STATISTICS AND DATA PRESENTATION

SUMMARY

Larger scale studies of the impact of study support can be suitable for statistical analysis. This guide helps you decide if such analysis is necessary and/or useful, and gives some guidance on choosing suitable statistical tests. It also explores how careful use of presentation techniques can substitute for statistical testing.

ARE STATISTICS NECESSARY?

Let's start by saying that, in most cases, measurement of the impact of study support in a school does NOT require statistical analysis. What is more, a lot of the data collected is not even suitable for statistical analysis! However, there are some occasions when statistical testing can be useful.

- Those pupils attending study support usually show improvement in one or more targeted outcomes. If the difference is large, it is clearly significant, but sometimes the difference is not large and you want to establish if it is real or not. Basically, is it big enough to really count? Statistical tests can be a measure of how significant an improvement is.

- Statistical significance can be a good tool to use to convince sceptics or funding providers that study support is of real benefit. Saying things like "These children improved their attainment after attending study support sessions and there is only a 1 in 10,000 chance that they would have done it without attending" can have some impact!

The purpose of statistics in the context of measuring impact is to convince an audience of some sort of beneficial effect. In many cases, presenting data in a thoughtful way can do this job without the need for statistics. Most people take in information more easily from graphs and charts more readily than from statistical analysis, so you should think carefully if you need to test your data.If the data cannot be presented in a way that looks impressive, then the chances are that the differences are not all that significant.

In most cases, statistical testing should be used when the effect reported may be open to some doubt from the audience to whom it is presented.

Many impact studies in schools are on a small scale, and are unsuitable for statistical analysis because sample sizes are too small for valid testing. **As a rule of thumb, statistical tests should not be used unless the sample size is at least 30 for each group tested.**

DATA PRESENTATION TECHNIQUES

Data can be presented as text, in tables, or pictorially as graphs and charts. Figures should not normally be put into text unless there are just two or three numbers. Tables and graphs are much clearer.

Tables are usually the best way of showing structured numeric information, whereas graphs and charts are better for showing relationships, making comparisons and indicating trends. Even where a graph or chart is used, it is usual to include a table to show the data from which it was drawn.

PRESENTING DATA IN TABLES

- Tables should be self-explanatory - they should not require detailed reference to the text.
- The title should clearly indicate what the table shows, and columns and rows should be clearly labelled.
- Include only essential data (for presentation purposes - fine detail can be given in a larger table in an appendix if necessary).
- Try to use relatively few significant digits. Too many decimal points can make data less clear (though sometimes they are necessary). Study support data often deal in whole numbers, anyway.
- If numbers are large, consider using percentages where applicable.
- Consider the orientation of the table. When you want to draw attention to a variable, it is better it is put as columns rather than rows (see below).

Tables showing outcomes related to study support participation

	Attended 2 or more clubs on more than 10 occasions each	Attended 1 club on more than 10 occasions	Attended no clubs on more than 10 occasions
Comparison with pre-dicted grades(mean +/- grades)	+1.91	+0.84	+0.14
Attendance at school between Sept - March(mean)	90.9%	86.3%	84.8%
Attitude score from teacher reports (mean, out of 10)	8.812	7.542	6.877

This table is poorly construct-ed. Headings are too wordy, the data is expressed to too many significant figures, and orientation is not the best.

	Comparison with pre-dicted grades(mean +/- grades)	Attendance at school (mean)	Attitude score (mean, out of 10)
Attended 2 or more clubs	+1.9	91%	8.8
Attended 1 club	+0.8	86%	7.6
Attended no clubs	+0.1	85%	6.9

This table is much clearer. Note in particular the increased clarity due to less significant figures, and the choice of orientation, which allows better focus on the effect of club attendance.

PRESENTING DATA AS GRAPHS AND CHARTS

Graphical presentation of impact studies is most usually presented as bar graphs, although on occasions line graphs and pie charts may be appropriate.

Bar charts give a clear display of simple results. They are used when the horizontal axis is composed of categories (e.g. male / female; those attending study support and those that don't; ethnic groups, individual pupils etc.). Impact studies very often compare categories, which is why bar charts are most often seen. A stacked bar chart can be used if some sort of improvement in a category needs to be displayed (see box below). If the bars are not separated by spaces, the chart is referred to as a **histogram**, rather than a bar chart.

Line graphs are appropriate when the horizontal axis is continuous rather than categories. In impact studies, they could be used to show progress over time (e.g. development of a measured skill each week over a ten-week course).

Pie charts are a visual tool to show proportions (e.g. percentages of pupils giving different responses when evaluating a course).

EFFECTIVE USE OF BAR CHARTS

One of the most important aspects of effective bar charts is the **choice of the scale** on the vertical axis. Examples are given in the box below. If the purpose of the bar chart is to draw attention to differences (which it usually is), then it is vital that the choice of scale does not tend to mask the scale of the difference. If the differences are so small that a slightly ridiculous axis scale needs to be chosen, then those differences are almost certainly not significant and the data is not really worth presenting!

