# Moderators' Report
## Principal Moderator Feedback

## Summer 2017

Pearson Edexcel GCSE
In Statistics
5ST02 (Controlled Assessment)

**Edexcel and BTEC Qualifications**

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

**Pearson: helping people progress, everywhere**

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2017
Publications Code 5ST02_01_1706_ER

**GCSE Statistics 2ST01**
**Principal Moderator Feedback – Controlled Assessment**

## General comments

The controlled assessment themes available this year, in order of popularity, were Employment, Weight and Music; Weight offering the greater opportunity for the collection of primary data.

In general the work from centres was well presented, neatly packaged, and arrived by the due date. Unfortunately some centres did not adhere to all the procedures surrounding the submission of samples. This year a significant number of centres did not ensure that all the authentication statements on the Student Record Forms had been signed by both the teachers and the students.

Centres are reminded that students are allowed to change their initial plans after teacher feedback. All changes to the initial plan must be agreed by the teacher and recorded on the Student Record Form (the initial mark for planning cannot be changed). Work not identified in the plan cannot be given credit in the assessment.

Some centres continue to submit work that has not been thoroughly annotated and checked for errors, eg crosses for calculation errors and crosses for omissions in diagrams. Unidentified errors in student work significantly increase the burden on moderators who effectively have to mark it.

Centres are reminded that internal moderation must take place if there is more than one teacher marking the work. This should be made clear for the moderator.

## Specific comments

## Strand 1 (Planning)

Many students produced investigations that they thought involved the use of interrelated variables but were actually a collection of mini investigations only loosely related by the theme. Often the choice of these mini investigations was to demonstrate the use of as many techniques as possible, rather than as a selection of the right techniques for a carefully specified problem.

Investigations attracting the highest marks were often those involving a degree of complexity. Complexity includes the analysis of interrelated variables and the sequencing of techniques.

Often students did not justify the choice of techniques at the notional difficulty of those techniques; more demanding techniques require a greater degree of justification than less demanding techniques. Simply stating that a technique will be used, or explaining how it will be used, is insufficient to access the higher marks. For example, students using scatter graphs to investigate the relationship between variables should explain which of the variables they consider to be dependent and which to be independent, and to justify their choice. When drawing histograms, students should justify their choice of class intervals and discuss their choice of sample size in relation to this. Furthermore, students should relate the choice of techniques to the type of variables being investigated, eg histograms for continuous data, box plots for discrete data.

Many students were able to anticipate possible problems in the collection of their data and were able to give a clear strategy for dealing with outliers and anomalies. Students should be encouraged to explain their reasons for removing poor data in the context of their investigation and to comment specifically on the possible impact the poor data could have on the reliability of their calculations and results. Simply stating that any outliers will be removed or replaced is a low demand activity, even if accompanied by sophisticated techniques for identifying them.

**Strand 2a (Data collection)**

Some centres continue to provide data sets for students. Centres are reminded that students are expected to collect their own data sets and not to subsample from data sets that have been provided for them. Students may collect data in groups if they wish, but each student must take part in the data collection, and each student must explain their role in the data collection.

Many students collected sufficient data for their investigations, identified the source of the data and commented on the possible presence of outliers.

Often the data collection was poorly described and ill defined. This was particularly the case when data sets had been provided for them. The vast majority of students had little or no idea of how it had been collected.

Students should be advised to give more detail when describing their sampling techniques, eg when selecting a random sample they should explain the unique numbering of the data items and the specific technique used to generate the random numbers, eg 30 × Ran#.

Students should be advised to think more carefully about the amount of data they are using in their analyses. The sample size should be justified in terms of the techniques employed. For example, using 15 items of data may be acceptable for some techniques, eg a box plot, but may not be appropriate for other techniques, eg a histogram, where more data is expected.

Many students are able to identify outliers formally and represent them in diagrams. But it should be noted that it is not only the calculation of the outliers that drives the assessment but also the purpose of the calculation. For example, the formal identification of outliers merely to represent them in a box plot is a low demand application of the technique, whereas the formal identification of outliers as a prerequisite to, eg calculate a reliable mean, is a more demanding application.

**Strand 2b (Processing, analysing and representing data)**

Most students were able to select appropriate techniques to analyse the data they had collected.

Moderators reported that the marks awarded in this strand were generally too high. To access the higher marks students are expected to produce a diagram and a calculation. These should be related activities. For instance, in a formal proof of normality it is expected that students will not only produce correct calculations for the comparison of normality but also a suitable representation for the comparison, eg a histogram.

It should be noted that the marks awarded in this strand should not be awarded in isolation. The notional difficulty of a technique does not mean that it can automatically be awarded a high mark. It is how the technique is being used that drives the assessment, eg the use of standard deviation to compare the spreads of data sets is, in principle, no more sophisticated than using interquartile range to compare the spreads of data sets, particularly when ICT is being used.

The quality of the interpretation of results also affects the marks that can be awarded in this strand. The greater the depth of interpretation of results the greater the demand of the technique, eg the interpretation of correlation in context is a higher demand activity than the identification of correlation, even though the calculation may be the same.

The assessment is also affected by the quality of the diagrams and accuracy of the calculations that are produced. Errors in calculations and omissions in graphs are penalised in this strand. For example, the omission of units in scatter graphs is usually penalised by the deduction of 1 mark.

The use of ICT to do the more arduous calculations and representations is to be commended, but students should be advised to be more critical of the accuracy of answers and the quality of representations. The diagrams and calculations are not always suitable for the purposes intended, eg the calculation of a line of best for a scatter graph when there is no correlation in the data.

**Strand 3 (Interpretation and discussion of results)**

Many students were able to draw their investigation together in a final conclusion and relate their findings to their initial hypotheses.

More demanding techniques should be accompanied by more demanding interpretations, eg a routine test of normality is not automatically a high demand activity. This needs to be justified beyond a desire merely to use a high demand technique.

Many students were able to comment on the reliability of their results in terms of the size of the samples they had taken; but it should be noted that a comment such as "I could have improved my results by taking a larger sample", is considered to be a low demand comment on reliability. Few students commented on the reliability of their results in terms of the techniques they had employed, eg the amount of data used to populate individual class intervals in histograms, the number of data points in scatter graphs and box plots, critical appraisal of the sampling regime, etc.

Only the best students were able to discuss the range of applicability of their results beyond the immediate sample or population.

**Key points**

Students should be advised to:

- Not to produce collections of mini investigations
- Give more detail when describing their sampling techniques
- Think more carefully about the amount of data they are using and explain their choice in the context of the techniques they employ
- Explain the reasons for removing/retaining outliers and the effect this has on their calculations and diagrams
- Be critical of the graphs and calculations produced by ICT
- Give practical interpretations of their results in the context of their aims

**Grade Boundaries**

Grade boundaries for this, and all other papers, can be found on the website on this link:
http://www.edexcel.com/iwantto/Pages/grade-boundaries.aspx

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom