

Editorial

John MacBeath · Lejf Moos

Received: 2 June 2011 / Accepted: 10 June 2011 /
Published online: 23 June 2011
© Springer Science+Business Media, LLC 2011

We start the journey in this edition in Sweden and end in Wyoming USA by way of the Netherlands and Germany, with a comparative study of the two German states, Hesse and Bremen, conducted by a Swiss based research team. In these very different cultural settings we are struck by the common strands – the tensions between attainment and learning, the qualitative and quantitative, the positivist and the ethnographic, the unfinished stories and the common plea for further research.

The first of these four contributions, Ruhe and Boudreau's paper, begins with programme evaluation in a medical practitioner context but ends up being about much more. The failure to find strong positive correlation data from programme reform leads to a far richer set of insights. The authors write:

The first lesson from our experience, then, is that the process of curricular innovation, educational measurement, and assessment/evaluation is not linear, but an iterative and dynamic work-in-progress. Second, educational measures of cognitive and affective variables are imprecise, often rendering good data elusive. This study has illustrated the difficulty of finding and/or creating a valid instrument aligned to specific program goals.

The problem, one of the researchers' own making, was to use a mode of evaluation that, in their own words, was too 'detached, experimental and positivistic' which, while generating some nuggets of knowledge, was bought at a high price, that is 'the loss of the opportunity to better understand the more subjective and inter-subjective effects of program change'.

One of the most telling insights from this study is the inter-relationship among the formal, informal and hidden curricula. What the evaluation failed to capture with its apparently non significant findings were the subliminal, and often less subliminal,

J. MacBeath
University of Cambridge, Cambridge, UK
e-mail: jecm2@cam.ac.uk

L. Moos (✉)
Danish School of Education, Aarhus University, Copenhagen, Denmark
e-mail: moos@dpu.dk

messages that are conveyed by hospital-based teachers in their relationship with patients, and medical students. ‘We recognized that the clinical (hospital and ambulatory) context, in particular the behavior of clinical teachers, was a powerful determinant’. They identified the incongruence between the expected, explicit, or advertised physicianship behaviours and the quality of physicianship that was actually demonstrated by faculty members. Whatever the teaching and theory the observation of what your teachers actually *do* weighs more heavily, while the ‘walk’ casts a shadow on the validity of the ‘talk’.

The incongruences and contradictions led Ruhe and Boudreau back to the drawing board to seek a more ‘nuanced appreciation’, to hear the unheard, to see the unseen.

Although we embarked upon our protocol for evaluation of the Physicianship Program with a definite bias towards an explanatory and outcomes-based model, we have subsequently come to appreciate the value of a more diverse mix of evidence. As already noted, we have acquired a more nuanced appreciation of the critical importance of the non-formal features of the curriculum, in particular, what is lived and experienced in addition to what is professed and taught. This realization has prompted us to be more receptive to the unexpected and attuned to the complex interplay between teachers and learners.

The authors write that as consequence of new insights they expanded their methodological toolkit and have now launched a multi-case study, using a blend of traditional methods, such as focus groups, and innovative techniques, such as solicited audio diaries. Their longitudinal study which will be conducted over four years has the objective of looking beyond statistically significant outcomes to focus on understanding some of the subtleties of the process by which the formal, informal and hidden curricula can be assessed.

While it may be argued that we would always opt for a sound clinician over one with a good bedside manner, the emotional affect and psychological well-being generated by the latter may count for more than is obvious at first sight. It is, of course not an either/or but students and neophyte medics would clearly benefit from witnessing the synergy between both, just as school pupils ‘catch’ as much from modeling by empathic teachers as they do from the subject matter conveyed to them. After all, acquisition of knowledge is embedded in relationships and the social context which makes it meaningful.

There are immediate resonances with Jaap Scheerens’ study in which the focus is on informal learning in a formal setting. The focus of the learning is citizenship, the goals of which are described as having three dimensions: a cognitive dimension with respect to knowledge about democratic institutions; a pragmatic dimension, in the sense of taking action and gaining experience; and an affective dimension, in terms of an attachment to the societies and communities to which one belongs. Citizenship is learned, not only through explicit teaching but is absorbed through the daily experiences inside and outside classrooms, through teacher modeling and through problem solving in which ‘students are confronted with real life manifestations of the goals of active citizenship education’.

Distinctions are made among the ‘formal’, ‘informal’ and non formal’. The first stands for the hierarchically structured, chronologically graded system, running from primary school through university. *Informal* education refers here to a lifelong

process in which attitudes, values, skills, and knowledge are shaped by day-to-day experiences and influences of family and neighbours, from work and play, from libraries, mass media and other spontaneous or ‘accidental’ sources. *Non-formal* education is defined as taking place outside the established formal system and may include extra-curricular activities, visits, exchanges, voluntary work, campaigns (what in Hong Kong would be termed Other Learning Experiences, or OLE for short). Reference is also made to the Council of Europe’s definitions which, confusingly include the ‘hidden curriculum’ within the informal curriculum.

Scheerens discusses aspects of learning which emerge through the classroom climate and school culture, the implicit and overt values that are promoted, the way in which conflicts are dealt with, the nature of cooperative decision-making, issues of trust and the establishment of clear behavioral rules. The nature of leadership, autocratic or democratic also adds powerfully to the mix. It is in this complex interplay, or ‘force field’, that distinctions between what we understand as the ‘informal’ and the ‘hidden’ curriculum can become easily blurred or problematic.

