Using test blueprint in classroom assessments: why and how

Chan Kan Kan
University of Macau
Faculty of Education
kankchan@umac.mo

Keyword: test blueprint, classroom assessments

This paper supports that developing test blueprint in classroom assessments is as important as that in the standardized test so as to make appropriate and meaningful decisions. This practice should not be left out in classroom assessments. With reference to the case in Macau, an educational system emphasizes assessments. It illustrates problems of the current practice of assessment in Macau educational context. Data are based on the convenience sample which the author meets in educational evaluation classes. Even though it is found that many teachers do not have the practice of developing test blueprint in classroom assessments, many judgements are made based on such practice. This implies that the assessment literacy is in need of attention. It discusses the factors which affect teachers' uses of test blueprint. It also provides suggestions on helping teachers to develop valid and reliable classroom assessments which are feasible to implement in practical environment without too much effort.

Introduction

Schools in Macau basic education often schedules tests for students. This paper describes the reasons and the consequence of frequent tests on students and teachers. It illustrates the problems of the assessment practices which lacks the use of test blueprint. It further argues for the employment of   test blueprint in the context of classroom assessment in basic education so as to achieve an assessment culture for the purposes of learning.

A test blueprint is also called table of specifications which shows the structure of a test. It usually includes the content areas along with the cognitive levels covered by the test. In a typical test blueprint, the content areas are listed in the left-hand column of the table while the cognitive levels are shown across the first row of the table. It has been highly recommended in assessment textbook to plan a test with a test blueprint (Suskie, 2009). In reality, do teacher develop a test using test blueprint? Do they understand its value? It seems that classroom teachers pay little attention to the design and developmetn of a valid and realible assessment. It is necessary for teachers and school leaders to understand the importance of test blueprints and adopt it in practice. The next section depicts the procedures of planning a test by test blueprints and the value of using test blueprints. After that, the assessment practice in Macau basic education is critically reviewed to understand the reasons behind the scene and the problems associated. The final section proposes means of adoping test blueprint in classroom environment which teachers can manage without too much effort.

Essence of using test blueprint in large-scale assessment

A test blueprint is crucial for the development of a sound assessment. It is quite common for large scale standardized assessments such as Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV), Third International Mathematics and Science Studies (TIMMS) to design, develop and implement assessment tasks with reference to the use of test blueprint (Garden R.A. and Orpwood, 1996). For instance, TIMMS, an international achievement test on the subject of science and mathematics, drafted a preliminary test blueprint using the TIMMS curriculum framework quite early in the process. The blueprint guides item collection and development. It provides a clear framework for the researchers to design items which assess the important concept or thinking skill listed in the test blueprint. This specification conducted in the early phase of the research processs ensure that the test focuses on important learning goals and makes writing test items much easier. It serves as the ground for improvement of the whole research process.

After the test is administered and scored, students' score is obtained by mapping students'

performance in the test with the test blueprint. It serves to document if students have achieved major learning goals. After that, interpretation of the assessment results is made with the evidence collected. It helps to ensure the appropriate and meaningful understanding of the assessment evidence. The reason for its application is to ensure that the results of the assessment is accountable to its stakehodlers. It also provide sources for the formative use of the summative assessment as suggested by Black and Wiliam (2001). From the perspective of quality assurrance , the test blueprint provides a lens for researchers to check if the test instrument fulfil its purpose by matching the content and goals. Therefore, the practice of drafting and revising test blueprint is a necessary and helpful practice to the researchers in many large scale assessments.

The assessment blueprint, a critical component of a good assessment, should be shared to the students. Sharing assessment nature, purposes and results to students is an ethical practice on one hand. Students have the right to know the details of the assessment so that they may respond to it in an informed manner. On the other hand, it demonstrates good pedadogy because because it help students focus their studies on learning goals. This is also one of the assessment for learning practice which Black and his colleagues (2003) recommend. Therefore, large-scale assessment usually provides clear assessment framework and sample items to the public. Teachers may use the provided information for the learning of students. Marking procedure is clearly shown and guidelines for the interpretation of test scores are also provided.

