Student Assessment that Promotes Student Learning: What Teachers Need to Know

Margarita Christoforidou[1]

Abstract

Formative assessment has been prevalent in the educational discourse over the past decades, shifting attention towards assessment practices that aid the learning and teaching process. The impact of formative assessment on student achievement has been widely documented; leading to the acknowledgment of formative assessment as a determining factor of educational effectiveness at both the classroom and the school level. This study uses a framework that allows us to measure teachers' assessment skills and examines whether these skills explain variation in student achievement in mathematics. The framework investigates teacher skills in using different techniques of assessment by taking into account the four phases of assessment (i.e., construction and administration of assessment instruments, recording and analyzing data, and reporting results to students and parents). It also takes into account the following five dimensions which describe the functioning of each effectiveness factor concerned with teacher behavior in the classroom: frequency, focus, stage, quality, and differentiation. Data of teachers' assessment skills and student achievement were collected at the beginning and the end of the school year. Based on the analysis of the assessment skills data, it was found out that these skills are grouped into four types of behaviour which are discerned in a distinctive way and move gradually from skills associated with everyday assessment routines to more advanced skills concerned with differentiation in assessment. Using student mathematics achievement data as criteria of effectiveness it was found out that teachers implementing more advanced types of assessment behavior were found to have better student outcomes. The methods and main results are presented and implications of findings are drawn.

1. Introduction

Given the pivotal role that teachers play in mediating assessment reform a large body of research has emerged on teacher education and professional development with particular reference to assessment. However, despite a series of research highlighting inadequate assessment training in teacher education [1], empirical support on how this assessment illiteracy can be addressed is vague. Although numerous attempts to improve assessment practice through professional development projects have been reported [e.g. 2, 3] teachers' everyday assessment practice still appears to be outcome - oriented [4]. This can be partly attributed to the fact that teachers' skills in assessment and how these can be developed were not taken into consideration.

This paper argues that prior to any attempt to improve classroom assessment practice; teacher skills in assessment must be examined. Thus, it draws on research on classroom assessment and teacher developmental theory [5] and examines whether developmental stages can be identified when investigating teachers' skills in assessment. In order to do so, assessment skills are defined and a framework for measuring these skills is proposed.

2. A framework for investigating teachers' skills in assessment

The proposed framework takes into account the dynamic nature of assessment and thus skills associated with each phase of the assessment process were examined. In addition, assessment skills were defined and measured in relation to teachers' ability to use various assessment techniques in measuring different types of learning outcomes. Traditional, as well as, alternative assessment techniques were taken into consideration. Moreover, a measurement framework developed within the field of Educational Effectiveness Research (EER) was adopted and both quantitative and qualitative characteristics of the assessment process were taken into account. Finally, teachers' skills in using assessment results not only for summative but also for formative purposes were taken into consideration. Figure 1 shows the theoretical framework that was used in measuring teacher assessment skills. Specifically, each of the four assessment phases was defined on the basis of

---

[1] Cyprus International Institute of Management, CIIM, Cyprus

assessment knowledge and skills involved across the five dimensions and in relation to the four most common assessment techniques.
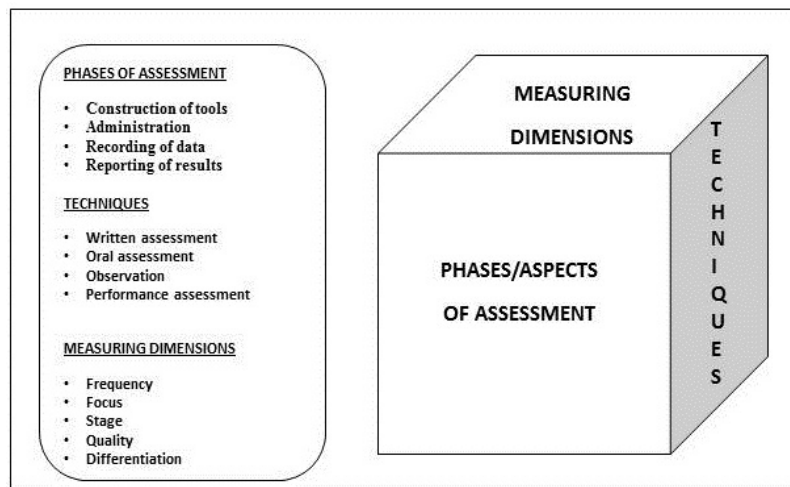

