Standardization and assessment practices

Stephan Gerhard Huber[1] · Guri Skedsmo[1,2]

Many countries have recently introduced reforms that aim to enhance education quality by implementing assessment tools to evaluate organizational, classroom, and personal practices (at school leader, teacher, and student levels). Standards and quality frameworks describe education processes and outputs with operationalized indicators, sometimes taking into account the conditions of educational practices. The evaluation process is aligned to these standards and standardized itself, particularly when used for accountability purposes.

On the one hand, standardization, and in particular defining minimum standards, can be regarded as convenient, as it establishes a common framework that focuses on the basic skills and competencies considered important. Standardized approaches to assessment and evaluation practices allow for comparison of educational quality across regions, states, and countries. On the other hand, standardizing assessment and evaluation practices is also constraining because it disregards contextual factors, such as local, regional, and national educational conditions, and their effects on individuals, classrooms, and organizations. Contextual aspects are important for fairly judging quality. Moreover, for improvement purposes—in contrast with accountability purposes—the contextual aspects of education quality are even more important. This issue raises the following key questions: How much contextual differentiation is needed? What are standardization's implications for education assessment and evaluation processes? What are the consequences of standardization, and are there any unintended negative side effects?

✉ Stephan Gerhard Huber
  stephan.huber@phzg.ch

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

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

1 Articles in this issue of EAEA 1/2017

In this issue, we present four articles that address topics related to the increasing standardization of assessment practices at different levels of the education system.

In the first paper, Dedering and Sowada explore school inspectors' work, and particularly the assessment processes built into inspection procedures. This research focuses on how inspector teams negotiate their individual assessments and reach a unanimous score. Based on qualitative interviews with 28 school inspectors in Lower Saxony, Germany, the authors identify three methods for negotiation: (1) the evaluative-positioning method, which begins with proposing and supporting an evaluation with arguments, thereby forming a point of departure for the team's negotiation; (2) the evaluative-thought-experimental method, which implies a discussion of assessment scores; and (3) the descriptive-evaluative method, which requires the inspectors to discuss the basis of their scoring preferences. There is a question as to whether or not the team members' affective relationship dimensions influence the method choice and evaluation process. The article makes a contribution by considering school inspection as a social practice. Further research on these relational aspects, as well as on the nature of negotiations and its possible implications for the professionalization of assessments, is needed.

Several German states have introduced state-wide upper secondary school exit exams to ensure a greater level of standardization. This provides a basis for increased comparability across schools within the state and educational quality benchmarking. Since both students and teachers have more at stake as a result of the state-wide exit exams, it is assumed that both groups will concentrate on efforts that support student learning. In the second paper, Maag Merki, and Oerke address the implications of implementing state-wide exit exams in two German states and point to somewhat ambivalent results: students' self-efficacy and motivation increases with more teacher support. The authors argue for more research on the complex interplay between student motivation, teachers' practices and standardized assessments.

In the third paper, Penk and Richter address student motivation and assessment practices in Germany—specifically, they focus on students' test-taking-motivation (TTM) with regard to low-stakes testing. Such tests have become increasingly important for evaluating educational quality in several countries. Due to research showing that motivated students outperform unmotivated students, TTM has received growing attention from researchers. It remains unclear whether students' scores represent their achievement level if they do not put their best effort forward during testing. The authors assessed students' TTM with a two-hour low-stakes tests, the German National Assessment Study (referred to as VERA in the German context), at three measurement points: before, halfway through and after the test. The authors found that students' perceived value of the test, and therefore their effort, decreased on average, while their success expectations remained stable. They also identified that the initial motivation predicted test performance better than the change in motivation during the test.

The last paper by Ebbeler, Poortman, Schildkamp, and Pieters, addresses schools' use of performance data to initiate education improvement efforts. Previous research has shown that educators struggle with using and making sense of test data. The authors reported on an intervention study conducted in secondary schools in the Netherlands. They found that the teachers taking part in the intervention developed new data literacy skills and showed a more

positive attitude towards data use. Based on these findings, they discussed important policy and practice implications and recommended further research on data use.

2 Increased awareness of implications is needed

The first three papers contribute to a better understanding of how different assessment practices within the standardization movement play out in different contexts. Since this movement has a global character, studying how such tools function within a given context (local, regional, and national) has both theoretical and practical significance.

The four papers emphasize the implications of using standardized tools for school inspection procedures, using instruments to measure student performance and using the data produced by the tools. Focusing on different perspectives, these studies point to certain implications related to the social practices that take place during or after educational assessment processes. While Maag Merki, Oerke, Penk, and Richter address the implications and the unintended consequences of standardized student assessment practices for student motivation and teacher-student relationships, Dedering and Sowada point to the role of professional judgment and negotiation among inspectors in assessing school quality. Additionally, Ebbeler et al. address how teachers can use test data to improve their practices and develop professionally. Their findings indicate the importance of focusing on professional judgment in ensuring assessments can improve education at different levels and in specific contexts.

It is important that further studies continue to explore the implications of assessment tools and the use of data derived by such tools, both in terms of unintended consequences and what role professional judgment should play in assessment procedures. Furthermore, policy and practice implications should be explored. We argue that attention must focus on the instrumental and symbolic functions related to the political governing of schools (Benveniste 2002). Assessment tools (e.g., standardized tests) and evaluation (e.g., school inspection) can be used to collect comparable data that can be used to support rational decision-making at the local, regional, and national levels. In other words, from a policy perspective, standardization might be an important approach, but we need to be aware that not all assessments and evaluations intend to improve education; some primarily aim to hold key actors accountable for achieved outcomes. Moreover, assessments and evaluations can also have a symbolic function, potentially implying that their primary purpose is not to highlight strengths and uncover deficiencies in education, but to legitimize policies.

Reference

Benveniste, L. (2002). The political structuration of assessment: Negotiating state power and legitimacy. *Comparative Education Review, 46*(1), 89–118.