A classification of questions from Irish and Turkish high-stakes examinations

Tugba Aysel[a], Ann O'Shea[a], Sinead Breen[b]

[a]*Department of Mathematics and Statistics, NUI Maynooth, Kildare, Ireland*
[b]*Department of Mathematics, St. Patrick's College, Drumcondra, Ireland*

In both Turkey and Ireland entrance to third level education is determined by performance on a high-stakes examination at the end of second level education. However, the examination systems in Ireland and Turkey are quite different from each other. In order to compare the examinations we attempted to classify the types of questions asked in 2009 and 2010. We used various classification systems including the Levels of Cognitive Demand Framework developed by the QUASAR Project (Smith & Stein 1998). We will report on the use of these frameworks and the results obtained for the Turkish and Irish mathematics examinations.

Introduction

This classification of examination questions is part of a larger project that studies the effects of examinations on the teaching and learning of mathematics at post-primary level in Ireland and Turkey. We administered questionnaires to Irish and Turkish students and teachers and interviewed teachers to explore the impact of the examinations on study habits, teaching methods, and attitudes to mathematics. In order to compare the examination systems in the two countries we decided to classify the mathematics examination questions in both countries.

Several studies have been done on the classification of examination questions. A group of researchers in Australia modified Bloom's Taxonomy (1956) to classify undergraduate mathematics examination questions in order to show how an examination should be constructed to assess a broad range of mathematical skills (Smith et al. 2007). Their modification used 3 classification groups as follows: Group A-factual knowledge, comprehension and routine use of procedures; Group B-information transfer and application in a new situation; Group C-justifying and interpreting, implications, conjectures and comparisons and evaluation. Schoenfeld (1992) details a Balanced Assessment Framework which was created to help examiners set a range of questions which cover different skills. There are 7 critical dimensions in this framework: content, thinking processes, student products, mathematical point of view, diversity, circumstances of performance and pedagogic-aesthetics.

Azar (2005) compared Turkish university entrance (OSS) physics examination questions and physics examination questions asked at schools with respect to Bloom's Taxonomy. His classification showed that OSS physics questions assessed the application, analysis, synthesis and evaluation skills described in Bloom's Taxonomy. The physics teachers used questions at the application and comprehension levels in school assessments to determine students' achievements. Close and Oldham (2005) mapped the mathematics questions from the 2003 Irish Junior Certificate (JC) examination onto the 3 dimensional PISA Mathematics Framework. From the model answers provided by the State Examinations Commission, they identified the skills involved and compared them to the 3 competency classes of PISA. These competency classes are *reproduction* (performing calculations, solving equations, reproducing memorized facts or “solving” well-known routine problems), *connections* (integrating information, making connections within and across mathematical domains, or solving problems using familiar procedures in contexts) and *reflection* (recognizing and extracting the mathematics in problem situations, using that mathematics to solve problems, analyzing and developing models and strategies, or making mathematical arguments and generalizations) (Close and Oldham 2005, 187). Their results showed that most of the 2003 JC questions belonged to the reproduction class and there were no 2003 JC questions that belonged to the reflection class.

Turkish Education System

After the first year of secondary school students are streamed into 4 groups: science, social, language and Turkish-mathematics. Students in the science and Turkish-mathematics groups can take the mathematics paper in the university entrance and placement examination (OSS). The examination system has two steps. The first examination (OSS1) takes place in April and second one (OSS2) takes place in June. Students need to reach a critical mark in the first examination to get a chance to take the second examination. The second examination determines entry to third level. All examination questions are multiple choice questions and they all have the same marks. In 2010 662,894 students graduated from secondary schools. The number of applicants to the OSS in 2010 was 1,587,993 and the percentage of those placed in universities was 55.06%. Up to 2010 the second examination took place on one day and lasted 3.5 hours. The examination system changed in 2010 and now takes place over 4 days. It comprises of mathematics, science, language and Turkish-social papers. The students sit one paper in one day. The number of questions on the mathematics papers has also increased.

Irish education system

Irish students spend 8 years at primary school, followed by 5 or 6 years at post-primary/second level. Students take an examination called the Junior Certificate after 3 years and an examination called the Leaving Certificate (LC) at the end of their time in post-primary school. These are state examinations held during the month of June. The results of the Leaving Certificate examination determine entry into third level education. Students typically take 7 subjects for Leaving Certificate and because of third level matriculation requirements most students study English, Irish and mathematics. In fact 96% of students who take the Leaving Certificate study mathematics. This means that approximately 82% of Irish students study mathematics until the age of (at least) 17. There are three different levels of mathematics at Leaving Certificate – they are Foundation Level (FL), Ordinary Level (OL) and Higher Level (HL). There is some fluctuation in the numbers taking the various levels from year to year but typically 11% of students take FL, 71% take OL and 18% take HL. Students who study mathematics at FL are not generally eligible for entry to third level. Leaving Certificate examination questions are partial credit questions and the questions may carry different marks.

