The need for evidence in education

C.P.M. VAN DER VLEUTEN, D.H.J.M. DOLMANS & A.J.J.A. SCHERPBIER
University of Maastricht, The Netherlands

SUMMARY *In this article a plea is made to use evidence in education. A remarkable difference in attitude is noted between university staff in their role as scientists in their discipline and in their role as teachers. Whereas evidence is the key to guide scientists in the development of their discipline, evidence on teaching and learning hardly affects their role as teachers. Teaching is, rather, dominated by intuition and tradition. However, particularly in education, intuitions and traditions are not always correct when they are submitted to empirical verification. It even often turns out that our intuitions are not justified or that assumed relations are far more complex. To illustrate the fallacy of our (implicit) intuitions and beliefs, a few of these assumptions are held against the available evidence. Two assumptions related to the learning of students and two assumptions related to the assessment of student achievement are discussed. The illustrations make clear that we do need to use evidence in education, just as we do in any other professional area. Being a professional teacher requires more than being an expert in a content area; it also requires familiarity, use, and perhaps production of educational evidence and theory.*

Introduction

There is a remarkable difference in attitude between university staff as teachers and as researchers. As researchers we critically read the newest literature, we think of new approaches and theories, look for empirical verification and submit our work to the critique of others through rigorous peer review. The scientific attitude lies at the heart of scholarship and is accepted by everyone in the field. We also have clear rules about becoming a researcher. Good researchers are carefully selected and trained before they are allowed to contribute independently to the research. We require degrees, expertise in methodology, a demonstration of scientific ability through output assessment, and so on.

A similar situation holds for our conduct in our professional practice area. For example, as doctors we have defined a long track of training before the profession may be carried out, we have certification procedures, and a system of follow-up training. We follow the literature, we are keen on the latest developments, and replace existing habits by new ones when appropriate.

The situation seems quite different in education. As teachers we seem to have a different attitude. We do the things we do, because that is the way we have been raised ourselves and that is the way it has been done for many years, even centuries. We hardly read the literature on education, or, more appropriately, are not even aware that such literature exists. It is difficult to change things in education, because as teachers we are highly convinced that what we do is appropriate and any challenge to one's convictions is an actual challenge to one's professional integrity. Becoming a teacher requires us to be licensed in a professional area, e.g. in medicine, and that is it. We are assumed to be good teachers, because we are qualified in a professional area. The better we are in that area, the better we are as teachers. Specific didactic training or other educational programmes are not required or, in many cases, even offered. Once we are teachers we have quite some autonomy in deciding what and how to teach. Peer review, quality control, follow-up training—quite common in research activities—hardly exist in education.

We realize that the above picture is drawn in black and white. We also realize that current practice in education is not 'inappropriate' or 'substandard'. The issue in education is not whether it is substandard, but whether it can be improved. There is at least a remarkable difference in attitude to which we would like to draw attention. In this article we would like to make a plea for the awareness and use of more evidence in education. We would like to argue that education is governed by tradition and intuition. We do not see much difference in critically scrutinizing our clinical approaches to treating patients and scrutiny in our approaches to teaching students. We would like to argue that, particularly in education, tradition and intuition can actually be misleading and that empirical evidence often contradicts our suppositions and beliefs. A few areas will be discussed in this article where this is the case. We will start with an apparent assumption as we implicitly make them in our teaching practice and subsequently look at the evidence related to that assumption. The intention is not to review evidence in each of these areas, but merely to illustrate how intuition can fail us in some instances, in order to draw attention to our central argument that it is wise to look for evidence in education. We will discuss two different areas: the learning of students, and assessment of student achievement. In each we will address two implicit assumptions.

Learning of students

Assumption: teaching is learning

Looking at educational practice, the most common method of teaching is having a teacher in front of a group of students. Classroom teaching, lecture-based teaching, is the most widely used method of teaching. The implicit theory is that knowledge is transmitted from the individual who has the knowledge to the group of learners who lack the knowledge, to some extent comparable with filling a bucket with water. Also typical in our approach in educational programmes is to offer the scientific body of knowledge from a number of

Correspondence: C.P.M. van der Vleuten, Department of Educational Development and Research, University of Maastricht, 6200 MD Maastricht, The Netherlands. Email: c.vandervleuten@educ.unimaas.nl.

