, the

³ *Biljet* in Russian

teacher can interrupt the student and tell him or her to move on to next question. In cases where students have misunderstood the question or got stuck on some smaller issue, the teacher can give an explanation or hint, or ask the student to solve a similar problem. Many teachers also ask some additional questions about central parts of the subject, to make sure that the student's knowledge cover the entire area that is examined, especially if the student is to get a five.

If the students feel that they got an "unlucky ticket", they can ask for a new one. This is usually granted, but then the grade is automatically lowered one step.

The teacher tells the student what grade he or she gets right after the presentation. If the student feels that he or she deserves a better grade than the one given, there is a possibility to ask for additional questions to show more knowledge. This opportunity is however not often used by the students, and it is up to the teacher to grant it.

It happens that students who were expected to perform better than they did on an exam get a chance to come back the following day to take the test again. This is used especially if the student is very nervous or ill in some way.

The student's notes are archived after the exam, to be used should there be any disagreement about the teacher's judgement.

Written Course Exams

Some teachers prefer to use written exams. In my sample, only a few mathematics teachers used them, and according to my interviewees, it is rather rarely used also at the universities in general. The control tests at the middle of the semester are however traditionally written.

The written exam contains a larger number of questions, usually between five and ten, with a higher percentage of problem-solving tasks than in an oral. It is done in an auditorium or classroom during a time span of three to five hours. The results are announced the following day.

Written tests are very often given out in multiple variations, to prevent students sitting next to each other from discussing the questions and helping each other.

State Exams

The state exams are always oral and are similar to the oral course exams in its performance, even though there are some differences.

The students receive the questions for the exams about two months ahead of time. The number of questions is about 100-120. The student presents the answers to a commission of five to seven teachers. The head of the commission is always a person from another university. After the students' presentations, the commission discusses the grades and the students are then called back to get their results. The process of deciding takes about ten to thirty minutes.

Cheating

There is a lot of cheating going on at the examinations, especially during the state exams, where a vast majority of the students use some kind of officially prohibited aid. There is no written policy against cheating at the university or in the national

guidelines and no strict policy for dealing with it. The most common form of cheating is to using crib sheets. There are two kinds: smaller notes with formulas or words for remembrance and bigger so-called “bombs”, a full-size answer sheet that is brought and switched with the paper for written notes, or copied over to the assigned sheet. I got to observe the preparation of such notes for the State Exams in mathematics and physics. The students wrote full answers to almost every theory question on strips of paper that were folded and packed into heaps and brought to the exam. Some notes were printed on computer and shared with other students.

Another classical form of cheating is to ask a friend for help. This has taken a new turn with the arrival of new technology. Students bring their mobile phones to the exam, and hide the headset under their hair or in their sleeves. A friend outside reads the answers and the student just copies it on the paper.

There are two main ways for the teachers to deal with cheating if it is discovered. One is to fail the student and let him or her retake the test at a later occasion, which means losing the chance of getting a scholarship. This action is however rather seldom used. The second and more common method is to investigate the student’s knowledge more thoroughly by asking additional questions during the presentation. Some also take action by telling the student to quit or take away the used material, but most of the teachers I interviewed don’t let the student know that they discovered the cheating. Consequences of the cheating and the opinions of teachers and students about it are further discussed below.

Teachers’ and Students’ Experiences

Perceptions of what is Measured, and on what Scale

To understand whether a measurement tool is adequate, one needs to know what it is supposed to measure. For this reason, one of the main focal points of the interviews was to find out what the exam is expected to reveal, and how teachers and students perceive the grading process.

According to the teachers, the primary content type of the exam is knowledge about and understanding of the material dealt with during the course. There is also some testing of problem-solving ability, but several teachers claim that oral exams are more focused on theoretical understanding, and the problems given are said to be rather basic. Little attention is said to be given to communication skills; it is only weighed in if the student is balancing between two grades, and it can only influence in a positive way.

The exam can be characterised as semi-oral, between Joughin’s poles of purely oral and orality as secondary. The students present their answers with the aid of a written paper, but it is not the written answer itself that is primarily judged. The teachers have different ways of dealing with this. While some teachers consider a “clean” solution to prove the students’ understanding and therefore don’t demand a complete presentation of the task, others see clean solutions as an indication of cheating and make sure that the students can explain what they have done. A badly structured sheet of notes can however never lower a grade, according to the teachers. Presentation skills are only judged in the methodology courses.

As mentioned above, there are no official grading criteria for the teachers to use, and so they need to create their own. How this is done differs greatly. The teachers of physics- and math didactics say that they discuss grading within the departments and feel that they have mutual criteria.

The teachers from the mathematics and physics departments on the other hand never discuss grading with each other. However, they claim to follow a silent tradition, based on their own experiences of how it has always been done. When asked about how they make their decisions of grading, they all have some kind of “rule of thumb” about how many tasks from the ticket that need to be correctly solved. However, there are no clear rules and lots of exceptions, and they claim to have a “feeling for” the grades, rather than formulated criteria. The teachers who had participated in state exams also claimed that there are usually no problems with agreeing about grading in the commission.

The same can be said about the students, they have no clear picture of what is expected for each grade, and the amount of information they get from their teachers vary greatly. Still some of them claim to have a feeling for how well they performed on the exams, and have some kind of silent understanding of the categories. Many teachers also have a tradition to explain to the student why a specific grade was given, and the teachers I talked to say that the student seldom objects. This can be seen as an indication of that students understand the criteria and find the judgement to be fair, but it could also be a result of the rather strict hierarchy between teacher and student that makes the student unwilling to protest.

