

**IMPROVING MATHEMATICS WRITTEN TESTS:
IMPACT OF RESEARCH ON STUDENT TEACHERS' CONCEPTIONS**

Rosa Antónia Tomás Ferreira

Faculdade de Ciências do Porto, PORTUGAL / Illinois State University, USA

Introduction: Current use of written tests – Despite current calls for diversifying classroom assessment tools (such as portfolios and classroom observations), teachers still rely on typical written tests for assessing student learning (e.g., Associação de Professores de Matemática [APM], 1998; Romberg, 2001). As in other countries, Portuguese teachers also consider written tests as the most objective and rigorous means of assessing student learning, ignoring education goals of an affective or attitudinal nature (e.g., Rafael, 1998). Relying solely on one form of assessment is bound to lead teachers to construct incomplete or incorrect pictures of what their students know and are able to do mathematically. Also, more and richer information on student thinking, beliefs, and knowledge can be gathered using more than one assessment tool (National Council of Teachers of Mathematics [NCTM], 1999, 2000). In fact, typical written tests suffer from a number of drawbacks. For example, they only provide the perspective of students' individual work on time-limited written tasks, and they reveal the results but not the processes of student thinking (van den Heuvel-Panhuizen, 1996). Usual written tests are inadequate for assessing students' ability to investigate and discuss mathematical ideas, or their perseverance and creativity, because students are not typically asked to construct their own answers, nor to explain their thinking or justify their responses (e.g., de Lange, 1993).

Nonetheless, written tests do have a role to play in school assessment as they allow the screening of a whole class, and teachers are very likely to continue emphasizing them in assessment. In addition, teachers are familiar with the design and grading processes of such tests, and students, parents, and the general society also value those tests (Thompson, Beckmann, & Senk 1997). Yet, written tests should be changed and improved rather than rejected as they can be a way to start changing the quality of school assessment (van den Heuvel-Panhuizen & Gravemeijer, 1993). In fact, "relatively minor changes in an item can have a major impact on the nature of a test" (Thompson et al., p. 59). For example, multiple-choice items can be transformed into more open-ended questions whose answers will offer teachers much richer information about the students' thinking.

Alternative written tests – Several alternative written tests, such as written essays and group tests, have been suggested (de Lange, 1987) and implemented in several countries (e.g., Abrantes, Leal, Teixeira, & Veloso, 1997; de Lange). De Lange's five *Basic Principles of Assessment* provide the theoretical grounds for the use of those tests: 1) "tests should be an integral part of the learning process" (de Lange, 1993, p. 199) so that they may improve teaching and learning; 2) tests should themselves generate learning situations; 3) tests "should enable students to show what they know rather than what they do not know" (p. 199); 4) tests should consistently address all educational and curricular goals of school mathematics; and 5) "the quality of the test ... [should] not be dictated by its possibilities for objective scoring" (p. 199).

The *2-phase tests* have two different phases, each one with its own goal. The first phase is similar to a usual written test, and it is aimed at finding out what students do not know or have difficulties in. The teacher's written comments on students' mistakes, suggestions, or requests for clarification provide students with an opportunity to deepen, elaborate, and improve their responses afterwards. In the second phase, students reflect on and complete or redo their answers at home within a few weeks. This phase is focused on what students know, encompassing more higher than lower level questions. Thus, while the first phase is more focused on mathematical procedures, the second one is more geared towards

The Mathematics Education into the 21st Century Project
The Future of Mathematics Education
Pod Tezniami, Ciechocinek, Poland
June 26th – July 1st, 2004

conceptual understanding. Both *written essays and written reports* can be accomplished individually or in groups, inside or outside the classroom. In a written essay, students are asked to elaborate on a certain topic or problem, with or without helping guidelines, being allowed to produce, integrate, and express their ideas. In a written report, students describe their work on a certain learning activity, and critically analyze it. *Student-generated tests*, covering a certain content topic for instance, may stimulate students to look for data, charts, graphs, etc. in order to design exercises, essay questions, or investigations to include in the test, encouraging students' metacognition. *Group tests* are group assessment tasks in which student-student communication and ability to express others' ideas and insights into personal words are the focus. Many curricular themes are suitable for *practical tests* which have a practical emphasis. For example, geometry or probability and statistics topics can be adequately assessed through practical tests. Students may use manipulatives or technological tools to solve problems, and they can also use their daily experiences as sources of information. *Take-home tests* are to be completed at home, and they typically encompass essay-like tasks and can be accomplished individually or in groups. Take-home tests are aimed at getting "a reasonable picture of the possibilities and capabilities of the students when confronted with tasks at a somewhat higher level" (de Lange, 1987, p. 233).

