Evaluating educational interventions
M Wilkes and J Bligh

*BMJ* 1999;318;1269-1272

Updated information and services can be found at:
http://bmj.com/cgi/content/full/318/7193/1269

*References*
This article cites 25 articles, 1 of which can be accessed free at:
http://bmj.com/cgi/content/full/318/7193/1269#BIBL

7 online articles that cite this article can be accessed at:
http://bmj.com/cgi/content/full/318/7193/1269#otherarticles

*Rapid responses*
You can respond to this article at:
http://bmj.com/cgi/eletter-submit/318/7193/1269

*Email alerting service*
Receive free email alerts when new articles cite this article - sign up in the box at
the top right corner of the article

*Topic collections*
Articles on similar topics can be found in the following collections

- Other Medical Education (337 articles)
- Other Statistics and Research Methods: descriptions (583 articles)
- Undergraduate (229 articles)
- Postgraduate: Academic (102 articles)
- Postgraduate: Family/General Practice (149 articles)
- Postgraduate: Residency / House staff (157 articles)

*Notes*

To order reprints of this article go to:
http://www.bmjournals.com/cgi/reprintform

To subscribe to *BMJ* go to:
http://bmj.bmjjournals.com/subscriptions/subscribe.shtml
These difficulties are not enough to excuse complacency. Rigorous research design and application are needed to attract research funding, to provide valid generalisable results, and to elevate the profile of educational research within the medical profession.

Competing interests: None declared.


Evaluating educational interventions

M Wilkes, J Bligh

Recent extensive changes have taken place in medical education at all levels in both the United Kingdom and the United States. These changes need to be assessed to measure how well reforms have achieved their intended outcomes. Educational innovation can be complex and extensive, and its measurement and description is made more difficult by the confounding and complicating effects of each later stage in the continuous curriculum. The radical curriculum reform at undergraduate level in the United Kingdom, managed care in the United States, and the increasing use of community sites for learning in both countries may greatly affect how medicine is practised and managed in the next century. We should know more about the educational processes and outcomes that result from the new courses and programmes being developed in medical schools and postgraduate training.

What is educational evaluation?

Educational evaluation is the systematic appraisal of the quality of teaching and learning. In many ways evaluation drives the development and change of curriculums (figure). At its core, evaluation is about helping medical educators improve education. Evaluation can have a formative role, identifying areas where teaching can be improved, or a summative role, judging the effectiveness of teaching. Although educational evaluation uses methods and tools that are similar to those used in educational research, the results of research are more generalisable and more value is invested in the interpretation of results of evaluation.

Evaluation can also be a hindrance to curricular change. In the United States, for example, enormous weight is placed on the standardised multiple choice type assessment (USMLE) that is taken by all medical students. Although many people believe in the exam, it has been a major barrier to curricular reform. Medical schools feel that any curricular change may sacrifice students' performance in this examination, which in some circles is still seen as the "gold standard." This reliance on conventional educational tools to compare a new innovative curriculum with the traditional curriculum has caused schools such as McMaster a great deal of angst.

Summary points

- Evaluation drives both learning and curriculum development and needs to be given serious attention at the earliest stages of change.
- Summative evaluation can no longer rely on a single assessment tool but must include measures of skill, knowledge, behaviour, and attitude.
- New assessment tools do not necessarily duplicate each other but assess and evaluate different components of a doctor's performance.
- Assessment needs to be part of an ongoing evaluation cycle intended to keep the curriculum fresh, educationally sound, and achieving its intended objectives.

At this point it is worth differentiating between monitoring, evaluation, and assessment. Assessment refers to the quality measures used to determine performance of an individual medical student. Monitoring is the gathering and recording of data about courses, teachers, or students and is regularly carried out at institutional level. Evaluation uses data gathered in the monitoring process to place a value on an activity. According to Edwards, evaluation seeks to "describe and explain experiences of students and teachers and to make judgements and [interpret] their effectiveness."

Approaches to evaluation

Recommendations intended to evaluate changing medical programmes have been made in the light of the extensive changes going on in medical schools in the United States. Four general approaches to educational evaluation have emerged over recent years. We have classified these as follows:

- Student oriented—Predominantly uses measurements of student performance (usually test results) as the principal indicator.
Programme oriented—Compares the performance of the course as a whole to its overall objectives and often involves descriptions of curriculum or teaching activities. This approach "closes the loop" of course or curriculum design by bringing together coherent accounts of how each element of the course—for example, use of teaching resources or choice of assessment methods—has contributed to the whole.

