

# CONCEPTIONS AND PRACTICES OF ASSESSMENT: A CASE OF TEACHERS REPRESENTING IMPROVEMENT CONCEPTION

**Astuti Azis**

(tuty\_azis@yahoo.com; astuti.azis@vuw.ac.nz)

*SMPN 4 Sungguminasa Kab. Gowa Sulawesi Selatan  
Jl. Mesjid Raya Sungguminasa Sombaopu Kab Gowa Sulsel*

*Victoria University of Wellington  
Wellington, New Zealand*

**Abstract:** Despite numerous quantitative studies on teachers' conceptions and practices of assessment, little research exists regarding the unique assessment environment of Indonesia. This study uses both quantitative and qualitative data to examine how Indonesian junior high school teachers understand assessment and how their conceptions of assessment relate to their assessment practices. This mixed methods study adopted a participant selection model in which quantitative data was analysed to select participants for the qualitative phase. Participants of this study believed that the purpose of assessment was to improve teaching and learning and also to demonstrate the accountability of students and school. They tended to disagree with the view that assessment is irrelevant. Further analysis of the data revealed that teachers' conceptions of assessment were conflicted. They were keen to use assessment practices to improve their classroom teaching, but felt that the state-wide examination policy requirements constrained their efforts. This suggests that government, policy makers, and curriculum developers must work to build a strong synergy among themselves in order to share consistent goals with teachers. If cultural expectations of school assessment and government policy were aligned, Indonesian teachers may be better able to resolve conflict between their beliefs and assessment practices.

**Keywords:** improvement, conceptions, assessment, teaching practice

**DOI:** <http://dx.doi.org/10.15639/teflinjournal.v26i2/129-154>

Assessment is crucial for both policy makers and practitioners. It could be conducted for purposes of accountability which includes establishing how well students have learned, or to inform the design of instruction (how to improve) in educational contexts (Danielson, 2008). These two purposes sometimes support one another, and sometimes compete or conflict with one another, which indicates that assessment is a complex process (Earl, 2003).

A conception is a lens through which a teacher views, interprets and interacts with his/her understanding of the world (Pratt, 1992). This means that conceptions influence the way an individual defines his/her work (Nespor, 1987) including teachers' teaching and students' learning (Calderhead, 1996; Pajares, 1992). Conceptions are related to shared social and cultural phenomena (van den Berg, 2002). People's beliefs and the norms of their social environment appear to be crucial in shaping their type of behaviour and practices (Brown, 2008). Griffiths, Gore and Ladwig (2006) report that beliefs affect teaching practices to a greater degree than teaching experience and socioeconomic school context.

An early study of teachers' conceptions of assessment was conducted by Wolf, Bixby, Glenn and Gardner (1991) who distinguished between assessment culture and testing culture. These authors believed that teachers' understanding of intelligence, the process of teaching and learning, the nature of tasks and the evaluation criteria influenced teachers' conceptions and practices of assessment. Delanshere and Jones (1999) also proposed three dimensions to identify teachers' conceptions of assessment. Those dimensions are (a) students' placement according to achievement level; (b) teacher's perceptions of curriculum and professional self-efficacy; (c) teachers' beliefs about teaching and learning and about students as learners. All these authors consider the relationship between assessment and teaching and learning.

Significant studies of teachers' conceptions of assessment have been conducted by Brown and his colleagues since the early 2000s. In his study of New Zealand teachers, Brown introduced four teacher conceptions of assessment (TCoA): a focus on improvement or teachers' views of assessment as a tool to improve teaching and learning; assessment as driven by school accountability purposes; assessment for student accountability; and perceptions of assessment as irrelevant or a meaningless practice in daily school life (Brown, 2002). This model has been subsequently validated in several studies (Brown, Lake, & Matters, 2011; Calveric, 2010; Segers & Tillema, 2011). Brown's model has also been adjusted and modified to fit Asian contexts by adding examination as

another crucial dimension for high-stakes assessment settings (Brown, Hui, & Yu, 2010; Brown, Kennedy, Fok, Chan, & Yu, 2009).

However, Remesal (2009) found that the TCoA model did not suit Spanish teachers. This challenged her to develop a continuum of assessment purposes. At one end she places pedagogical conceptions (assessment for monitoring of teaching and learning) and at the other extreme she identifies a societal-accreditation view (assessment for teachers' accountability and certification of achievement) with some mixed conceptions between the two poles (Remesal, 2011). The continuum is based on four roles of assessment according to how teachers used assessment in learning, in teaching, in the certification of learning and for accountability when students' achievement is used as an indicator of the teachers' professional capacity. Despite the differences in their studies, both Brown and Remesal agree with previous authors that assessment could and should benefit both teachers and learners.

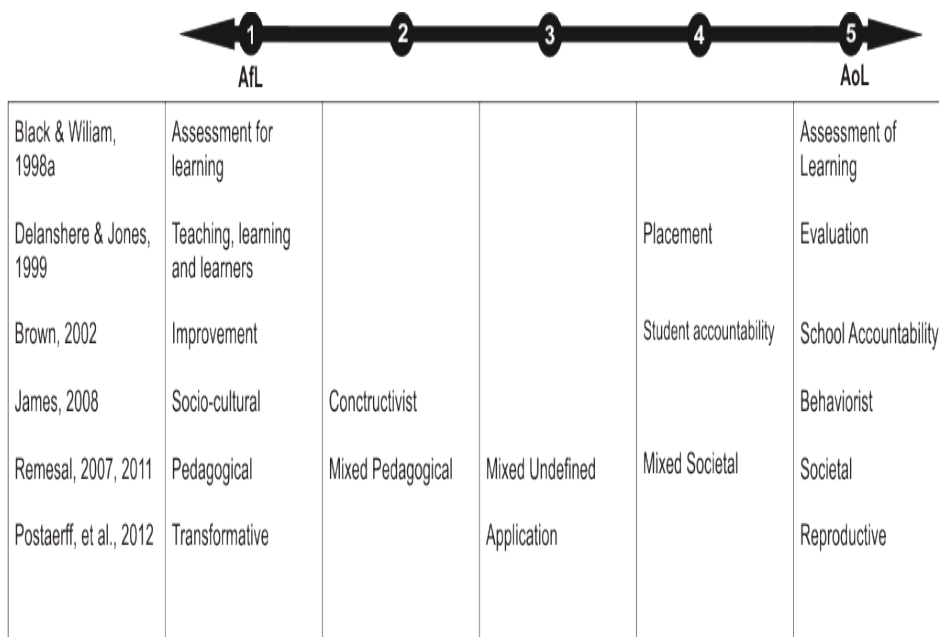
A more recent study on the issue of teachers' conceptions was conducted in Helsinki, Finland (Postareff, Virtanen, Katajavuori, & Lindblom-Ylänne, 2012). These authors identify conceptions as being either reproductive or constructive. The reproductive conception emphasises students' understanding through memorisation of the content of the study module and how it is applied in real life. The constructive conception on the other hand, looks at deeper understanding including reflection and justification for an argument. Unlike former studies, Postareff and her colleagues focus on the impact of assessment on learning.

Although these researchers use different terms to address assessment conceptions, they appear to refer to the two classic purposes of assessment: summative (Broadfoot, 2007; Wiliam & Black, 1996) or formative purposes (Black & Wiliam, 1998b).

Based on the literature on assessment purposes I have drawn a continuum of assessment with two different focuses (see Figure 1). At one pole is assessment for learning (AfL) while assessment of learning (AoL) is at the other end of the continuum.

Assessment for learning (AfL) or formative assessment (FA) focuses both on students' learning and teacher's teaching and these functions are inseparable (Black & Wiliam, 1998b; Harlen, 1998). In other words, the purpose of using FA to assess students' learning and its application to make beneficial changes in instruction creates a tight link with instructional practices (Boston, 2002).

These assessment purposes reflect both improvement conceptions (Brown, 2002) and pedagogical conceptions (Remesal, 2011).



**Figure 1: Continuum of Assessment**

Assessment for learning (AfL) or formative assessment (FA) focuses both on students’ learning and teacher’s teaching and these functions are inseparable (Black & Wiliam, 1998b; Harlen, 1998). In other words, the purpose of using FA to assess students’ learning and its application to make beneficial changes in instruction creates a tight link with instructional practices (Boston, 2002). These assessment purposes reflect both improvement conceptions (Brown, 2002) and pedagogical conceptions (Remesal, 2011).

Assessment of learning or summative assessment (SA) is a way of looking back or summing up learning (Broadfoot, 2007; Wiliam & Black, 1996). SA involves marking and grades which require unified procedures that enable comparability among the results of all students (Harlen, 1998). This assessment type focuses on measuring the extent to which students reach or do not reach

required standards (Firestone, Mayrowetz, & Fairman, 1998). It is used as a means to record the nature and level of students' achievement throughout their academic careers (Hill, 2000). SA or AoL can be conducted by teachers as well as external agencies. External agencies often use the results for accreditation, accountability and monitoring (Harlen, 2005), selection, placement and certification (Black & Wiliam, 1998a), as well as for public reporting (Guthrie, 2002).

Brown's (2002) student accountability and school accountability conceptions or Remesal's societal conceptions reflect AoL. This conception type focuses on institutional goals or communal and societal perspectives and interests. Accountability conceptions seems to be preferred by teachers in educational settings in which priority is given to summative assessment (Chan, 2007), the condition that is similar to the context of this study.

Nevertheless, teachers may hold various combinations of conceptions like those falling between pedagogical and societal (Remesal, 2011), improvement and accountability (Brown, 2002), transformational and reproductive (Postareff, et al., 2012). This condition potentially brings internal disagreement or negative impressions of the purposes of assessment among teachers which Brown identifies as irrelevance (Brown, 2002). There is very limited study on how Indonesian teachers understand assessment and how their understanding is reflected in practice.

This study is conducted to explore junior secondary school English teachers' conceptions of assessment and to investigate how and why teachers believed in particular assessment conceptions. It is carried out to answer this main question: How do Indonesian junior secondary school teachers understand the role of assessment and how do they perceive that assessment impacts upon student learning?

## **METHOD**

The participants of the study are 107 English junior high teachers in one region of South Sulawesi, Indonesia. They were recruited by accessing the database held by the Department of Education to identify teachers of English who were actively participating in a professional development (PD) programme currently being conducted in the Gowa Region, South Sulawesi. Teachers working in the Gowa region were specifically targeted as Gowa is fairly typical in size and population of a region in the South Sulawesi province. There are nineteen

districts in this region, and each district has a minimum of two junior secondary schools. Recruitment of participants for the study was through their PD groups which were organised according to district.

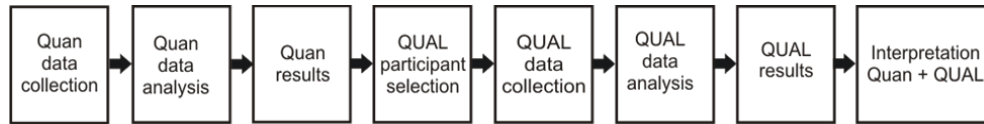
Table 1 summarises the demographic characteristics of participants including age, gender, qualification level and teaching experience.

**Table 1. Basic Demographic Information of Participants**

Characteristics	Category	Number	% Total N=107
Gender	Female	78	72.9
	Male	29	27.1
Age	Under 23 years	1	0.9
	23-30 years	32	29.9
	31-40 years	30	28.0
	41-50 years	40	37.4
	Over 50 years	4	3.7
Education level	Diploma	2	1.9
	Bachelor	100	93.5
	Master	5	4.7
	Doctor	0	0
Teaching experience	Less than 3 years	15	14.0
	Between 3-10 years	46	43.0
	Between 11-20 years	27	25.2
	Over 20 years	19	17.8

It was notable that the largest group of participants were female teachers. Participants mostly belonged to the 41-51 year age group and the majority of them had attained a Bachelor's degree. Almost half of the participants had been teaching between 3-10 years and only 14% of the total respondents reported having less than three years teaching experience.

In this mixed methods study, I connected the quantitative and qualitative phases during stage two by selecting participants for the qualitative phase from those who responded to the survey in the quantitative phase. Secondly, I connected the phases by using the quantitative survey findings to inform the development of relevant interview questions to explore in greater depth the participants' beliefs. The third integration is conducted during the discussion of findings. A visual model of how the study was integrated is presented in Figure 2.



**Figure 2: Visual Diagram of Sequential Explanatory Design: Participant Selection Model (Creswell & Plano Clark, 2007, p. 73)**

In the quantitative phase, participants completed the Hong Kong model of teacher conceptions of assessment (TCoA) questionnaire (Brown, Hui, & Yu, 2010) during October 2012 to January 2013. I chose to use the validated Hong Kong model of TCoA (Brown, et al., 2010) because I anticipated that Indonesian teachers might hold similar conceptions to teachers in Hong Kong and China.

I prepared a translation of the Hong Kong model of the TCoA survey, which was then approved by Brown, and piloted with individuals similar to the intended participants in Indonesia. Feedback from the trial was used to revise the questionnaire into its final form and this was checked with Brown.

The main part of the questionnaire consists of several subscales. Improvement subscales comprise eleven items that evaluate teacher agreement against three dimensions. Five items address assessment and student development and three items measure teachers' responses to the function of assessment to improve students' learning. The remaining three items belong to the dimension of accuracy in assessment. There is only one dimension asking about irrelevance. The third subscale (accountability) contains the most items. There are eighteen items which investigate three dimensions: eight items belong to the examination category, two items address error in assessment and the five remaining items measure the accountability purposes of assessment in terms of illustrating the quality of teacher and school. Table 2 summarises the meta-factors, factors, and examples of items in the TCoA survey.

**Table 2. Dimensions of the Teachers' Conceptions of Assessment (TCoA) Survey**

Conceptions of assessment	Dimensions	Examples of Items
Improvement	Students development (5 items)	<ul style="list-style-type: none"> <li>▪ Assessment is used to provoke students to be interested in learning.</li> </ul>

<b>Conceptions of assessment</b>	<b>Dimensions</b>	<b>Examples of Items</b>
		<ul style="list-style-type: none"> <li>▪ Assessment cultivates in students a positive attitude towards life.</li> </ul>
	Help learning (3 items)	<ul style="list-style-type: none"> <li>▪ Assessment helps students improve their learning.</li> <li>▪ Assessment determines if students meet qualification standards.</li> </ul>
	Accuracy (3 items)	<ul style="list-style-type: none"> <li>▪ Assessment results are trustworthy.</li> <li>▪ Assessment results can be depended on.</li> </ul>
Irrelevance	(5 items)	<ul style="list-style-type: none"> <li>▪ Assessment has little impact on teaching.</li> <li>▪ Assessment forces teachers to teach in a way that is contrary to their beliefs.</li> </ul>
Accountability	Examination (8 items)	<ul style="list-style-type: none"> <li>▪ Assessment prepares students for examination.</li> <li>▪ Assessment selects students for future education or employment opportunities.</li> </ul>
	Error (2 items)	<ul style="list-style-type: none"> <li>▪ Teachers should take into account error and imprecision in all assessment.</li> <li>▪ Assessment results should be treated cautiously because of measurement error.</li> </ul>
	Teacher and school control (5 items)	<ul style="list-style-type: none"> <li>▪ Assessment results contribute to teachers' appraisal.</li> <li>▪ Assessment is an indicator of a school quality.</li> </ul>

In the qualitative phase, qualitative semi-structured interviews and document analysis were utilized to construct three case study groups to understand how and why teachers believed in particular assessment conception. The interviews were conducted during January to April 2013 involving 12 teachers.

Data obtained from the questionnaire was analysed using SPSS version 19. Descriptive statistics were employed to analyse respondents demographic information. Percentages and frequencies were calculated to ascertain the



extent to which respondents believed in the function of assessment to improve learning, the extent of their confidence in assessment being irrelevant and the extent that they believed in the accountability purposes of assessment. Cronbach's Alpha was conducted to determine the internal consistency and reliability of the questionnaire. In addition, the validity of the questionnaire was also tested by executing exploratory factor analysis. Qualitative analysis involved both manual thematic analysis and Nvivo 8. The researcher identified main themes from participants' interviews, categorized, coded and grouped them according to the themes that emerged.

## FINDINGS AND DISCUSSION

### Findings

#### *Overall Results of Teachers' Conceptions of Assessment*

Findings regarding overall results of participant responses to the TCoA questionnaire can be seen in Table 3.

**Table 3. Agreement Level of Teachers for Components in TCoA (N=107)**

Components	Mean	SD
Improvement	4.99	1.09
Irrelevance	1.94	1.27
Accountability	4.66	1.25

Table 3 reveals that assessment for improvement received the highest agreement (M=4.99) followed by the accountability conception (M=4.66). The two conceptions were not widely different and participants tended to answer *mostly agree*. On the other hand, the irrelevance conception gained the lowest response (M=1.94) and participants generally chose *mostly disagree*. All subgroups indicated similar standard deviation above 1 %. The irrelevance subgroup received the highest score (SD=1.27), while accountability received 1.25 and improvement 1.09 respectively. This suggests a minimal variation in comparison to the other two subgroups.

This table also shows that participants in the study believed in conducting assessment for improvement purposes as well as to describe accountability of

students and school. Teachers also revealed that assessment was important by giving a low rating to items testing for irrelevance. This response pattern is consistent with previous studies using the same instrument such as those conducted in New Zealand (Brown, 2002), Virginia, USA (Calveric, 2010) Ankara, Turkey, (Vardar, 2010) and the Netherlands (Segers & Tillema, 2011).

The analysis of key ideas involved coding, merging codes, and naming and renaming themes. The interview data revealed that teachers' understanding of assessment purposes fell into two groups. They differentiated between internal assessment and external assessment. Teachers' understanding of these assessment types are illustrated along with supporting quotes in the following section. In this paper I only present and discuss finding of one case study group: Improvement (IM). Participants of this group gave the highest agreement toward improvement conception.

Intan, Emma, Andin and Lisa (pseudonyms) are examples of teachers holding a conception of assessment which valued improving teaching and learning. Their responses to the questionnaire reflected their strong advocacy for using assessment for students' development, to help learning and develop accuracy. In addition, they agreed with accountability purposes in assessment and were unsure whether assessment could be considered irrelevant.

### ***Internal Assessment***

IM teachers' values of assessment fall into the two main categories of internal and external assessment. Internal assessment in the improvement case consists of the themes of openness to change, authentic formative assessment practices, and grading to show achievement.

IM teachers favoured internal assessment where they could adjust teaching, develop values and practise with numerous assessment strategies. Their statements about assessment revealed that they were open to change, believed that assessment was an effective way to teach students good values and indicated an agreement to score students.

### ***Openness to Change***

IM teachers believed that the main purpose of assessment was to inform teaching. Through internal assessment teachers could measure the quality of teachers' teaching. IM teachers reported that it was crucial for them to reflect upon students' assessment results.

I question myself, is my teaching effective? Do my students understand the lesson? (Lisa)

These teachers worried about the efficacy of their teaching and signaled a willingness to modify their teaching continuously. IM teachers tended to articulate a belief in formative assessment because they used their classroom assessments to inform their teaching. In their consideration of students' results, IM teachers involved their students in the reflection process. This two-way communication tracked the efficacy of teachers' teaching:

Every semester I ask my students to comment on my teaching...do they like the strategies I use, how they want the learning process run. (Lisa)

To IM teachers, internal assessment could also bridge a better relationship with students. IM teachers seemed to hold themselves accountable for students' motivation in addition to their academic success. They took responsibility for encouraging low-achieving students through personal conversations:

I try to motivate them, I give them feedback, I ask what causes such unsatisfactory results, what is the problem, which part is hard and so on...So I assist them to realise reasons behind their failure. (Intan)

IM teachers appreciated that a low score could affect students' confidence negatively so these teachers encouraged students not to be pessimistic:

I convince them that they have many opportunities to improve and I give them those chances. (Emma)

Teachers' reflections and communication with students resulted in changes to their teaching approaches:

This strategy might work well in class A or B but does not necessarily work in Class D. Sometimes I almost give up, but at the same time, it motivates me to continuously search for the right strategies. (Emma)

This is an example of how IM teachers adjusted their teaching in order to meet the students' needs. IM teachers responded to assessment results positively in the sense that they were encouraged '*to learn more in order to be a good teacher*' (Intan). This might explain why they valued classroom assessment.

My assessment is 99.9% reliable...I assess my students based on their proficiency and I am confident that it is accurate; I am opposed to those who manipulate students' scores. (Emma)

Thus, IM teachers favoured internal assessments as an evaluation tool in order to improve the quality of their teaching. They trusted that internal assessment generated trustworthy results.

Another way that IM teachers showed their openness to change was in supplementing the set textbook. IM teachers relied on the textbook as the main source of teaching materials. IM teachers were confident in the validity of the textbook as a teaching tool. The textbook published by the Ministry of Education contains all key competencies and tasks that are accessible to both students and teachers. However, they were willing to include additional materials.

I use the textbook as the main resource of my classroom activities but also combine with other resources especially for listening activities. (Intan)

This extract signals that IM teachers were open to supplementing the textbook with extra resources:

I love going to bookstores to find additional resources. I use all [kinds of] relevant materials in my teaching including CDs and cassettes. (Intan)

Teachers from this group also showed me some teaching materials they used in the classroom. Few of these were taken from the textbook. Most were colourful, laminated and taken from a newspaper, magazine and other authentic sources. These teaching materials and the portfolio assignment suggest that IM teachers are interested in the use of authentic materials to assess students' learning.

### ***Authentic Formative Assessment Practices***

IM teachers based their teaching and assessment on the four micro skills: reading, writing, listening and speaking. They recorded these scores in their modified scoring grid document. They reported that authentic English teaching requires all four modes of the language to be taught. In the current curriculum, called the school level curriculum, English language scores are composite. This means that scores for listening, speaking, reading and writing are collapsed into a single overall score and some teachers prefer to concentrate their teaching on reading and writing at the expense of skills that are not examined. Nonetheless, teachers in the improvement group reported teaching and assessing students according to all four language skills.

...I measure speaking, listening, reading and writing. So students are scored based on these skills. (Lisa)

IM teachers reported that they specifically taught each of these skills and tried to involve their students in assessments that would inform their learning. These teachers reported using oral tests in speaking, listening and some reading classes, and employed written tests in writing and reading comprehension classes.

IM teachers wished to use assessment that matched each skill which would ensure the validity of the test. These teachers insisted that it was necessary to use several criteria to determine students' scores such as looking at students' responses in listening, as well as whether or not they followed the commands given by teachers. Speaking skills were challenging in that there were multiple aspects to evaluate.

I use some criteria such as content, fluency, intonation, and pronunciation". (Emma)

IM teachers also reported developing particular criteria and rubrics in assessment to assist student learning. These teachers insisted that sharing the assessment criteria with students was important in improving students' learning because it helped them to realize the measurement aspect.

This understanding indicate that teachers tried to communicate their expectations in order to guide students towards performing work that would meet the set criteria (Martins, 2008). It also suggests that the IM teachers tried to share responsibility for the learning and assessment process with their students. IM teachers illustrated authenticity in assessment by assessing students per skill. They also utilised what they believed as the appropriate pedagogy in teaching English. They felt that contextualising teaching materials and creating interesting learning experiences could support student achievement.

### **Grading to Show Achievement**

IM teachers believed in authentic classroom-based assessment that would inform teaching and learning. However, they also indicated a strong interest in grading practices. IM teachers believed that assessment should be useful in revealing student learning. They argued that testing was a preliminary instrument to measure the quality of students' work. One confidently stated that it helped

...to measure students' proficiency, how well they comprehend the lesson.  
(Intan)

IM teachers appeared to believe that students have different levels of competency; assessment assisted them to recognise students' position among their fellows. IM teachers assumed that the higher the score, the more learning a student had achieved. IM teachers seemed to be happy in using assessment results to map student achievement levels. These teachers appeared to believe that assessment and scoring were inseparable.

...it is a sign, whether students master the teaching materials or not, if they get 10 (ten), it means they understand the lesson well. (Andin)

In other words, IM teachers viewed a score as a symbol of achievement. This understanding appeared to be strong and was reflected in these teachers' teaching and assessment practices. Teachers also contended that a good score was important for students' self-actualisation and provided an indication of their further education level.

...when students have a good score or when they have good achievement, they will be recognized in our society. (Intan)

It seemed that both teachers and students believed that scores reflected students' achievement, suggesting that scoring and grading were a valued part of learning. Students were '*happy when their work is scored*' (Andin) and requested every task to be rated by asking '*will you mark this or not?*' (Lisa). Students appeared to '*study only to get a good score*' (Lisa). IM teachers believed that assessment could be used to identify students' position against the standards, acknowledge students' positions among their fellows and build positive impressions among community and society.

This phenomenon indicates a communal assumption that grading is crucial for communicating the quality of learning. This shared perception of the value of grading suggests that scoring and grading is a significant part of the participants' culture. The IM teachers did not find an emphasis on grade to be incompatible with a focus on formative assessment.

### ***External Assessment***

IM teachers believed that exam-based external assessments are an important means of evaluation to certify students' learning, teachers' teaching,

and the accountability of a school. However, they also felt that external examinations had a negative impact on a teacher's autonomy, equity among students and they doubted the credibility of external tests. This ambivalence was reflected in the themes that arose which related to conflicting conceptions, teacher autonomy and the credibility of external assessment practices.

### **Conflicting Conceptions**

IM teachers believed that external assessment could assure teachers' accountability. These teachers believed that the collection of students' assessment results was an indication of the quality of teachers' teaching.

It tells me whether I am able to transfer knowledge to my students. (Andin)

This understanding suggests that IM teachers supported conventional or behaviourist views of learning. These teachers assumed that students' success in learning was visible when they could reproduce teachers' knowledge. In short, IM teachers considered that assessment was a powerful way to communicate the standard or quality of students' learning and teachers' teaching.

...they [parents] look at the result. When it is good, it means the teachers are qualified. (Emma)

In addition, IM teachers agreed with the government's use of external assessment to evaluate the quality of a school. IM teachers understood the main function of this high-stakes external assessment was to assign each school a rank in the region, province and country.

The students' proficiency indicates the quality of a school (Lisa)

This extract illustrates that IM teachers acknowledged the role of society in assessment. The strong confidence placed in examination scores as a measure of students' proficiency by students, teachers and the community appeared to be underpinned by the National Examination. While IM teachers supported the use of high-stakes assessment to fulfill the above purpose, they appeared to have mixed feelings about it. In contrast to their support for assessment as a means to hold schools accountable, IM teachers also argued that using a score as the indicator of quality placed schools in a difficult situation. IM teachers reported that some schools felt obliged to engineer acceptable scores to fulfil the demands of society, or to encourage new enrolments.

We are required to report scores which are acceptable or above the standards, otherwise teachers' and the school's reputation is contested". (Emma)

The effect of student achievement in external examinations on the wider community was obvious to this teacher.

Students' examination results determined the reputation of a school, influencing prospective parents seeking to identify and select prestigious schools for their children. IM teachers seemed to be in doubt about the validity of using external assessment to hold a school accountable in this way. They suggested that this also led to several unintended consequences for teachers' autonomy and the credibility of tests.

IM teachers reported assessment practices that accorded with their understanding of assessment. Being unable to mitigate against the dominant role of external assessment in their context, they conformed to practices that fit assessment of learning. In informal classroom assessment situations where their autonomy was acknowledged, IM teachers confidently implemented constructivist assessment practices. However, these appeared to only receive a very small portion of the teaching time because teachers used the opportunities presented by other internal assessments like teacher tests or mid-term semester tests to familiarise students with the format used in external assessment.