ISSN 0142-159X (print)/ISSN 1466-187X (online)/00/030246-05

disciplines. We assume that in later professional practice the students are capable of integrating this knowledge and are able to use it in coping with the problems they encounter. However, herein lies the problem: students are actually not able to able to integrate this knowledge. According to Camp (Camp, 1996) people are disappointed with traditional education because too many students memorize, forget, fail to apply or integrate knowledge and resist further learning. By understanding the principles of learning we can understand why this is the case. The question really is whether teaching is similar to learning, or, more appropriately, what learning is about. Let us look at the following example (Van de Wiel, 1997, p.42):

> *An older lady gets easily tired and is sometimes short of breath. In the evening her ankles are swollen. She visits her GP, but is confronted with an advanced undergraduate medical student. Although the student has sufficient knowledge, he cannot immediately come up with the right diagnosis. The student asks the GP for advice. The GP asks the student to explain to the old lady what her problem is and why she has these symptoms. The student explains that the heart cannot adequately pump blood into the blood vessels. He also explains about the oxygen transport in the blood, but he has difficulty explaining why this mechanism leads to the symptoms. The GP gives a fluent explanation in which he easily makes connections between underlying mechanisms and the symptoms of the patient, because his knowledge is encapsulated (condensed and automated) and integrated (your muscles do not receive sufficient blood, due to which you feel easily tired, some blood remains in the heart causing an accumulation of fluid in the lungs causing less oxygen to be absorbed in the blood and you feel tired when you walk up the stairs).*

This example shows that the doctor has a wide knowledge network which is quite flexible. It has knowledge elements that have been aggregated into 'encapsulated' or chained concepts allowing flexible connections between concepts, symptoms and causes. Learning apparently has to do with building up of these semantic networks. Isolated knowledge in itself is of little use. Knowledge needs to be structured through relationships between knowledge elements and concepts in a network. A network therefore consists of a number of related concepts. These networks differ from person to person, depending on the individual's learning experiences and acquired experience. The knowledge networks are constantly being adapted and changed as a result of new learning experiences. New knowledge or experience is assimilated into the network and may reconfigure it. Hence, acquiring knowledge is more than consuming information. To understand and use the information students need to structure, organize and restructure the information. In other words, the learner is the architect of his/her own learning. As a consequence, education should not only provide the information, but should also provide an environment to help build these networks.

A second example involves the way we structure curricula. We usually train conceptual knowledge first with a sequence of basic sciences preceding clinical sciences. Sometimes we offer perceptual and procedural knowledge through skills-training programmes. Essentially we structure education in isolated components. We train knowledge separately from skills and separately from attitudes. It is at the end of the programme that students are (suddenly) required to integrate what they have learned by working with real patients. It is again assumed that the provision of information suffices to be able to apply the information. However, it does not. The cause is that when the information was not learned in (a meaningful) context we have difficulty in accessing that information to be used for a particular situation. Our conventional educational programmes do not teach our students to identify in which situations what knowledge is required.

There is ample evidence that context matters for learning. A classical experiment required scuba divers to learn a number of words in two different contexts (Godden & Baddeley, 1975, Emmerson, 1986). One group learned the words under water and the other group above water. Subsequently, half of both groups changed to the other location and their recall of words was tested. Recall was best when the condition of learning was similar to the condition of recall. Apparently, the situation or the context of learning is relevant for the recall of information. In daily practice we realize and use this, more or less unconsciously. For example, when we forget something we have the tendency to return to the spot where we thought of the original action (back in the kitchen we remember to collect the garbage bin). Another example of the same phenomenon is the difficulty we have in recognizing a person in the supermarket who, afterwards, we realize is a colleague from university. Within the university you would probably not have had any difficulty in recognizing the person instantaneously. These are examples of the context specificity of learning. There is also evidence for processing specificity, i.e. the fact that learned material can only be applied in the way it was learned (Regehr & Norman, 1996). An example is that we all know the letters of the alphabet and we can easily list them. However, we would have great difficulty if we were to list the alphabet in reverse order. Although all the elements are there we are only able to use them in the way they were learned. Acquiring new knowledge in the context of some professionally meaningful problem or situation will lead to more accessible knowledge, because the situational cues that activate the knowledge are stored within the same cognitive structures (Brown *et al.*, 1989).

