

## Medical education research: a vibrant community of research and education practice

Cees P M van der Vleuten

**OBJECTIVES** Medical education research is thriving. In recent decades, numbers of journals and publications have increased enormously, as have the number and size of medical education meetings around the world. The aim of this paper is to shed some light on the origins of this success. My central argument is that dialogue between education practice (and its teachers) and education research (and its researchers) is indispensable.

**REFLECTIONS** To illustrate how I have come to this perspective, I discuss two crucial developments of personal import to myself. The first is the development of assessment theory informed by both research findings and insights emerging from implementations conducted in collaboration with teachers and

learners. The second is the establishment of a department of education that includes many members from the medical domain.

**CONCLUSIONS** Medical education is thriving because it is shaped and nourished within a community of practice of collaborating teachers, practitioners and researchers. This obviates the threat of a fissure between education research and education practice. The values of this community of practice – inclusiveness, openness, supportiveness, nurture and mentorship – are key elements for its sustainability. In pacing the development of our research in a manner that maintains this synergy, we should be mindful of the zone of proximal development of our community of practice.

*Medical Education* 2014; 48: 761–767  
doi: 10.1111/medu.12508

Discuss ideas arising from the article at  
[www.mededuc.com 'discuss'](http://www.mededuc.com/discuss)



Department of Educational Development and Research,  
Maastricht University, Maastricht, the Netherlands

*Correspondence:* Cees P M van der Vleuten, Department of Educational Development and Research, Maastricht University, PO Box 616, Maastricht 6200 MD, the Netherlands. Tel: 00 31 43 388 6725; E-mail: [c.vandervleuten@maastrichtuniversity.nl](mailto:c.vandervleuten@maastrichtuniversity.nl)

*Editor's note: This article is published as part of our celebration of the careers of those individuals who have won the Karolinska Institute Prize for Research in Medical Education.*<sup>1,2</sup>



*Dr. Cees P M van der Vleuten, Professor, Chairman of the Department Health, Medicine and Life Sciences, Educational Development and Research, Maastricht University, The Netherlands*

---

## INTRODUCTION

Research in medical education is thriving. Despite the explosion of international journals in the domain (I lost count after 20 or so), all medical education journals continue to be swamped with research manuscripts. Acceptance rates have dropped to below 15% or 10%, whereas impact factors are steadily moving upwards. Medical education meetings are held all over the world, attended by ever-increasing numbers of participants. The annual meeting of the Association for Medical Education in Europe (AMEE), for example, consistently attracts over 3000 participants every year. What lies behind the success of all of this? I do not claim to have a comprehensive explanation, but in this paper I will illuminate my perspective on research in medical education and the reasons for its success.

In his overview of the history of medical education research, Geoff Norman identified three generations of researchers in medical education, who varied in the degrees of their preparedness for conducting education research.<sup>3</sup> The first generation came to medical education in the 1950s and 1960s, from a mix of academic backgrounds, and relatively naïvely began to conduct research in medical education. The second generation was better prepared. They had degrees and doctorates in parent sciences (e.g. psychology, sociology and general education). They brought different theories and methodologies from their training backgrounds and pursued scholarly careers in medical education. As a result, medical education research professionalised. The third generation is trained within the medical

education community. They typically have backgrounds in the health sciences, and master's degrees and doctorates in medical education.

