

Comparison of Classroom Assessment Practices: A Case of Selected Ghanaian and Japanese Mathematics Lessons

Katsunori HATTORI ⁽¹⁾
Adzifome Nixon Saba ⁽²⁾

International Cooperation Center for the Teacher Education and Training
Naruto University of Education ⁽¹⁾
Peki Training College, Ghana ⁽²⁾

Abstract : This was a case study that examined and compared the classroom assessment practices of teachers of some selected Ghanaian and Japanese primary school mathematics lessons based on the National Council of Teachers of Mathematics (NCTM) assessment standards. The objective was to determine whether there were any differences and/or similarities in assessment practices between the two nations and how they measured up to the NCTM standards. It involved the use of a thirty-seven-item evaluation sheet, developed based on the NCTM standards to assess eleven video lessons. Three conveniently sampled Ghanaian lessons and eight Japanese lessons, four of which were conveniently sampled and the other four were already available specially recorded lessons used for in-service training in Japan. Both qualitative and quantitative data were obtained and analyzed item by item and section by section. The sectional analysis of the means and standard deviations values of the ratings of the two raters revealed that both categories of Japanese lessons showed more evidence of the tying of assessment to curriculum, integration of assessment and instruction, and effective blending and use of formal and informal assessment strategies than the Ghanaian lessons. However, the Ghanaian lessons promoted and were more sensitive to equity in assessment than the Japanese normal lessons. One-way analysis of variance and subsequently Tukey test yielded corroborative results. The means for the three categories were statistically significantly different from one another. The conclusion was drawn that, though there were a few similarities, there were also substantial differences, qualitatively and quantitatively, in the classroom assessment practices of teachers of the two nations.

Keywords : psychometrics, equity in assessment, cascade model, lesson study

Introduction

Assessment is one area of education that continues to generate discussions in professional and research cycles both from the perspectives of philosophy and practice. Assessment has become an instrument of change (Ruthven, 1994), a means of quality control and

an instrument of educational reform (Stakes, 1998). Education agencies and professionals have contributed to the pool of information, all with the one aim of reaching an assessment ideal which so far seems illusive in most educational systems. Gipps (1994) acknowledges assessment has taken on broader definition and purpose and gone through a “paradigm shift from psychometrics

to a broader model of educational assessment” (p1), characterized by the wide range of assessment presently in use now. It called for “an assessment for learning” (Gipps, 1994) instead of an assessment of learning.

Mathematics assessment is defined as “the process of gathering evidence about students’ knowledge of, ability to use and disposition toward mathematics and of making inference from that evidence for a variety of purposes” (NCTM, 1995, p.3). There is the need to appraise how teachers of Ghana and Japan obtain evidence of students learning. Thus, a comparative study of classroom assessment practices of the two nations is significant and worthwhile since they have essentially same system of education. Until recently, they had same 6-3-3-4 structure of education, which is six years of primary education, three years of lower secondary education, three years of upper secondary education and four years of university education. Ghana’s education system is an adoption of the Japanese and American model (Quist, 2003). Both emphasize the teaching and learning of mathematics and use students’ performances in national examinations to measure effectiveness of teaching in the classroom. The point is, whether a test taken at a sitting sufficiently samples and measures all the skills, knowledge and disposition required. Assessment is now viewed in the context of what primarily goes on in the classroom. Stakes (1998) indicates that “the validity of measurement of achievement is not the same as validity of those same scores as an indicator of quality of teaching and learning conditions”. This approach views assessment as a process-based on what students are exposed to and which informs what they produce.

The role of mathematics cannot be over-emphasized in a globalized world. Namukasa (2004) describes globalization as a phenomenon which among other things link education to international curriculum harmonization. Thus, education within national boundaries is now being influenced by factors without than within. For the two nations, it becomes imperative for the teaching and learning of mathematics be thoroughly assessed in line with national standards designed in consideration of globalization and its consequent internationalization of education. This should be done so that, it provides clear and accurate information on the extent to which students have achieved targets in terms of mathematical knowledge, skills and disposition. Having glowing curriculum philosophy and recommended assessment practices is one thing and their implementation in the

classroom is another. The extent to which these are reflected in the day to day classroom experience needs to be appraised. Therefore, the challenge is how to close the gap between the intended assessment and implemented and/or attained assessment in the mathematics classroom.

At least, calls for reform and the perception of fallen standards in both countries are ample testimonies it remains a daunting challenge.

The National Council of Teachers of Mathematics (NCTM) addresses these concerns by proposing standard-tied assessment reforms that calls for the matching of assessment to curriculum; integration of instruction in assessment; promotion of equity in assessment; the use of both formal and informal strategies to obtain assessment information and the central role that teachers should play in assessment. These underlie its six assessment standards for judging the quality of mathematics assessments practices (refer to the Assessment Standards for School Mathematics).