All these examples illustrate that teaching is far from being equal to learning. Insight into the nature of learning has progressed substantially in recent decades and better insight is gained on how students and doctors acquire and accumulate expertise (Regehr & Norman, 1996). Learning is more than providing information. Learning is facilitated by allowing learning to take place in meaningful contexts, by activating prior knowledge, by requiring the student to engage actively in the learning process, by operating on (similar) knowledge in a number of (different) contexts, by arousing intrinsic motivation, etc.

In traditional curricula the emphasis is on knowledge transfer from teacher to student and is based on a conception where knowledge is considered as 'absolute', based on 'facts' and being 'objective' (Williams, 1992). Knowledge in this conception is the sum of information to which the student has been exposed. Learning is a matter of transfer of 'truths' on what has been scientifically proven. However, from the above, it is clear that this conception is naive.

Current philosophical views on human learning are therefore based on a view in which knowledge is not 'absolute', but is constructed by the learner based on previous knowledge and overall views of the world. Learning is a process that results from interactions with the environment. It is the learner who constructs new knowledge and who is at the centre of the educational process. This view is called constructivism (Savery & Duffy, 1995). From the evidence on learning this theory seems a better view on education than our intuitive naive one.

Understanding of the principles that facilitate learning provides a challenge to think about instructional strategies that take these principles into account. With the learning principles in mind we need to think about how education could be linked to a professional context in order to provide meaning to learning (e.g. by working with authentic problems), about how to activate prior knowledge in students (e.g. by giving assignments), about how to engage students in active involvement (e.g. through debate with others), etc. These learning principles also point to the limitations of the lecture as the dominant format of instruction (although it is not excluded that the lecture will adhere to some of the learning principles) and the effectiveness of others (e.g. small-group work, learning by doing, teaching others, etc.). Most training programmes would look quite differently if more attention were to be paid to the insights of the learning process and expertise development.

Assumption: the more we teach, the more students learn

Looking again at educational practice it can be observed that most curricula are quite full. The week of a student is scheduled with lectures, practicals and other educational events. Some programmes have 25 to 35 hours of scheduled educational activities within a week. Implicitly we assume that all these activities will lead to more learning.

Some studies have actually shown that the relationship between scheduled teaching activities and learning is far more complex (Van der Drift & Vos, 1987, Gijselaers & Schmidt, 1995). For a number of curricula the number of hours scheduled on teaching activities was related to the number of hours on self-study by the student and a regular pattern was found.

Learning activities indeed increase when there are not too many teaching activities planned, but beyond a certain level of planned activities the opposite occurs: learning activities decrease despite an increase in teaching activities. The optimum is reached at approximately 40% of scheduled time for teaching activities, leaving about 60% of self-study time. Achievement scores of students actually increase significantly when this balance is reached (Gijselaers & Schmidt, 1995). An acknowledgement of this relationship would have major consequences for quite a few undergraduate medical training programmes. It shows again that our intuitive beliefs and actions are not always empirically justified.

Student assessment

Assumption: competence consists of distinct competences

The typical view on clinical competence is one where competence is seen as a constellation of a number of different attributes or entities, which are fairly distinct from each other. These attributes are stable across time and situations, and expertise is matter of (monotonous) growth of these attributes. If one has more of a particular attribute, one is able to handle more situations that require this attribute. However, evidence shows that this view is too naive. The assessment of problem solving or clinical reasoning skills is a good (though but one) example to demonstrate this (Swanson *et al.*, 1987). Problem solving has typically been measured by confronting examinees with authentic problems, e.g. patient cases. An examinee had to work through a case and responses to questions, the decisions that were taken, or the pathways that were used were taken as an indication of problem-solving ability. However, research showed that once a score was derived on a particular case, that score turned out to be hardly predictive for a score on another case. Apparently, problem solving was not a generic skill, but highly dependent on the clinical context. Even small changes in the context had a substantial influence on performance. This phenomenon has been called the content specificity problem and is one of the best documented empirical findings in the assessment literature (Van der Vleuten, 1996). Content specificity turned out not to be limited to problem solving, but was found in virtually all entities of competence assessment. Another consistent research finding was also disturbing. When a large enough sample of cases was taken to overcome the generalizability problem caused by content specificity, it turned out that the scores on these problem-solving measures were highly predictive for tests of knowledge using 'simple' multiple-choice questions. Apparently, problem solving was neither a generic skill nor very distinct from knowledge. These research findings had serious theoretical implications in our thinking on the nature of problem solving. Through a wealth of research of a more fundamental nature looking at the cognitive psychological changes while learning and becoming experienced, we now better understand the nature of problem solving. It is beyond the scope of this article to explain these insights and theories (cf. Schmidt & Boshuizen, 1993, Regehr & Norman, 1996), but clearly the view of generic independent skills has turned out to be a false one. The realization of this has many practical implications for assessment (Van der Vleuten, 1996).

