Assessment and evaluation: incentives, sanctions and power relations

Guri Skedsmo^{1,2} · Stephan Gerhard Huber¹

Published online: 9 August 2018
© Springer Nature B.V. 2018

Assessment and evaluation can represent constructive processes that promote student learning and teacher professionalism if individuals grasp the opportunity for reflection and growth that they bring about (please see Ford in this issue). In this case, the primary purpose of assessment and evaluation is to highlight strengths and uncover deficiencies in encouraging ways that promote learning and development (Scriven 1991). Scholars have also indicated other functions in terms of enabling transparency and providing insights into the performance of students, teachers and the school organisation to control and hold key actors accountable based on various formulated standards (Strathern 2000). Such procedures are sometimes tied to incentives and sanctions to induce desired behaviours. On the one hand, incentives or sanctions could boost performance; on the other, they may have consequences in terms of certain choreographed behavioural patterns adopted by individuals to comply with performance expectations (Webb 2006).

1 Articles in this issue of EAEA 3/2018

In the first article, Zeng, Huang, Yu and Chen present findings based on a critical review of the student assessment literature. By exploring various assessment concepts, such as assessment of learning, assessment for learning and assessment as learning, they discuss how these concepts relate to each other in terms of the inherent formative and summative elements and the possible tensions that work to enhance learning. Further, they explore the potential of an integrated approach, namely, learning-oriented assessment (LOA), which has emerged as an alternative assessment concept

✉ Guri Skedsmo
guri.skedsmo@ils.uio.no

Stephan Gerhard Huber
stephan.huber@phzg.ch

¹ Institute for the Management and Economics of Education, University of Teacher Education Zug, Zug, Switzerland

² Department of Teacher Education and School Research, University of Oslo, Oslo, Norway

combining different components of the earlier approaches. Accordingly, the authors also clarify the historical foundations of this approach, and using this as a basis, they suggest the dynamic framework of LOA, which embraces the arrangements of learning and assessment by focusing on method implementation, mindset-changing and capacity building among the different actors involved. Based on their review findings, they also arrive at the implications for practice, policy and further research in this area.

The distinction between the summative and formative purposes of assessment also lies at the core of the second article, where Ford reports on a study of a teacher evaluation system known as ‘Compass’ in the state of Louisiana, USA. The summative purpose of teacher evaluation can be expressed in terms of holding teachers accountable for the results achieved, which is often measured by the students’ performance on standardised tests. The formative purpose would indicate a system designed to provide information to be used for professional development and to improve instruction. Compass aims to integrate both purposes. The system was applied for the first time in the academic year of 2012–2013 and was implemented state-wide the year after. Ford investigates teachers’ use of information generated to drive ongoing instructional improvements. The current version of the system includes two different components: growth measures of student achievements and teaching observations, which jointly determine the teachers’ overall Compass performance score. On the one hand, the system provides teachers with a framework to identify and track their students’ yearly performance; the teachers are expected to use this information to improve the quality of instruction. On the other hand, if a teacher receives an ‘ineffective’ rating, he or she is subject to a remediation plan and has to achieve an ‘effective’ rating in 2 years. In case this result is not achieved, their teaching certificate is not renewed, and this could result in contract termination. The findings showed that teachers had concerns about the validity of the measures; further, they struggled to make use of the information provided or used it superficially to boost performance scores. Thus, they would require support to improve their capacity to utilise the information. However, the teachers’ expanded locus of control over the data seemed to create better conditions under which to apply the data in the early phase of the trial.

Teacher evaluations linked to economic incentives is also a topic addressed in the third article by Mintrop, Pryor and Ordenes. The authors report on their analysis of six evaluations of the Teacher Incentive Fund (TIF), a federally funded programme in the US that aims to reform teacher and principal compensation systems by providing districts and schools with monetary support to create innovative teacher evaluation systems linked to bonus pay for high performance. The programme was first initiated under the ‘No Child Left Behind’ act and was later retained by the democratic Obama administration. The crux of the article relates to the ways in which evaluators can meaningfully analyse such programmes when the overarching policy agenda and grant requirements are based on a simple and linear logic, which may be at odds with the local context and thereby fail to factor the important aspects of the local implementation reality. The authors analyse the degree to which they manage to capture the complexity of the social system in implementing teacher evaluation and bonus pay systems. Based on this, they introduce a Complex Adaptive System (CAS) approach to programme evaluation, which is designed to explore the unintended consequences and multiple pathways of implementation in the context of local interpretations and adaptations. They apply this approach to their own evaluation study, discussing and comparing

programme logics and the associated implications in terms of the outcomes and conclusions of evaluation studies.