If there is more than one bar per category, then it is more effective to **separate the categories**

Presenting data for a reading club

A school wanted to report on the progress in reading age made by a group of 8 pupils that had attended a ten-week reading club. All pupils made gains in reading age, some more significant than others.
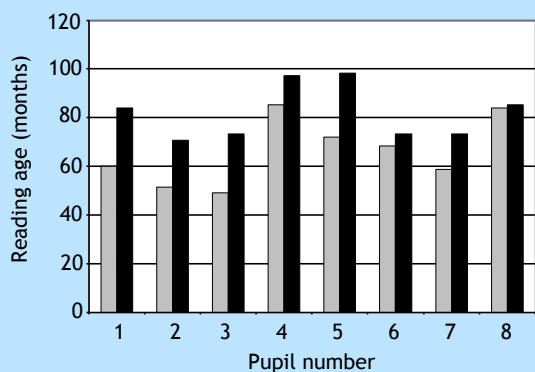

Straightforward bar chart. Notice that the pupil data is separated by a space for additional clarity. The grey bar indicates the initial reading age, and the black bar the reading age after attending the club. The range of the vertical axis scale could have been reduced (there is no need to go past 100) which would have emphasised the gain a little more as the bars would be extended.
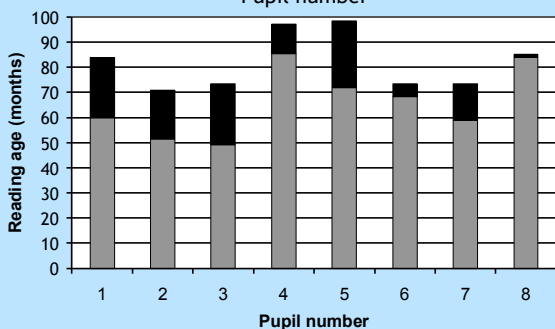
In this stacked bar chart the vertical axis has been chosen carefully and the graph emphasises the gains (gains are the black areas). This sort of chart can be visually effective but the gains are not so easy to measure as on a standard bar chart. There is a tendency for the gains to be greater in pupils with a lower initial reading age, and this could be emphasised by re-ordering the pupils in order of initial reading age, as shown by the second stacked bar chart.
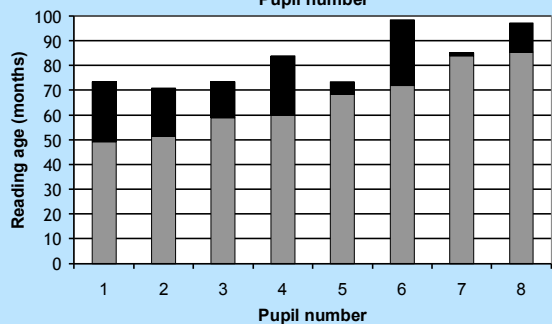
The final method plots the gains only to give a very clear focus. The vertical axis has negative numbers on it which tends to emphasise that all pupils made *gains*, however small. The fact that the reading club benefits lower reading age pupils most cannot be seen in this bar chart.
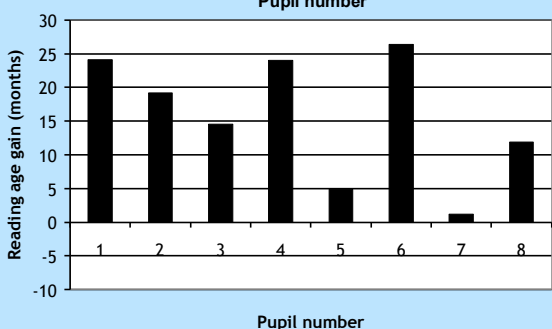
Note that all of the above methods have strengths and weaknesses, and the method chosen would be influenced by the message that was being conveyed. The fact that gains were consistent and have visual impact in the graphs means that statistical testing may not be necessary. The only weakness is the absence of a control group that did not attend the club.
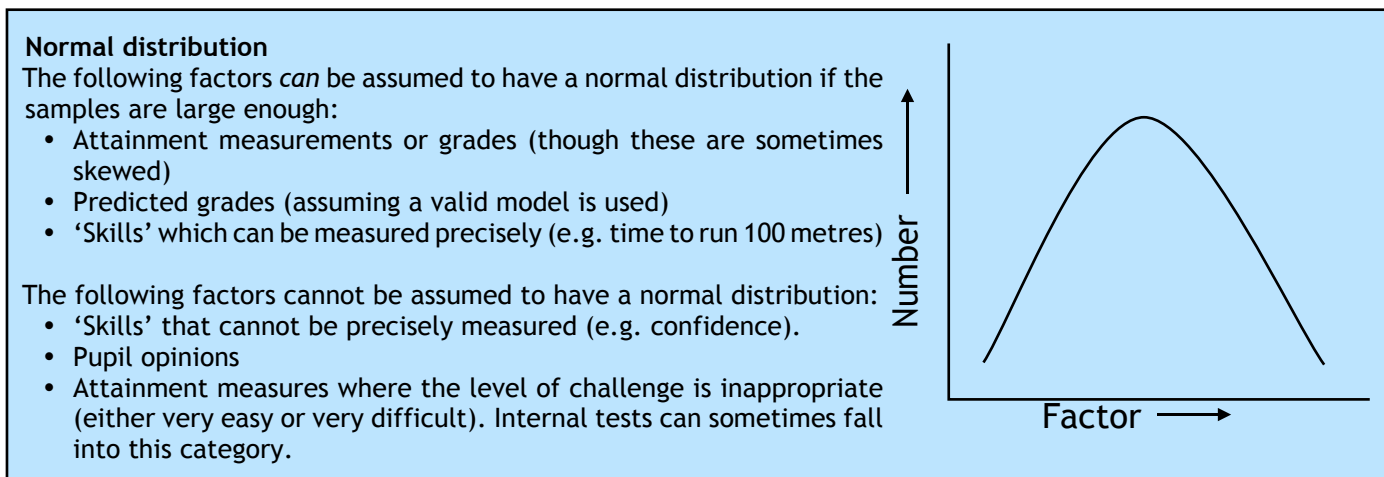
SOME STATISTICS BASICS