However, their focus on external examinations seemed to conflict with their other, formative practices. These teachers reported that assessments were an exercise to '*familiarise students with the final examination*' (Andin). For this reason, the semester test was intended to shadow the final examination and teachers created internal tests that resembled models used in external assessment.

I use matching, multiple choice, particularly for the ninth grade students. (Intan)

IM teachers' decision to familiarise students with external assessment formats indicated their focus on external assessment. In extreme cases, teachers engaged in intense examination preparation providing extra classes and exercises similar to the examination formats.

In the second semester of the third year, we give extra lessons to students which are conducted after school. Moreover, four or five weeks before the examination, the curriculum section [of the school] will design a new lesson



schedule so that students will learn only the subjects tested in the examination. (Intan)

Teachers seemed to be unaware that such practices limited the subject matter content as well as potentially limiting students' engagement with the content. It seems that the National Examination influenced both these teachers and their schools to maintain a focus on summative assessment practices.

### ***Case Summary***

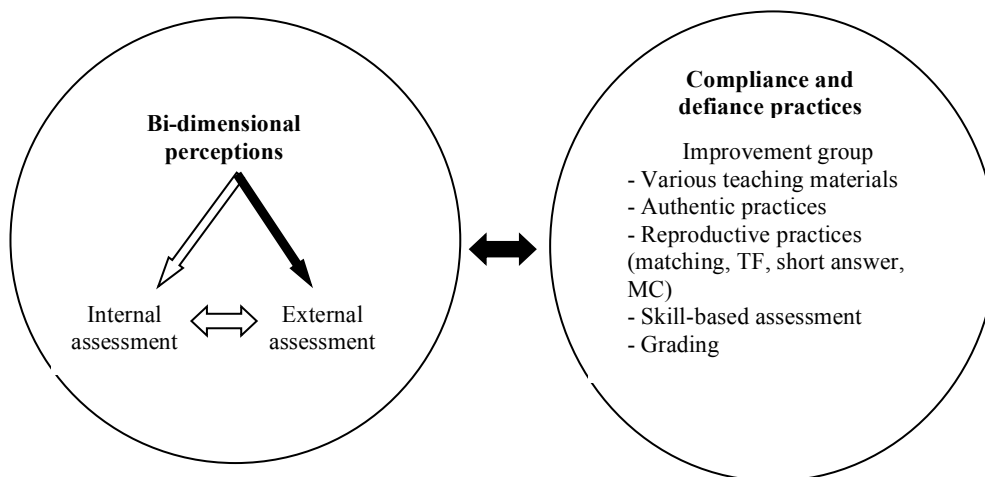
IM teachers reported divergent understandings of assessment purposes that could be categorised as both formative and summative. They favoured teacher-driven assessment (formative) where they had room to develop and implement strategies to improve teaching, learning, and felt that their own assessments were likely to be effective and reliable. In contrast, they reported less faith in exam-based assessment conducted by external agencies. Although they appreciated that exam-based assessment (summative) was necessary to certify students or to make teachers and schools accountable to the community, they argued that external assessment could impact negatively upon teachers, students and the credibility of tests. Their understanding and practices of assessment illustrated a conflict between internal and external assessment. IM teachers demonstrated formative assessment practices to reflect their understanding of internal assessment and more summative types of assessment to reflect their focus on external accountability. The emphasis placed on summative types of assessment despite their strong belief in assessment for learning revealed their conflicting conceptions.

### **Discussion**

The pattern of TCoA responses shown by IM teachers in the quantitative phase of this study is not consistent with their perceptions and practices. In their response profile they favoured improvement conceptions the most followed by accountability conception and they tended to disagree with irrelevant conceptions. This pattern should represent a strong commitment to the use of assessment result to improve teaching and to enhance learning. Teachers with this conception should also maximize the utilization of classroom and authentic assessment. However, data from the qualitative phase reveal that they actually held mixed conceptions which eventually influence their assessment practices.

***IM teachers' perceptions and practices of assessment***

IM teachers appeared to have bi-dimensional perceptions and practices of assessment which resulted from these socio-ecological influences. Figure 3 illustrates the bi-dimensional perceptions of internal and external assessment and compliance and defiance assessment practices. I use the term bi-dimensional to signify coexisting but inconsistent perceptions of assessment which existed within every teacher. As teachers revealed conflicting perceptions based on their consideration of internal (teacher-driven assessment) and external assessment (examination-based assessment), their assessment practices combined elements deriving from both. I call this compliance and defiance practice and subdivide this into several components: authentic vs. reproductive practice, using additional teaching materials and grading practices.



**Figure 3: IM Teachers' Perceptions and Practices of Assessment**

Perceptions and practices are interrelated as denoted by the two-headed arrow. The strong arrow connecting perceptions to external assessment indicates teachers' focus in this assessment. The compliance and defiance practices consist of several mixed practices combining authentic and reproductive. Internal assessment refers to teacher driven assessment during or after instruction which is conducted for quality assurance; this assessment is also called classroom as-

assessment (CA) (Harianti, 2005). To assure quality, teachers reported conducting CA both to evaluate the process and results of learning. This understanding was aligned with the MoNE expectation (Ministry of National Education, 2005c) which recommended evaluation through classroom tasks, teacher tests, a mid-term semester test and semester tests. This means that teachers' perceptions of CA reflected the published fitness of purposes of the assessment system in Indonesia. The focus of internal assessment shared between students as well as teachers might suggest a similarity to the improvement conception (Brown, 2002) or a pedagogical conception (Remesal, 2007, 2011).

