How to evaluate and improve a medical education programme

Gwyndaf Roberts[1] and Rich Bregazzi[2]

1. Lecturer in Medical Education, School of Medicine, Newcastle University, UK
2. Lecturer in Medical Education, School of Medicine, Newcastle University, UK; Visiting Research Fellow in Healthcare Education, St John's College, Durham University, UK

This series of papers is part of the global outreach work of the School of Medicine, Newcastle University, UK.

Correspondence:
Gwyndaf Roberts
Gwyndaf.roberts@newcastle.ac.uk

ABSTRACT

The evaluation of any medical education programme is an important and continuous task. The information gathered will inform strategic decision making and programme improvement. Using an established and widely recognised evaluation model helps structure and support the process.

Key words: evaluation, framework, evidence.

INTRODUCTION

Improving the quality of medical education programmes contributes to higher quality healthcare and an improved patient experience. A well-structured evaluation process will ensure programmes evolve and improve based on carefully sourced evidence of actual practice. This article introduces a model for evaluation and provides some suggestions for how it could be used. It is the sixth in a series of articles aiming to offer practical guidance to healthcare educators.[1]

PURPOSE AND PRINCIPLES OF EVALUATION

The challenges facing medical programmes in sub-Saharan Africa range from issues with resources to changes in political priorities,[2] but these are not exclusive to this region.[3] All programme leaders have a responsibility to make sure that they are providing the best educational experience possible, and one that meets the needs of their local situation. Programme evaluation is a continuous process that involves exploring a programme's attributes from several perspectives, encompassing the quality of the education offered, and the extent to which it meets its wider role in its community.[4]

Evaluation models provide a framework to analyse the effectiveness of a programme using a number of standardised metrics. For example, they may focus on aspects such as the quality of the students' learning experience, the validity and reliability of assessment approaches, or how well the curriculum delivers the desired outcomes.

There are a number of established evaluation models used across a range of professional environments[5] but the model developed by Kirkpatrick in 1959[6] is one of the most widely used, particularly in medical education.[7] Kirkpatrick's framework has four levels for evaluation, namely reaction (the learner experience), learning (the quality of the learning), behaviour (to what extent the learning is applied), and results (the wider impact of the learning).[8] Whether Kirkpatrick's levels are fit-for-purpose for evaluations within education has been widely debated[7] with some arguing that it is too simplistic for something as complex as medical education.[9] Others have developed the model further in search for a better alignment, including Kaufman, in 1994, who considered that Kirkpatrick's model gave insufficient emphasis on the influence of the teacher, and the programme's impact on the institution and wider community. He proposed a six-level model based on the original, to address these issues.[10]

KAUFMAN'S MODEL OF LEARNING EVALUATION

It is important to recognise that while the organisation of Kaufman's model (Figure 1) suggests a hierarchy of the different levels, they can be approached in any sequence. There may be some aspects considered within one level that also align closely with another. It is only when the conclusions drawn from each level are brought together that the overall evaluative picture will emerge.

Level 1 - Input

Focus: Programme resources and their impact on the learning experience

Suggested areas to consider

Arguably the most important resource to be considered is the quality of the teaching team. Students will participate in the programme for a relatively short time, but the teaching staff could be involved for decades, and so have a significant influence on the programme at every level of the Kaufman Model. The programme's management not only needs to recruit and retain inspirational teachers, but they also need to ensure that they are used where they are most effective.

- Do staff have the expected qualifications and experience?
- Do staff teach within their specialisms?
- Are staff active in improving their own teaching?
- Is the staff to student ratio appropriate?
- How good is staff morale and teamwork?
- How long do staff tend to stay on the programme?
- When we recruit, what is the quality of the candidates?

It is also important to review the materials used during teaching sessions, and the resources that are available to support students during self-study.

- Are there sufficient training materials in quantity and quality to meet the needs of the teaching?
- Do the materials meet the learning objectives, and encourage active learning?
- Are the materials planned around student needs and interests?
- Are the materials accessible for all students?
- Is the content relevant to local needs and contexts, including clinical priorities, social and economic factors?
- What are the strengths and limitations of the materials?

Level 2 - Process

Focus: Teaching

Suggested areas to consider

The students' views of their own learning experience are a widely used, valuable and cost-effective source of information, but we should be aware that cultural and hierarchical traditions may limit the sharing of some views. There is some debate about whether students have the experience and perspective necessary to make considered value judgements about their teachers,[11] but others argue that it plays an important part in establishing the overall picture.[12,13]

- What do students feel about their method of instruction, formative assessment strategies, and mentoring?

