Success and Failure in Statistical Education - A UK Perspective

Anne Hawkins - London, England

1. Introduction

Research into Statistical Education is a *relatively* new field, less extensive and coherent than, say, research into Mathematics Education or Science Education. Therefore success and failure may be less easily defined. A medical analogy may help, with failures as symptoms caused by "wrong" statistical education. We see, and treat, many of them in our statistical consulting rooms. They occur in epidemic proportions, with students aged 10 or 11 years upwards being particularly vulnerable. Untreated symptoms tend to worsen with age. The underlying disease appears primarily to be *infectious*, being readily transmitted from teachers to students, although probabilistic reasoning difficulties may have a *hereditary* factor.

Real success would be the eradication of the disease, but successes featured at ICOTS 3 will more likely be new diagnostic insights and recommended treatments. Some will merely be "sticking plasters". Some will go nearer to the heart of the problem. Some will be preventative, others will be remedial. They will not always have been subjected to full clinical trials, and most of them will not be widely available on prescription, through the NHS, the National H'education Service.

Some examples of common symptoms of failure

Failure/presenting symptom	*Success/suggested treatment*
"My MPhil sample size is 36. How many more do I need for a PhD?"	Counselling in concepts of (im)-precision, bias, (in)accuracy, etc., and research efficiency. Treat with confidence to ensure significant results.*
Compulsive averaging, characterised by very early onset of reflex response to any set of data.	A chronic mental condition. Very resistant to treatment.

Impaired vision, uni- or bi-variate blindness.	Multi-variate therapy with graphical and exploratory techniques (not necessarily methods). Preventable if treatment started early enough.
Statistical delusions, e.g. beliefs that statistics is only surveys, or only sums, or only graphic art, or only number kleptomania, etc.	Prevention better than cure, otherwise long-term behaviour therapy necessary.
"This is my data, what shall I do with it?"	Probably too late, terminal case. Painkillers and tranquillizers. In event of remission, pre-planning therapy to prevent recurrence.

*Note: A recent sign that successful progress is being made. The *British Medical Journal* is now demanding that confidence intervals be quoted, together with the point estimates, rather than only P-values.

Statistical educators do not always agree on whether the *real* roots of statistics lie in theory or applications, and so will not always concur on what constitutes success and failure in Statistical Education. Students' understanding *should* be developed to include both aspects, but statistical education programmes still emphasise the *manipulation* rather than the *development* of relevant statistical models, and students therefore have great difficulty relating theory to application.

Changes in technology, and hence in professionals', and now increasingly lay-persons', practice of statistics, have made our views of statistical education unclear. Delegates at the International Statistical Institute's Round Table Conference, "Training Teachers to Teach Statistics" (Budapest, 1988), had difficulty determining training priorities because what teachers should or would teach was not clearly defined.

2.   Statistics' place in the curriculum

Statistics is choosing how best, how most efficiently, to investigate and communicate the *probable* truth. The Schools Council Project on Statistical Education (POSE, 1981) stated to the Cockcroft Committee (1982) that "statistics is not just a set of techniques, it is an attitude of mind in approaching data. In particular it acknowledges the fact of uncertainty and variability in data and data collection. It enables people to make decisions in the face of this uncertainty." The basic principle that Statistics is *statistics plus probability* is often underplayed. Intuitive likelihood estimates, not necessarily quantified, can and should be extensions of data representation, exploration and comparison exercises from the earliest Primary levels. As the Schools Council Project on Statistical Education concluded, much of what is being done at Primary level is quite good. However, the intuitive substrata of probability are still under-utilised and under-developed, and the UK lags behind countries such as Japan, Hungary and the US in this respect.

Views of statistics which emphasise its *apparent* dichotomies (for example, the statistics/probability divergence), provoke arguments about who should teach statistics, in what way, and whether it should be taught as a self-contained discipline, or with mathematics, or in the context of "user"-disciplines. Zidek (ICOTS 2, 1986) considered the extents to which a core statistics exists, and to which the subject is fragmenting into specialist areas of other disciplines (e.g. econometrics). How these two aspects are to be managed within the school curriculum has still to be resolved. However, successful education does not only depend on identifying the curriculum. In the UK, the intended, taught and learned contents of that curriculum lack equivalence, and research is needed at national, school, teacher and student levels to rectify this.