I am a member of the second generation. As an undergraduate psychology student I developed a fascination for research. After graduating in clinical psychology – the field that had attracted me to psychology in the first place – I worked for a year in a clinical environment. Over the course of that year, I became disenchanted with this line of work because I was disappointed to find that many regular clinical activities lacked a sound basis in research evidence. So I decided to go back to university to pursue my second love, and obtained a degree in the science-oriented area of personality psychology and psychometrics, which was all about measuring human qualities. In 1982, by sheer coincidence, I found myself in medical education at a university that had implemented a fully problem-based education programme. It was a thrilling environment in which everyone participated in lively and often heated discussions about education. The debate was highly rhetorical though, based on ideas and beliefs about problem-based learning. Once again, I found myself in a working environment in which rhetoric trumped evidence. Having been well prepared for scholarly work, I ventured into medical education research by investigating education practice at our university (my first work was around objective structured clinical examinations [OSCEs]). The resulting evidence was subsequently used to change or adapt existing practices, and the work was carried out in close collaboration with the teachers who were responsible for those practices. The debate and the dialogue continued with undiminished fervour, only now there was (a tiny bit of) evidence in the discussion. Such dialogue between teachers engaged in education practice and education researchers has shaped (and even today continues to inspire) the professional mission that has given direction to my scholarly life in medical education and the choices I have made. In this reflective article, I will highlight two specific examples that illustrate this interaction and draw some conclusions with regard to what, in my view, makes medical education successful.

---

## ASSESSMENT IN EDUCATION

My psychometric baggage made it relatively easy for me to engage in assessment research in medi-

cal education. Existing methodology was transferred to an emerging field. Medical education was climbing Miller's pyramid<sup>4</sup> by striving for more authentic assessment, with relevance and connections to the real world. In the 1970s and 1980s, standardised performance assessment was introduced with the OSCE in the UK, and standardised patient (SP)-based testing was initiated in the USA. The necessity of dealing with multiple sources of variance in these quite complex assessment instruments fit with and extended my familiarity with generalisability theory. My most cited paper was that I wrote with Dave Swanson on SP-based testing and OSCEs.<sup>5</sup> At that time validation research was pretty classic and consisted largely of correlational research. We made the discovery that (reliable) test scores predicted other test scores, and for me, coming from personality research, this was nothing short of a revelation. While the conveyor belt of publications was running, I was also responsible for the assessment programme of our medical school. In other words, the dialogue and debate between theory and practice represented a personal experience that continued as enthusiastically and vigorously as ever, rapidly enlightening me to the finding that assessment in education was something completely different from the validation of psychological tests.

Testing in education serves a purpose that differs from that of psychological tests in that reliability and validity are only small pieces of a much larger puzzle. This realisation is reflected in my first assessment overview, which I wrote in 1996, and in which I introduced additional criteria for assessment, such as acceptability to learners, costs and the educational impact on the education system, the learner and the teacher.<sup>6</sup> Other criteria were later added.<sup>7</sup> Being in the midst of education practice, I was able to observe first-hand that any assessment involved a compromise in all assessment criteria. This resulted in what is now often presented in the literature as the 'van der Vleuten formula': utility of assessment defined as the multiplicative relationship between different assessment criteria. The purpose and the context of assessment determine which compromise will best fit a concrete situation. For example, if teachers or students do not accept a certain assessment method, the utility of the method is bound to be low, irrespective of the strength of the evidence to support its validity. Acceptability, learning impact and resources are only marginally addressed in publications and handbooks on assessment. Yet, in

the reality of education practice, we overlook their impact at the peril of dysfunctional assessment.

That is why during the remainder of my assessment career I have chosen to work and struggle on the educational side of assessment (which does not imply that I disengaged from psychometric studies). In doing so, my ideas around educational assessment have developed along two lines: formative assessment, and programmes of assessment.

In reference to the first line of thought, many publications have reiterated the adage that 'assessment drives learning'. More often than not, however, assessment can be a driving force for undesirable types of learning.<sup>8</sup> In many assessment regimes, learners peak from test to test. Feedback is scarce. Test content is hardly ever disclosed to learners after test administration, and performance information is mostly restricted to grades. In the workplace, feedback for learners is equally scarce.<sup>9</sup> Although we may have extensively validated the contents of our assessment toolbox, the prevailing summative culture favours a reductionist approach to learning. Albeit that such an approach may fit quite well with an (out-dated) mastery learning-based model of education, it is no longer relevant to modern training programmes.