As a summary, this section outlines the reasons of using test blueprint in large scale test. Test blueprint is beneficial to test developers, researchers and students in different phases of the assessment. Its ultimate goal is to make sound and appropriate judgement. Besides, it is an ethical approach of conducting assessment.

Crtical review of Assessment practices in Macau

The educational system in Macau emphasizes tests very much to an extent which Morrison and Tang (2002) described as "state-wide mentality of testing". Tests are scheduled as frequent as once a week. Sometimes it may sometimes be twice a week. Teachers construe testing as a driving force for students to study and learn. Test provides a way for teachers to know if how much students have learnt. Test is considered as an effective strategy to ensure the learning of a large class size. School authorities use students' performance in test to conduct teachers' appraisal. If students' test score are low especially in university entrance examinations, teachers may be under pressure. This is because students' performance in public examination is one indicator of their value in school. They have to provide remedial classes for students. Students are often overloaded with tests or homework. They are always under pressure, especially for those who score poorly in tests. These students may be suffering considerable blame from

teachers and parents. They have to join extra classes with the aim to catch up their study but this educational practice has limited effects in reality.

Despite the high mentalitity of testing in Macau schools, there is no state-wide examination. Schools have the power to set their own standard of graduation. Tests in schools are mainly controlled by teachers. Subject teachers have the freedom to set the purpose, contents, scope, nature, timing, question types and frequency of the test. Teacher may adjust the test score or difficulty of test items depending on the performance of students. In some schools, tests are devised collaboratively while in smaller schools, test is the sole effort of subject teacher. Programme coordinator usually endorses teachers' design of test nature and details. Tests make students focusing on topics which teacher assigns. Students are required to practice rote learning instead of higher order thinking skills. Tests seem to dampen students' interest in learning but highlight them to seek high marks. Over testing engenders the above problems which teachers are aware of. However, teachers still recyle their usual practice. Their reasons include the pressure of the university entrance examinations,   the hugh class size and their perceptions of the positive effect of test. Teachers have the knowledge of alternative assessment and understand the value of alternative assessment in providing evidence of students' complex skills. However, they point out that current situation in schools such as school policy, workload and examinations favour more on traditional assessment.

With respect to the development of test, more than 100 in-service teachers who studied the course "Educational Evaluation" in the Faculty of Education, University of Macau were asked in class about their practice of test developments. These teachers have teaching experience ranging from two to more than 10 years. They teach different subjects in different primary and secondary schools in Macau. They could not represent all the teachers in Macau school. But their educational practice is quite prevalent in the real situations.

Almost all teachers reported that they had no prior knowledge of test blueprint. They seldom design test items using test blueprint. Some schools require teachers to state learning objectives when they plan instruction for students. However, teachers do not refer to these objectives when they are devising a test. They usually modify items in test papers of last academic year to form a new version of test paper. They may also use item banks provided by textbook publishers. They do not analyze the response pattern of students after marking and scoring. They usually correct students' mistakes in class to help students undertand their mistakes. If students' performance is within their expected one, teachers are relieved otherwise they are under pressure of students' performance and the limited time to cover huge amount of knowledge. They also indicated that they have no time to review the test items using students' responses. Even though they recongize the value of item analysis after having study it, they still consider the process as time consuming and think that it is impossible to do item analysis in real situation.

Once students receive the marked test paper from teachers. They get a rough estimation of their achievements in test. Meaning of scores are not properly explained to students. Students with high scores usually tend to believe that they are capable. If their scores are low, the effect may be adverse. Interpretation of the meaning of scores seems to be superficial and inappropriate. Students may generalize their weakness or strength on one particular test to mean their overall performance. Yet, as shown above, the design of a test might not be as reliable as it supposes to be. It also exposes to different threats of validity. Similarly, subject teachers typically pass the mark list to class teachers and class teachers have to write comments on students' performance, achievements, strength, weaknesses and interest in subjects. Their comments are usually ambiguous and shallow because they have no real knowledge of students' achievements in specific subjects. Messages transmittted through report cards or parent teacher meetings is not clear enough (Chan, 2009).