What is appropriate to the teaching of statis*tics* as a discipline in its own right (whether theoretical and/or applied) is not necessarily appropriate to statis*tical* teaching within other disciplines. We need to decide whether this distinction is to be maintained. So far, unifying terms such as "stochastics" have failed to be universally understood or accepted. Most of students' exposure to statistics tends to be via mathematics courses, but controversy persists over whether this should be statistics or statistical education. Deciding on the basis of whether students are being trained as specialists in statistics will not suffice for *school* mathematics because, for most students, this precedes specialist statistics training. Here the objective is presumably three-fold:

(i)   to provide academic credibility for a course in *mathematics*;
(ii)   to provide sufficient theoretical background for those who might become specialists; and also
(iii)   to ensure sufficient applied skills to "serve" the immediate requirements of students' other disciplines.

Current debates in the UK about how to teach *mathematics* may result in ways of satisfying objective (i) that will benefit *statistics* teaching, but it is too early to say.

"Success is relative; it is what we can make of the mess we have made of things" (T S Eliot, *The Family Reunion*, 1939). Recommendations on statistical education in the Cockcroft Report, "Mathematics Counts" (1982), provide a yardstick against which to assess whether we are making a "mess" of statistical education. This UK report, contemporary to ICOTS 1, is recent enough for its recommendations to be currently meaningful, as well as being old enough for them to have been acted on.

3.   Statistical education now, compared to the 1982 recommendations

3.1   *The nature of statistics and statistical education*

The Cockcroft Committee stated that statistics should be taught as a *practical* subject, with real-life applications. "Practical" implies significantly more than the term "applied". Applied statistics teaching fails when many of the "practical" statistical methods are still taught in a theoretical way. This failure is compounded when students are then assessed on the basis of their practical coursework.

Many teachers have no experience of applying statistics, let alone of managing others' practical projects. They therefore tend to resort to stereotype projects, "for the

sake of doing projects". This defeats the purpose as much as did the "theoretical" simulation projects in Pure Mathematics and Statistics courses which taught students nothing about the *use* to which a particular distribution could be put. Students (and their teachers) must have *skills training* in statistics, and there is not enough research into how to do this. Moreover, if such research continues to lag behind curriculum development projects, these latter will fail because they simply will not reliably yield enough "successful" learning.

The Cockcroft Committee identified a need for the development and provision of teaching materials emphasising a practical approach. While there is a never-ending flow of new materials, *research* into their design and use *is lacking*, many are inadequately field-tested, and some are clearly misleading.

Indeed, there is too little research into statistical education in general, and what there is tends to use limited, and sometimes less than robust, methodology, with a general lack of synthesis, cohesion and *on-going* development of productive areas of research. Attitudes and expectations need changing so that relevant research is *assumed to be essential* for statistical education to progress; where it exists it must be used in planning new research or educational strategies; where it is lacking it must be undertaken.

3.2   *Teacher-training and recruitment issues in statistical education*

The Cockcroft Committee confirmed that "some teachers, especially those who completed their degree courses some years ago, do not have sufficient knowledge of probability and statistics to teach in these areas effectively". Furthermore, the committee found that few teachers, including those whose degree or other courses had included the study of statistics, had received training in how to *teach* statistics in schools. Committee members were pleased to hear about "one PGCE course in which ... an inter-disciplinary curriculum option on the teaching of statistics" had been introduced. They recommended "the provision of a similar option within other initial training courses".

It is a failure that teachers of statistics can still come through the system with *no* statistical training beyond (early) secondary school level. Hawkins (1989) found that of 367 teachers of statistics at primary and secondary levels, 1 in 4 owned to having "no" (8%) or "only school level" (16%) statistical education themselves. Far from Cockcroft's recommendations resulting in an upsurge of relevant training, however, the succession of "crises" in education in England and Wales has made it very difficult for teachers to avail themselves of either pre- or in-service training. It was also a conclusion of the ISI's Round Table Conference (1988) that, although initiatives were being taken by trainers, with few exceptions their impact was relatively localised. Centralised directives and finance are needed to change this.