In recent years we have begun to more broadly study the learning aspect of assessment. One model explaining how assessment impacts learning has been proposed and validated,<sup>10</sup> and feedback in the clinical workplace has been extensively studied.<sup>9,11-13</sup> The results tell us when feedback from assessment is likely to be or not to be used by learners. Scaffolding feedback by educational measures that stimulate learners to use the feedback appears to be a crucial condition. An example of a scaffold is to have in place a mentoring system in which feedback is analysed and discussed. The same holds for social relationships, learning climate and a feedback culture. The summative nature of many of our assessment approaches, however, creates barriers rather than incentives for learners to make use of the feedback they receive.<sup>14</sup>

The resulting dialogue with workshop participants and training programmes has shaped education practice and fuelled new research. It has even inspired some teachers to engage in education research by joining research teams or by setting up research projects of their own. Although I am a professed advocate of competency-based education, I am not blind to its many downsides.<sup>15</sup> Despite these,

it is an approach that challenges us to think beyond the cognitive domain. It encourages us to move towards teaching and assessing performance in the real, non-standardised world of 'messy' education practice, in which assessment cannot rely exclusively on simple grades, but can benefit from the use of words. Real-world assessment compels us to look beyond our traditional methodologies and statistical approaches and explore new avenues.<sup>16,17</sup> To me, these are fascinating times which will make assessment more meaningful to learning. This is the time of 'learning that drives assessment'.

From the insight that any individual assessment is a compromise, and therefore suboptimal, emerged the notion of combining individual assessment methods into a purposeful arrangement. In other words, a shift is proposed from individual assessments to a coherent programme of assessment.<sup>18,19</sup> Such a programme allows us to compensate for the inevitable shortcomings of any individual assessment method or single data point by combining information across many data points. From this perspective, we are dealing with an optimisation problem: where do we optimise what in order to improve an assessment programme? In my view, each individual data point is not sufficiently robust to justify high-stakes pass/fail decisions. It thus follows that individual data points should be optimised to facilitate learning and the provision of feedback that is meaningful to learners. Information from combined data points, however, will be robust enough to inform pass/fail decisions. With the right checks and balances,<sup>20</sup> this can lead to an assessment approach that optimises both the formative and the summative aspects of assessment. Programmatic assessment was first proposed as a theoretical framework,<sup>21</sup> based on a set of assessment principles derived from earlier research<sup>22</sup> and scaffolded by a set of guidelines for how to put the theory into practice.<sup>23,24</sup> Today, it is implemented in a number of places<sup>25,26</sup> and the first scientific evaluations have been published.<sup>27</sup> Although this account of the emergence of programmatic assessment may give the impression that it has a sound scientific grounding, in reality it began with the trying out of vague ideas in assessment practices, such as occurred in the Maastricht University graduate-entry programme and at Cleveland Clinic Lerner College of Medicine.<sup>25</sup> It was these early practices that modelled the theory, rather than the other way around. Obviously, the 'vague ideas' that were tentatively used in practice had their origins in earlier insights derived from assessment research findings.

I hope to have made it clear through these reflections that, to my mind, education practice and research can and should be inextricably interwoven. Personally, I find inspiration in both. For example, implementing programmatic assessment requires intense dialoguing with teachers and students. As a researcher, I find as much inspiration in discussions with practitioners in the course of implementing and monitoring how things function in practice as I do in the outcomes of the ongoing research endeavour.

As an administrator, I have similarly tried to establish contexts that enable such dialogue rather than allowing each group to retreat from the battlefield to its own ivory tower.

---

#### SETTING UP A DEPARTMENT

In 1996, I was appointed department chair. The remit of the department is teaching, research, education development and service to all education programmes of the Faculty of Health, Medicine and Life Sciences at Maastricht University. True to my philosophy that we should strive for interconnectedness of research and practice, I felt that education development and research should be very close to the domain of health and medicine and to the shop floor of the education enterprise. I therefore recruited many staff members from the education domain (medical doctors, biologists, basic scientists), in addition to educationalists and psychologists who were more grounded in the research domain. As a result, we became embedded in both realms as firmly as was possible.