The intention of the Indonesian model of classroom assessment is to monitor the on-going achievement of students and to summarise achievement at a particular time (Harlen, 2005; Segers & Tillema, 2011). This conceptions of classroom assessment is similar to the exclusively normative use of assessment in the USA (Shepard, 2000). In the normative use of assessment, students usually do tasks and perform 'to please the teacher or to get good grades rather than to pursue a compelling purpose' (Shepard, 2000, p. 31). This argument perfectly illustrates the condition of classroom assessment in the context because teachers believed that students' understanding of teaching materials was reflected in their grades. They believed that the higher the grade, the better the quality of teaching, so higher grades reflected more effective teachers. When interpreted through such conceptions, the intended purposes of classroom assessment were forgotten. Shepard (2000) contended that the compelling purpose of classroom assessment was to find shared understanding between teachers and students on what makes sense and what doesn't, and this could be used to design learning strategies. In other words, in their assessment practices, teachers in the context of the study appeared to only focus on what works rather than on what does not. This finding implies teachers were interested in evaluating the results of teaching rather than the process of learning.

Nonetheless classroom assessment policy was set by the MoNE and teachers simply performed their role as policy implementers. Indeed, the Indonesian classroom assessment model appeared to only partially support the empirical prototype of classroom assessment suggested by Resnick and Resnick (1992) or Airasian (1991) who proposed that classroom assessment should focus on teachers and students. Participants' classroom assessment evaluated instruction but results were rarely used to identify students' personal needs; classroom assessment seemed to only inform instructional changes but not students' learning. My participants supported internal assessments and demonstrated a prefer-

ence for policies in which their assessment practices promoted classroom tasks and teachers tests. Teachers argued for a version of classroom assessment in Indonesia focused on process where they could use and develop several assessment strategies to improve teaching. This understanding reflects formative purposes of assessment where teachers could continuously make changes during instruction (Boston, 2002) in order to improve teaching quality and students' learning (Black & Wiliam, 1998a; Harris & Brown, 2008).

However, my participants were constrained to evaluate and summarise students' achievement at particular times. These assessment practices included traditional assessments like paper and pencil tests with items covering multiple choices, true/false, matching, fill in /completion, short answers and essay tests. Any teaching adjustments were included in a remedial programme made after the teacher test which involved re-teaching the same materials with or without significant changes in strategies or giving students another chance to re-do the tests. This conception and practice contradicts earlier interpretations of classroom assessment (Resnick & Resnick, 1992; Stiggins, 1999). These authors contended that changes in teaching instruction should be conducted in day to day operation in order to maximise the diagnosable function of assessment to improve teaching and learning.

Teachers' confidence in the impact of CA on teaching was interesting because they tended to differentiate between the purposes of assessment for teaching and learning, components that are inseparable to some scholars (Black & Wiliam, 1998b, 2007; Stiggins, 1999). However, the findings from my study indicated a different understanding of connections between modifications to teaching and a possible improvement in learning. These teachers may have agreed in theory with the relationship between teaching and learning as explained in formative assessment (FA) but they interpreted these factors differently in practice. Teachers appeared to concentrate on one aspect only (teaching) assuming that teaching would automatically impact the other side (learning). This reasoning shared by participating teachers might signal the need for further investigation into the impact of assessment on teaching and learning. In other words, there is potential to separate the purposes of assessment for teaching and learning into two different components. This means that researchers need to clearly classify or separate the purposes of assessment such as for teaching and for learning in order to avoid confusions among teachers (Frey & Schmitt, 2010). This idea (which suggests a need for PD) seems not to have been considered by policy makers in Indonesia. Therefore, my participants'

conceptions that classroom assessment impacts on teaching and thus on learning could not truly be categorised as formative assessment (FA) or assessment for learning (AfL).

## CONCLUSIONS AND SUGGESTIONS

IM teachers superficially demonstrated an interest in FA or AfL in their classroom practices but a deeper examination of their beliefs revealed a closer affinity with summative assessment (SA). Even though these perceptions may seem inconsistent, they fit the purposes of the existing assessment system in Indonesia (Ministry of National Education, 2005a) that focuses more on external assessment. In other words, teachers' blurred perceptions that did not fully fit within parameters set by international scholars (Airasian, 1991; Black & Wiliam, 1998a; Resnick & Resnick, 1992) may have developed as a result of the Indonesian education system. Therefore, although participants in this study reported that the purpose of assessment was to inform teaching and improve students' learning, their understandings do not precisely align with an improvement conception (Brown, 2002), a pedagogical conception (Remesal, 2011), or a transformative conception (Postareff, et al., 2012). These participants' perceptions of assessment fit more closely within a societal or mixed societal conception (Remesal, 2011) or to accountability conceptions (Brown, 2002). At the same time, unlike participants in other studies, participating teachers also balanced improvement, pedagogical and irrelevance conceptions. As a consequence, their assessment practices signify a conflict which resulted in a combination of authentic and reproductive assessment. This might signify teachers' confusion of assessment conceptions which requires serious consideration from policy makers. The government should attempt to understand teachers' perceptions, knowledge, readiness and responses to any planned reform particularly on assessment. Otherwise, the result could be mixed if not useless.

## REFERENCES

- Airasian, P. W. (1991). Perspective on measurement instruction. *Educational Measurement: Issues and Practice*, 10(1), 13-16.
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.