Institution oriented—Usually carried out by external organisations and aimed at grading the quality of teaching for comparative purposes. A wide range of information and evaluation models is used in this approach. For example, the recent round of visits to university departments in the United Kingdom by the Quality Assurance Agency on behalf of the Higher Education Funding Council used observation of teaching and examination of course materials to assess teaching quality.

Stakeholder oriented—Takes into account the concerns and claims of those involved and affected by the course or programme of education including students, faculty patients, and the NHS (in the United Kingdom) or managed care organisations (United States).

In addition to these broad approaches, evaluators are concerned with other outcomes from the educational intervention. These outcomes may include the goals and organisation of the course; whether participants achieved the learning objectives of the course; whether learning led to long term behavioural changes as the result of new knowledge or skills, and whether the program achieved longer term effects intended to improve the health of society (improved health outcomes, decrease costs, improved test ordering, etc).

Current medical education programmes are often complex with teaching spread over many disciplines, delivered in many different locations (hospitals, clinics, classrooms, laboratories, etc), and spanning several years. A wide range of learning and teaching styles are used and students therefore have varied experiences during their training. Using just one approach to assess learning is fraught with difficulties. It is not surprising that many of the innovative courses set up over the past 20 years have used a pragmatic approach with an eclectic choice of assessment methods. There are four possible strategies for evaluating an educational programme, beginning with the most basic tools—structural issues. Did the students attend lectures? Did the lecture follow the intended outline, etc? The next step is usually some sort of before test/after test multiple choice examination to measure gains in knowledge. This approach may be acceptable for the early stages of an undergraduate course or for a targeted intervention in continuing medical education. But when the aim of an educational intervention is to combine knowledge and acquisition of skills to form competence or change practice behaviour this approach is too simplistic. In this case an objective structured clinical examination, computer based examination, or videotape of an actual patient (or standardised patient) encounter may be better. The most important, and most difficult, level of evaluation examines not at the level of the intervention or provider but seeks to evaluate whether the intervention actually had a benefit on the health of society.

Traditionally, medical education was primarily concerned with the delivery of knowledge. It is therefore unsurprising that assessment tools in this area are well developed. However, over the past decade medical educators have developed various new techniques intended to better assess skills, attitudes, and behaviour. Medical schools the world over are using a variety of practical examinations (for example, the objective structured clinical examination) intended to evaluate interactions between student and patient (including history taking, physical examination, and clinical reasoning). The feedback of students and house officers is important in evaluating curriculums, and methods for gathering their views have been described, although the questionnaire remains the commonest form of feedback in medical courses.

Medical schools in the United States are experimenting with computer based examinations intended to move beyond the simple assessment of knowledge by integrating features of clinical reasoning and data acquisition.

As medical schools include more small group teaching (case based or problem based learning), often using materials from multiple disciplines, they are finding that multiple choice questionnaires, short
answer exams, and even computer based exams no longer capture the goals of the course. Instead, self
assessment, peer assessment, and written essays or
critiques are playing an increasing part in evaluation.
But barriers still exist to the wider implementation of
these assessment tools. Their development requires a
sizeable commitment of faculty time and institutional
resources. Furthermore, as with any curricular change, it
requires an institutional shift in thinking such that the
goal of assessment is not a precise numerical grade but
a global assessment with specific narrative feedback. In
addition, the medical school infrastructure (governance,
leadership, reward systems, allocation of teaching time,
promotion, and intense competition among faculties)
often poses insurmountable barriers to the develop-
ment of innovative evaluation programmes.

Evaluation of process and outcome
Medical education is a complex combination of
systematic teaching and learning activities within a
professional environment where unplanned learning
is an important objective part of clinical learning. How
students learn is as important as what they learn, and
understanding how they learn can contribute much to
improving what they learn.