Unfortunately, this strategy got me into major trouble. The success of the department and of the faculty's education programmes – the medical training programme invariably ranks highest in league tables in the Netherlands – caused the leaders of our university to copy it to other faculties. At one point it was suggested that my department should leave its organisational anchoring place within the medical faculty to serve all faculties. Because I felt that this would jeopardise our effectiveness, which depended on our being embedded in one faculty, I refused. The repercussions for our internal reputation were severe and lasted for years. This period was a difficult one for my department and for me personally, but with hindsight I am happy to have stuck to my convictions. I strongly believe that our embedding in the educational and clinical domains we were established to support, with our feet in the mud of practice, has contributed to

the forging of strong and respectful relationships among those working in each realm and has been a key factor in our success.

A number of years later, we submitted our research activity to an audit by an international panel. One member of the panel, Brian Hodges, said: 'You are awfully productive, but it is a lot of the same.' This inspired me to broaden the disciplines in the department. We now have staff members with backgrounds in statistics, cultural science, mathematics, computer science, economics and even econometrics. The mixture of all these disciplines has had tremendous benefits in both research and development. Regardless of an individual's background, however, all staff members are expected to make connections with the professionals and teachers working in the domain. In my view, this interdisciplinary mix and domain embedment are essential to successful innovation, as well as to successful research.

---

## CONCLUSIONS

In the field of general education a recurring topic of debate concerns the gap between education research and education practice.<sup>28</sup> One frequent observation is that research is disengaged from practice. Researchers and teachers are living in separate worlds. Consequently, research evidence has little impact on practice and vice versa. Teachers and researchers speak different languages and it seems almost inevitable that these two worlds will never meet. This state of affairs is deplorable. I would argue, however, that this condition bears no resemblance to the ambience of medical education. In my view, the situation in medical education is quite the opposite and this is one of the secrets of its success. I would argue that it is the mix of teachers, education practice and research we have nurtured that makes us what we are today. We have created a community of practice with tight bonds between practice and research, between teachers and researchers, and between research production and implementation in practice. Teachers are active participants in this community, many as users of research information, others as scholars contributing to it. We have created a lasting and mutually stimulating bond between educational innovation and research. With many barriers to be surmounted, change in education is slow to evolve,<sup>29</sup> but within my career in medical education I have seen a tremendous professionalisation. I predict this will continue as it goes without saying that our community of practice provides us with easy access to research data, and these

research data focus on and are collected by the very people who are positioned to use the lessons learned. The bringing together of education research and education practice is the heartbeat of medical education and the source of its success. We should therefore cherish this community and ensure its sustainability. We should also realise that this goal has important consequences for both the research and the researchers in medical education.

In order to sustain a vital community, we must welcome novices. We need to embrace an open, welcoming and supportive culture. This means that senior members should be responsive to junior members and should actively seek out and invite new members. We should always be on guard against our own retreat into an ivory tower of scholarship that is disconnected from the field of real action. We should nurture, mentor and guide novices to become education experts, using all our knowledge about how learners learn best. We should engage them, motivate them and challenge them in steps of increasing complexity, should provide them with plenty of feedback in the context of safe and meaningful relationships with other members in the community, and should gradually decrease our support to allow learners to become independent professionals. In this learning trajectory we should explicitly include the values of a community of openness and mutual support. In this way we will build on and strengthen our community.