- Black, P., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. London: Kings College.
- Black, P., & Wiliam, D. (2007). Large-scale assessment systems design principles drawn from international comparisons. *Measurement, 5*(1), 1-53.
- Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research & Evaluation, 8*(9). Retrieved from <http://pareonline.net/getvn.asp?v=8&n=9>
- Broadfoot, P. (2007). *An introduction to assessment*. New York: Continuum International
- Brown, G. T. (2002). *Teachers' conceptions of assessment*. (Doctoral Dissertation, University of Auckland, Auckland, NZ).
- Brown, G. T. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York: Nova Science Publishers.
- Brown, G. T., Hui, S. K. F., & Yu, F. W. M. (2010, July). *Teachers' conceptions of assessment: Developing a model for teachers in Hong Kong*. Paper presented at the Biannual conference of the International Test Commission, Hong Kong, China.
- Brown, G. T., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conception and practices of assessment. *Assessment in Education: Principles, Policy & Practice, 16*(3), 347-363.
- Brown, G. T., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education, 27*(1), 210-220.
- Butterfield, S., Williams, A., & Marr, A. (1999). Talking about assessment: Mentor-student dialogues about pupil assessment in initial teacher training. *Assessment in Education, 6*(2), 225-246.
- Calderhead, J. (1996). Teachers: Belief and knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 709-725). New York: McMillan.

- Calveric, S. B. (2010). *Elementary teachers' assessment beliefs and practices*. (Doctoral dissertation, Virginia Commonwealth University, Virginia, USA).
- Chan, J. K. (2007, May). *We have various forms of assessment but only summative assessment count: Case studies of the implementation of an innovative assessment policy in Hong Kong*. Paper presented at the Redesigning Pedagogy-Culture, Understanding and practice, Singapore.
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks: Sage Publications.
- Danielson, C. (2008). Assessment for learning: For teachers as well as students. In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 191-213). New York: Taylor & Francis.
- Delanshere, G., & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: A case of assessment paralysis. *Journal of Curriculum and Supervision*, 14(3), 216-240.
- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximise student learning*. Thousand Oaks: Corwin Press.
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-111.
- Frey, B. B., & Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal*, 5(3), 107-117.
- Griffiths, T., Gore, J., & Ladwig, J. (2006, November). *Teachers' fundamental beliefs, commitment to reform, and the quality of pedagogy*. Paper presented at the Australian Association for Research in Education (AARE). Adelaide, Australia.
- Guthrie, J. T. (2002). Preparing students for high-stakes test taking in reading. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction* (3rd ed., pp. 370-391). Newark: International Reading Association.
- Harianti, D. (2005). *Model penilaian kelas*. [Classroom assessment model]. Jakarta: Departemen Pendidikan Nasional.

- Harlen, W. (1998). Classroom assessment: A dimension of purposes and procedures. In C. Ken (Ed.), *SAMEpapers* (pp. 75-97). Hamilton New Zealand: Center for Science, Mathematics and Technology Educational Research, University of Waikato.
- Harlen, W. (2005). Teachers' summative practice and assessment for learning-tensions and synergies. *The Curriculum Journal*, 16(2), 207-223.
- Harris, L., & Brown, G. T. (2008, December). *New Zealand teachers' conceptions of the purposes of assessment: Phenomenographic analyses of teachers' thinking*. Paper presented at the Australian Association for Research in Education (AARE), Brisbane, Australia.
- Hill, M. (2000). *Remapping the assessment landscape: Primary teachers reconstructing assessment in self-managing schools*. (Doctoral dissertation, University of Waikato, Hamilton, NZ.)
- Ministry of National Education. (2005a). *Indonesian Government Policy Act No. 19. 2005: Standard of Education*. Jakarta: Ministry of National Education.
- Ministry of National Education. (2005b). *Peraturan pemerintah Republik Indonesia nomor 19 tahun 2005 tentang standar nasional pendidikan*. [Government regulation number 19 year 2005 on national standards of education]. Jakarta: MoNE.
- Nespor, J. (1987). The role of beliefs in the practice of teaching. *Curriculum Studies*, 19(4), 317-328.
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307-332.
- Postareff, L., Virtanen, V., Katajavuori, N., & Lindblom-Ylänne, S. (2012). Academics' conceptions of assessment and their assessment practices. *Studies in Educational Evaluation*, 38(4), 84-92.
- Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly*, 42(4), 203-220.
- Remesal, A. (2007). Educational reform and primary and secondary teachers' conceptions of assessment: The Spanish instance, building upon Black and Wiliam (2005). *The Curriculum Journal*, 18(1), 27-38.



- Remesal, A. (2009, December). *Spanish students teachers' conceptions of assessment when starting their career*. Paper presented at the Perceptions and conceptions of assessment in the classroom: Different national perspectives, Amsterdam.
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment : A qualitative study. *Teaching and Teacher Education*, 27, 472-482.
- Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessment: Alternative views of aptitude, achievement, and instruction* (pp. 37-75). Boston: Kluwer Academic Publishers.
- Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, 37(1), 49-54.
- Shepard, L. A. (2000). *The role of assessment in teaching and learning*. Santa Cruz, CA: CRESST/CREDE, University of California, Los Angeles.
- Stake, R. E. (2000). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 435-454). Thousand Oaks: Sage Publications.
- Stake, R. E. (2010). *Qualitative Research: Studying how things work*. New York: Guilford Press.
- Stiggins, R. J. (1999). Assessment, students, confidence, and school success. *Phi Delta Kappan*, 81(3), 191-198.
- van den Berg, R. (2002). Teachers' meaning regarding educational practice. *Review of Educational Research*, 72(4), 577-625.
- Vardar, E. (2010). *Sixth, seventh, and eighth grade teachers' conception of assessment*. (Master's Thesis, Middle East Technical University, Ankara, Turkey).
- Wiliam, D., & Black, P. (1996). Meaning and consequences: A basis for distinguishing formative and summative functions of assessment. *British Educational Research Journal*, 22(5), 537-548.

Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well : Investigation new forms of students' assessment. *Review of Research in Education, 17*, 31-74.