Where the informal and hidden curriculum meet needs to be clearly explicated. For example, a school may give careful attention to its culture, building congruence and coherence into conventions and structures, into ‘the way we do things round here’. This approach is one which consciously promotes and evaluates the values it seeks to instil in order to avoid the incongruences and contradictions that Ruhe and Boudreau point to. When the walk and talk are at odds and when teaching and learning fail to meet, hidden and unintended learning come into play. A signal lesson that is too often learned in school is parodied by Schultz’s Charlie Brown – ‘what you sign up for and what you get are two different things’. It is a powerful message that characterises too many schools as well as many other social institutions.

It is into this ‘fuzzy reality’ that Jaap Sheerens has attempted to reach a more structured understanding through the development of indicators. This is a field in which Jaap has been a leading international figure for four decades. The main focus of that work has been on attainment and value-added in systems where tests and exams lend themselves more readily to comparative data internationally than this very fuzzy domain. Undaunted, Scheerens concludes that specifying the issues in terms of indicators has the advantage of allowing for a more data driven approach in the further development of citizenship education, while at the same time bringing citizenship out of the cold into the mainstream of measured achievement. This is, as discussed elsewhere in this volume, work in progress with the final caveat, ‘the complexity of the issue would keep requiring qualitative, more in depth description’. While the following aphorism has now gained the status of a cliché, it is no less true for that - ‘We must learn to measure what we value rather than valuing what we can easily measure’.

The value of examination attainment is at a premium but how do students interpret their examination success, or lack of it, and to what extent is that attributed to perceptions of their own ability or to the abilities of their teachers? This is the complex territory into which Oerke and colleagues from the University of Zurich venture, comparing two German Lander in which external exams had been introduced at different periods, so allowing a comparative analysis of students’ attributions for their success or failure. The context for this, in common with the other three papers in this volume, is one which there has been ‘a shift from

curriculum-focused input regulation to output regulation, focusing on the attainment of learning outcomes'. With the impetus from TIMSS and PISA showing wide cross-country variation in student achievement, as well as within decentralised countries such as Germany, it provided the platform for a study, attempting to unravel the impact of external exams on attributes such as self efficacy, self esteem and ability.

The introduction of national achievement standards accompanied by central examinations was seen by German states (Lander) as an economical means of creating incentives to improve student and teacher performance, 'given that many high-achieving states had a long tradition of central exams whereas many low-achieving states did not'. With support from the German Research Foundation (DFG), the Zurich-based research team examined relative exam success against self report questionnaires of students in their last year of school in the German states of Bremen and Hesse. Until that point, it is argued, the direct effect of centralized testing on attributions has not been subject to rigorous analysis, particularly in relation to ways in which achievement is moderated by motivation and self perception.

Drawing on attribution theory, with particular reference to Weiner, the hypotheses to be tested were whether external exit exams produce negative effects on motivation and instruction and the extent to which this might differ from state to state or from subject to subject. Weiner's classification of causal attributions, identified three dimensions; the locus (internal or external to the person); stability (enduring or changing from situation to situation); and control, referring to 'the volitional influence that can be exerted over a cause'- in plainer English, the extent to which people feel they have the power to change things.

It is ultimately problematic to disentangle cause and effect – the extent to which poor performance affects self efficacy and/or the extent to which self efficacy affects performance but that there are differential effects exist depending on perceptions of self and of teachers is clear. In common with other papers in this volume we are left with an fascinating set of issues, an interesting methodology and conclusions which, not unsurprisingly, point to the need for further research.

The significance of teachers' self perceptions and attributions are revisited in the paper by Range and colleagues which concludes that teachers' sense of ownership relies on their buy-in to the premises on which school improvement plans are based and the criteria by which teacher effectiveness is measured. Also to be factored into the equation is the spectre of accountability, with the admonition that 'Accountability should not be feared but welcomed as an endeavor that monitors student achievement and teacher success'. This does beg the question of accountability for what, for whom, and by what criteria? 'Ownership in the improvement plan, write the authors, is vital in supporting long-term growth as opposed to superficial compliance', but by what argument and on what grounds? Can teachers be blamed for superficial compliance in a high stakes context where compliance is the name of the game?

The context is, once again, the key consideration. This is not Germany, Canada, or the seven countries of the Scheerens' study but Wyoming USA where supervision and evaluation of teachers appears to be in a continuous state of flux, where the resolution between formative and summative evaluation is unclear and contested and where 'holistic accountability' may be no more than a pious hope. We are reminded

that ‘principals’ supervisory beliefs and actions are not only influenced by their values, but also by rules which may create a power struggle among beliefs, actions, policy, and statues concerning supervision and evaluation’.

Echoing much that has gone before in this volume, the paper’s conclusion is that the selection of student measures that accurately and fairly assess teacher performance ‘has been problematic’, ‘creating fear in teachers, which inhibits candid conversations about teachers’ concerns in the classroom’, a fear that ‘stymies the inherent goal of the supervision and evaluation process’. Drawing on other studies the authors argue that summative evaluation is likely to undermine collegial, constructive feedback and that teachers are ‘less likely to listen and admit error, thereby eroding the professional growth process’.

The buck stops, unfortunately and unequivocally, with the principal. He/she is charged with not only collecting formative data to support teachers’ professional growth, but also summatively rating their performance. This role dichotomy requires principals to build non-threatening instructional climates and then act as the evaluator when retention or dismissal decisions are made concerning teachers.

‘As Wyoming schools adopt WDE mandates to incorporate student achievement data into teacher evaluations, it is important for both policy makers and practitioners to understand that utilizing student achievement data to rate teacher effectiveness has proven to be a complex issue’. Have we heard this somewhere before? However long and however much we have been revisiting this issue in this journal over the three years of its existence, as the French are wont to say ‘plus ca change, plus c’est la meme chose!’