There is another consequence of promoting such close collaboration between knowledge producers and knowledge users (and of blurring the boundary between the two) that I believe may not be easily shared by others. In my view, it is important that the academic community should develop at a pace that makes it attractive for practitioners to stay connected to the research community. In recent years, there has been much more emphasis on theory in medical education.<sup>30</sup> Some journals consider a theoretical orientation a prerequisite for the acceptance of papers. I agree that theories are the cement of education and that they help to explain why certain effects are found under certain conditions. However, I also see how novices in research, particularly those from the health sciences, struggle to use our theories or read and understand our publications. They have been nurtured in quite a different epistemological and methodological climate. To illustrate this, I cite the following anecdote: I was once asked by a clinician (with a degree in medical education) to translate the messages from a paper written by a sociologist in one of our medical education journals. He said he did not understand one word of it. Again, I strongly

favour the use of theory, but I wish to point to the mindful reconciliation of our advancements with the 'zone of proximal development' of our community.<sup>31</sup> We should beware of marching too far ahead of the troops if our aim is to keep everyone engaged in an optimally functioning unit.

Before I conclude these reflections on medical education as a special community, a couple of disclaimers are in order. I am not suggesting that research must always be practice-based or should always have practical implications. I am very much in favour of fundamental research and of testing theories in laboratory experiments. Nonetheless, I also think we should be mindful of preserving our connection to the medical education community at large. We should be mindful that theoretical and methodological rigour can lead to arrogance and may intimidate and discourage precisely those who need to use the evidence. I also have nothing against psychometrics. Psychometrics provides a great toolbox with which to tackle measurement issues, but when reality is more complex than our models can handle, we should not mould reality to fit the model. I have been involved in accreditation procedures in which the audit team looked simply at whether Cronbach's alpha for course-related assessments exceeded 0.80, a measure of horrible simplicity that will ultimately harm the quality of our education.

I am very grateful to have witnessed the growth and success of our medical education community. I think it is phenomenal and quite unique in comparison with other professional domains. In a recent editorial, Geoff Norman described what he called the 'third wave in health sciences education'.<sup>32</sup> I would prefer to see it as a potential fourth generation consisting of research scholars from outside medical education who join our community. I will give them a warm welcome, but I will simultaneously hope they will grasp and embrace the bigger picture of medical education.

---

*Acknowledgements:* none.

*Funding:* none.

*Conflicts of interest:* none.

*Ethical approval:* not applicable.

---

## REFERENCES

- 1 Eva KW. Lessons learned by those who were prepared to greet chance head on. *Med Educ* 2014;**48**:738–9.
- 2 Eva KW. The shoulders of giants. *Med Educ* 2011;**45**:760–1.
- 3 Norman G. Fifty years of medical education research: waves of migration. *Med Educ* 2011;**45**:785–91.
- 4 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;**65** (Suppl): 63–7.
- 5 van der Vleuten CPM, Swanson D. Assessment of clinical skills with standardised patients: state of the art. *Teach Learn Med* 1990;**2**:58–76.
- 6 van der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract* 1996;**1**: 41–67.
- 7 Baartman LKJ, Bastiaens TJ, Kirschner PA, van der Vleuten CPM. The wheel of competency assessment. Presenting quality criteria for competency assessment programmes. *Stud Educ Eval* 2006;**32**:153–70.
- 8 Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP. The mechanism of impact of summative assessment on medical students' learning. *Adv Health Sci Educ Theory Pract* 2010;**15**: 695–715.
- 9 Watling C, Driessen E, van der Vleuten CPM, Lingard L. Learning from clinical work: the roles of learning cues and credibility judgements. *Med Educ* 2012;**46**: 192–200.
- 10 Cilliers FJ, Schuwirth LW, Herman N, Adendorff HJ, van der Vleuten CP. A model of the pre-assessment learning effects of summative assessment in medical education. *Adv Health Sci Educ Theory Pract* 2012;**17**:39–53.
- 11 Daelmans HE, Hoogenboom RJ, Donker AJ, Scherpbier AJ, Stehouwer CD, van der Vleuten CP. Effectiveness of clinical rotations as a learning environment for achieving competences. *Med Teach* 2004;**26**:305–12.
- 12 Sargeant J, Mann K, Sinclair D, van der Vleuten C, Metsemakers J. Challenges in multisource feedback: intended and unintended outcomes. *Med Educ* 2007;**41**:583–91.
- 13 Watling C, Driessen E, van der Vleuten CPM, Vanstone M, Lingard L. Understanding responses to feedback: the potential and limitations of regulatory focus theory. *Med Educ* 2012;**46**:593–603.
- 14 Harrison CJ, Könings KD, Molyneux A, Schuwirth LWT, Wass V, van der Vleuten CPM. Web-based feedback after summative assessment: how do students engage? *Med Educ* 2013;**47**:734–44.
- 15 Morcke AM, Dornan T, Eika B. Outcome (competency) based education: an exploration of its origins, theoretical basis, and empirical evidence. *Adv Health Sci Educ Theory Pract* 2013;**18**:851–63.
- 16 Schuwirth LWT, van der Vleuten CPM. Changing education, changing assessment, changing research? *Med Educ* 2004;**38**:805–12.
- 17 Govaerts MJ, van de Wiel MW, Schuwirth LW, van der Vleuten CP, Muijtjens AM. Workplace-based assessment: raters' performance theories and constructs. *Adv Health Sci Educ Theory Pract* 2013;**18**: 375–96.