Literature supports the integration of assessment and instruction and the tying of assessment to curriculum. On assessment and instruction, there are calls for the integration of teaching, learning and assessment (Bolte, 1999; Ruthven, 1994). However, the situation is that many a times the inability to establish the relationship between learning and assessment results “in a mismatch between the high quality learning described in policy documents as desirable and the poor quality learning that seems likely to result from associated assessment procedures’ Willis (1992b, p.1, cited by Gipps, 1994, p.4). Gipps (1994) urges the recognition of the improvement of teaching and learning as the prime purpose of assessment.

On the mathematics curriculum, educators, policymakers, and parents are beginning to recognize that minimums and basics are no longer sufficient (Winking & Bond, 1995) and that students should be able to think critically, analyze and make inferences (Bond, 1995). The primary aim of assessment is to foster learning of worthwhile academic content for all students (Wolf, Bixby, Glenn, & Gardner, 1991). While NCTM (1995) calls for the teaching of important mathematics contents like: “algebra, geometry, trigonometry, statistics, probability, discrete mathematics, and even calculus” (p.2) and calls for instruction that emphasizes mathematical inquiry and conceptual understanding and stimulates intellectual learning. Ruthven (1994) argues that public assessment influences “curriculum and pedagogy

colloquially expressed as 'WYTIWYG' syndrome: namely 'what you test is what you get' (p.433). Therefore Classroom assessment should not be limited to what gets tested alone but foster the teaching and learning of worthwhile academic content and provide meaningful mathematics experience.

The relevance of both formal and informal assessments in obtaining valid and reliable inferences of students learning has been recognized. NCTM (1995) call for a reliance on the professional judgments of teachers and judgment of students' achievement on more than formal sources, arguing that teachers are primary assessors and in best position to judge students' progress. Taylor & Nolen (1996) reveals the classroom context is one of fairly constant formal and informal assessment. However, demands for valid and reliable assessments have made the formal means dominate assessment practice. But, Marony & Olssen (1994) argues informal assessments in the form of teachers' observations, peer-assessment and students' self-assessment have legitimate and valid place in obtaining valuable inferences of students' learning. The real progress during instruction can better be appreciated through teachers' observations. Also, students' self-assessment enables them to focus any reflection on their learning (Marony & Olssen, 1994) and foster their confidence and independence in learning mathematics (NCTM,1995) since they are capable of assessing their own performance (Kasanen & Raty, 2002; Brookhart, Andolina, Zuza & Furman, 2004)). Watson (2000) argues that, objective statements of mathematical attainments are possible and provide yardsticks against which to judge the reliability of informal assessment.

The inability of some students to meet expected educational goals brings the issue of equity to the fore. There are disparities in mathematics performances with respect to gender (Beller & Gafni, 2000), language or cultural background (Evans, 2006) and then economic and social backgrounds (Namukasa, 2004 NCTM). NCTM's view of equity is that all students are exposed to same content and given opportunity to demonstrate their knowledge and supported to attain expected levels. The unique background and ability of each learner must be recognized (Levin, 1993). NCTM (1995) argues that the uniqueness of each student's background in terms of experience, physical condition, and gender, ethnic or cultural and social has been ignored. Also, Malloy & Malloy (1998) have called for culturally relevant

mathematics teaching. It's imperative that equity is ensured in mathematics teaching, learning and assessment through fair means.

Professional development and training in assessment in particular has been recognized as very essential for the success of any assessment reforms. NCTM (1995) stresses the use of teachers' professional judgment to obtain adequate and relevant inferences of students learning. Noting, the validity of such inferences depends on their expertise. However, teachers have not been adequately prepared to create and conduct valid assessments (Taylor & Nolen, 1996; Novick, 1996) and lacked professional training (Stigler & Hiebert, 1999; Winfield and Woodard, 1994). This calls for mathematics teachers to be trained to teach important mathematics and use strategies consistent with the current vision of assessment to obtain relevant inferences of students' mathematics learning.

Theoretical and empirical literature support the current assessment paradigm which calls for mathematics assessment matched to what is considered important content, integrates assessment in instruction, promotes equity and involves teachers, equipped with the expertise, to obtain valid and reliable inferences of students' learning. The implication is for needed changes in what mathematics is taught, the way it is taught, and how its learning is assessed to ensure all students learn what they should.

This research does not only reveal the state of affairs by identifying the differences and similarities between Ghanaian and Japanese lessons and the extent to which they meet standards but by implication, suggest another means of assessing the effectiveness of teaching and learning in class than the use of students' performance on tests. Teachers, educationists and other interested persons in education will therefore find this study useful.