Figure1. A framework for measuring teacher assessment skills

3. Methodology

By taking into account the theoretical framework, a questionnaire was developed and administered to a representative sample of 10% of Cypriot primary teachers at the beginning the school year. Of the 240 teachers approached, 178 responded, a response rate of 74.2%. The teacher-sample was found to be representative of the teacher population of Cyprus in terms of gender ($x2 = 0.81$, d.f. = 1, p = 0.42) and years of experience (t = 1.21, d.f. = 2578, p = 0.22). The questionnaire consisted of 87 items, designed to measure teachers' assessment skills in mathematics across the three aspects of the framework presented above (i.e., phases of assessment, techniques of assessment, measurement dimensions). In order to test the impact that teachers' skills in assessment have on student learning outcomes, student achievement (n = 2358) of the same teacher sample at the beginning and at the end of the school year was measured using a battery of tests in mathematics.

4. Results

The extended logistic model of Rasch was applied to the whole sample of teachers and all 87 measures concerned with their assessment skills, using the computer program Quest [6]. The results of the various approaches used to test the fit of the Rasch model with our data also revealed that there was a good fit with the model when teachers' performance in these assessment skills was analysed. Finally, it was found that local independence was not violated for any item [7].

Having established the reliability of the scale, the procedure for detecting pattern clustering in measurement designs [8] was used. Applying this method to segment the assessment skills on the basis of their difficulties that emerged from the Rasch model showed that they are optimally clustered into four clusters. A description of the four different levels/types of teacher assessment behavior is given below.

Type 1: Using written tests to measure basic skills in mathematics for summative reasons

The assessment skills included in this stage, reveal that teachers demonstrating this type of behavior use everyday assessment routines. Type 1 teachers enrich or alter ready-made written tests and use a variety of types of written questions to assess students' performance. However, they don't use in a systematic way oral assessment and observation to assess their students' performance. Records are kept only in relation to written assessment results whereas results are reported only to parents for summative purposes. Finally, type 1 teachers appear to be consistent to homework check.

Type 2: Measuring basic skills in mathematics for formative reasons

Teachers demonstrating this type of behavior are able to measure basic skills in mathematics for formative purposes. Specifically, skills concerning the quality dimension of written assessment

construction are included. Therefore, type 2 teachers develop representative written tests by creating a specification table before developing their own written tests. They also construct quality test items (measuring not only the product but also the process) and take into consideration their students' ability in order to do so. Clarification comments are offered during assessment administration whereas planned oral assessment and observation are used both formally and informally. Teachers of this type move beyond homework check and use homework information to assess their students for formative reasons. Recording of assessment data also takes a more formative character with the use of descriptive comments rather than overall scores. Reporting of assessment information is also done for formative reasons however it only addresses parents.

Type 3: Using assessment techniques to measure more complex educational objectives

Teachers demonstrating this type of behavior are able to use assessment techniques to measure more complex educational objectives in mathematics such as their ability in communicating by using mathematics. Thus, observation is used in a systematic way by setting specific goals and creating observation tools in relation to these goals. Recording is done for data deriving from all assessment techniques and not merely written assessment and takes the form of goal and/or exercise specific documentation. Reporting for formative reasons is expanded to cover all assessment techniques and addresses both parents and students. Finally, group assessment is done in a systematic way and is concerned with each student's contribution to the team work instead of the team's overall performance.

Type 4: Differentiation in Assessment: Applying assessment in and for different occasions and students.

Based on the assessment skills included in this type of behavior it appears that type 4 teachers are able to differentiate assessment procedures and tools based on their students' needs. Teachers therefore use differentiated written tests and appear more flexible during the administration process (they also give extra tasks to those who finish earlier and give more time to slow learners). They also differentiate reporting of assessment information to both parents and students (e.g., more often to those needed; using different forms/language that are in line with the educational level of parents) and pursue teacher-parent communication especially when the last are not visiting the school.

We then examined the extent to which the classification of teachers into the four stages explains variation in achievement in mathematics. Due to the hierarchical structure of the data (students within classes within schools), multilevel analysis was carried out using MLwiN (see Table 1). Based on the analysis of data it was found out that the developmental stage at which a teacher is situated was found to have a statistically significant effect on student achievement. Specifically, students of teachers at stage 1 had the lowest achievement, whereas students of teachers at stage 4 had higher achievement than students in the first three stages.