There are some basic terms that you need to know before entering the realm of statistical analysis.

Probability

In statistical tests that are looking for differences between two sets of data, the end product of the test is a probability value. This is an indication of how likely the results are to have occurred *purely by chance*. So, a low probability is what you are usually looking for, because that means that the results are probably NOT due to chance, but to something real. Probability is expressed as a decimal, so a probability of 0.1 (1/10) indicates that there is a 1 in 10 (10%) chance that the results are due to chance, whereas 0.005 (5/1000) indicates a 5 in 1000 or 1 in 200 chance. Statistics can never prove anything, only give probabilities. In scientific research, a probability of less than 0.001 (i.e. less than 1 in 1000 chance) is sometimes taken to indicate statistical significance, but that is arbitrary and used when there are possible life and death implications of the research. Lower probabilities could also be regarded as significant, with decreasing confidence as the probability value gets higher. Social sciences commonly regard a possibility of less than 0.05 as significant.

Parametric and non-parametric tests

Parametric data makes assumptions about the population being tested, mainly that the population follows what is called a normal distribution. A normal distribution curve is shown below:

Normal distribution
The following factors *can* be assumed to have a normal distribution if the samples are large enough:
- Attainment measurements or grades (though these are sometimes skewed)
- Predicted grades (assuming a valid model is used)
- 'Skills' which can be measured precisely (e.g. time to run 100 metres)

The following factors cannot be assumed to have a normal distribution:
- 'Skills' that cannot be precisely measured (e.g. confidence).
- Pupil opinions
- Attainment measures where the level of challenge is inappropriate (either very easy or very difficult). Internal tests can sometimes fall into this category.

Non-parametric data makes no assumptions at all about the population being tested.
Any parametric test is invalid if the data is non-parametric. Non-parametric tests can be valid for any type of data, though they are less powerful than parametric tests.

USEFUL STATISTICAL TESTS

This pamphlet will not give step by step instructions of how to do statistical tests. It simply tries to ensure that you choose the most appropriate test to do. All of the tests below can be done automatically by Microsoft Excel, and explanations can be found in the help section of that program under "Statistical functions"

Standard deviation

This give a measure of how variable your results are. As a rule of thumb, increasing variability means decreasing reliability.
Excel function: STDEV
Web link: http://www.robertniles.com/stats/stdev.shtml

Student's t test

This is used to compare two groups with each other to see if there is a difference between them (e.g. Comparing pupils attending study support with those that have not). It is a parametric test (see above) and cannot be used when there are more than two groups.

Excel function: TTEST
Web link: http://www.biology.ed.ac.uk/research/groups/jdeacon/statistics/tress4a.html

The chi-squared test
This is used to compare results that you have measured with an 'expected' result. This can obviously be used to compare pupil performance with predicted grades, but it can also be used to compare one group with another (your 'expected' result here would be that there would be no difference between the groups - even if your *actual* expectation is that there would be). This is a non-parametric test.
Excel function: CHITEST
Web link: http://www.biology.ed.ac.uk/research/groups/jdeacon/statistics/tress9.html

Spearman's rank order correlation
This non-parametric test tests links (correlation) between two factors. Remember, though, that a correlation does not necessarily imply a cause and effect relationship.
Excel function: CORREL
Web link: http://www.revision-notes.co.uk/revision/181.html

There are a great many more statistical tests and techniques available. A useful guide to how to choose statistical tests can be found in Cohen. L. et al(2007): *Research Methods in Education*, Routledge.

ACKNOWLEDGEMENTS