Purpose of the study – The student teaching experience of teacher education programs is ideal for eliciting feelings of need for change in teaching and assessment, as student teachers begin to explore these activities in real classrooms, and may become increasingly aware of the value of the various assessment instruments. As part of a larger research endeavor (Tomás Ferreira, 2003), I investigated the impact, if any, of the reading and discussion of selected research studies and reform texts on Portuguese student teachers' conceptions of classroom assessment in general, and of the role and value of written tests in the assessment of student learning.

Methodology – The participants were enrolled in a 5-year secondary mathematics teacher education program for grades 7 through 12, offered by a public university in a large urban community in northern Portugal. This teacher education program is characterized by an emphasis on mathematics content coursework with no current mathematics education course offerings. The student teaching phase lasts for a whole school year and is designed as a group experience. Each group of student teachers works with a cooperating teacher and with a university supervisor (supervising an individual or collective monograph, addressing topics hardly related to mathematics education). Student teachers work as full-time teachers with a reduced teaching load. Twenty student teachers completed a survey which sought information on their beliefs and practices of classroom assessment and written tests. Based on variability of responses and placement schools, and on availability, 9 student teachers were selected to further participate in this study. The participants were placed in 3 schools. Groups A, B, and C had 3, 4, and 2 members, and were teaching in a suburban secondary school, in a small rural junior high school, and in a secondary school in a mid-sized urban community, respectively.

Data were collected in three time periods: October 2002, and January and March 2003. Besides the survey, written reflections upon two packets of selected readings (including reform documents and research articles), group interviews, and written tests constructed by the participants were the data collection instruments. The selected readings were written in English and in Portuguese, and all were accompanied by an abstract in Portuguese. The first reading packet included guidelines to facilitate written reflection but the second packet purposefully excluded reading guidelines. The semi-structured interview protocols, based on reading guidelines, written reflections, and survey responses, were audio-tape recorded for further analysis. The interviews, each lasting between 60 to 90 minutes, were conducted in my university office.

The Mathematics Education into the 21st Century Project
The Future of Mathematics Education
Pod Tezniami, Ciechocinek, Poland
June 26th – July 1st, 2004

Results and Discussion – In this paper, I report on the survey, written reflections, and interviews data of the study. Many survey items (dealing, for instance, with students’ reactions to their tests) were left blank because the respondents had not yet given any written tests to their students. However, Group C responded to all items, including those mentioned above. All participants reported using various assessment tools such as written tests, and students’ participation, behavior, and homework completion. Yet, many of them saw written tests as comfortable and rigorous means of quantifying student learning and teaching effectiveness. This is not surprising given the results of previous research conducted in Portugal (e.g., Rafael, 1998).

All respondents indicated they planned to use written tests as means of identifying students’ learning difficulties, but no further elaboration was ever offered in the survey. Only one student teacher in Group C mentioned her concern about the informative role of written tests for both students and teachers: to find out students’ learning difficulties, and to use that information to revise and adapt teaching to meet the different needs of all students. With a few exceptions (namely Group C), the participants’ conceptions about school mathematics were consistent with the typical society perspective “of mathematics as a set of discrete hierarchically arranged facts and skills; ... [its] view of learning mathematics as replication and repetition; ... [its] view of teaching mathematics as exposition and practice; and ... [its] view of assessing mathematics as paper and pencil testing for the sole purpose of grading and ranking” (Herrington, Herrington, & Glazer, 2002, p. 1105).

Group A never completed written reflections. Clear distinctions between the other two groups were evident regarding the quality of their written reflections as seen in Table I.

