

English summary
of Fafo-rapport 2020:01

On equal terms?
Evaluation of a 10th year
mathematics exam 2017–
2019. Final report

On equal terms?

Evaluation of a 10th year mathematics exam 2017–2019. Final report

The theme of the report is the 10th year mathematics examination in the period 2017–2019. This is the last of three reports, each of which has looked at the relevant year's exam. In addition, we have given priority to highlighting a particular theme each year. In the first report (Andresen et al. 2017), the focus was on the impact of the language and concepts used in the exam questions, while in last year's report (Bjørnset et al. 2018), we analysed the significance of the exam's strong emphasis on the candidates' digital skills. The focus of this year's report is a comprehensive assessment of the mathematics exam held in the period in question. In addition, we place a particular emphasis on the poorest performing pupils, i.e. those who attained an average grade of '1' or '2' for the year, pupils who scored the lowest 10–30 per cent of points in the exam, as well as minority language pupils and pupils with poor literacy skills.

Common to all reports is the question of whether the exam is fair and whether it is perceived to be fair, and as such enables all candidates to be assessed on equal terms. The project seeks to answer nine questions.

1. Is there a close correlation between the exam syllabus and what is actually taught?
2. Are the assessments consistent across examiners?
3. Does the exam include questions of varying degrees of difficulty that can measure all levels of competence?
4. What do the pupils think of the amount of work required in the exam in relation to the time available to complete the exam and the time they can spend on parts 1 and part 2 respectively?
5. Is the design of the exam paper suitable for assessing the pupils' mathematics skills?
6. Assessment of how the exam and the exam marking have developed in the three-year period.
7. Is the exam paper comprehensible in terms of text and illustrations?
8. What kind of teaching did the pupils receive in using digital aids, and how were they prepared for the use of digital aids in the exam?
9. How well does the exam work for the poorest performing pupils?

The data collection for the survey was complex. We sent an electronic questionnaire to mathematics teachers whose 10th year pupils were taking the exam. In

English summary of Fafo-rapport 2020:01

On equal terms?

Evaluation of a 10th year mathematics exam 2017–2019. Final report
Mathilde Bjørnset, Aina Fossum, Jon Rogstad og Bjørn Smestad

© Fafo 2019

addition, we conducted qualitative interviews with teachers and pupils at four schools, and undertook classroom observations. We also added questions to the Directorate of Education's questionnaire for the examiners and analysed the exam answers.

Below are some key conclusions from this year's report and from the project.

A god-quality and fair exam

The main conclusion of the report is that the exam was, on the whole, fair and of a high quality in all three years. This implies that the various questions in the exams were closely correlated to the teaching that the pupils had received and the expressed competence objectives. Mathematics exams have a high level of legitimacy among pupils and teachers. More specifically, almost all parts of the syllabus were tested in the years we studied. This conclusion also matches the teachers' own assessments, while an increasing proportion of the examiners believe that some parts of the syllabus are never tested in the exam.

Pupils and teachers further find that the content of the exam is mostly in line with what is actually taught. We have observed that the exam is closely correlated to the content of the pupils' textbooks, and that there are no systematic differences between the various learning materials. When we looked at digital tools specifically, however, we found a variation in the teaching that pupils had received. In particular, few had received training in CAS¹. There is therefore a systematic difference in the candidates' opportunities for making use of digital aids in the exam. The clear conclusion is that these kinds of differences challenge the ideal that candidates should be able to answer exam questions on equal terms.

The dominant impression that the exam has maintained a high quality is also reinforced by the analyses of degree of difficulty and amount of work entailed in the exam based on the assessment forms. The degree of difficulty was, for the most part, sufficiently varied to enable pupils at all levels of achievement to demonstrate their competence – with the exception of the poorest performing pupils, whom we will return to. Our analyses also show that relatively few pupils felt they did not have enough time to complete the exam, and there are few indications that the pupils are systematically failing to complete the final questions in the exam.

With regard to language and the use of illustrations, the main concern is that the vast majority of questions entail linguistic challenges. Although reading is a basic skill, including in mathematics, it is not necessary to test this skill in almost all of the exam questions. We have analysed a range of linguistic features that we

1 CAS – Computer Algebra System. For example, the popular program GeoGebra contains a CAS section.

English summary of Fafo-rapport 2020:01

On equal terms?

Evaluation of a 10th year mathematics exam 2017–2019. Final report
Mathilde Bjørnset, Aina Fossum, Jon Rogstad og Bjørn Smestad

© Fafo 2019

know can make questions more difficult for pupils with poor reading skills, and recommend that further steps are taken to reduce the incidence of these.