It is important to note that many teachers see the exam only as a part of the examination process as a whole. Apart from what is shown on the exam, performance during the course plays a great role, both concerning knowledge and level of participation. Many teachers reward hard work and development.

“If a student works hard and tries his best, I am more loyal to him. /.../ To some extent it is benefit for hard work; for the other, if the student is hard working but he hasn’t the ability to... his head is not as good as it could be, I am loyal to that. Another point is that if a student works hard, he must have something in his head. If he can’t show that, maybe he is nervous, maybe he has a headache or something else, but I know for sure he knows something.”

(Teacher B)

Teacher A says that if she has two students with different levels of ability but with the same results, she would give a higher grade to the weaker student, since “they grow up compared to students who come with knowledge and leave with the same knowledge”. She thinks that the exam is more of a formality since she has already formed a good picture of the student’s knowledge during the semester. She also adds that there is less risk of forgetting what is learned if the student is working during the semester, than if he or she is “cramming” the last week, and for that reason, knowledge acquired during a longer period of time is valued higher.

Teacher B also takes the level of the class into account when deciding on criteria.

“You should take into account the average level of students. Students are averagely very weak, and you should be much less demanding, and if they are clever, you should be more demanding. But it is very difficult to estimate, even for experienced teachers.”

(Teacher B)

When I asked if he did not think it was unfair that a student could get a lower grade and thereby lose the scholarship or the red diploma for being in a strong class, he said that he had never thought of that.

Several students claimed that previous grades can play a role at the examination. Student G tells me about a saying amongst the students:

“The first and second year, you work for the zachotka⁴; the third and fourth year, the zachotka works for you.”

(Student G)

Showing the teacher nice grades can influence the teacher to think twice about giving a low grade: because it shows that the student is normally very good, or because he or she doesn't want the student to miss receiving a red diploma or a scholarship.

Perceptions of Strengths and Weaknesses of Oral Examination

During the interviews, oral examination was discussed in terms of benefits and drawbacks, mainly in comparison with written exams since that is the only alternative used here.

The first thing mentioned as a benefit by a majority of the teachers was that the oral exams gave a deeper picture of the students' understanding. Several teachers think that the students have a better chance of showing their knowledge, partly because of lacking communicational skills. It is difficult to determine from a written exam if they have understood the material or not, and if faults in the answers are due to lack of knowledge or just mistakes and misunderstandings.

Teacher D answers like this on the question of why he prefers oral examination:

Teacher D: Well, I think it is better for me and it is better for the students, because our students don't know Russian. No, it's a terrible thing, yes. /.../

Eva: How do you mean they don't know?

Teacher D: They don't know how to write good questions...good sentences, they don't know where to put commas, and other things yes, and it is usually very poor, very poor and it is very poor answers, they speak better.

So for me, I can ask them things that I think they forgot to put on the paper but they know, I can... It is better for me and it is better for them, I count.

But it is poor Russian. Our students in the whole don't have a good knowledge of Russian. O, I don't mean that everybody, yes, but most of the students have a low level of Russian.

Eva: So you feel that you get a better picture of what they know?

⁴ The zachotka is a notebook where all the student's grades during the program are noted.

Teacher D: Yes, I feel a better picture and they can express themselves better also.

Teacher B had a different motivation:

”When you have an oral examination, you can get to know how they understand the subject, cause you can ask them questions from this area, from that area, and you can get the total picture. Another point is that you can check the theory – how they understand it. When they solve the problems, they use the standard method, algorithm, to this, this and this, while they can absolutely not understand the theory.

(Teacher B)

The possibility of “*little questions*”, as the interviewees call them, about the solution or about other areas of material, allow them to detect if the student has cheated or memorised the method or definition without understanding it. They can also find cases where the student has misunderstood the question, made a calculation error or got stuck on some smaller detail. In these cases, the teachers can *help* the student, by giving a hint, explaining something or reformulating the problem. The possibility of giving students that perform worse than expected a second chance is also seen as a benefit. The teachers are aware that an exam is a very stressful experience for the students and don’t want it to be the reason for failing.

One teacher who uses written exams has added an oral part to get around problems with cheating and memorising. When the students come to get their grades, she asks them to explain their solutions if she suspect cheating, and lowers the grade if they can not answer her satisfyingly. She also discusses the solutions with students that have poor results, and if they can correct their calculations or solve a similar problem, she lets them pass.

The students also see the high level of interaction between the teacher and student during the presentation as positive. Several of them brought up the possibility of getting help and hints as one of the big advantages with oral exams. One student even use it as a strategy:

“And another method is like “warm and cold”. Do you know this game? You hide something and you say “warm, warmer, oh it is hot, oh, you find it!” And you can do this with the teacher and he ask me and I: “so if this is ok, so that it will be... it will be... it will be...” and the teacher answer and I, not I answer his question.”

(Student A)

Teachers that I discussed this with were however doubtful of if it would work, and claimed to find it quite amusing to see how the students tried to get hints like this.

The students also like the possibility of getting a second chance, ask for a new ticket or raise the grade by asking for more questions. However, the interaction also causes more stress to a majority of the students than a written exam, according to both teachers and students. This is not only seen as something negative – it is also

presented as one of the benefits, since it gives valuable practise in dealing with stressful situations.