Methodology

Initially we classified the mathematics questions on the Irish and Turkish examination papers according to Smith et al. (2007) and to Schoenfeld’s Framework for Balanced

Assessment (Schoenfeld 1992). However, we ended up with most mathematics examination questions falling into the same categories – in the case of Smith's classification categories, it was 'routine use of procedures' (Smith et al. 2007). Recall that Schoenfeld's Framework for Balanced Assessment has 7 dimensions, which are further subdivided into categories (Schoenfeld 1992). In each of the dimensions the questions usually fell into just one or two categories. For example in 'students' product' dimension all of the examination questions fell under 'exhibition of techniques' or 'problem solutions'. So these classification systems failed to discriminate between questions both within the Irish mathematics examinations and the Turkish mathematics examinations, and between the two countries' examinations. We decided to use a different classification method called the Levels of Cognitive Demand Framework (LCD) developed by the QUASAR Project (Stein & Smith 1998). The reason for using this classification method (LCD) was that almost all questions under inspection involve using procedures and this system distinguishes different levels of procedural questions. We classified all questions on the 2009 and 2010 Irish Leaving Certificate Higher and Ordinary Level Mathematics papers and the Turkish OSS Mathematics examination papers. First, each of the three authors classified the questions independently and then we resolved our disagreements through negotiation.

The levels of cognitive demands framework

Here we reproduce the description given by Stein and Smith (1998, 349) of the four levels of cognitive demand used in the QUASAR Project (the emphasis is ours):

Lower-level demands (memorization) (LM)

The questions involve either *reproducing* previously learned facts, rules, formulas, or definitions or committing facts, rules, formulas or definitions to *memory*. They cannot be solved using procedures because a procedure does not exist or because the time frame in which the task is being completed is too short to use a procedure. They are not ambiguous. Such tasks involve the exact reproduction of *previously seen material*, and what is to be reproduced is clearly and directly stated. They have *no connection to the concepts or meaning* that underlies the facts, rules, formulas, or definitions being learned or reproduced.

Lower-level demands (procedures without connections to meaning) (LP)

The questions are *algorithmic*. Use of the procedure either is specifically called for or is *evident* from prior instruction, experience, or placement of the task. They require *limited cognitive demand* for successful completion. *Little ambiguity* exists about what needs to be done and how to do it. They have *no connection to the concepts or meaning* that underlies the procedure being used. They are focused on producing correct answers instead of on developing mathematical understanding. They require no explanations or explanations that focus solely on describing the procedure that was used.

Higher-level demands (procedures with connections to meaning) (HP)

The questions focus students' attention on the use of procedures for the purpose of developing deeper levels of understanding of mathematical concepts and ideas. They suggest explicitly or implicitly pathways to follow that are *broad general procedures* that have close connections to underlying *conceptual ideas* as opposed to narrow algorithms that are opaque with respect to underlying concepts. They usually are *represented in multiple ways*, such as visual diagrams, manipulatives, symbols, and problem situations. Making connections among multiple representations helps develop meaning. They require *some degree of cognitive effort*. Although general procedures may be followed, they *cannot be followed mindlessly*. Students need to engage with conceptual ideas that underlie the procedures to complete the task successfully.

Higher-level demands (doing mathematics) (HD)

They require *complex and nonalgorithmic thinking* – a predictable, well-rehearsed approach or pathway is not explicitly suggested by the task, task instructions, or a worked-out example. They require students to explore and *understand the nature of mathematical concepts*, processes, or relationships. They demand self-monitoring or self-regulation of one's own cognitive processes. They require students to access relevant knowledge and experiences and make appropriate use of them in working through the task. They require students to analyze the task and actively examine task constraints that may limit possible solution strategies and solutions. They *require considerable cognitive effort* and may involve some level of anxiety for the student because of the unpredictable nature of the solution process required.

Creation of an intermediate level

While classifying questions using the LCD framework, we found some problems that we felt fell between Lower Level Demands (Procedures without connections to meaning) and Higher Level Demands (Procedures with connections to meaning). So we decided to create a new category called Intermediate Level Demands (Procedures). The description of this category follows:

Intermediate-level demands (procedures) (IP)

The questions are *algorithmic*. Use of more than one procedure may be evident from previously learned information. Algorithms with *multiple steps* may need to be used. Although there is a well-defined procedure to be used, students may need to make an educated choice of starting point. Also students may have to make *some connections from different areas* of mathematics to the underlying concepts. The questions require *moderate cognitive effort*. A complicated but routine calculation is involved in the questions.

Examples

Here we give some examples of questions in each category. The questions are taken from the Irish State Examinations Commission (2011) and Turkish Student Selection and Placement Examination Higher Education Council (2011) websites.

Lower-level Demands (Memorization) 2009 HL

Prove that the measure of one of the angles between two lines with slopes m_1 and m_2 is given by $\tan\phi = \frac{m_1 - m_2}{1 + m_1 m_2}$. (This is one of a very small number of proofs that students are expected to be able to reproduce.)