Table I: *Contrast among the written reflections*

Groups	Quality of written reflections	Reactions to written reflections
A	No written reflections.	Only one student teacher seemed to have read some texts.
B	Generally superficial, and reduced to a set of responses to reading guidelines (1 st packet) or to a poor summary (2 nd packet) of the readings.	Failed to identify key ideas in the readings. Hardly ever connected readings to own classroom practice, but, towards the end of data collection, two members started thinking about and questioning their assessment practices, including test quality.
C	Deep, extensive reflections, expanding on the texts’ ideas that caught their attention the most.	Almost always related readings to own classroom practice and personal perspectives and ideas. Showed great concern for equity and social justice in testing.

The group interviews served mainly as debriefing sessions on the selected readings and completed reflections. One theme emerging from the first set of interviews was related to the role and use of quizzes. The participants’ use of this assessment instrument is summarized in Table II. It was never clear whether Group A agreed with their cooperating teacher, nor what Group B would do after identifying students’ difficulties. Group C was the sole group of participants mentioning that the information collected on quizzes or tests would be used to adapt instruction accordingly.

Table II: *Different uses of quizzes*

Groups	Use of quizzes	Purpose of using quizzes
A	No use	(Cooperating teacher was against using quizzes)
B	Frequent use, right before the <i>real</i> test	To identify students’ learning difficulties
C	Frequent use one or two weeks before the <i>real</i>	To give students enough time to overcome their difficulties, and to use quiz information to revise lesson plans and classroom

The Mathematics Education into the 21st Century Project
The Future of Mathematics Education
Pod Tezniami, Ciechocinek, Poland
June 26th – July 1st, 2004

	test	teaching in order to accommodate for students' needs.
--	------	---

The interviews provided more opportunities for reflection than the survey or the written reflections. Table III provides a picture of the main results of both sets of group interviews.

“Alternative assessment creates a climate of unpredictability because we can never be sure what students are going to say when we ask about their mathematical thinking” (Cooney, Bell, Fisher-Cauble, & Sanchez, 1996, p. 485). This certainly represents a challenge for teachers who prefer a more organized classroom where students' and teachers' expectations and demands are well known by all. This seemed to be the case in Group A. Student teachers from Group B had low expectations for their students and they complained about student underachievement, misbehavior, and lack of motivation. Also, parents overvalued typical written tests, and were skeptical of other assessment instruments. These two factors seemed to account significantly for Group B's resistance to using alternative written tests, and to these student teachers' sense of helplessness about implementing new forms of classroom assessment. *Content coverage* also hinders teachers from implementing alternative forms of assessment, including alternative written tests (Cooney et al., 1996), and this did seem to be the perspective of Group C's cooperating teacher, who even prohibited her student teachers from doing group activities and using technological aids for instruction, even not using class time. However, her student teachers were determined to overcome those barriers and implement alternative written tests, finding ways that would please her.

Table III *Emergent themes from student teachers' comments in group interviews*

Groups	Written tests	Cooperating teacher
A	Main ideas in texts not entirely new; interviews helpful for clarification and expansion of those ideas; alternative written tests difficult to do with misbehaved students.	Seen as very knowledgeable and prepared for the job; had no significant influence on test construction; gave no support to student teachers, contributing to their anxiety about classroom observations.
B	Generally uninterested in doing alternative tests: seen as non-doable; recognition of overemphasis on written tests, but only a few consider possibly changing this situation in the future.	Gave relative autonomy in test construction, and significant support in terms of materials and incentives to improving classroom teaching (under her own perspective).
C	1 st contact with alternative assessment, though had several similar, intuitive and personal ideas for classroom assessment; great concern about explaining to students the role and value of tests and quizzes; afraid of implementing own instructional and assessment ideas, but excited about using alternative written tests somehow.	Overly concerned about covering the curriculum and avoiding anything that could interfere with this goal; gave some support in terms of materials and help in classroom teaching; had already hindered student teachers from proposing different activities for their students

For all groups, dealing with *increasing demands on time* seemed to be a major obstacle to implementing alternative written tests. Though this endeavor is less demanding than implementing other forms of assessment (Cooney, Badger, & Wilson, 1993), significant effort and commitment are necessary to accomplish this goal. One member of Group C referred to this issue as follows: “The general unwillingness of teachers to change and their widespread lack of passion for the profession itself, largely contributes to their passive and accommodated attitude towards their mission as teachers” (Int. 2, March 03; my translation). This student teacher suggested that younger teachers such as herself and her fellows, being more willing to change current classroom practices, “should get together to discuss and reflect on issues of reform-based instruction and assessment as a first step to start changing the state of affairs in Portuguese junior high and secondary schools” (Int. 2; my translation). In this regard, “alternative written tests could work as ideal starting points since changes in these

The Mathematics Education into the 21st Century Project
The Future of Mathematics Education
Pod Tezniami, Ciechocinek, Poland
June 26th – July 1st, 2004

tests are not too difficult to do and can make a huge difference regarding the quality of information about student thinking that can be collected through such tests” (Int. 2; my translation).