To evaluate outcome it is essential to develop a lon-
gitudinal database to allow long term follow up to deter-
mine the validity of selected outcomes. Possible long
term outcomes of medical education include the quality
of clinical care provided by doctors, cost effective
decision making, professional satisfaction, and patient
satisfaction. Measurement of these variables is notori-
ously difficult, partly because of a lack of standardised
tests and partly because of ethical and professional con-
cerns surrounding public identification of differently
performing clinicians. Recently, Tamblyn et al in Quebec
reported a strong association between high passing
scores in a national qualifying examination and subse-
quent good practice in primary care as measured by a
range of clinical indicators, including prescribing
and referral. Other schools, including the University of
California at Los Angeles, have found little or no corre-
lation between standardised examination scores and
other objective indicators of clinical excellence (course
grades, objective structured clinical examinations, etc) or
subjective measures (faculty assessments, peer assess-
ments, patient assessments, or allied health profession-
als’ assessment).

International perspectives and
experiences
The box on the previous page shows a snapshot of
recent reports of evaluation of medical education. A
wide range of approaches have been tried in each of
the four international regions. Reports include a
mixture of outcome and process evaluations, with the
emphasis on student oriented and programme
oriented approaches.

At its core, educational innovation is about introduc-
ing and implementing change. Successful change
requires assessment and feedback so that alterations can
be made, mistakes corrected, and momentum main-
tained. The management of change itself needs to be
monitored. Frequent evaluation is essential, and a range
of methods is usually necessary to obtain the best infor-
mation from many sources. Such formative evaluation is
different from the summative evaluation that takes place,
often formally, once a new course is implemented. The
box above shows the range of assessment tools currently
used to evaluate the individual, the course or curriculum,
and the institution.

Curriculum governance and evaluation
Curriculum governance is the term that describes the
responsibility of medical educators to establish and
maintain high standards of teaching and learning.
Evaluation lies at the heart of this process and, like
clinical governance, it uses audit, self review, and peer
review as its principal methods. Like clinicians,
educators are accountable to their faculty, students,
patients, and society. Internationally, great attention
has been devoted to curricular reform and assessment
of change. Unfortunately, the development of sound
evaluation and assessment tools is often an after-
thought, occurring when funds are depleted, the
faculty is exhausted, and students are frustrated and

Indicators used in evaluating educational
innovations

**Structural evaluation measures**
- Attendance at class
- Number of applications to medical schools
- Assessment by national body

**Outcome evaluation measures**
- Career choice or preference
- Nature of practice
- Quality of care indicators
- Student achievement compared with other schools and
  national norms
- Cost effectiveness measure
- Effects of different curriculum tracks on assessment
  and career choice
- Patient satisfaction
- Peer assessment
- Quality of care

**Process evaluation**
- Group work characteristics (such as tutor and
  student styles)
- Entry and selection policies
- Assessment practices
- Psychometric measures including learning styles,
  stress, etc
- Student satisfaction with medical school

**Evaluation tools**
- Questionnaires
- Focus groups
- Objective structure clinical examination
- Multiple choice questions
- Viva
- Thesis project
- Qualitative written assessment
- Patient assessment
- Allied healthcare professionals’ assessment
- Peer evaluation
- Self assessment
Trend spotting: fashions in medical education

Joseph K Campbell, Cindy Johnson

This article explores the nature of fashions in medical education and identifies some of the questions about medical education that such fashions raise. What constitutes a fashion in a professional or academic discipline? What do such “fashions” contribute, both positively and negatively, to the development of medical education? What steps can practitioners and academics in medical education take to change a trend into an important step forward in the development of a discipline?

We define a fashion in medical education as an approach to education that is based primarily on social influences, in contrast to approaches based on established educational principles and theories, critically evaluated experiences, or the results of valid research. An analogy is the distinction between fashion in clothes (colour and style) and the quality and functionality of clothing (material and comfort). We chose multi-professional learning and multimedia computer aided learning as case studies, and we have illustrated why each should currently be characterised as a fashion rather than informed practice. Both have received attention internationally in medical curriculum reform. 1


11 Broomfield D, Bligh J. Validation of the CEQ. Medical Education 1998;32.
14 Rolle IE, Pearson S-A, Barnsley L. Attitudes of doctors toward community medicine: differences between graduates from innovative and traditional medical schools. Teaching and Learning in Medicine 1996;8:77-82.

Summary points

Unexplored assumptions about the knowledge underpinning practice underlie some fashions in medical education

Definitions of a topic may be vague or differ among practitioners

In a fashion, educational reasoning and justification are implicit at best and at worst absent

Qualitative information to substantiate empirical data is often lacking

The feasibility of implementing interventions is ignored or overlooked

Method

A Medline search was conducted and all available abstracts of articles addressing multimedia computer...