To summarize, the above illustrates several problematic practices in the area of assessment in Macau educational context. Reliability and validity in assessments are quite hard to achieve. Data are based on the convenience sample which the author meets in educational evaluation classes. Practices which enhance validity and reliability are seldom followed by classroom practitioners. Yet many educational judgements such as promotion to next grade level are made based on such practice. The irony here is that teachers and school authorities do not question on their own practice. This implies that the assessment literacy of stakeholders in education is in need of attention. That is why the Macau government has launched initiatives such as training teachers, conducting school evaluation to help school to understand their problems and provide suggestions.

Towards a meaningful and ethical assessment practice

It is not an easy task to change the exising assessment culture. Every stakeholders in the Macau school education should be moving towards an assessment for learning culture. It is stated in a recent revised educational law[1] that assessment should be used for the purpose of learning using multiple forms of assessment. In order to improve the assessment practice in Macau schools, both long and short term strategies are necessary. For long terms one, teachers must be given professional opportunities to enhance their assessment notions and item analysis. The change in mindset of teachers and the overall educational system is far beyond the control of individual schools. It requires the leadership of government, the cooperation of educational practitioners and social acceptance.

---

[1] http://bo.io.gov.mo/bo/i/2006/52/lei09_cn.asp

	Recall		Comprehension		Application		Total
	%	items	%	items	%	items	%
Chapter 1	15%	4	8%	2	10%	2	33%
Chapter 2	19%	6	22%	6	10%	2	51%
Integrated Knowledge	------	------	16%	11	------	------	16%
Total	34%	10	46%	19	20%	4	100%

Table 1. A test blueprint for a test in Chinese subject.

An easy but effective policy is to request classroom teachers to prepare a test blueprint.What do classroom teachers really need to do? In fact, it is not much. Teachers have to know what exactly they want to assess in a given test and its scope. After that they create a table to clearly distribute the cognitive level to be assessed and the topic involved. An example is listed in Table 1. It shows that the test mainly assess students' recall of knowledge in chapter 1 and chapter 2 and their understanding of the terms and rules learnt in these two chapters. Number of items used in test is also shown obviously. In actual situation, questions in test paper might incline to certain area. This table helps to prevent teachers from having too many questions in the recall section of chapter 1. It also enables teachers to develop alternative forms of test paper while maintaining the reliability of the test. Teachers are empowered to create a suitable test for their students while maintaining the standard. The same table provide teachers a clear framework to do quality assurance in terms of the representatives of items to the achievement domain. It may be used to list the difficulty of questions which teachers may improve the quality of test in future. Reference to the the cognitive levels assessed listed in test blueprint limits teachers to interpret students' score to the scope of the test. In order words, the test buleprint helps to achieve the validity from the sources of content, response and consequences evidence(Downing, 2003)

Overall, this paper argues that test blueprint is critical to make appropriate interpretation of students' performance in test. Test blueprint helps to ensure that the development of items in test is aligning to certain requirements. Assessment procedure becomes rigorous and sound. As the impact of testing is high on both students, it is necessary to make proper judgement on students. It also provides suggestions on helping teachers to develop valid and reliable classroom assessments which are feasible to implement in practical environment without too much effort.

Reference

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.

Black, P., & Wiliam, D. (2001). *Inside the black box: Raising standards through classroom*

assessment. London: School of Education: King's College.

Chan, K. K. (2009). *How can a report card facilitate assessment for learning?* Paper presented at the 35th International Association for Educational Assessment (IAEA) Annual Conference, Brisbane, Australia.

Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education, 37*(9), 830.

Garden R.A. and Orpwood, G. (1996). Development of the timss achievement tests. In M. O. Martin & D. L. Kelly (Eds.), *Third international mathematics and science study (timss) technical report, volume i: Design and development.* Chestnut Hill: MA:Boston College.

Morrison, K., & Tang, F. H. J. (2002). Testing to destruction: A problem in a small state. *Assessment in Education, 9*(3), p.289-317.

Suskie, L. (2009). *Assessing student learning: A common sense guide 2nd edition*: Jossey-Bass.