How does Umalusi standardise the final National Senior Certificate examination results?

Introduction

The standardisation process conducted by Umalusi is complex. It comprises more than just the adjustments process which is controversial as it is perceived to be 'fiddling' the marks.

The standardisation process includes

- the external moderation of examination papers;
- a consideration by independent experts as to whether examination papers from one year to the next are comparable in demand;
- the review of results against the historical performance of candidates in each subject. It is at this stage that consideration is given to adjustments of marks.

The general public's perception is that any adjustments to raw results is 'tampering' with marks and unjustified, while leaving raw marks as is, is the ideal. While understandable, this is a rather naïve view. This short explanation is intended to make you aware of some of the complexities of the process and equip you with the necessary understanding to explain the process to others, if the need arises.

Rationale for mark adjustment

Society expects from the education system that results indicate a consistent level of learner competence from year to year e.g. a learner scoring an A this year is comparable in respect of competence within a subject to a learner who scored an A last year and will be comparable in respect of subject competence to a learner scoring an A in next year's examination. By the same token, there is an expectation that a learner who fails this year would have also failed if they had written last year's examination and would fail if they write next year's examination (assuming no additional learning interventions).

When universities select learners for courses, they need to know that learners who enter their institution with a particular set of results will have comparable competence and ability to learners with that same set of results achieved in last year's examination and next year's examination. This is so that they have a consistent idea about the level at which they should pitch their first year courses. It also ensures some consistency in selection, so that a learner from this year's cohort who applies for entrance to an institution can be compared fairly to a learner from last year's cohort, who might have taken a 'gap' year, and hence is only applying for university selection this year.

To adjust or not to adjust?

Consider the following example:

> An examiner sets a question which he, the internal moderator as well as the external moderator believe will be of a particular difficulty level. When writing the paper, however, learners do not 'read' the question in the same way as this panel of 'examiners' e.g. they do not see the difficulty intended by the panel and hence perform much better than

intended by the panel; alternately, they find the question to be much more difficult than the panel intended and hence their overall results are far worse than the panel intended.

It is not fair for the learners in this specific examination to be unduly advantaged i.e. they had an easier paper than the learners in the previous year, or disadvantaged i.e. they had a more difficult paper than the learners in the previous year. It would be unfair if learners from one year benefitted or were disadvantaged by an error of judgement on the part of members of the examining and moderation panel.

However, one cannot expect that in all cases where adjustments are made on the basis of the history of performance, the examining panel had made an error in judgement. Consider the following examples:

For whatever reason, a decision is taken to be stricter with Grade 11 pass rates. The effect of this would be that the learners entering Grade 12 are of a better quality (in terms of competence) than in previous years. One would then naturally expect that the Grade 12 results of that cohort to be better than results of previous cohorts – the weaker learners did not sit for the Grade 12 examination, as they were prevented from entering Grade 12. It would not be fair to expect the same failure rate from the new cohort of learners because they are in fact a better group of learners.

or

A decision is taken to conduct a meaningful intervention with teachers to upgrade their ability to teach a specific concept in a subject discipline. There is an expectation that if the intervention was successful, there would be an improved performance by learners who have benefitted from the intervention with the teachers.

Hence the standardisation process is intended to:

- limit the 'positive' or 'negative' impact of situations outside the control of the learners e.g. an error in the paper, a mis-reading of the standard required, a change in the quality of the learning or teaching cohort;
- at the same time, mindful of societal expectations of the competence levels of learners scoring specific symbols.

There is a perception that adjustments are wrong (whether they be upwards or downwards) and keeping raw marks is a good thing.

Given the examples above, consider the following:

The historical performance of learners in a subject are as follows:

- Mean: 50%
- % scoring a rating of 7 (80%+): 10%
- % fail i.e. scoring 1: 10%

The examiner makes a serious mistake in the paper and the results for the year are as follows:

- Mean: 40%
- % scoring a rating of 7: 2%
- % fail: 20%

It would not be fair for these students to keep the raw marks as clearly there are learners who would have scored As had they written in a different year, but now score Bs because of the error. Similarly there are learners who fail who would have passed, if not for the error. In this instance, an upward adjustment is appropriate. If it was not done, this cohort of learners would have been unfairly disadvantaged in comparison with learners from previous years.

Looking at another example,

The historical performance of learners in a subject are as follows:

- Mean: 50%
- % scoring a rating of 7 (80%+): 10%
- % fail i.e. scoring 1: 10%

The examiner makes a mistake in determining the overall standard of the paper and the results for the year are as follows:

- Mean: 65%
- % scoring a rating of 7: 25%
- % fail: 2%

It would not be fair for these students to keep the raw marks as clearly there are now learners who have scored As who would have scored Bs if they had written the examination last year or next year because this paper is too easy, and there are learners who failed last year and others who probably will fail next year who would have passed if they had written this year's paper. In this instance, a downward adjustment is appropriate.

If, however, there is a compelling reason to believe that this cohort of learners should have done better e.g. a change in the capacity of the learners or the teachers, there may not need to be as big an adjustment as the results suggest. If for example, entry from Grade 11 to Grade 12 was closely controlled and a much better group of learners progressed to Grade 12, then one would need to consider whether this strategy would have affected the pass rate as dramatically as this and as a result, the downward adjustment might not be as great.

This example illustrates how leaving raw marks could also be unfair. If the situations described in this last example were to arise and no compelling reason could be found to explain such a dramatic change in performance, it would be unacceptable to leave these raw results without adjustment. While one might believe intuitively that marks should not be adjusted, the learners in this cohort would definitely be unfairly advantaged, if the raw marks were left as is.

Conclusion

These rather simplistic examples serve to illustrate that standardisation decisions are complex and require consideration of a wide variety of pieces of information. Adjusting or leaving results as per the raw results, are not in themselves good or bad decisions. They should be seen in context and that context may not be as simple as the general public may assume.

This process is sometimes called 'norming'. It should be noted that all mainstream examining authorities that do not make use

of pre-standardised tests, apply 'norming' in some form. In some UK examining boards, the chief examiner together with officials from the examining body and the regulatory authority determine the cut-off scores at key points e.g. a distinction, the pass/fail cut-off. While in these systems, marks are not adjusted, the principle that a distinction is not 'fixed' at a specific mark (80%) from year to year, does apply. Marks in South Africa are adjusted because historically South Africa made use of an aggregate and hence, the expectation of comparability not only of standard or demand, but also marks themselves.

This process means that marks reported to learners and higher education for a subject in different years can be compared and are fair, valid and reliable. Furthermore it maintains and establishes standards – what is known as a 'nation-wide' assessment currency.

Moderation of School-based Assessment

Introduction

School-based assessment is part of the subject assessment requirements in the NSC certification process. While the Independent Examinations Board acknowledges the immense potential of school-based assessment (SBA) in terms of its positive benefits in holistic assessment, the IEB has to guard against difficulties that may arise in terms of reliability, fairness and validity. The methods that guard against these difficulties are explained below and ensure that the standard of the NSC certification is not compromised by unreliable SBA.

Method 1: Social Moderation

Social moderation is used to validate teacher judgments in SBA within the 'guild of professionals' i.e. teachers moderate each other's work and in so doing, establish a common understanding of the standard required. The various stages, at which a common understanding of the SBA requirements stipulated in the Subject Assessment Guidelines of each subject can be arrived at, include:

- Departmental meetings at the school level;
- Cluster meetings;
- Regional moderation;
- National moderation.

SBA is a compulsory component of the final examination result. Should there be non-compliance with requirements on the part of the learner, the learner will forfeit the marks for the tasks not completed. In order to write the final examination, learners must have an SBA result. If the learner does not have an SBA mark, or if the result is zero because of non-compliance, the learner will not receive a final NSC result in that subject, even if s/he writes and passes the final examination. Should it be found that there has been non-compliance with requirements on the part of the teacher, the SBA marks may be disregarded and students are compensated accordingly.

Method 2: Socio-Statistical Moderation

The IEB Assessment Specialists in consultation with management review various statistics which describe the relationship between the SBA and the final examination marks at a centre. These include the mean score, the standard deviation and the correlation co-efficient. The mean score gives an indication as to whether the school's overall standards are comparable to that of the examination; the standard deviation provides an indication as to whether the spread of marks in the SBA is comparable to the spread of marks in the examination and finally the correlation co-efficient provides an indication as to whether learners have been ranked appropriately. These statistics together with other qualitative data inform IEB decisions for learner file selection for regional and/or national moderation. It also informs schools of potential problem areas in respect of an understanding of the appropriate standard required in each subject offered at the school.

Method 3: Statistical Moderation (Computer-Based Adjustment of SBA Marks)

This process is carried out by the IEB under instruction from Umalusi and is carried out regardless of the actions of Methods 1 and 2 above. Methods 1 and 2 are aimed at developing the capacity of the teacher to conduct SBA as required and at the appropriate standard. This final process is aimed at ensuring that the SBA of learners is in line with the performance of the centre in their examination.

Hence all SBA is statistically moderated per subject, per school in line with the requirements and specifications for the process, as outlined in Umalusi directives. The IEB adheres to all Umalusi policy requirements in this regard.

The adjustment process in essence aligns the school's SBA results with the school's performance in the final examination. Hence individual learners benefit from working hard during the year and getting good results while other learners who do not do as well during the year, are similarly affected. The key statistics in this process are:

- The mean of the SBA in comparison with the mean of the group in the examination;
- The standard deviation of the SBA in comparison with the standard deviation of the group in the examination.

The linear transformation used in the statistical adjustments does not change the examination rank order of students or the relative differences between them. Small groups of candidates (less than 8) which can be affected significantly by the standard deviation are standardised using a different formula from larger groups.

All methods used to validate the SBA are intended to protect both the learner and the integrity of the assessment and resulting process.