“I think oral tests prepare you for that life is an exam. During our life, we introduce ourselves to people, we meet with people and we talk to unknown people and later this unknown people become our friends. And so oral exams are not only the answer to the question, it is a psychological exam too. It shows how the student can keep from showing how he is worrying, to present his answer. Even if we don't know the material as well as we want, if I talk good, if I speak good, I can speak with the teacher, and with the help of some methods, I can pass the exam, and I will get a not so bad ball⁵, like not three but four or maybe five, even if I am not ready.”

(Student A)

The interviewees also talk about other aspects of the oral exam as a *learning opportunity*. Above all, the students get an important chance of practising to communicate their knowledge orally. Teacher E claims that it not only practises their oral skills *in* math, but also *through* math, to become better at explaining in a more general sense.

Several teachers talk about the possibility to explain misunderstandings and errors, and give instant feedback to the students, who also appreciate this possibility. Teacher H says that she liked oral exams as a student because the opportunity to discuss with a skilled teacher gave her a sense of “self-realisation”, as well as a broader view on the subject.

Teachers D and F stress the importance of personal contacts with the students as a tool for teaching. They assert that the oral exam provides a valuable and rare opportunity for direct dialogue with the students, not only about the subject, but also about other things, as studying technique and life in general.

“Any contact between teacher and student is a wonderful opportunity for upbringing”.

(Teacher D)

Teacher F uses this opportunity mostly with the more gifted students; he doesn't spend so much time checking their answers and use the time for dialogue about more sophisticated parts of the topic or solution.

Teachers and students address several problems with oral examinations, apart from the stress mentioned above. One such problem is exhaustion:

“When you are alone taking that exam, you get awfully tired. By the end, you have no power to get to know how, to investigate what he knows. That's kind of a pattern: three out of four [correct solutions], four [as grade]. Go away.”

(Teacher B)

⁵ *Ball* is Russian for “mark”

Teacher E, who is a very experienced teacher, says that this was a bigger problem twenty years ago, but thinks that it works fine now since there is a new rule that says examinations can not be held for more than six hours per day. Even if she still gets tired, she thinks that she keeps her focus.

The time aspect is a problem also for organisational reasons. An oral exam takes a lot of time to do and is rather hard work for the examiner. This is, according to Teacher F, the reason why the control tests at the middle of the semester almost always are in writing. It is however also a matter of tradition; in lower education, written control tests are mandatory.

There is some disagreement among the teachers about whether the oral exam has the ability to test the entire curriculum or not. While an oral exam has the ability to “dig deep” into the students’ understanding of the discussed concepts, a written exam has more width and can cover a larger area of the course.

One teacher who uses written exams, answered the question of what information he gets about the knowledge of the student like this:

”Oral examination usually give an insufficient information. However oral exams allow us better understand how well a very strong student knows intricate details. /.../ Taking into account that oral exams are 5 times shorter than written exam, a person who conducts an exam has to cut number of topics discussed.”

Quote from written answer from teacher I.

Teacher E on the other hand contends that written tests give too little information about each student and is better to use as a formative tool during the semester to check the level of the group.

Some teachers, and also students, contend that there is a “gambling” problem and stress the fact that there is a lot of luck involved. If a student knows only parts of the curriculum but is lucky and gets a “good” ticket, he or she can get a five anyway. According to some students, this is also used as a strategy: to learn only half of the questions, and hope to get a good ticket or know the answers to half the ticket and pass the other half by cheating or getting help from the teacher.

However, not all teachers consider this to be a problem. They think that they can check the important areas asking some additional questions. Teacher A even claims that students learn the general parts of the curriculum more thoroughly before an oral exam than before a written, because that know that they can be asked about it.

It is worth noting, however, that the use of tickets itself, as well as using different versions of written tests, are not commented on as unfair by anyone. If someone gets an easier ticket than someone else does, that is just a matter of luck. Whether the examination is equitable or not between the students, does not seem to be a big concern.

The flexibility and open structure of the oral exam leads, as already noted, to a great diversity in how the students are assessed. Due to this, the exam is affected by the subjectivity of the teacher. This was one of the important foci in the interviews, since it has obvious consequences to the equity of the assessment.

This characteristic of the oral exam was addressed as both a strength and a weakness. For the few teachers that chose to have written examinations, objectivity was the most important argument.

“I prefer written examinations because they are more objective than oral examinations. /.../ It is very difficult to maintain consistent grading during a six hour long oral exam. /.../ I believe that a discussion on the subject is the most important way of learning mathematics. However I believe that examinations should be objective, and as a consequence, written.”

(Quote from written answer from Teacher I).

The presentation and grading of the oral exams are as we have seen sensitive to many contextual factors, such as the teacher’s previous impressions of the student. It is also impossible to compare the answers of several students before deciding the mark.

Some of the students also addressed the level of subjectivity as a problem. All my interviewees agree that the teacher’s opinion of you will have consequences for how much help that is given, and what questions that will be asked. However, the perception is that it is more common that liked students get benefits than that disliked get a harder judgement. Whether it is a problem or not is a question for discussion. Some teachers and students claim that the subjectivity of oral exams is what makes it fair. They think that it gives teachers a chance to assess more than just the performance at hand, but to take a broader view into account. Helping students by giving them extra questions, hints and second chances is not liberalism, but a way of getting more information about the student. It is only fair according to this group that students who have worked hard or has some kind of personal problems can get help on the exam, in the same way as it is fair that those who have worked poorly or is suspected to cheat get more questions.