Methodology of the study

The study was case study and comparative study. For the purpose of this study, qualitative and quantitative data were collected through observation (notes taken and raters' remarks) and through the use of an evaluation sheet (a Likert-type scale) respectively to assess video-recorded lessons. 11 video-recorded lessons from Ghana and Japan were conveniently sampled. Three from primary schools in Ghana (The intended fourth lesson

from Ghana could not be video-recorded due to administrative and technical difficulties), eight from Japan, four of which were conveniently sampled video lessons (recorded during *kenkyukai* (research meeting) and the other four were already available video lessons used for in-service training in Japan, considered ideal mathematics lessons. The lessons were video-recorded at schools in Tokushima prefecture in Japan and in South Dayi/ Kpando district in Ghana. Data was collected in Ghana in June, 2007 and between April and June, 2007 in Japan.

The evaluation sheet was categorized under four 4 assessment themes that underpinned and covered the NCTM assessment standards. The first section had 10 items and covered the mathematics standard, which dealt with the assessment and curriculum (content). The second had nine items and covered the learning standard, which dealt with the integration of assessment in instruction. The third section had items nine and covered the openness, inferences and coherence standards, which principally dealt with the use of teachers' expertise to obtain valid evidence of students' learning and the fourth section, had nine items and covered the equity standard. The four-point scale was used to avoid the centralization of rating.

Instrument validity and reliability was ensured through expert scrutiny and try-outs using neutral lessons. Three assessors used the instrument to rate a neutral lesson and data obtained was used to examine the reliability of the instrument. The calculated cronbach's alpha coefficient of reliability of the instrument yielded an alpha (α) value of .85, indicating a high reliability. Alpha values of above 0.8 are considered highly reliability (Cohen, Manion & Morrison, 2007).

Two raters finally evaluated the video lessons-categorized into three and coded. The transcribed lessons, syllabus and/or textbooks, teachers' lesson plans and students' work samples were provided. While the assessment was done independently, the raters discussed lessons or aspects of it to reach consensus and/or deepen their perception on any thorny aspects. The average for an item in each of the three categories of lesson was assigned as the extent to which that item was evident in the category, thus the NCTM assessment standards.

The data collected was analyzed using spreadsheet for graphical analysis and the Statistical Programme for Social Sciences (SPSS) software, Version 13 for the analysis of variance, Tukey test and t-test analysis. The

analysis was in the form of item by item, section by section and lesson category by lesson category.

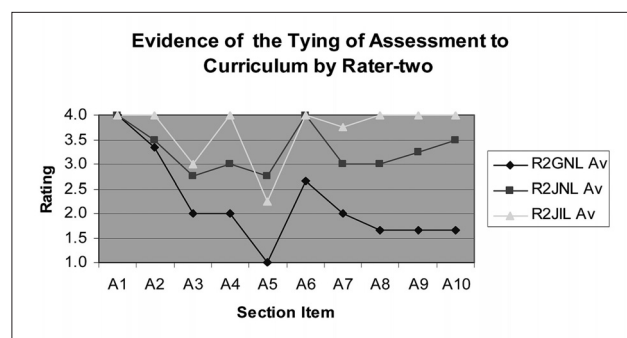
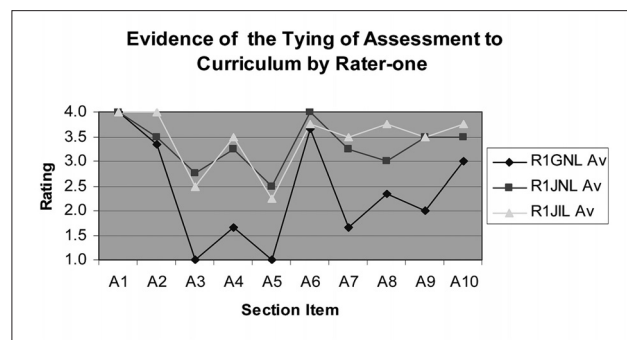
Findings

1. Graphical analysis

The line graphs give visual impression of lessons across groups and allow for comparison. The qualitative analysis segment is an attempt to complement and/or justify the quantitative data.

The flow of the JI and JN lessons though at different

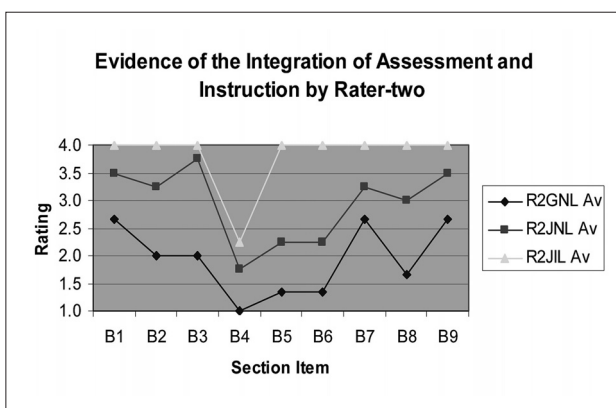
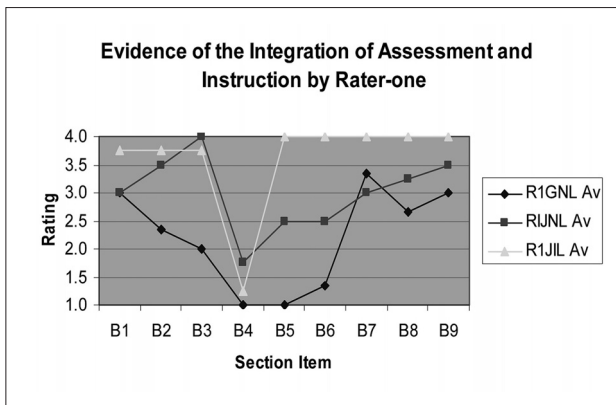
a) Evidence of the tying of assessment to curriculum