The Pre-service Statistical Education courses, featured in the Cockcroft Report, recruited well in the early to mid 1980s, when PGCE students of all disciplines wanted to enhance employment prospects by offering statistics as an additional teaching subject. With more training resources and an increased demand for teachers of core curriculum subjects, however, students now choose courses in mathematics or science education. Statistical education cannot compete with this. The course that gained the Cockcroft Committee's praise has now ceased to exist. Difficult as it is to recruit *statistics* post-

graduates, the problem is compounded in an institution specialising in "user"-discipline such as Education, where recruitment to *statistical education* may be even harder hit.

Statistical education generally depends on the initiative of *isolated* individuals. At taught Masters level, larger numbers of mathematics education lecturers can sustain more options, encouraging students to specialise solely in *mathematics* education, rather than *statistical* education. It is not always mathematics (or other disciplines) which are to blame for our recruitment problems. At tertiary level, *for students majoring in other disciplines*, statisticians tend to set pre-requisites too high, and to offer statistics courses which are inappropriate for the levels and needs of the potential students. Also, *statistics* lecturers are often unaware of, or not interested in trying, teaching innovations which studies in statistical education have shown to be more effective. *Students see our statistics colleagues failing to practise what we statistical educationists preach.* This highlights the case for statistical education being established as a substantive area of study in its own right, and for tertiary level teachers of statistics being trained in how to teach, rather than merely being recruited on the basis of their statistical qualifications and research experience. In fact, I am actually prevented from recruiting higher degree students who lack classroom experience, even though they are well qualified in statistics. The fact that few statistics graduates take up teaching is a problem in itself, of course, but I am concerned that such specialists are also deterred from moving into the academic study of, and research into, statistical education.

The failure to recruit sufficient numbers of students, and hence to justify sufficient numbers of academic staff, is a major impediment to the development of a pedagogic discipline or research basis for statistical education in its own right. I see this "chicken and egg" situation as a critical state of affairs. The ISI's Round Table Conference (Budapest, 1988) recommended that the crisis in statistical education should be tackled first at the in-service level. Lack of finance or time allowances for in-service release make this difficult in the UK, although the cross-curricular emphasis on data handling in the new National Curriculum may help. Distance learning approaches, teacher networks, journals, newsletters, and further exploitation of electronic communication can make in-service training more accessible, as can action research programmes, giving teachers responsibility for the development of curricula, teaching methods and materials, especially if combined with cascade models where the trainees go on to train their colleagues. Another possibility, recommended by both the Cockcroft Committee and the Schools Council Project on Statistical Education, is that of a statistical coordinator in each school who could also promote in-service statistical education. Although some progress is now being made, this recommendation has still not been widely implemented.

3.3   Computers in statistics teaching

The Cockcroft Committee suggested that "the use of computer simulation might enable essential ideas ... in probability and statistics to be assimilated more rapidly ...". Research into this possibility is still needed, e.g. into how students deal with the different levels of abstraction required, and the emphasis on speed rather than on qualitative aspects of the learning outcomes is unfortunate. ICOTS 2 delegates were treated to "101 ways of prettying up the Central Limit Theorem on screen", but if students are not helped to see the *purpose* of the CLT, and if the software does not take

them beyond what is still, for them, an abstract representation, then the software fails.

Recent successes in statistical software development would include the decline in numbers of home-grown packages re-inventing the wheel to compute averages, standard deviations, etc. Plenty of examples of bad software are still being produced, however. A particularly dangerous kind is the "reduced" package, designed for the "statistically naive", in which the range of available analyses is often limited. More dangerous, however, is the removal of essential decision making processes from the user, precluding sensible, informed analyses and interpretations, and giving the misleading impression of statistics as a mechanistic process. A better approach would be to concentrate efforts into designing robust user-friendly front-ends for reputable statistical software, and to emphasise in our teaching the evaluation and use of such software.

Some assert that students can learn about statistical concepts by writing programs to compute them. Whether this enhances students' statistical understanding, as distinct from algorithmic knowledge, has yet to be properly researched. In any event, this approach is unlikely to be successful for students who are not mathematically inclined. The *real* advantage of computers is that they give the opportunity to invest-igate the characteristics of statistical concepts in the face of changing data characteristics, to explore by trial-and-error the effects of concepts like random variation and bias, and to experiment in solving problems using statistical (or stochastic) models, representations and simulations. Too few examples of this kind of software are being developed or used, and too many mathematics teachers still see the computer as an opportunity for students to do fifty of the same type of statistical computation where formerly there was only time for three, or worse still, to drill students in the arithmetic steps for computing statistical concepts (Hawkins, 1989). We fail if we cannot convince ourselves and others that such uses are essentially redundant in real statistical education.