There are however also sources of bias that are generally seen as problems. One is the high stakes that is given by the scholarship issue. There is a tradition in the Russian school system always to judge in the favour of the student if there is hesitation, and these factors increase that tendency. Some students use the scholarship as an argument for higher grades, and the teachers experience it as stressful. They know that losing the scholarship can have great consequences for the students.

Clearer grading criteria was a method suggested by several teachers, to achieve a less sensitive assessment. The teachers think that more experienced teachers with clearer criteria also make more objective and balanced judgements.

Opinions on Cheating

Since the cheating going on during the exams, influences to a high extent how teachers arrange the presentation, and since the possibility of discovering cheating is presented as one of the advantages with oral exams, this was one of the areas for discussion in the interviews.

As mentioned, there is a lot of cheating on exams, and teachers normally do little to prevent it.

Teacher D: You know, life changes. Twenty years ago you usually said to such a student that you should go away. He received a bad mark and that is all. Now something happened, I don't know, you know. We don't send them away but, I use the method I told you about, I ask many more questions. When I put him his mark, I usually don't take into attention the answer to the tickets, because they are cheated. Sometimes I can tell.

Eva: Do tell the student: since I know you have been cheating...

Teacher D: You know, it is very inconvenient for me to tell them you know. I go, I say, It's shame for me to say such a thing, I would become red you know. (Laughs.) When I see such things, it is not my... so I don't say anything but usually they understand.

Teacher D is rather representative in his answer. Both students and teachers contend that cheating is a tradition that is difficult to break, and therefore there is little point in punishing the cheaters.

Several teachers don't see cheating as a problem since it is easy to discover. They often see who is cheating during the preparations, or when the students are presenting their answers, and by probing the student's knowledge extra deep, they see if he or she knows the material or not. A cheating student can still get a five, if he or she proves to have the necessary knowledge.

Even though students are rarely formally punished, they still fear getting caught. They know that it is not allowed but don't have any clear picture of what could happen if discovered.

Some teachers and students think that cheating is good, in the sense that students learn a lot from preparing crib sheets, and many have them with them to feel less nervous but never use them.

Cheating is also seen as a way of dealing with a system that is not functioning. Teacher A means that some courses and especially the state exams would be impossible to pass without cheating, and that the teachers accept it because they know this. The content of the exams can not be changed because of the teaching plans, and so the rules of the examination are bent instead.

“So in our situation in Russia, in our system of education, we can't live without cheating because we have criteria from our state that students need to know, but no student knows it. Cheating is just a part of the system. /.../ It is very bad for the country in general, because if you start to cheat in school or university, of course you will continue to cheat and it is not so good of course. But we need to change our system of education before we make a taboo on cheating.”

(Teacher A)

One student thinks that the state exam commission allows cheating because they and the university would look bad if too many fail. Two students with experience from working in compulsory school said that they themselves allowed some types of cheating and did not punish cheaters by failing them because a large number of twos

would give them problems with the principal. Other interviewees say that teachers look away to avoid extra work.

The general opinion is that school has no moral responsibility to stop the cheating; students should not learn that it is ok, but cheating must be dealt with in a bigger perspective.

I hate cheating, it is like stealing, but it is not my job to stop it.
(*Teacher E*)

One student even posed the opinion that cheating was the most valuable thing that he learnt from his education, since cheating is a way of looking after oneself and an important skill to succeed in the real life.

There is a considerable amount of practicality in the reasoning of some teachers. One teacher expressed that it is good for students to learn things properly instead of cheating, but that they had to be “careful” when they were not cheating in less important subjects, since there was a risk of learning unimportant things. I was also told about teachers leaving the room during the preparation, in order to give the students a chance to help each other. The reason for this was said to be that the teacher knew that the students could not pass the course without cheating, and that it did not matter anyway since the course was not so important to this group of students.

Analysis and Discussion

The purpose of this thesis is to investigate how teachers and students perceive oral examination as a tool for assessment, in order to make their experiences available as a source of information for others who want to try the method. It also enables a scientific evaluation of the method, which can help its users to improve or replace it if necessary.

This chapter begins with a summarising analysis of the results with focus on the research questions asked in this thesis, followed by implications of the results. After that, the method is discussed and some questions for further research are proposed.

Analysis and Implications of the Results

Characteristics of Oral Examination

Based on the results of my research I have been able to extract some general characteristics of oral examination in the Russian teacher education, of which the most interesting are the flexibility, subjectivity and individualisation within the method. These contain both the strengths and the weaknesses of the method.

Oral exams and other forms of oral assessment, as they are used at the faculty I studied, contain a high degree of *flexibility*; the teachers have possibility to choose and change the rules and procedure of the assessment. This characteristic is highly appreciated by both teachers and students. The open structure of the exams makes it possible for the teachers to adjust the level and content of the interaction to the student at hand, taking factors as the student's previous performance, level of stress and suspicions of cheating into account. Actions as letting a student have a second try is an example of flexibility that goes outside the dimensions of structure and interaction, as Joughin (1998) defines them.