wavelengths (orientations) is similar suggested some commonality in the way assessment is connected to curriculum. All the three, had their lowest rating on item A3 (exploration of connection between topic and other areas in the curriculum) and/or item A5 (The use of current and available technology). While item A1 (whether lesson objectives tally with curriculum objectives) and A6 (currency and relevance of the content taught in today's context) had the highest ratings. From both graphs, in the JI lessons assessment was more tied to curriculum than the JN lesson and that was also more tied to curriculum than the GN lessons. The irregular pattern of the GN lessons on both graphs though shows differing views attest to lack some consistency.

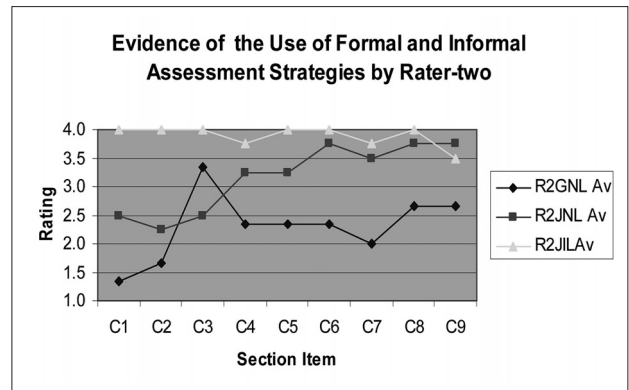
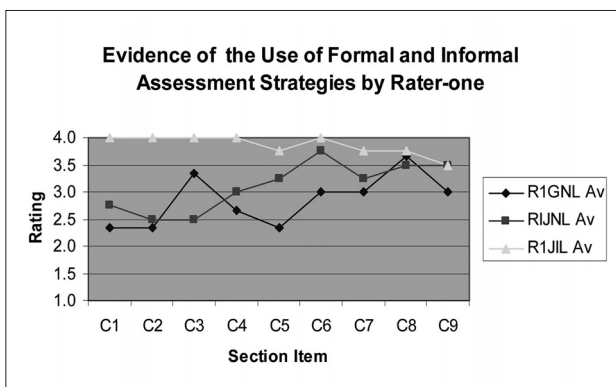
The graphs reveal some level of consistency and

b) Evidence of the integration of assessment and instruction



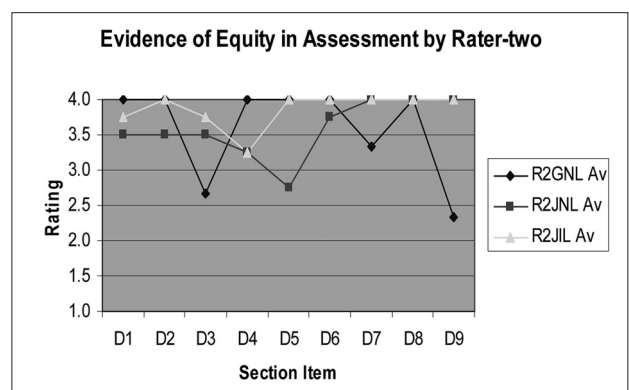
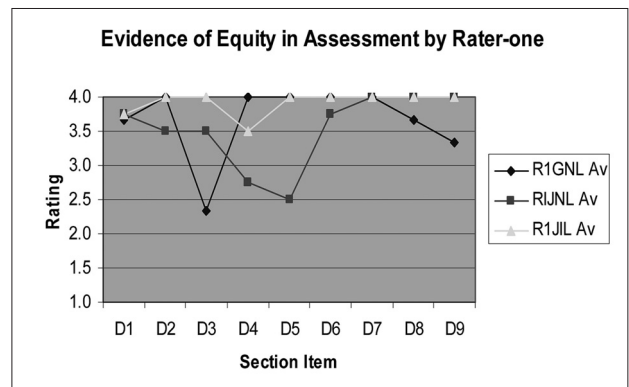
showed that the lowest rating of each category was on item 4 (problem creation after conceptualization). By far the JI lessons saw an almost complete integration of assessment and instruction than the JN lessons which also saw same much more than the GN lessons. GN lessons had the low ratings on items 5 (provision of open ended assessment tasks) and item 6 (fair and adequate assessment of various responses provided to open ended tasks) as well.