Factors	Model 0	Model 1	Model 2
Fixed part (Intercept)	2.19 (0.40)	1.20 (0.12)	0.66 (0.10)
Student Level			
Context			
Prior achievement in maths		0.64 (.12)	0.64 (.11)
SES		0.41 (.14)	0.41 (.14)
Gender (0=boy, 1=girl)		0.12 (.04)	0.11 (.03)
Classroom Level			
Context			
Average achievement		0.40 (.10)	0.40 (.10)
Average SES		0.21 (.10)	0.21 (.10)
Percentage of girls		N.S.S.	N.S.S.
Quality of Assessment			
Stage 1			-.34 (.07)
Stage 2			-.19 (.07)
Stage 4			.18 (.07)
School Level			

Context			
Average achievement		0.10 (.04)	0.10 (.04)
Average SES		N.S.S.	N.S.S.
Percentage of girls		N.S.S.	N.S.S.
Variance components			
School	9.0%	8.1%	7.1%
Class	16.7%	14.5%	10.5%
Student	74.3%	45.1%	44.1%
Explained		32.3%	38.3%
Significance test			
X^2	1033.4	813.1	708.0
Reduction		220.3	105.1
Degrees of freedom		6	3
p-value		.001	.001

Table1. Parameter estimates and (standard errors) for the analysis of student achievement in mathematics (students within classes, within schools)

5. Discussion

In this study, a specific measurement framework was used to describe both quantitative and qualitative characteristics of classroom assessment to help us define and measure specific skills associated with assessment practice. The proposed measurement framework could help clarify the hazy area of classroom assessment by directly associating it with specific dimensions. These dimensions would permit the measurement of classroom assessment's effectiveness not only in terms of its formative purpose but also in terms of all aspects of the assessment process. Furthermore, by moving away from the commonly applied summative-formative distinction, the four stages of assessment behaviour could represent an integrated approach to assessment practice, including the various functions and purposes of assessment.

Looking at the description of the four stages, we can see that they move from relatively easy to more advanced types of teacher behaviour in terms of assessing student knowledge and skills in mathematics. Starting from skills associated with everyday classroom routines with a mainly summative orientation, we can observe a gradual movement towards skills associated with the use of assessment for formative purposes. Using student achievement data, it was found that teachers situated at a higher stage of assessment are more effective than those situated at the lower stages. These findings are in line with recent literature that supports the view that effective teachers use formative–oriented assessment in everyday classroom practice [9]. Specifically, students of teachers in stage 1 had the lowest achievement, whereas students of teachers in stage 4 had higher achievement than students of teachers in the first three stages. Therefore it was found that teachers exercising more advanced types of assessment behaviour had better student outcomes. This finding confirms the impact that assessment practice has on student outcomes. Thus, assessment is not only necessary for evaluating learning but is also a means for achieving it, placing assessment at the heart of the learning process. Furthermore, the fact that the content of each stage is distinctively defined allows the identification of specific assessment skills that have a greater impact on student achievement. This suggests that more extensive usage of assessment skills that were found to have a bigger impact on student outcomes should be encouraged among teachers. These findings can be used not only to determine what constitutes effective assessment, but also how it translates into action.

References

[1] Popham, W.J. (2004). All about accountability / Why assessment illiteracy is professional suicide. *Educational Leadership, 62*(1), 82-83.

[2] Birenbaum, M. (2007). Assessment and instruction preferences and their relationship with test anxiety and learning strategies. *Higher Education, 53*(6), 749-768.

[3] Webb, M.E., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice, 16*(2), pp. 165-184.

[4] Herman, J. L., et al. (2006). *The nature and impact of teachers' formative assessment practices.* CSE Technical Report #703. National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

[5] Dall'Alba, G., & Sandberg, J. (2006). Unveiling professional development: a critical review of stage models. *Review of Educational Research, 76*(3), 383–412.

[6] Adams, R.J., & Khoo, S.T (1996). *Quest - the interactive test analysis system*. Victoria, Australia: Australian Council of Educational Research.

[7] Yen, W.M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement, 30*(3), 187-213.

[8] Marcoulides, G., & Drezner, Z. (1999). A procedure for detecting pattern clustering in measurement designs. In M.Wilson & G. Engelhard, Jr (Eds.), *Objective measurement: Theory into practice* (Vol. 5). New Jersey: Ablex.

[9] Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81-112.