The teachers see the flexibility as a crucial property of the method, and claim that it gives them the possibility to probe the students' knowledge and abilities in a deeper and more profound way. Factors that could threat the validity of the exam, such as low performance due to misunderstandings or nervousness, or a possibility to get high marks through memorisation and skilled communication of shallow understanding, are to a large extent neutralised according to the teachers.

On the other hand, the exam becomes rather sensitive to the teacher's intended or unintended *subjectivity*. Both teachers and students express worries that personal relations can have influence on the assessment. It affects how much help the student gets, as well as how the performance is perceived, and some teachers feel pressured by the fact that their marks can cause a student to loose a scholarship or a red diploma. These results rhyme well with Abrahamsson's results about "bias that is introduced in the interaction" (Abrahamsson, 1983, p.34, in Joughin, 1998, p. 370). The interviewees suggest clear criteria for grading as one of the possible ways to get around such effects, and indicate that experienced teachers are better equipped to withstand them.

The interviewees present oral examination not only as summative assessment but also a *learning opportunity*. The interaction enables a high level of *feedback* to the

student, a highly appreciated possibility even amongst the critics of the method. Discussions with the teacher give the student a broader view on the concept or problem that is discussed, correct misunderstandings and make the student aware of lacks in his or her knowledge. This adds a *formative function* to the exam. The examination itself also provides practise in communication skills and performing under pressure.

All these characteristics of the oral exam create good preconditions for *individualisation*, as the teachers can stimulate different students in different ways. While nervous students can get support and help, high-achieving students get to show their wits and be challenged. The teachers can also adjust the questions to get exactly the information that was needed to have enough data for grading the student. However, it is interesting to note that no teacher or student mentioned the concept of individualisation explicitly.

Influential Contextual Factors

There are a large number of contextual factors affecting the examination, and it is very difficult to map them all. However, there are some factors that affect the assessment more than others.

The most widely affecting factor is the *cultural context*. Whether the examination is satisfying in its performance and results or not, is depending on what is expected, and that is highly influenced by culture. First of all, it is important to note that the use of oral exams is based on a strong tradition, and is therefore not really an active choice on the part of the teachers. They can choose to use another means of assessment, but oral examination is the norm. The main features of the exam are similar across its users, and it is reasonable to assume that it is the result of a rather stable scheme of what an oral exam is, that is reproduced by teachers and students.

One part of the cultural context is views on what should be assessed and when. At this university, the grades depend not only on the performance on the exam, but also how the student has worked and performed during the semester. Hard work is almost as appreciated as knowledge and skills, and is therefore taken into consideration when the assessment is done. As a consequence, only a few teachers and students find it unfair or problematic that hard working students get more help at the exams, while people missing classes get more questions.

The Russian culture also seems to contain an attitude where people are expected to look out for their own interest and accept a certain amount of arbitration from the power-holders. This has several side effects. One is that people don't see the possibility of drawing a "lucky" ticket as problematic, giving some students an easier exam. The students, as well as the teachers, don't seem to put the assessment of one student in relation to the assessment of another. When I asked if they saw it as unfair, many did not even understand the question. If you are lucky, then that is good for you, and it does not affect me in any way. In the same way, cheating at exams was only seen as negative since it made the student learn less and made it harder for the teacher to assess the true level of knowledge. No one addressed it as getting an unfair advantage over the classmates, in terms of higher grades and scholarships.

The *organisational context* is also highly influential. The teachers have guidelines within their profession to adjust to. A very distinct factor is the teaching plan, since it states what material that should be dealt with at the exam. A teaching plan that many consider to be too hard in its demands, together with the fact that teachers fear getting problems with their authorities if too many students fail, seems to make teachers more willing to overlook incidents of cheating. If they feel that the student knows the material good enough, there is no need for disciplinary punishment. However, there are exceptions in this matter.

Another important part of the context is the *frame factors*, for example number of students per class and available time. Since the classes often are big and an oral presentation takes time, only a few questions can be given. This decreases the width of the exam and makes some teachers concerned that they get an unsatisfying picture of the knowledge of the student. Other teachers mean that this is not a problem since they can ask additional questions, and also have with them rather much information on the student gathered during the semester. The fact that the teaching during the courses contain a lot of “checking up” activities like black-board presentations, homework and control exams make the narrowness of the actual exam less problematic that it would be in a course purely based on lectures.

Obviously, the time intensity on the exam also limits the teachers’ perception. After listening for several hours, it is hard to put the same amount of energy and thoroughness into every assessment. This lowers the reliability of the exam.

Implications of the Results

The results of this study draw a picture of oral examination as a flexible assessment tool with many strengths, as well as weaknesses. The major source of both pros and cons is that the teacher plays such an active role in the process compared to other forms of examination. This makes it much more complex.

If we assume that the teachers’ perception of getting an in-depth view of the students’ knowledge is correct, this makes oral assessment a valuable tool for all teachers. In the teaching situation, this probing through questions and discussions should take place continuously and not be saved for the final evaluation. However, in an academic environment such as universities, there is often little time for such activities.

One of the critiques against the method is that it probes a too narrow area of knowledge to be considered as a valid assessment of the entire course. As I see it, this depends a great deal on how the teaching during the course is organised, as well as on how the teacher choose the tasks and handle the presentation. The Russian teaching tradition contains a lot of activities where the students can be assessed, such as blackboard presentations, homework and control exams. If the teachers use this information during the exam to probe areas that have not yet been assessed in other ways, this method might prove to be more informative than an ordinary standardised exam, where all students answer the same questions. The focus on theory that is perceived by the teachers is also not a big problem since the control exams, labs and practise lessons deal with those areas. Consequently, oral exams can be said to be a

suitable method in the Russian system of education, given that the teacher uses the information available.