c) Evidence of the use of formal and informal assessment strategies



The graphs reveal high level of consistency in the lessons and shows that in the JI lessons, there was very high use of both formal and informal assessment strategies than the JN which in turn saw same more than the GN lessons, even though rated above average. Considering both graphs, the high ratings for the JN and GN were from items 6-9, the same items were the lowest for the JI lesson though higher as compared to the former. The JI lesson had its highest rating from item 1-4 and these happen to be the low points for the JN and GN lessons.

d) Evidence of equity in assessment



Considering the JI lessons, the least rating, though high, was on item 4 (Even distribution of questions

across student's geographical position, differing abilities and gender). For the JN, items 4 and 5 (gender and socio-cultural sensitivity in questioning and throughout instruction). These items were highly rated in the GN lessons. The GN had its least rating, though rated averagely; on items 3 (assistance given to individuals based on their needs), 7 (building of lessons on prior knowledge and experiences common to all students) and 9 (fairness in the scoring and/or interpretation of students responses) but these were highly rated in JI and JN lessons. There were differences on few items but to a large extent all three showed high level of sensitivity to the issue of equity in assessment.

2. Observed differences and similarities in assessment practices.

a) Differences in assessment practices

Japanese lessons promoted conceptual understanding and problem solving. The Ghanaian lessons remained essentially traditional in approach-based on the behaviorists view of a teacher as the dispenser of knowledge. The teacher dominated and used tailored questions that eliciting specific answers.

In the Ghanaian lessons contents were shallowly treated and obvious connections with other areas were not exploited. Specific procedures dictated by teachers were followed in investigation or problem solving. There were no prompts about alternative solutions neither were they elicited from or suggested by students. In many instances facts arrived at following rigid procedures were emphasized at the expense of the mathematical processes involved.

Ghanaian teachers mainly asked facts-eliciting questions that demanded students to make simple logical mathematical deductions from procedures and not that which challenged to investigation. Unlike the Japanese lessons, the Ghanaian teachers scarcely used the skill of observation to identify and exploit students' mistake and/or error to deepen and reinforce their understanding.

One distinguishable feature missing in the Ghanaian lessons was the in-depth mathematical discourse in which the Japanese teachers engage their students. These exchanges gave students' opportunity to evaluate the alternative responses and/or suggest view points on problem at stake. Through these students' deepened their understanding, peer-assessment, self-assessment were also promoted.

In two-thirds of the Japanese lessons, teachers heavily used sociometry. This was where a student who had responded to a question identified the student to respond to the next question. Teachers appeared hesitant in posing questions directly to students, limiting the answering of questions to a few students and paved the way for other learners to be ignored. It affected the effective distribution of questions across geographical, gender and ability wise. It created the impression that the teachers did not want to engage students in an emotional clash. Posing questions to students could be a means of external motivation that could make them come out of their shells. In the Ghanaian lessons, teachers directly posed questions to students.

b) Similarities in assessment practices

The contents were relevant and consistent with what NCTM considers as important mathematics. Lessons were built on previous knowledge common to all students. There was systematic and on-going assessment throughout the lessons. In most lessons, teachers showed appreciable sensitivity to the issue of equity. Students were given about the same amount of attention depending on the progression of the lesson and exposed to same challenging content. Except in one lesson (JNL2), questions asked by teachers and illustrations were generally gender neutral.

The use of technology for lesson delivery was not fully exploited though most of lessons lent themselves to it. Except in one lesson, problem creation by students was not observed. Though group work were designed as part of the lessons and formed, students' work individually (in the Japanese case) without visible interaction amongst the students and dominated by a few (in the case of Ghana).

3. Statistical analysis

a) Means and standards deviations (section by section)

The means and standard deviations values of the section by section analysis of the ratings of the two raters (Rater-two results in parenthesis) yielded the following corroborative results for Section A, which showed that the Japanese Ideal lessons (JIL) (M= 3.45 (3.70) SD= 0.599 (0.599)) and the Japanese Normal lessons (JNL) (M= 3.33 (3.28) SD=0.487 (0.463)) were more tied to curriculum than the Ghanaian Normal Lessons (GNL) (M= 2.37 (2.20) SD=1.082 (0.892)).

Also, section B analysis showed that the JNL

($M=3.00$ (2.94) $SD= 0.637$ (0.693)) and the JIL ($M= 3.61$ (3.81) $SD= 0.894$ (0.583)) saw more of the integration of assessment in instruction than the GNL ($M= 2.19$ (1.93) $SD= 0.899$ (0.641)).

Again, Section C analysis showed that both JNL ($M= 3.11$ (3.17) $SD= 0.475$ (0.588)) and the JIL ($M= 3.86$ (3.89) $SD= 0.182$ (0.182)) lessons effectively blended the use of both formal and informal assessment strategies than the GNL ($M= 2.85$ (2.30) $SD= 0.475$ (0.599)).

However for section D, the JIL ($M= 3.92$ (3.86) $SD= 0.177$ (0.253)) promoted and were sensitive to the issue of equity in assessment than the GNL ($M= 3.67$ (3.59) $SD= 0.553$ (0.662)) but the GNL also did so more than the JNL ($M= 3.53$ (3.58) $SD= 0.551$ (0.415)) though all three categories were appreciably rated high.)

b) Analysis of variance

One-way analysis of variance for Rater-one ratings: ($F = 16.675$, $p < 0.001$) and Rater-two ratings ($F = 35.500$, $p < 0.001$) showed that there were statistically significant differences among the three categories. The Tukey tests confirmed the means for the three categories Rater-one: (GNL = 2.75; JNL = 3.26 and JIL = 3.72) and Rater-two: (GNL = 2.50; JNL = 3.27 and JIL = 3.82.) were statistically significantly different from one another.

c) Inter-rater reliability

As a measure of the inter-rater reliability, dependent (paired) t-test was calculated for the two raters (using their ratings for each category of lessons). The t-test values revealed the observations made by the two raters about the JN lessons ($t = -0.223$, $p = 0.825$) were about same since there was no statistically significant difference in their ratings. However, there were statistically significant differences between their ratings of the GN ($t = 3.035$, $p = 0.004$) and JI ($t = -2.727$, $p = 0.010$) lessons. This means that the two raters differed widely on their observations. However, considering the high correlation values (GN: 0.854, sig. 0.000; JN: 0.923, sig. 0.000 and JI: 0.925, sig. 0.000), it can be interpreted that, though they did not place the same value on their observations (ratings), there was relative agreement between the two raters in their observations.

4. The general image or pattern of lessons

The observed teaching pattern in the Japanese lessons were essentially what is traditional about the teaching of

mathematics in Japan cited by Shimizu (1999) that Japanese lessons follow this order : 1. Presentation of problem; 2. Individual problem solving; 3. Whole-class discussion about methods for solving the problem; and 4. Summing up by the teacher (Exercises/Extensions). Except one (JNL1), all Japanese lessons followed this pattern of teaching. The Ghanaian teaching pattern can best be described as teacher-lead. It was neither lecture-based nor teacher-centered since the teachers went along with the pupils, assigned them tasks and some how involved them in active whole-class solution to problems. However, the distinguishable features were teachers' dominance, teachers dictating the pace and direction of lessons, approving and disapproving off students' responses to teacher-guided solutions to problems.

5. Reasons for uniformity

The high level of uniformity of assessment practices revealed in the Japanese lessons could be attributed to lesson study. (Lewis & Tsuchida (1998) and Aun et al. (2006) attributed the uniformity of education standards in Japan to lesson study and the use of the method identified as the traditional method of teaching in Japan (Shimizu, 1999).

Conclusion

The qualitative data obtained through observation reveal some measure of similarities but also revealed that there were large differences in the assessment practices used by the Ghanaian and Japanese teachers. The analyzed quantitative data revealed that, based on the NCTM standards, there was significant difference between the classroom assessment practices used by Ghanaian and Japanese primary school teachers. Therefore, the rejection of the null hypothesis and the acceptance of the alternate hypothesis that, there was statistically significant difference in the classroom assessment practices of the teachers of the two countries.

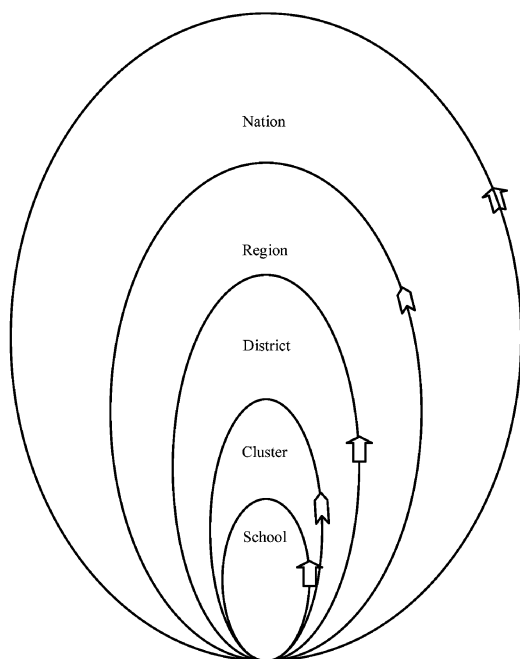
Recommendations

The findings outlined and the conclusion drawn from this study calls for necessary action. Therefore, the researcher makes the following recommendations. For Japan, the three areas of use of computer and current technology in the planning and lesson delivery; use of the technique of problem creation in the assessment and

exploration of obvious linkages of the topic to other curriculum areas are taken up as themes for lesson study meetings and mathematics fora to find ways of equipping teachers to deal with them.

