OPEN ACCESS
EURASIA Journal of Mathematics Science and Technology Education
ISSN: 1305-8223 (online) 1305-8215 (print)
2017 13(6):1981-2004
DOI 10.12973/eurasia.2017.01210a

Portfolio of Evidence: An Assessment Tool in Promoting Geometry Achievement among Teacher Education College Students

Hailu Nigus Weldeana
Mekelle University, ETHIOPIA

Desta Berhe Sbhatu
Mekelle University, ETHIOPIA

Received 19 November 2015 ▪ Revised 8 August 2016 ▪ Accepted 28 October 2016

ABSTRACT

Background- This article reports contributions of an assessment tool called Portfolio of Evidence (PE) in learning college geometry.

Material and methods- Two classes of second-year students from one Ethiopian teacher education college, assigned into Treatment and Comparison classes, were participated. The assessment tools used in the Treatment and Comparison classes were PE and paper-and-pencil, respectively. Data sources were scores of: Self-/Teacher-Assessment Rubric (STAR); Mathematical test measuring Skills, Properties, Uses, and Representations (SPUR); and Learning and Study Strategies Inventory-High School Version (LASSI-HS).

Results- Comparison of students' Self- and Teacher-Assessment data showed that students in the Treatment Class were able to assess their own learning and progress as authentically as the teacher. Analyses of SPUR data revealed that the learning gains among the students in the Treatment Class were significantly greater than that of the Comparison class in tests requiring higher order thinking ($p \le 0.05$). Analyses of LASSI-HS showed that students in the Treatment Class made more statistically significant shifts towards demonstrating supportive learning behaviors and towards abandoning inhibiting behaviors than those in the Comparison Class ($p \le 0.05$).

Conclusions- Thus, effective integration of PE in the instructional process helped students develop reflective thinking and other metacognitive skills and solve real-life problems that demand higher order thinking.

Keywords: self-assessment, portfolio of evidence, authentic assessment, teacher assessment

INTRODUCTION

Paper-and-pencil assessment in mathematics – where students are assessed and penalized for what they don't know instead of acknowledging what they know – was dominant for so long (van de Walle, 1998). Paper-and-pencil assessment strategy influences one's philosophy

Correspondence: Desta Berhe Sbhatu, *Mekelle Institute of Technology, Mekelle University, Mekelle Institute of Technology Campus, 1632 Mekelle, Ethiopia.*
shumageba@gmail.com

State of the literature

- Portfolios are employed for authentic assessment using student self-assessment, peer-assessment, teacher assessment, and group assessment to measure students' performances.
- Portfolios are excellent means of communication among stakeholders since they provide access to vital presentation of students' performances and understanding more clearly than letter grades.

Contribution of this paper to the literature

- Effective integration of Portfolio of Evidence in teaching college level mathematics courses helped students develop reflective thinking and other metacognitive skills and solve real-life problems that demand higher order thinking easily and effectively.

and vision of assessment. It leads mathematics teachers and educators to treat assessment as an isolated component, thus they fail to integrate assessment and instruction. Paper-and-pencil assessment obliges teachers and educators alike to rely on testing for lower-level mathematics skills that gives little room for students' reflection and transfer.

Over two decades have lapsed since the calls for redefining mathematics and mathematics learning and for revisiting the ways students' successes are measured have caught momentum (Jones, 1994; Schmidt & Bronsnan, 1996). Quality student learning is ensured by integrating assessment and instruction, where the former becomes routine classroom activity rather than an interruption (Cathcart et al., 2001; National Council of Teachers of Mathematics [NCTM], 2000; Reichel, 1994; Robinson & Bartlet, 1993). Quality student learning can be achieved by the use of authentic assessment procedures, that are more practical, realistic, and challenging compared to traditional paper-and-pencil tests (Kantrov, 2000; Marsh, 2004). Authentic assessment procedures engage students in meaningful, context-bound activities, putting their energies on "challenging, performance-oriented tasks that call for analysis, integration of knowledge, and invention" (Darling-Hammond et al., 1995, p. 2). Research on authentic assessment methods shows that they require and promote realistic, performance-based, and cognitively complex tasks; establish clearly defined roles of learners; and begin with clearly set criteria of measuring performance (Frey et al., 2012). They boost students' motivation and self-esteem towards their own learning and progress (Birgin, 2011; Dandis, 2013).

From the perspective of school teachers, assessment in mathematics instruction is used for several purposes, including: a. monitoring student progress; b. evaluating programs; c. making instructional decisions; and d. evaluating student achievement (NCTM, 1995). As has established by Frey et al. (2012), authentic assessment methods help in realizing teachers' goals by enhancing students' intrinsic motivation, sustained and creative engagement, and ownership and commitment in and for their own learning and progress. One of the assessment strategies in mathematics that involves authentic assessment is student portfolio of evidence. This strategy engages students in reflective practices such as self, peer, group, and teacher assessment. Student portfolio of evidence focuses on providing feedback with constructive comments and future directions. Involving students in such authentic assessment practices enhance self-regulation strategies in mathematics learning (Tanner & Jones, 1994; van den Berg, 2004).