The issue of narrowness becomes more critical if we apply the method to a teaching tradition where the exam constitutes the only basis for assessment, as according to my experience is the case for many math courses at Swedish universities. In such an environment, width in the assessment is much more important. However, the written tests often don't demand more than memorisation of proofs and standard algorithms, and one could question if not deep probing of some areas would facilitate learning better than no probing at all. Joughin (in Brown and Glasner, 1999) argues that oral assessment make the students learn in a different way since they know that their understanding can be tested and good memory is not enough. Data in this study support this.

When considering assessment, one needs to think about reliability and validity, but also about the purposes of the assessment and perceptions of what is a fair judgement. There is often a high focus on reliability when discussing assessment (Davis, 1998); the tasks should be as objective as possible and easy to evaluate. However, reliability often conflicts with validity. At the Russian oral exam, the teachers are more concerned with getting a profound understanding of the knowledge and skills of each particular student than with reliability issues. One basic part of reliability is that all respondents should be measured on the same scale, i.e. answering the same questions. However, it is not enough that all students read or hear the same question, they also need to understand the question in the same way, unless it is the ability to interpret questions that is tested. The Russian teachers argue that the flexibility within the method allow them to give the students hints and additional questions in such a way that they can show their knowledge even if they misunderstood the initial question or got stuck on some irrelevant detail in the task. In this way, they argue that asking *different* questions actually increase reliability. This is also a matter of validity – the assessment becomes focused on the understanding itself, rather than on the ability to interpret and deliver well-articulated answers.

However, there are other reliability issues that deserve particular attention within the oral examination procedures, tied to the issue of equity. When assessment is done with a selective purpose, giving students grades or results to compete with, or when assessment is done for comparative purposes to evaluate an educational program or school, it is very important that the assessment is equitable (Black 1997). One aspect of this is to what extent the assessment is consistent with the set criteria. For example, as an oral part was introduced on the Swedish national test in mathematics, that is used both as a tool for teachers in grading and as a control of the quality of the schools, large emphasis was put on creating clear guidelines for grading (Skolverket, 2003). An evaluation of the initial attempts with the tests showed that the assessment was facilitated by teachers having studied and internalised clear examples for comparison, and discussed the criteria with the students before the tests. In this perspective, the Russian oral exams are problematic. Both teachers and students feel that there is a risk of large variation in how students are judged, partially because many teachers have very vague standards for grading. It would be desirable to have

an extended co-operation and discussion amongst the teachers in this matter, to reduce the differences.

Another factor influencing the equity of the assessment is the fact that students get different amounts of support and questioning during the exam. Cheating is also a part of this problem. This is however an area where the perceptions of the purposes and fairness of the test must be considered. It should be noted that “fairness” is a highly subjective and culturally determined concept that depends on what one expects and sees as correct judgement. For example, many of my interviewees see it as a natural thing that exams are not held as something separated from the course itself. There is nothing strange with giving automatic exams and extra chances to hard working students or questioning students with low participation and cheaters extra hard, rather the opposite; it is “fair” that earlier performance as well as discipline is weighed in when the grade is set.

There are also other moral considerations to be made. The Russian schools system has a tradition of judging in favour of the student. This is understandable since there are high stakes at play when grades are set. Failing at an exam can cost a student an entire school year, and for boys, it also means a risk of being drafted to the army⁶. Other less dramatic consequences of low performance are losing the scholarship and the chance of getting a red diploma. When judging this goodwill, we need to take the society as a whole into consideration. In the Russian system, one could argue that getting students through the education is more important than judging everyone according to the same standards. On the other hand, there is a risk of lowering the demands too much, and thereby hurting the quality of the education. For other schools and systems, such considerations might be unacceptable, but in this case, they are often possible to motivate. As some of the teachers and students point out – the system needs to change before the school can do it.

However, some measurements could be taken to improve the equity. For example, some teachers allow the students to use books or bring notes to the exam. Then the conditions are equal to all students, and the teacher does not need to silently allow breaking the rules.

It should be noted that the teachers who promote written exams do this mainly from an objectivity perspective. This should be seen in the light of the Russian history. According to Chen (1994) and Vershik (1994), the entrance exams at some institutions for higher studies during the Soviet time were very arbitrary and were used to exclude politically unwanted students. According to my interviewees, this type of inequitable selection still takes place in some faculties, but on other grounds than political ones. This is of course unacceptable in a democratic society.

To summarise, this study gives a picture of a method with great possibilities, if used in a proper way. The teachers’ active participation in the assessment process creates possibilities of reaching high validity and compensating for several sources of bias, such as misunderstandings and memorisation. However, it also gives great

⁶ Military service is compulsory for all men over the age of eighteen in Russia. If they attend higher studies, the drafting is postponed.

responsibility to the teacher. If a fair and equitable assessment is to be reached, the teacher must use all the available background information on the student, have clear criteria for grading and make sure to leave all personal relations that are not connected with the assessed qualities aside.