- 18 van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. *Med Educ* 2005;**39**:309–17.
- 19 Schuwirth LW, van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach* 2011;**33**:478–85.
- 20 Driessen E, van der Vleuten CPM, Schuwirth L, van Tartwijk J, Vermunt J. The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study. *Med Educ* 2005;**39**: 214–20.
- 21 van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, van Tartwijk J. A model for programmatic assessment fit for purpose. *Med Teach* 2012;**34**:205–14.
- 22 van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol* 2010;**24**:703–19.
- 23 Dijkstra J, van der Vleuten C, Schuwirth L. A new framework for designing programmes of assessment. *Adv Health Sci Educ Theory Pract* 2010;**15**:379–93.
- 24 Dijkstra J, Galbraith R, Hodges BD, McAvoy PA, McCrorie P, Southgate LJ, van der Vleuten CP, Wass V, Schuwirth LW. Expert validation of fit-for-purpose guidelines for designing programmes of assessment. *BMC Med Educ* 2012;**12**:20.
- 25 Dannefer EF, Henson LC. The portfolio approach to competency-based assessment at the Cleveland Clinic Lerner College of Medicine. *Acad Med* 2007;**82**:493–502.
- 26 Driessen EW, van Tartwijk J, Govaerts M, Teunissen P, van der Vleuten CP. The use of programmatic assessment in the clinical workplace: a Maastricht case report. *Med Teach* 2012;**34**:226–31.
- 27 Bok HG, Teunissen PW, Favier RP, Rietbroek NJ, Theyse LF, Brommer H, Haarhuis JC, van Beukelen P, van der Vleuten CP, Jaarsma DA. Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Med Educ* 2013;**13**: 123.
- 28 Broekkamp H, van Hout-Wolters B. The gap between educational research and practice: a literature review, symposium, and questionnaire. *Educ Res Eval* 2007;**13**: 203–20.
- 29 Bland CJ, Starnaman S, Wersal L, Moorhead-Rosenberg L, Zonia S, Henry R. Curricular change in medical schools: how to succeed. *Acad Med* 2000;**75**: 575–94.
- 30 Gibbs T, Durning S, van der Vleuten C. Theories in medical education: towards creating a union between educational practice and research traditions. *Med Teach* 2011;**33**:183–7.
- 31 Vygotsky LS. Zone of Proximal Development. In: *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press 1987;52–91.
- 32 Norman GR. The third wave in health sciences education. *Adv Health Sci Educ Theory Pract* 2013;**18**: 319–22.

*Received 12 March 2014; editorial comments to author 24 March 2014; accepted for publication 26 March 2014*