For Ghana, teachers should deepen their content knowledge, teach content more tightly; employ methods that promote conceptual understanding and actively involve students in rich and cognitively demanding tasks; employ problem creation by students in assessment; explore the connection between topics and other areas in the curriculum; use open-ended assessments that call for variety of solutions and acquire ICT skills.

The General recommendations are for Ghanaian teachers to get exposure to the NCTM assessment standards and other professional publications to broaden perceptive and create means of exchanging ideas amongst themselves to improve assessment practices. The findings of this study are used as themes for in-service training programmes and encourage teachers to be classroom researchers. Based on the experience of the Japanese education system, we recommend the intromission of lesson study, using the cascade model, into the in-service education and training (INSET) programme in Ghana. The adoptions of lesson study rapidly rolls out the benefits and help improve teachers' assessment practices, teaching and learning in schools and promote professional growth.



The cascade model for the institutionalization of lesson study in Ghana

The cascade model ensures effective dissemination, enrichment, evaluation and continuity.

References

- Bolte, L. A. (1999). Enhancing and assessing preservice teachers' integration and expression of mathematical knowledge. *Educational Studies in Mathematics*, 2 (2),167-185.
- Bond, L.A (1995). Critical issue: Rethinking assessment and its role in supporting educational reform. Oaks Brooks IL: North Central Regional Education Laboratory. Retrieved 15/12/2007 from <http://www.ncrel.org/sdrs/areas/issues/methods/assment/as700.htm>
- Bond, L.A., Herman, J., & Arter, J. (1994). Rethinking assessment and its role in supporting education reform. In Laboratory Network Program, A Tool Kit for Professional Developers: Alternative Assessment. Portland, OR: Northwest Regional Educational Laboratory.
- Brookhart, M. S., Andolina, M., Zuzza, M., & Furman, R. (2004). Minute math: An action research study of student self-assessment. *Educational Studies in Mathematics*, 57,213-227.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*.5th Ed. New York: Routledge.
- Fernandez, C., & Yoshida, M. (2004). *Lesson study: A Japanese approach to improving mathematics teaching and learning*. New Jersey: Laurence Erlbaum Associates.
- Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. Philadelphia: RoutledgeFalmer.
- Kasanen, K., & Raty, H. (2002). "You be sure now to be honest in your assessment" : Teaching and learning self-assessment. *Social Psychology of Education*, 5,313-328.
- Levin, B. (1993, May 4). Students and education productivity, *Education Policy Analysis Archives*, 1 (5). Retrieved 2007/9/27 from <http://epaa.asu.edu/epaa/v1n5.html>
- Lewis, C.C., & Tsuchida, I. (1998). A lesson is like a swiftly flowing river: How research lessons improve Japanese education. *American Educator*, 2-17 & 50-52.
- Malloy, E.C., & Malloy, W. W. (1998). Issues of culture in mathematics teaching and learning. *The Urban*

- Review, 30 (3), 245-257.
- Morony, W., & Olssen, K. (1994). Support for informal assessment in the context of standards referenced reporting. *Educational Studies in Mathematics*, 27(4), 387-399.
- Namukasa, I. (2004). School mathematics in the era of globalization. *Interchange*, 35(2).
- National Council of Teachers of Mathematics (1995). *Assessment standards for school mathematics*: Reston, Va.: National Council of Teachers of Mathematics.
- National Council of Teachers of Mathematics (2005, Jan/Feb). PISA results show international standing of math students, *NCTM news bulletin*, 41(6), 2, & 6.
- Novick, R. (1996). Actual schools, possible practices: New direction in professional development. *Education Policy Analysis Archives*, 4(14). Retrieved 25/10/2007 from <http://epaa.asu.edu/epaa/v4n14.html>
- Quist, H.O. (2003). Transferred and adapted models of secondary education in Ghana: What implications for national development. *International Review of Education*, 49(5), 411-431.
- Ruthven, K. (1994). Better judgment: rethinking assessment in mathematics education. *Education Studies in Mathematics*, 27(4), 433-450.
- Shimizu, Y. (1999). Aspects of mathematics teacher education in Japan: Focusing on teachers' roles. *Journal of Mathematics Teacher Education*, 2, 107-166.
- Stake, R. (1998). Some comments on assessment in US education. *Education policy analysis archives*, 6(14). Retrieved on 2007/9/26 from <http://epaa.asu.edu/epaa/v6n14.html>
- Stigler, J.W., & Hiebert, J. (1999). *The Teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York: The Free Press.
- Taylor, S.C. & Nolen, B.S. (1996, Nov.11). What does the psychometrician's classroom look like? : Reframing assessment concepts in the context of learning. *Education Policy Analysis Archives*, 4(17). Retrieved 2007/9/27 from <http://epaa.asu.edu/epaa/v4n17.html>
- Watson, A. (2000). Mathematics teachers acting as informal assessors: Practices, problems and recommendations, *Educational Studies in Mathematics*, 41 (1), 69-91.
- Winfield, L.F., & Woodard, M.D. (1994, March). Assessment, equity and diversity in reforming America's schools. *Education Policy*, 8(1): 3-27.
- Winking, D. (1997). *Critical issue: Ensuring equity with alternative assessments*. Oaks Brooks IL: North Central Regional Education Laboratory. Retrieved 15/03/2007 from <http://www.ncrel.org/sdrs/areas/issues/methods/assment/as800.htm>
- Winking, D., & Bond, L. (1995). *What you and your school should know about alternative assessment*. Oaks Brook, IL: North Central Regional Educational Laboratory.
- Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. *Review of Research in Education*, 17, 31-74.