Geometry is one of the important applications of mathematics in every aspect of life. The International Commission on Mathematical Instruction (ICMI, 1995) describes the role and contributions of geometry as "a tool for understanding, describing and interacting with the space that every one of us lives − the most intuitive, concrete and reality-linked part of the mathematics" (p. 28). Unfortunately, an extensive body of research reported that students experience difficulties in learning this very important field of mathematics at all levels (e.g. Clements et al., 1998; Craime & Rubenstein, 1993; Movshoritz-Hadar et al., 1986; Gutierrez & Jaime, 1998; Senk, 1989; Stallings-Roberts, 1994). The difficulties may, in part, be due to the lack of appropriate assessment strategies. This study investigates the contributions of Portfolio of Evidence (PE) as assessment tool in learning geometry among college students.

This article reports the contributions of PE in improving pre-service teacher education college students' geometry achievement based on empirical research. In particular, it reports the roles of PE in: improving students' understanding in geometry; promoting students' reflective thinking; and, enhancing students' self-regulating skills. The guiding question for the research was: "What is the effect of using PE as authentic assessment tool on pre-service teachers' mathematics achievement and performance?" The specific questions were: a. How much does using PE promote pre-service teachers' reflective thinking? b. How much does using PE improve pre-service teachers' understanding in geometry? c. How much does using PE enhance pre-service teachers' self-regulating skills?

Authentic assessment tools, including PE, have been championed in revolutionalizing assessment of student learning and progress as well as instructional practices. The works of Birgin and Baki (2007), Dandis (2013), Davies and Le Mahieu (2003), De Fina (1992), Frey et al. (2012), Jonsson and Svingby (2007), Kuhs, (1994), NCTM (1995, 2000), Melograno (2000), Slater (1996), and Taki and Heidari (2011) are good testimonials in that regard. Unfortunately, limited studies have been conducted on the effects of PE in mathematics teaching and learning in general and on pre-service teachers' performance and achievement in geometry in particular. The rationale of this study is, therefore, to provide some insight into the contributions of an authentic assessment referred to as PE on pre-service teachers' achievement and performance in geometry, as well as in developing the students' self-regulating skills.

REVIEW OF LITERATURE

Current assessment practices follow George Polya's problem solving stages, which include: understanding concepts and procedures required to accomplish tasks; planning to select appropriate procedures and tools to accomplish tasks; carrying out the tasks in class;

and looking back to interpret assessment data (van de Walle, 1998). Assessment informs each stage of the teaching and problem solving cycle. Mathematics programs need to assess students' knowledge and dispositions (attitudes, interests, beliefs), equity (assessing students' opportunities in learning), and effectiveness to meet the goals (Webb & Welsch, 1993). In assessing students' dispositions towards mathematics, teachers need to systematically observe students': a. confidence in applying mathematics; b. flexibility in exploring mathematical ideas; c. perseverance in mathematical tasks; d. interest, curiosity, and inventiveness in doing mathematics; e. inclination to monitor and reflect on their own thinking and performance; f. valuing of the application of mathematics, and g. appreciation of the role of mathematics in their culture (NCTM, 1989). These require having authentic assessment tools.

Portfolios are used as authentic assessment tools. They contain a meaningful collection of artifacts that provide evidences showing students' progresses in learning mathematics. They contain samples of written descriptions of results of mathematics investigations, extended analyses of problems, diagrams of problem solving, reports and photographs of projects, etc (Webb & Welsch, 1993). Therefore, portfolios help maintain records of students' progresses demonstrated through different mathematical competencies – mathematical activity behaviors and social interactive behaviors (Black, 1999; Gardner, 1999). They are excellent means of communication among stakeholders (van de Walle, 1998) as they provide access to vital presentation of students' performances and understanding more clearly than letter grades (Lee & Silverman, 2001; NCTM, 2000). Moreover, they promote student self-assessment and enable students to communicate their understanding with high level of proficiency (Lambdin & Walker, 1994). Portfolios are employed for authentic assessment purposes through: student self-assessment (Cathcast et al., 2001; Stenmark, 1989; van den Berg, 2004), peer-assessment (Race et al., 2005; Tanner & Jones, 1994), teacher assessment (Kyriacou, 1998), and group assessment (Reicher, 1994; van de Walle, 1998) to measure students' performances as well as learning and progress (Lambdin & Walker, 1994; Lee & Silverman, 2001; NCTM, 2000; van de Walle, 1998). As portfolio assessment is broad and open ended, rubrics or scoring guides have to be set (Kuhs, 1994; Reichel, 1994).