The same survey found that *concept* simulation software, despite its prolifer-ation, was surprisingly little used, and computer simulations of *real-life* applications rarely featured at all. About 40-45% of respondents taught using computers, the highest usage being in GCSE and 'A'-level statistics courses and in the "user"-disciplines. A far higher percentage (74% overall) use mathematical or scientific calculators, but the use of statistical calculators was essentially limited to students taking specialist statistics courses. A noticeable feature was that mathematics students often observe their teacher using a computer, instead of gaining "hands-on" experience themselves. Overall, a rather restricted view of computers and their possible uses is being presented to students.

Projects based on the use of Exploratory Data Analysis techniques and materials are an alternative attempt to "accelerate" statistics teaching, or to improve the quality of understanding achieved. Such studies fail, however, unless they take enough account of teachers' training needs. Teachers' own weaknesses and misconceptions do not necessarily disappear just because they are teaching using EDA techniques.

3.4   *Curriculum and syllabus development*

The Cockcroft Committee observed that "the increasing use of statistical techniques, not only in the study of other subjects but also in business, commerce and manufacturing industry, provides many students with a strong reason for wishing to study probability and statistics". At the time, the UK Government was interested in establishing a broader-based Sixth Form course by the introduction of "half"-'A'-levels.

The Cockcroft Committee suggested that practical (applied) statistics could usefully make such a "half"-course, and that "rather than a full 'A'-level course in statistics, such a course would serve the needs of a much greater number of students". Such a course was produced by modifying the statistics part of an existing Pure Mathematics with Statistics 'A'-level. It was a complete failure. As well as extra applied aspects being introduced, excessive and indigestible amounts of theoretical content were added by the mathematician designers of the syllabus, supposedly to support the applied aspects. The syllabus became far more extensive than a typical first year statistics degree course, totally conflicting with the Cockcroft Committee's view that overloaded syllabuses prevent students from obtaining "a secure grasp of ... the applied elements of their course".

3.5   *The public image of statistics and statisticians*

This failure to make a course in statistics fit the requirements and levels of the students is akin to the "service-teaching dilemma" in tertiary level education. Who is better qualified to say what is required, the student and/or their "user"-discipline teachers or the statistics specialist? In fact, as statisticians, *we* often alienate potential students of statistics by our "purist" approaches to both statistical content and methods of teaching. The problem is exacerbated by those statistics specialists who feel that service-teaching is somehow beneath them, or that statistics is for (intending) statisticians only. This provides other departments with the excuse to undertake the required statistics teaching themselves, or to reduce that aspect of their students' courses, or to eliminate it altogether. Ironically, this elitist view, held by some statisticians, is still found alongside equally inappropriate classroom-filling exercises based on the principle that "anyone can teach statistics".

The Cockcroft Committee stated that it received "surprisingly few" submissions which made direct reference to the teaching of statistics. Certainly the Royal Statistical Society and the Institute of Statisticians did contribute. So too did the Schools Council Project on Statistical Education. However, the ambiguous role of statistics within the school curriculum, and the emphasis on mathematics, together mean that statisticians are not always involved in the decision making concerned with statistical education, and "user"-discipline specialists may not be consulted because they, or others, assume that they are not concerned with statistics.

Current computer software and the availability of cheap hardware contrives to mislead "the World and her husband" into thinking that they can be statisticians, so they fail to appreciate the need for specialist statistical training, or to see the advantages of seeking help from a professionally-trained statistician, or the risks of not doing so. In these respects, the medical profession has done a far better job with health education than we have with statistical education. "The Statistical Needs of Non-Specialist Young Workers" (1985), a report produced by the Centre for Statistical Education at Sheffield, highlighted this problem in the "statistical generation gap" it observed between young employees and their managers.

For many years, statisticians at all concerned with statistical education have criticised non-specialists for teaching statistics in a "cook-book" manner. However, statisticians must not fall into the converse trap of giving the impression that statistical communication is, like a cordon-bleu cookery book, dependent on fluency with a foreign

language. Statisticians must learn to speak the language of the people, and more particularly of their potential students. If students do not speak the language of mathematics and calculus, then we must find other ways of teaching them statistics. To start by expecting fluency in the language of mathematics loses us vast numbers of students, and fails with others by delaying their *statistical* education until after their mathematical fluency is established, which it may never be for some students.