Attachment**EVALUATION SHEET**

Using a scale of 1-4, judge the extent to which you have seen the stated descriptions (A-D) in the video-lesson viewed.

Key: 1- poor (not observed) 2- below average (faintly observed) 3- average (partially observed) 4- excellent (clearly and completely observed)

Section A	Evidence of the tying of assessment to curriculum	
1.	The lesson objective(s) tally with the one stated in textbook/syllabus	
2.	Lesson was developed based on students' previous knowledge and this connection is maintained throughout the lesson when necessary	
3.	The obvious connection between topic being taught and other areas in the curriculum was established	
4.	The lesson aimed at the development of high order thinking skills-investigation, communication and problem solving	
5.	Current and/or available technology was employed during lesson	
6.	The content was current and presented in a way that is relevant to today's context	
7.	The assessment afforded students the opportunity to deal with meaningful problems that provide worthwhile educational experiences	
8.	Classroom assessments that valued the ability of applying knowledge by reasoning and solving novel problems was developed and used	
9.	The assessment adequately samples the breadth and depth of possible important content	
10.	The assessment-tasks/exercises tally with the objective(s) of the lesson	
Section B	Evidence of integration of assessment and instruction	
1.	There was evidence of on-going and systematic assessment throughout the lesson	
2.	Teacher's questions were probing and aimed at understanding pupils' mathematical thinking	

3.	The teaching method used promoted conceptual understanding (i.e. development of concepts from concrete to abstract and vice versa)	
4.	Pupils were given opportunity to create their own problems after conceptualization	
5.	The assessment tasks or exercises were open ended - call for a variety of ways of being solved	
6.	There was fair and adequate discussion of the various responses provided to open ended tasks and exercises	
7.	There was provision of quick and timely feedback to students	
8.	Students were given opportunity to evaluate, reflect on and improve their own work	
9.	The assessment suits the teaching methodology employed	
Section C	Evidence of formal and informal assessment strategies	
1.	Opportunity was provided for student self-assessment	
2.	Peer assessment was encouraged	
3.	Group work or co-operative learning was encouraged	
4.	There was conscious use of the skill of observation to identify students' strengths and weaknesses	
5.	Students' weaknesses were remediated and strengths reinforced	
6.	Students were engaged mathematical discourse for the purpose of gaining an appreciation of their level of understanding	
7.	Feedback was given on an ongoing basis to students during lesson.	
8.	Exercises or tasks, during or after the lesson, were provided with clear instructions on what is expected of students	
9.	There was provision of needed assistance when students were doing assigned task	
Section D	Evidence of equity in assessment	
1.	Attention was given to all students and all were involved in the lesson.	
2.	Both genders were equally and actively involved in the classroom interactions	
3.	Assistance to individual students based on their needs (Aptitude Treatment Interaction Skills).	

4.	Questions was evenly distributed to students- across students geographical position, differing abilities and gender	
5.	Teacher's presentation and questioning (the framing of) showed gender, social and/or culture sensitivity	
6.	All students were exposed to the same challenging content or activities.	
7.	Lesson was built on prior knowledge and experiences common to all students	
8.	All students had equal opportunity to participate in and demonstrate their abilities	
9.	There was fairness in the scoring and interpretation of pupils' responses	

要 約

今、世界各国で教育改革が進められている。その中心的な課題の1つとして授業改革が求められている。本研究では、NCTM(全米数学教育者協議会)の授業分析項目を参考にして開発した授業分析スケールを用いて、日本において教員研修等で用いられている算数科の授業ビデオ、徳島県公立小学校における算数科授業のビデオ及びガーナにおける典型的な算数科授業のビデオをもとに、算数科授業を分析し、よりよい授業展開について考察した。

(訳責：教員教育国際協力センター)