The possibility of individualising the assessment to suit the different needs and qualities of the student makes oral assessment an effective option. The study also present some interesting variations of the method, for example the colloquium where the students are assessed in groups, and are expected to give feedback on each other's presentations. Here, a larger number of items are dealt with and the students can learn from each other. Another interesting idea is provided by the teacher who added an oral part to a written exam. This is a possible way of uniting the benefits of two methods and thereby reducing the lacks of both.

Discussion of Method

Since the object of study was not oral exams per se but rather the users' perceptions of it, the results should be interpreted carefully. It is also important to keep in mind that perceptions are closely tied to culture, which make the results difficult to generalise directly onto other cultural settings. However, the results are still interesting in an international perspective.

The ethnographic approach used in this study gives a broader picture of oral examination, placing it in its sociocultural context. This makes it possible to see what it is that makes certain characteristics an asset or a problem. The possibility to probe the student's understanding of a concept in a profound way is for example a general quality, while the perception of fairness in giving students different amounts of help is culturally bound. It is also important to keep in mind that the interviewees' judgements should be viewed with regards to their points of reference. They speak of oral examinations in a context where that is the traditional choice and thereby the norm.

The fact that this study is dealing with examination in mathematics and physics do not limit the generalisation possibilities to those subjects only. The discussed characteristics are rather fundamental, and can be applied to most subjects.

When conducting an ethnographic study, the researcher is the measuring tool and his or her preconceptions and ideas will affect the results. This study focused on not just looking on what happened practically and how people perceived this, but also how the routines and perceptions were tied to surrounding factors as culture. This helped me to adjust the questions and foci of my investigation as it went, and compensate for my initial preconceptions. For example, using standardised interviews instead of semi-standardised would have caused big errors in validity since my initial questions were based on incorrect assumptions.

The participatory observations were also an important source of data since they allowed me to discover aspects that I would not have found through interviews. The observations and interviews were used as triangulation in the study. They were mutually confirmatory; things I observed were confirmed in the interviews and vice versa. This gives the results rather high validity.

It is important to acknowledge that the fact that I don't speak Russian was an obvious obstacle in the study. Mastering the language would have made it possible for me to get more information from the participatory observations as well as from the interviews. It is however a bigger problem for the reliability than the validity of the results, they would have become more diverse and detailed if I knew Russian, but there is little reason to question the validity of the results that have been produced.

Questions for Further Research

Since this study was focused on giving a description of subjective experiences of oral exams, it mainly speaks of what teachers and students *think* they do, and not of what actually happens. This information is valuable in itself, but it would also be of interest to compare it with a study of the exam itself. For example, it would be valuable to observe actual exams to see how the teachers and students use the tool, and to find aspects that they themselves are unaware about:

- What kind of questions is asked? Do the teachers manage to move beyond standardised algorithms and definitions as they feel they do?
- What does the communication look like? Is the student really “talking” math or physics when explaining solutions or is it just a description of what is on the paper? Is it true that the quality of the written notes and the oral skills of the student do not affect the grading?
- Do teachers manage to put their personal relations aside when questioning the student?
- What criteria do teachers have, and how well do the students understand them? How well does the grading follow the criteria?

It would also be interesting to conduct a study on gender and hierarchy. The Russian universities are extra interesting from the perspectives of mathematics and natural science, since there is a large number of female teachers as well as students within these subjects at the pedagogical universities. At the same time, the gender structure in society at large is rather rigid. It would be interesting to see what kind of influence this has on the assessment process.

Closing Words

This has been a very interesting thesis to write, and it has given me new insights on assessment. Oral assessment in general and the forms of examinations studied in this thesis provides a useful tool for any teacher. As a Swedish teacher, I will be working in an environment where individualised teaching is expected, and communication has a high priority. In that perspective, oral examination has a natural place, and I will definitely try the method.

References

- Airasian, Peter W. (1994). *Classroom Assessment*. Singapore: McGraw-Hill Book Co.
- Berg, Bruce B. (2004). *Qualitative research methods for the social sciences*. Boston: Pearson.
- Brown, Sally & Glasner, Angela (editors) (1999). *Assessment Matters in Higher Education. Choosing and Using Diverse Approaches*. Suffolk: The Society for Research into Higher Education & Open University Press.
- Brown, Sally & Knight, Peter (1994). *Assessing Learners in Higher Education*. London: Kogan Page Limited.
- Chen, Alexander (1994). Entrance exams to the Mekh-mat. *The Mathematical Intelligencer*, vol 16, no.4, pp 6-10.
- Cheser Jacobs, Lucy & Chase, Clinton I. (1992). *Developing and using tests effectively. A guide for faculty*. San Francisco: Jossey-Bass Publishers.
- Davis, Andrew (1998). *The limits of educational assessment*. Oxford: Blackwell Publishers.
- Gilbert, John (editor) (2004). *The RoutledgeFalmer Reader in Science Education*. London: Routledge Falmer.
- Hamilton, David & Roos, Bertil (2005). Formative assessment: a cybernetic viewpoint. *Assessment in Education*, 1, pp. 7-20.
- Joughin, Gordon (1998). Dimensions of Oral Assessment. *Assessment & Evaluation in Higher Education*, 4, pp. 367-378.
- Kullberg, Birgitta (1996). *Etnografi i klassrummet*. Lund: Studentlitteratur.
- Nyström, Peter & Näsström, Gunilla (coming). *Validering av en modell för bedömning av muntlig redovisning i gymnasieskolans matematik*. Umeå: Institutionen för beteendevetenskapliga mätningar.
- Skolverket (2003). *Nationellt kursprov i matematik, kurs D, våren 2003. Muntligt delprov. Information till lärare, uppgifter och bedömningsanvisningar*. [www document]
URL: <http://www.umu.se/edmeas/np/information/np-info-muntligD.html>
- Vershik, A. K. (1994). Admissions to the Mathematics Faculty in Russia in the 1970s and 1980s. *The Mathematical Intelligencer*, vol 16, no.4, pp 4-5.