Observations of teaching by line managers can provide valuable information about the quality and effectiveness of practice, but they must be organised carefully to avoid undermining the teacher's confidence. Different observers will evaluate different criteria unless careful guidance is used to direct their attention to criteria that reflect the programme's values. A single observer will provide their own subjective viewpoint, and a single observation provides only a snapshot of reality, but thoughtful planning (for example, using two observers) can limit these biases.[13]

- Is the teaching approach (for example, a formal lecture or active learning) appropriate to the context, and does it meet institutional standards?
- Is the subject matter appropriate to achieve the learning objectives?
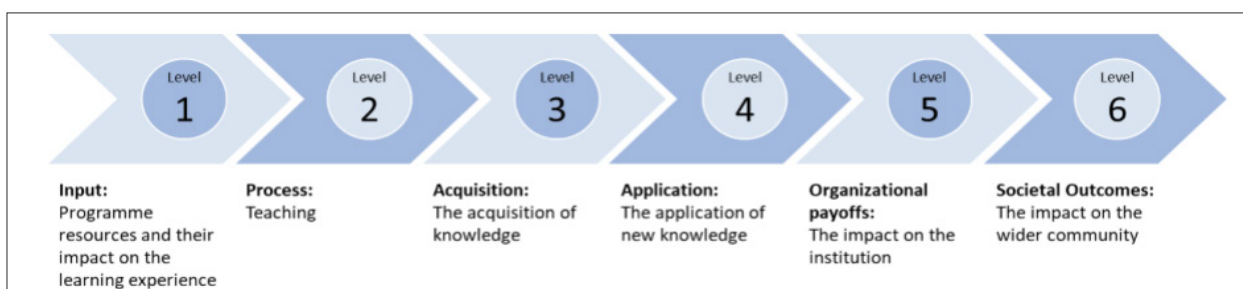- Are the particular learning needs of the students considered?

Figure 1. Kaufman's Model of Learning Evaluation (Credit: Gwyndaf Roberts).

- Has assessment data informed the planning of the session?
- Are there gaps in staff training that should be addressed?

Level 3 - Acquisition

Focus: The acquisition of knowledge

Suggested areas to consider

At its most straightforward, this aspect can be measured using the summative assessment data at the end of each set of formal exams. Set against established professional standards, these will help you judge how successful the teaching team have been at meeting the programme's learning objectives.

- Are there sub-groups of students that share areas of weakness?
- Are there aspects of the programme that have underperformed and might need further resources?

It is important to remember that effective teaching is supported by an assessment strategy that monitors the student's progress. Careful monitoring of progress will help identify if additional support is needed by individuals or groups of students, or whether a teaching approach needs to be adjusted.

- How well is the student doing compared to where they started?
- Can you identify a student, or a group of students, that need more support?
- Are there differences between students from different backgrounds, between the genders, or other sub-groups?
- Are there categories of student who need extra support?

Level 4 - Application

Focus: The application of new knowledge

Suggested areas to consider

The early years of a medical programme require our students to take in vast amounts of information, but it counts for little if they are unable to apply this knowledge later in their training. The teaching programme should be relevant to the realities of their future practice.

- Do the students apply their knowledge in the clinical environment?
- Do they have the confidence to contribute to workplace discussions?

The Hidden Curriculum is the name given to the unspoken rules and values expected within the professional environment that are learnt informally by students. As with all learning, the rate at which they display these attributes may depend on their background and past experiences, but they can be helped by effective role modelling from teaching and clinical staff.[14]

- Is the students' professional behaviour developing at the expected rate?
- Do they interact with other healthcare workers and patients appropriately?
- Do they conduct themselves professionally away from patients?
- Do staff act as good role models?

Level 5 - Organisational Payoffs

Focus: The impact on the institution

Suggested areas to consider

There are many different elements of a programme that contribute to building an institution's reputation and collective pride, both of which are important to sustain the motivation to deliver high quality education. It may be possible to compare your current evaluation with a previous one, to recognise improvement. It may also be possible to compare the evaluation with those of similar sub-Saharan institutions, to recognise comparative strengths, and to identify weaknesses that need to be addressed. A continually improving programme will lead to a vibrant learning community that will help to develop and maintain a positive institutional ethos.[15]

- How do student outcomes compare to similar institutions, and what can be learned to improve outcomes?
- How much of a problem is academic misconduct?
- What links have staff established with other institutions?

Level 6 - Societal outcomes

Focus: The impact on the wider community

Suggested areas to consider

Every medical education programme exists within the context of its own local and national healthcare priorities. How well it addresses these can influence its reputation and long-term success.

- How well is the programme, and the opportunities and experiences it provides for students, aligned with national healthcare realities and priorities?
- Do the programme's graduates take up posts that serve the healthcare needs of the country?