Lower-level Demands (Procedures without connection to meaning) 2009 OL

Let f(x) = $x^3 + x^2 - 4x - 4$. Verify that f (-2) = 0.

Intermediate-level Demands (Procedures) 2009 OSS2

Suppose α and β are the roots of $x^2 - 2x - 4 = 0$. Which of the following equations has $1/\alpha$ and $1/\beta$ as roots?

A) $2x^2 - x + 4 = 0$ B) $2x^2 + x + 1 = 0$ C) $4x^2 + 2x - 1 = 0$ D) $4x^2 + 3x - 4 = 0$
E) $8x^2 - 3x + 4 = 0$

Higher-level Demands (Procedures with connection to meaning) 2010 OSS2

How many units is the area between the $y = x^3$ curve and the $y = x$ line?
A) 1/2 B) 3/2 C) 1 D) 1/3 E) 2/3

Higher-level Demands (Doing mathematics) 2010 HL

Let f be an affine transformation. The point M is the mid-point of the line segment $[AB]$. Show that f (M) is the mid-point of the line segment $[f(A)f(B)]$.

Results

Tables 1 and 2 below show the percentages of examination questions that were classified as belonging to the five categories in our scheme.

	LM	LP	IP	HP	HD
2009 OSS1	-	40%	20%	40%	
2009 OSS2	-	23.33%	26.67%	46.67%	3.33%
2010 OSS1	-	61.54%	20.51%	10.26%	7.69%
2010 OSS2	2.5%	32.5%	27.5%	31.25%	6.25%

Table1: Classification of 2009-2010 Turkish university selection and placement examination questions

	LM	LP	IP	HP	HD
2009 HL	7.45%	39.36%	24.47%	23.40%	5.32%
2010 HL	-	50.51%	26.80%	19.59%	3.10%
2009 OL	1.55%	79.07%	9.30%	10.08%	-
2010 OL	1.75%	84.21%	5.26%	8.77%	-

Table2: Classification of 2009-2010 Irish terminal/end-of school examination questions

	LM	LP	IP	HP	HD
2009 HL	8%	34%	27%	26%	5%
2010 HL	-	42%	28%	27%	3%
2009 OL	4%	74%	10%	12%	-
2010 OL	4%	80%	8%	8%	-

Table 3: Percentages of total marks for each level of Irish terminal/end-of school examination questions

All questions in the OSS examination carry equal marks but this is not true in the Irish examination. Table 3 gives the percentages of marks in each category.

Conclusion

From Tables 1 and 2, we can see that the OSS2 examinations have more high level cognitive demand (HP-HD) questions than the Irish examinations. The OSS1 (2010) has a lot of LP questions. However, the OSS1 examination plays only a limited role in determining entry to third level. In Ireland, the OL examinations have a very large proportion of LP questions with about 80% of marks being awarded for questions with lower-level demands (Table 3). The HL examinations have less LP questions but they also do not have many HP questions and about 30% of the marks are awarded for questions involving higher-level demands. The OSS2 and HL examinations have a similar percentage of IP questions. According to these results, we can say that most OL examination questions have limited cognitive demand and little connection to concepts that underlie the procedure being used. The OSS2 questions use broad general procedures, are represented in multiple ways and have some degree of cognitive demand. The HL examination requires moderate cognitive effort and some connections from different areas of mathematics.

Azar's (2005) results concerning physics questions in the OSS examinations have similarities with our analysis of OSS results. We too found that questions with high levels of cognitive demand were asked in the OSS examinations. Also we have seen similarities with Close and Oldham's (2005) work, who found that most Irish Junior Certificate examination questions do not require higher order thinking skills.

Most questions on all of these examinations inspected involve the use of procedures as can be seen from the small numbers of HD tasks. For this reason the LCD classification method seems to be a good method for our situation.

References

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, 68-74

Bloom, B. S., M. D. Engelhart, E. J. Frust, W. H. Hill, and D. R. Krathwohl 1956. Taxonomy of Educational Objectives: Cognitive Domain (New York: Mckay).

Close, S., and Oldham, E. 2005. Junior cycle mathematics examinations and the PISA mathematics framework. In Proceedings of the First Conference on Research in Mathematics Education, ed. S. Close, T. Dooley and D. Corcoran, 185-204. Dublin: St Patrick's College.

Schoenfeld, A. H. 1992. A Framework for Balance: Balanced Assessment Project, Unpublished document, prepared for NSF-sponsored Calculus assessment workshop.

Smith, G., L. Wood, M. Coupland and B. Stephenson. 1996. Constructing mathematical examinations to assess a range of knowledge and skills. International Journal Mathematics Education Science and Technology, 27, 65-77.

State Examinations Commission. 2011. Accessed April 20, 2011. http://www.examinations.ie

Stein, M. K. and M. S. Smith. 1998. Selecting and creating mathematical tasks: From research to practice. Mathematics Teaching in the Middle School, 3(5), 344-350.

Turkish Student Selection and Placement Examination Higher Education Council. 2011. Accessed April 20, 2011. http://www.osym.gov.tr