Appendix 1. Interview Questions

Background (About Assessment and Testing in General)

Teaching plans (both university and lower stages)

Tell me about the teaching plan in math.

What does it look like and how does it affect your work?

Does it focus on any specific areas?

Do you see it as a guide or a limitation?

What does it say about grading?

What does it say about assessment and assessment methods?

Assessment in compulsory and upper grade school

How do the teachers do to assess the pupils knowledge during the lesson or course, for teaching reasons?

What kind of tests are used as material for grading?

How is oral assessment used? And for what purpose?

Assessment at exams

What types of assessment are used on the university/school? Why?

Are you free to choose type of assessment?

Tell me about the oral state exams.

- Have you participated as a teacher at this kind of exams?
- Who writes the questions?
- How are they done, practically?
- How much time does the student have for preparation before the test?
- How are the questions for each student selected?
- How long time do they have for preparation/ presentation of answers?
- How is the order of appearance of the students decided?
- How is the decision of grading made?
- What feedback is given to the student?
- How is the test documented?

What is the university policy on cheating?

How is cheating dealt with in practical life?

What level of cheating would require that actions are taken, and what would that action be?

Interviews with Teachers

- Tell me about your history in the profession. (How long, what subjects, what levels?)
- What kind of examination do you personally prefer? Why?
- Do you choose different methods of assessment for summative purposes and for assessment of development during the course?

- What pedagogical gains and losses do oral assessment methods have?
- What kind of knowledge and skills are you looking for?
- What kind of questions do you ask and why?
- How is the presentation made? What is your role in it?
- Do you look at the notes of the student? Should they look in a certain way?
- Is there any difference between more and less experienced teachers in how they do the examination?
- How do you make your assessment (in terms of grading)?
- Is it a problem that different teachers grade differently? Are there big differences?
- How do pupils act during an oral test? How do you think it feels for them?
- Do the pupils have any special tricks or methods to pass the test?
- How do you deal with cheating? Is it a problem?
- Do you think that students would study in a different way if they could not cheat?
- Do you think that students cheat more on State Exams? What do you think about that?
- Is there a risk that you help some pupils too much in giving them clues on finding the answer?
- How much is the test affected by the relation between the pupil and the teacher? Is an oral test more sensitive in this way? Is it a benefit or a loss?
- How does it feel when you give grades? Are you nervous? How does it feel to fail a student? Is it worse than on a written exam?
- Could these tests be done in a better way? How?
- Is there something that the student learns from doing a test orally that he would not learn from a written test?
- When you test the students, do you in any way take into account the fact that they study to be teachers? Are good explanation skills rewarded?

Interviews with Students

- What subjects/ on what level do you study?
- Do you consider yourself to be a good student?
- Is education important to you? Why/why not?
- What is your goal with your studies?
- How does the teacher do to find out what pupils know during the lessons?
- What kinds of tests are used?
- What kind of exams do you have?
- How do you feel about oral tests in comparison with written ones?
- What is better or worse compared to written tests?
- Do you feel that you get to show what you can? What kind shows it better?
- Is there something that you learn from doing a test orally that you would not learn from a written test?
- How does it feel to do an oral exam or present something in front of the class?
- Tell me about how the test is done, practically (Exam/test/presentation).

- What is the teacher looking for at the test?
- How do you think that the teachers make the decision of grading?
- Do you feel that the assessment of your math skills have any relation to the fact that you study to be a teacher?
- Is it possible to get help from the teacher in how to find the answer?
- How much is the test affected by the relation between the pupil and the teacher? Is an oral test more sensitive in this way?
- Do you have any special strategies for passing the exam?
- Does it happen that pupils/students cheat? In what ways? How do the teachers react to it?
- How would you deal with cheating if you were the teacher? Do you consider cheating to be a problem?
- Do you think that people would study in a different way if cheating was impossible?
- Could these tests be done in a better way? How?

Appendix 2. Information to the Interviewees

My name is Eva Andreasson and I am a student at Umea University in Sweden. I study to be a teacher in math and psychology and will teach in grades 7 to 12. The reason for my trip to [name of the city] is to do my practice and write my diploma work.

The preliminary title of my diploma work is "*Pros and cons of oral assessment - A qualitative study in Russian teachers' and pupils' experiences of oral assessment methods*". The goal is to understand more about different kinds of testing and assessment methods, with special focus on oral examination. This is an unusual method in Sweden, and I would like to learn from the rich experience that Russian teachers and students have from it.

This interview will be a part of my research material for my diploma work. The interview is totally voluntary, and you can choose to end it at any time, without any explanation. The interview is confidential. All material from the interview will be coded to make it impossible to identify you and I will destroy all audio recordings and notes when my work is finished.

If you have any questions about my investigation, please contact me.

Thank you for participating!