Assumption: the curriculum dictates learning

In educational practice we tend to ignore a very strong and lawful relationship between student assessment and student learning. The lawful relationship is that assessment drives learning. From a student's perspective, success is defined by being able to pass the examinations. The student will exert maximum effort to optimize chances of success. Students will do whatever the examination programme tells them to do and they will not do whatever the examination programme does not reward. For the students, the examination programme *is* the curriculum. There is no problem with this lawful relationship if there is a good match between curriculum objectives and the (realized) objectives of the assessment programme. However, this is not unproblematic (Frederiksen, 1984). The question is whether this is really true in many practical situations. Our training programmes strive for high academic competences for our graduates. However, our examination programmes not uncommonly

focus on the measurement of skills of a rather limited nature. Many of our examination programmes consist of series of hurdles, often testing no more than small facts which could only be reproduced correctly when studied a short time before the examination. Typical programmes require examinations to be passed at the end of courses. Often there are multiple courses and multiple examinations running concurrently. Students go from hurdle to hurdle, they postpone learning until just before the examination, and after the examination they 'wipe clean their hard disk' in order to be best prepared for the following examination. When the content of the examination is not readdressed at later stages in the curriculum, the implicit assumption is that the student will remember sufficiently for life from this single (superficial) learning occasion. We know that unused knowledge rapidly decays (Semb & Ellis, 1996). The consequence is that these programmes tend to reinforce short-term knowledge and a surface approach to learning, quite in contrast with the noble academic goals which are set out in the curriculum as such.

Clearly the above assessment illustration is again rather black and white, perhaps even a caricature. However, a fact is that we tend to ignore the lawful relationship between assessment and learning. As teachers we are inclined to engage ourselves primarily with the training programme itself, with the curriculum, and less with assessment. By realizing the relationship between assessment and learning we might invest more in assessment programmes, or, even better, use the assessment strategically to achieve desirable learning outcomes (the tail wagging the dog).

Discussion

The illustrations above demonstrate that tradition and intuition are sometimes misguided when held against an empirical mirror. It is shown that evidence could play a major role in making decisions about the architecture of instructional programmes. There is a lot of evidence available, particularly in the area of medical education. Medical education is actually a unique professional area with at least eight international journals fully dedicated to teaching and learning (*Academic Medicine*, *Medical Education*, *Medical Teacher*, *Education for Health*, *Evaluation & the Health Professions*, *Advances in Health Science Education*, *Health Professional Education*, *Medical Education Online*). It is unfortunate that, with ample evidence available, use of educational evidence hardly plays a role when decisions about medical education are being made (Nelson *et al.*, 1990).

The plea for using educational evidence by no means suggests that we have sufficient understanding of teaching and learning: quite the contrary. Compared with medicine, educational science is still in its infancy. But this is no excuse to dismiss or ignore the available evidence. Furthermore, it is clear that the level of professionalism in educational research is also different from medical research in general. The quality of much of the research produced in education, only in part due to the methodological intricacies of educational research, is of poor quality. However, this again does not discharge us from looking for good evidence and using it. The limitation in the body of evidence on education also implies that we cannot fully rely on evidence in every decision, or that we have consensus about which educational strategy is best. The effect of knowing more about educational theory and empirical evidence probably leads to increased modesty about what we are currently doing in practice. This is, in our conviction, not all too different from any clinical discipline, and again does not provide an alibi for ignoring educational evidence.

Granted the limitations of the current state of affairs in education, there is an available body of knowledge. If we used this body of knowledge, current educational practice would probably change dramatically. We should stop teaching knowledge in isolation and requiring students to cram and memorize for the occasion of the upcoming examination. We should help students to build up flexible networks of knowledge by activating prior knowledge, by stimulating elaborations, by embedding new information in a meaningful context. The learning environment and the application environment should be as similar as possible. We could make use of situations from the future profession to integrate knowledge from different relevant disciplines. The knowledge acquisition and the application should preferably take place at the same time. We should make use of numerous and varied situations in which knowledge should be applied, therefore providing repeated practice in retrieving information. How this is translated into concrete instructional strategies and designs for curricula is not a fixed pathway, and there are probably many ways that lead to Rome. Traditional education, however, seems not well rooted in what we know about effective learning. Moreover, with educational insights in mind, it is not surprising that new instructional methods such as problem-based learning have turned out not to be another modernism or a passing fad (Camp, 1996). It is a way (and only one way) of tailoring the educational process to the learning process. That, indeed, is the challenge.