In the fourth article, Zhou, Kallio, Rinne and Suominen present their analysis of reform and change with respect to the school inspection system in China, where inspection procedures can be traced back to the 900s. Since 1977, and concurrent with the introduction of the market economy in China, the school inspection system has undergone several changes. The authors identify three stages of institutional changes with respect to school inspections: the restoration stage, the formalisation stage and the transition stage. The authors describe how the practice of school inspections was restored during the 1980s and became centrally managed by the newly established National Office of Education Supervision, a unit under the Ministry of Education. In the early 1990s, the inspection system expanded, becoming formalised by legislation, the development of procedures and the establishment of education offices. Finally, from 2007 onwards, it transformed further to include evaluation practices, and this development was strengthened by the integration of large-scale assessments in 2015. Interestingly, this analysis demonstrates how inspection authorities and their practices change in response to external pressure, including global ideas and trends as well as internal demands and the needs of institutional self-preservation. It also highlights how the inspection logic changed from emphasising obedience in schools to generating evidence for decision-making.

2 Reflections on perspectives in the contributions

In particular, two interesting themes arise from the four articles described in this issue. The first relates to the interplay between the formative and summative aspects of assessment, especially when assessment practices are linked to consequences for individuals in terms of sanctions and incentives. The second concerns various forms of assessment and evaluation, which place emphasis on matters of power and power relations.

By reviewing the different approaches and how they have developed chronologically, Zeng et al. argues for an integrated approach of student assessment, which places student learning, and how to enhance their learning, at the centre of assessment practices. This approach would entail the active involvement of students and teachers; it also requires that they develop the necessary professional capacity. At the same time, performance-oriented accountability is toned down to a minimum. In contrast, Ford describes the case of a teacher evaluation system characterised by performance-oriented accountability linked to negative sanctions if the teachers do not prove to be ‘effective’. Such conditions do not seem to promote the professional capabilities of teachers, and with little support from teachers, the formative use of data to inform and transform instruction for the benefit of student learning—as measured by standardised testing—seems difficult. This type of system seems to give rise to asymmetric power relations, where the teacher becomes the object of the evaluation instead of an active knowledge producer and learner, which could be the case for students as suggested in the first article’s integrative assessment approach.

Earlier issues of EAEA (e.g. Skedsmo and Huber 2017) have discussed the importance of evaluation designs and have problematised evaluation designs. Mintrop et al.

take this discussion further by pointing out the limitations in the programme logic of policies as well as in the logic and design of several evaluation studies that examine TIF. Evaluations sometimes adopt the programme logic of policies and are not designed to capture the complexity of how these policies work in practice while instead exploring possible unintended consequences; in such cases, their results may present an erroneous picture and legitimate wrong decisions on various levels of the education system. As such, this article raises the issue related to the power of the programme logic, which can focus the attention of evaluators on certain desired aspects of the implementation process and cause other aspects to be overlooked.

The last article lays emphasis on evaluation in terms of institutional power. Zhou et al. present a case where the school inspection process changes and transforms over several phases. While this is partly in response to external demands and pressure in the last phase, it is perhaps chiefly due to the strengthened position of the educational inspection authority in adapting to new situations and thereby gaining influence in the education system.

References

- Scriven, M. (1991). Beyond formative and summative evaluation. In M. W. McLaughlin & D. C. Phillips (Eds.), *Evaluation and education: At quarter century (Ninetieth yearbook of the National Society for the Study of Education)*. Chicago: University Press.
- Skedsmo, G. & Huber, S. G. (2017). Governing education quality—challenges of perspective and methodology. *Educational Assessment, Evaluation and Accountability*, 29(4), 315–318. <https://doi.org/10.1007/s11092-017-9271-3>.
- Strathern, M. (2000). The tyranny of transparency. *British Educational Research Journal*, 26(3), 309–321. <https://doi.org/10.1080/713651562>.
- Webb, P. T. (2006). The choreography of accountability. *Journal of Education Policy*, 21(2), 201–214. <https://doi.org/10.1080/02680930500500450>.