Conclusions – The participants in this study held very different conceptions of classroom assessment in general and of written tests in particular, ranging from traditional perspectives of assessment based almost exclusively on written tests to more reform-oriented standpoints of assessment based on a plethora of instruments and practices that are integrated and aligned with curriculum and instruction. Group C seemed more homogeneous in their reform-based conceptions of classroom assessment, while the other groups evidenced internal differences, especially group B.

Besides providing the participants with useful information on alternative assessment and alternative written tests, this study, especially its interview component, constituted an opportunity for all to understand and/or clarify some of the ideas addressed in the readings, and to broaden their perspectives on classroom assessment and the role of written tests. However, this study seemed to have impacted differently the participants’ conceptions about assessment and written tests, as summarized in Table IV.

Table IV: *Impact on student teachers’ conceptions*

Groups	Impact
A	Minor impact: lack of engagement and interest in the study’s activities, and lack of willingness to use alternative tests.
B	Relatively minor impact: the study provided interesting information about alternative classroom assessment, but there were different degrees of willingness to use that information
C	Relatively major impact: serious engagement in the study’s activities, and significant excitement about reforming assessment practices and improving written tests.

The student teachers’ dispositions towards reform-based teaching and their excitement about the teaching profession seemed crucial for embracing a reform-oriented practice of classroom assessment and a wiser and fairer use of alternative written tests. The participants whose perspectives on classroom teaching and learning were more reform-oriented were the ones who engaged more seriously in this project and reflected more deeply on the topics suggested, going much beyond the reading guidelines. Also, they did not let the barriers posed by the cooperating teacher or school department policies prevent them from trying new practices nor from keeping their perspectives and ideas, at least for a near future. The student teachers with less reform-based perspectives on classroom teaching and learning were concerned with covering the curriculum and used this goal, together with lack of time, and student misbehavior and underachievement, as excuses for not using alternative forms of assessment.

Limitations and Implications for Future Research: There were a number of limitations in this research project, most of them due lack of opportunities for data collection (since I was in Portugal for shorts periods of time, also collecting data for my doctoral dissertation). For example, I was not able to conduct classroom observations, nor to interview the three cooperating teachers, nor to analyze any kind of student work on written tests. Due to legal constraints, I could not ask the participants to try out alternative written tests in their classrooms either, thus failing to situate “the change process in the actual teaching and learning contexts where the new ideas will be implemented [which] is an effective strategy for helping teachers change their practices” (Borko, Mayfield, Marion, Flexer, & Cumbo, 1997, p. 267). In addition, many texts were in English, which may have caused some participants to feel discouraged about reading them. Some of the major suggestions for future research include: the actual use of alternative written tests in the participants’ classrooms, observations of their teaching and assessment practices, meetings to construct and/or analyze alternative

The Mathematics Education into the 21st Century Project
The Future of Mathematics Education
Pod Tezniami, Ciechocinek, Poland
June 26th – July 1st, 2004

written tests, analyses of samples of students' work on written tests, and interviews with students and cooperating teachers.

Based on the data collected, the three cooperating teachers differed greatly in terms of the constraints posed in test construction. Surprisingly, the cooperating teacher who posed more serious obstacles to alternative assessment did not seem to have affected her student teachers' dispositions to actually implement their non-conventional ideas for assessment. On the contrary, the cooperating teacher who seemed more knowledgeable about alternative written tests and who actually used some of these tests did not encourage his student teachers in using alternative written tests. Future research should take into account the preparation and contribution of cooperating teachers in order to facilitate changes in the student teachers' assessment practices, especially regarding the improvement of typical written tests. Other questions were left unanswered, such as the following: 1) Why did students' lack of motivation seem to prevent the participants from assessing differently but not from teaching differently? 2) How did students react to the alternative assessment tasks proposed by group C? and 3) What role could the university play to help cooperating teachers better accomplishing their mission?