- Do the programme's graduates continue to develop their clinical skills, and teach others, in the postgraduate setting?

USING THIS FRAMEWORK

This framework can be adopted and adapted to meet your needs and context. The goal is improvement. The process is a rolling schedule of evidence collection, review, and improvement action. This can be organised at the level of the individual programme, department, or institution. Your particular context may prompt you to ask different or additional questions relating to each of the six levels described. Having settled on the right questions to ask, the next step is to determine what evidence is available to answer those questions, and to organise how and who to collect it. Reviewing the evidence will allow you to more fully understand your strengths and weaknesses, and to set priorities for action planning, informed by your knowledge of education and by the good practice of others in similar contexts. Reference to earlier articles in this series[1] may help inform your action planning.

CONCLUSION

While there are a number of established evaluation frameworks[5] available to help programme leaders, we have described one that reviews the involvement of all stakeholders involved in medical education. Programme evaluation should be an ongoing process, where different aspects of the framework are updated as new data is collected. Programme leaders will need to compile a timetable for data collection, and they will be responsible for bringing the different aspects together to establish the overall picture. However, it can often be an advantage to delegate the responsibility for the data collection and its interpretation, so that there is a wider sense of ownership of the overall judgement.

Excellence maybe the ultimate goal, but along the way an effective evaluation process will help embed a culture of continuous improvement. It provides the opportunity to identify and celebrate achievement, and provides a structure to identify and examine areas that require further development. Some of the questions we must ask, and the answers we uncover, are difficult and challenging. However, an honest and comprehensive evaluation process will help a medical education programme stay responsive to the needs of the community it serves.

References

1. Bregazzi R. What do teachers do? A framework for improving teaching in healthcare education, South Sudan Medical Journal 2022;15(2):80-81

2. Gibbs T. Medical education in Africa: Not always a level playing field. Medical Teacher 2007;29(9–10):853–4. https://doi:10.1080/01421590701814260

3. Majumder AA, D'Souza U, Rahman S. Trends in medical education: challenges and directions for need-based reforms of medical training in South-East Asia. Indian J Med Sci. 2004;58(9):369-80

4. Fink, A. Evaluation Fundamentals: Insights into the Outcomes, Effectiveness, and Quality of Health Programs. 2nd ed, SAGE Publications, Inc., Thousand Oaks, CA;2005:38

5. Preskill HS, Russ-Eft DF. Evaluation Models, Approaches, and Designs. In: Building evaluation capacity: Activities for teaching and learning. SAGE Publications, Inc., Thousand Oaks, CA;2016:102–81. https://doi.org/10.4135/9781412983549

6. Kirkpatrick DL. Techniques for evaluating training programs. J ASTD 1959;13:3–9.

7. Lovato C, Peterson L. Programme Evaluation in: Swanwick T, Forrest K, O'Brien BC, editors. Understanding Medical Education: Evidence, theory, and Practice. John Wiley & Sons, Inc., Hoboken, NJ;2019:443–55.

8. Kirkpatrick DL. Evaluating Training Programs: The Four Levels. Berrett-Koehler Publishers, San Francisco;1994.

9. Yardley S, Dornan T. Kirkpatrick's levels and education 'evidence.' Medical Education 2011;46(1):97–106. https://doi:10.1111/j.1365-2923.2011.04076.x

10. Kaufman R, Keller JM. Levels of evaluation: Beyond Kirkpatrick. Human Resource Development Quarterly 1994;5(4):371–80. https://doi:10.1002/hrdq.3920050408

11. Vanacore A, Pellegrino, MS. How Reliable are Students' Evaluations of Teaching (SETs)? A Study to Test Student's Reproducibility and Repeatability. Soc Indic Res. 2019;146,77–89. https://doi.org/10.1007/s11205-018-02055-y

12. Felder RM, Brent R. How to Evaluate Teaching. Chemical Engineering Education 2004;38(3),200–202. https://www.engr.ncsu.edu/wp-content/uploads/drive/1Ga-yVaVS1qj87QavBs0igWCosXnOgY-L/2004-r_teacheval.pdf

13. Berk RA. Survey of 12 Strategies to Measure Teaching Effectiveness. International Journal of Teaching and Learning in Higher Education 2005;17(1):48–62. Retrieved May 22, 2023, from https://www.isetl.org/ijtlhe/pdf/IJTLHE8.pdf

14. Lempp H, Seale C. The hidden curriculum in undergraduate medical education: Qualitative study of medical students' perceptions of teaching. BMJ 2004;329(7469):770–3. https://doi:10.1136/bmj.329.7469.770

15. Mora J. Governance and management in the New University. Tertiary Education and Management 2001;7(2):95–110. https://doi:10.1080/13583883.2001.9967044