4.   **Recent developments in the UK**

Mathematics in the core of the National Curriculum in England and Wales *appears* to reflect the Cockcroft emphasis on real-life applications, but appearances can be deceptive. It is certainly too early to say how effective its implementation will be. Three out of fourteen targets of attainment are dedicated to *data handling*:

12.   Pupils should collect, record and process data.
13.   Pupils should represent and interpret data.
14.   Pupils should understand, estimate and calculate probabilities.

Other areas of the core National Curriculum such as Science Education also feature data-handling, and statistics is thereby seen as a medium for cross-curricular collaboration, and its breadth of application is demonstrated. It is less clear how coherence and useful transfer between contexts will be ensured.

The success for statistical education is that what might have been implicit before, but was nevertheless under-taught, is now made explicit and its teaching will be directly and regularly monitored from primary level upwards. In all cases, "Knowledge, skills, understanding and use of ..." are to be included in the teaching, but *how* to achieve them is less obvious. Probability, although under the heading "data-handling", is still in reality presented as distinct from statistics. "Data-handling" seems to be being used as a new term for what some of us have tried to call "stochastics", but since probability is still limited to the usual axiomatic rules, no new consideration of the role of probability in data-handling is stimulated.

A further shortcoming is the omission of *investigations* of the statistical concepts themselves, tending to create the impression that their application can be mechanistically and computationally determined. Social survey and questionnaire investigations are over-emphasised, with little to encourage experimental methods or simulation investigations, thereby promoting the restricted view of statistics which is already too prevalent in schools, as has been seen from their entries to the Annual UK Applied Statistics Competition.

5.   **Assessment issues**

The frequent lack of personnel with statistical expertise on relevant public examination boards adversely influences not only the initial syllabus development, but also the subsequent setting, and marking, of examination questions, thereby distorting teachers' and students' views of the underlying skills of statistics. Ill-chosen examples

and inappropriate data sets given as textbook or examination exercises can create long-lasting bad thought habits in students, just as can ill-designed syllabuses, and the mathematicians' emphasis on "the one right answer". Examiners without appropriate statistical expertise, do not always realise how to set a question which encourages candidates to display *statistical reasoning*, rather than merely arithmetical competence. The problem is particularly acute in the assessment of less able candidates who are invariably assumed to be best occupied with guided, step-by-step *arithmetical* processing.

Changes in assessment and changes in teaching are inter-dependent. One solution may be to involve more teachers in the moderation process. Despite the "I won't interfere with your students' grades if you don't interfere with mine" syndrome, the use of teams of teacher moderators provides an opportunity for informal insights into what colleagues in other schools are trying and achieving, providing an important in-service training function. Sadly the costs, in time and money terms, are generally prohibitive.

6.   Concluding remarks

1982, or thereabouts, is a convenient watershed when it comes to defining criteria of success in statistical education. Prior to 1982, many students developed a "have data, will analyse" view of statistics. Statistical training had most likely focussed their attention on computational aspects. For many people, criteria of success in statistical education depended on students' competence to perform analyses. By 1982, we were becoming concerned that statistical education should encourage students to develop and exercise the skills underlying such computations. Success became measured in terms of how aware students could be made of broader data-based/application issues. ICOTS 1 and 2 saw us working hard to develop teaching methods and materials to do this. Now, in 1990 (ICOTS 3 and beyond), criteria of success in statistical education must reflect its development as a substantive discipline in its own right, with an adequate research infrastructure.

Returning to my medical analogy, ICOTS 3 will show important advances in diagnostic skills, new treatments and preventative measures, but a real cure for statistical education will not be found until a more *systematic* approach to its study is adopted. Meanwhile, people that we educate today; pupils, students, teachers, trainers, colleagues, policy-makers, employers or the general public, will be better equipped to appreciate and to contribute to future developments in statistical education, if they receive the *better* statistical education that we should now be able to offer. Ironically, success in statistical education may at present be measured by the extent to which we can confront such people with its failures, in order to gain on-going commitment to, and support for, finding the *best* statistical education.