Acknowledgements: I wish to thank Michelle L. Wallace and Professor Norma C. Presmeg, from the Mathematics Department of Illinois State University, for their comments on other versions of this paper. I also thank the student teaching supervisors of my Department, in Portugal, for their cooperation in this research project.

Note: This work was partially funded by the Grant PRAXIS XXI/BD/19656/99/ from *Fundação para a Ciência e a Tecnologia*, Portugal, European Union.

References

- Abrantes, P. Leal, L. C., Teixeira, P., & Veloso, E. (1997). *MAT 789: Inovação curricular em matemática [Curricular innovation in mathematics]*. Lisboa, Portugal: Fundação Calouste Gulbenkian.
- APM (1998). *Matemática 2001 – Diagnóstico e recomendações para o Ensino e Aprendizagem da Matemática: Relatório final [Mathematics 2001 – Diagnosis and recommendations for the teaching and learning of mathematics: Final report]* Lisboa, Portugal: APM & IIE.
- Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching and Teacher Education, 13*(3), 259-278.
- Cooney, T., Badger, E., & Wilson, M. R. (1993). Assessment, understanding mathematics, and distinguishing visions from mirages. In N. L. Webb & A. F. Coxford (Eds.), *Assessment in the mathematics classroom*, 1993 NCTM Yearbook, (pp. 239 – 247). Reston, VA: NCTM.
- Cooney, T. J., Bell, K., Fisher-Cauble, D., & Sanchez, W. B. (1996). The demands of alternative assessment: What teachers say. *Mathematics Teacher, 89*(6), 484 – 488.
- De Lange, J. (1993). Assessment in problem-oriented curricula. In N. L. Webb & A. F. Coxford (Eds.), *Assessment in the mathematics classroom*, 1993 NCTM Yearbook, (pp. 197 – 208). Reston, VA: NCTM.
- De Lange, J. (1987). *Mathematics, insight, and meaning*. Utrecht, The Netherlands: OW & OC.
- Herrington, A., Herrington, J., & Glazer, E. (2002). Authentic approaches to learning assessment strategies: Beginning teachers' practice in classrooms. In D. Mewborn (Ed.), *Proceedings of PME-NA 24*, (pp. 1105 – 1114). Athens, GA. October, 2002.
- NCTM. (1999). *Normas para a avaliação em matemática escolar [Assessment standards for school mathematics – Portuguese translation]*. Lisboa, Portugal: APM.
- NCTM (2000). *Principles and standards for school mathematics*. Reston, VA: NCTM.
- Rafael, M. A. P. (1998). *Avaliação em Matemática no ensino secundário: Concepções e práticas de professores e expectativas de alunos [Assessment in secondary mathematics: Teachers' conceptions and practices, and students' expectations]*. Master's Thesis, University of Lisboa, Portugal. Lisboa, Portugal: APM.
- Romberg, T. A. (2001). Mathematics goals and achievement in the United States. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the PME 25*, (Volume 1, pp. 180 – 185). Utrecht, The Netherlands. July 12 – 17, 2001.
- Thompson, D. R., Beckmann, C. E., & Senk, S. L. (1997). Improving classroom tests as a means of improving assessment. *Mathematics Teacher, 90*(1), 58 – 64.
- Tomás Ferreira, R. A. (2003). *Improving mathematics written tests: Impact of research on student*

The Mathematics Education into the 21st Century Project
The Future of Mathematics Education
Pod Tezniami, Ciechocinek, Poland
June 26th – July 1st, 2004

- teachers' conceptions*. Unpublished manuscript. Illinois State University. Normal, IL. USA.
- Van den Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education*. Doctoral Dissertation. Utrecht, The Netherlands: Center for Science and Mathematics Education.
- Van den Heuvel-Panhuizen, M. & Gravemeijer, K. (1993). Tests aren't all bad: An attempt to change the face of written tests in Primary school mathematics instruction. In N. L. Webb & A. F. Coxford (Eds.), *Assessment in the mathematics classroom*, 1993 NCTM Yearbook, (pp. 54 – 64). Reston, VA: NCTM.