Knowledge of some of the concepts is linked to factors that are not related to mathematical competence, but which will largely vary depending on whether pupils are born in or outside Norway and their socio-cultural and socio-economic background. This may in turn impact on the fairness of the exam. The deciding factor in assessing fairness is whether the candidates can perform on equal terms and whether the assessments made are equitable. Given that we have concluded that the exam has consistently been of a good quality during the project period, we have also mainly concluded that it is fair. However, the linguistic complexity of some questions raises the question of whether some pupil groups are unfairly disadvantaged.

For the exam to be fair, Bokmål and Nynorsk pupils also need to be given questions that entail the same linguistic challenges. Each year, there have been weaknesses in the translations that have led to concerns being raised at the individual question level, but for the mathematics exams as a whole, there is no systematic bias between Bokmål and Nynorsk pupils.

Examiners' assessments

A growing share of examiners are reporting that they do not encounter challenges when ensuring fair marking. The guidance documents are perceived by the examiners to be better this year than three years ago, but there is still a desire for the advance marking report to be issued at an earlier stage. One element in this context is also that digital submissions have increased during the project period.

Several examiners highlighted how the poorest performing pupils are given little opportunity to show their overall competence, since many of the questions are multiple choice questions and questions that only require a simple answer. Furthermore, our analyses of assessment forms from the 2017 exam showed that there was some significant variation in examiners' assessment of the questions that required the use of digital aids, the questions in which the pupils themselves choose an appropriate method, and the questions that place higher demands on communication and reasoned answers.

The poorest performing pupils

In this year's report, we have looked specifically at pupils who performed poorly in the exam. The question was whether they had the same opportunity to demonstrate their skills as the other pupils. The tenth of pupils with the poorest performance consistently failed to provide the correct answers to many of the ques-

English summary of Fafo-rapport 2020:01

On equal terms?

Evaluation of a 10th year mathematics exam 2017–2019. Final report
Mathilde Bjørnset, Aina Fossum, Jon Rogstad og Bjørn Smestad

© Fafo 2019

tions in Part 2 of the exam. The questions they managed were mostly multiple choice, and of a level corresponding to the competence objectives for 4th and 7th year in primary school. Section 3-3 of the assessment regulations (Regulations to the Education Act, 2006) stipulates that exams must be marked on the basis of the competence objectives. Our analyses indicate that the pupils seldom meet the criteria for a grade '2' when the outcome achievement standards are applied. However, the statistics show that many of these pupils attain the points needed to secure a grade '2' according to the thresholds set, and therefore end up with a grade '2'.

The teachers' responses indicated that the questions in this year's exam were not considered very difficult. On the contrary, the general view seemed to be that the exam could have been somewhat more difficult.

Developments in the period 2017–2019

Based on the data we collected, we will highlight some trends during this three-year period. First, our informants generally found that the exam in the period was of a fairly high quality. In this context, it is relevant to point out that the design of the questions and, in some cases, the actual formulation of the questions, can be similar from year to year. However, this is not necessarily consistent. Several informants referred to the exam in 2015 as an example of an exam that did not work very well.

There have been significant changes in two areas during the period: the weighting between Part 1 and Part 2 has changed considerably, and the proportion of multiple-choice questions has increased dramatically. With regard to the overall degree of difficulty in the exam questions, the IRT analyses in Bjørnsson (2020) show that the pupils' competence changed 'very little' in the three-year period 2017–2019 (*ibid*, p. 21) and that there were 'very small' changes in the degree of difficulty of the questions (*ibid*, p. 15). Nevertheless, the average grade has increased from 3.4 to 3.6. When we remember that an increase of one tenth means that every tenth pupil has moved up one grade, an increase of two tenths must be characterised as a significant increase. However, the IRT analyses do not take into account the weighting of the questions. This implies that the improved exam results are mainly due to the change in the weighting between the different questions. When Part 1 is given a higher weighting and the more difficult Part 2 is given a lower weighting, and the grading thresholds remain the same, this has a positive impact on the grades. Attaining a grade '2' in the exam has therefore become considerably easier.

Throughout the three-year period, the correlation between the exam syllabus and what is actually taught has remained strong, the amount of work entailed in

English summary of Fafo-rapport 2020:01

On equal terms?

Evaluation of a 10th year mathematics exam 2017–2019. Final report
Mathilde Bjørnset, Aina Fossum, Jon Rogstad og Bjørn Smestad

© Fafo 2019

the exam has changed little and the examiners are satisfied with the improvements in the guidance for examiners.

Within the parameters set, the exam questions for 2017–2019 have mainly been of a high quality and fair, viewed in light of the framework set by the Curriculum for Knowledge Promotion (LK06). The curricular reform has a new focus that will also require changes to the form of exams, but many elements from the current form of exam and from our reports will also be relevant in the design of future exams.

English summary of Fafo-rapport 2020:01

On equal terms?

Evaluation of a 10th year mathematics exam 2017–2019. Final report
Mathilde Bjørnset, Aina Fossum, Jon Rogstad og Bjørn Smestad

© Fafo 2019