We are convinced that the governance of tradition and intuition in education is for a large part based on ignorance. Most teachers probably just do not know what is out there. Teacher programmes, faculty development courses, and—not to be forgotten—incentives to engage in these activities are therefore imperative to deliver a change in the attitude of teachers. Being a good teacher requires more than being a content expert. In our view a teacher is someone with a certified competence in a professional area, but a *professional teacher* is someone who bases his teaching activities on a good understanding of what has been done before and what makes sense and what not. Perhaps the highest level of professionalism in being a teacher is to contribute to the accumulation of evidence, to 'problematize' one's own educational situation, to start investigating and to engage in educational research. These teachers will soon find out that the difference between educational research and clinical research lies mainly in the object of measurement.

Notes on contributors

C.P.M. VAN DER VLEUTEN is Professor and Chair of the Department of Educational Development and Research, University of Maastricht.

D.H.J.M. DOLMANS is Assistant Professor in the Department of Educational Development and Research, University of Maastricht.

A.J.J.A. SCHERPBIER is Professor and Director of the Educational Institute, Faculty of Medicine, University of Maastricht.

References

BROWN, J., COLLINS, A. & DUGUID, P. (1989) Situated cognition and the culture of learning, *Educational Researcher,* 18, pp. 32–42.

CAMP, G. (1996) Problem-based learning: A paradigm shift or a passing fad? *Medical Education Online,* 1.

EMMERSON, P.G. (1986) Effects of environmental context on recognition memory in an unusual environment, *Perceptual and Motor Skills,* 63, pp. 1047–1050.

FREDERIKSEN, N. (1984) The real test bias: Influences of testing on teaching and learning, *American Psychologist,* 39, pp. 193–202.

GIJSELAERS, W.H. & SCHMIDT, H.G. (1995) Effects of quantity of education on time spent on learning and achievement, *Educational Research and Evaluation,* 1, pp. 183–201.

GODDEN, D.R. & BADDELEY, A.D. (1975) Context-dependent memory in two natural environments: On land and underwater, *British Journal of Psychology,* 66, pp. 325–331.

NELSON, M.S., CLAYTON, B.L. & MORENO, R. (1990) How medical school faculty regard educational research and make pedagogical decisions, *Academic Medicine,* 65, pp. 122–126.

REGEHR, G. & NORMAN, G.R. (1996) Issues in Cognitive Psychology: Implications for Professional Education, *Academic Medicine,* 71, pp. 988–1001.

SAVERY, J.R. & DUFFY, T.M. (1995) Problem based learning: an instructional model and its constructivist framework, *Educational Technology,* 35, pp. 31–37.

SCHMIDT, H.G. & BOSHUIZEN, H.P. (1993) On acquiring expertise in medicine. Special Issue: European educational psychology, *Educational Psychology Review,* 5, pp. 205–221.

SEMB, G.B. & ELLIS, J.A. (1996) Knowledge Taught in School: What is Remembered? *Review of Educational Research,* 64, pp. 253–286.

SWANSON, D.B., NORCINI, J.J. & GROSSO, L.J. (1987) Assessment of clinical competence: written and computer-based simulations, *Assessment and Evaluation in Higher Education,* 12, pp. 220–246.

VAN DE WIEL, M.W.J. (1997) *Knowledge encapsulation: Studies on the development of medical expertise.* Maastricht: Doctoral dissertation, University of Maastricht.

VAN DER DRIFT, K.D.J.M. & VOS, P. (1987) *Anatomie van een leeromgeving (Anatomy of a Learning Environment).* (Lisse, Swets & Zeitlinger).

VAN DER VLEUTEN, C.P.M. (1996) The assessment of Professional Competence: Developments, Research and Practical Implications, *Advances in Health Science Education,* 1, pp. 41–67.

WILLIAMS, S.M. (1992) Putting case-based instruction into context: Examples from legal and medical education, *Journal of the Learning Sciences,* 2, 367–427.