Self-Assessment

Students' self-assessment is authentic or part of an authentic assessment tool that requires learners to evaluate themselves and identify the areas of their strengths to build and the areas of their weaknesses for improvement. Generally speaking, it requires students assess their own learning activities (Cathcart et al., 2001). Self-assessment is carried out using a priori set standards or criteria. It can be used in evaluating essays, reports, presentations, performances, projects, interviews, reflective journals, rubrics, writing, and exams as well as in assessing a comprehensive portfolio of evidence. It is vital formative assessment tool that results in admissible summative assessment (Ethiopian Ministry of Education [MoE], 2006).

Many authors attest that self-assessment: increases students' responsibility and autonomy to learn and assess their progress, promotes students' competency, uplifts students' roles and status from passive to active learners, improves students' critical reflection skills, enhances students' critical and higher order thinking capacities, and promotes students' metacognitive skills thus identify what to learn and study more (e.g., Jonsson & Svingby, 2007; MoE, 2006; Purwanti, 2015; Slauijimans et al., 1998; Stenmark, 1989; Struyven et al., 2005; Topping, 1998; van den Berg, 2004). It is also helpful in integrating assessment with instruction (Slauijimans et al., 1998).

Students' self-assessment procedures are designed to help students assess mathematical contents as well as their dispositions and behaviors toward mathematics learning (Burns, 1995; Gardner, 1999; van de Walle, 1998; Williams, 2000). However, whereas some workers have reported mismatches between students' self-assessment scores and that of their tutors (e.g., Orsmond & Merry, 1997), others have observed strong match (e.g., Mires et al., 2001). Since such mismatches are believed to happen due to inadequacies in preparing students for self-assessment, due care needs to be taken in implementing the tool.

Peer Assessment

Student peer assessment is a strategy where students make assessment decisions on other students' works based on mutually agreed criteria. Like self-assessment, the quality of peer assessment depends on the adequacy of the preparations among the student-assessors to use peer assessment for improving learning. Peers have to be prepared and be aware of what, how, and why to assess as well as how to use the procedures and why. The presence of a well -developed and well-understood peer-assessment rubric is very critical (Jonsson & Svingby, 2007; Lockhart & Ng, 1995; Race et al., 2005; Topping, 1998; van den Berg et al., 2006).

Peer assessment offers multiple benefits to students and teachers, like: reducing students' dependency, reducing teachers' workload in grading, encouraging students' participation in decision-making, clarifying assignment outcomes and expectations, and improving students' reflective thinking capacity. Also, it enhances students' satisfaction by their contributions, promotes students' collaborative learning, enhances students' self-regulating skills, increases the degree and quality of students' communication and interaction, and increases the volume, depth and quality of students' learning (e.g. Brown & Knight, 1994; Chang & Chou, 2011; MoE, 2006; Somervell, 1993; Sluijsmans et al., 1998; Segers & Dochy, 2001; Tanner & Jones, 1994; van den Berg et al., 2006).

Teacher Assessment

Authentic teacher assessment is an ongoing feedback system to monitor and record students' learning and outcomes. It helps teachers provide constructive comments on portfolios of evidence (Huetinck & Munshin, 2000). Teachers' feedback motivates students to learn and to make tangible progress (Kyriacou, 1998). Proficient teachers of mathematics implement multifaceted quantitative and qualitative assessment techniques, like: open-ended questions, constructed response tasks, selected-response items, performance tasks, observations and conversations, reflective journals, interviews, and portfolios as appropriate (Kyriacou, 1998; NCTM, 1989, 1995, 2000; Race et al., 2005). The selection of assessment tools depends on various factors including teaching goals, teachers' and students' skills and experiences, and the overall instructional environment.

In any case, however, authentic teacher assessment has to employ checklists and formats to follow-up students' degree and quality of participation (Huetinck & Munshin, 2000). Such checklists and formats would help teachers in differentiating students' responses into varying complexity and in determining their levels of proficiency. The checklists enable teachers to provide students with proper feedback following any assessment to help them in setting their own goals for learning to become more independent learners. Also, teachers use assessment data to: ensure that students are growing in the right direction, decide and implement assessment for learning (formative) and off learning (summative), ensure quality learning is taking place among students, make appropriate instructional decisions and choices, support students' progress toward learning goals, and draw valid inferences through the convergence of valid results of assessment (NCTM, 2000).

Group Assessment

Group assessment through observations with the help of some instruments can also provide authentic student progress data (Reichel, 1994). Group work, where all students are involved, makes the collection of authentic assessment data much easier and more complete. Group work provides teachers and students with much more data and information as compared to other instructional settings, while reducing the amount of resources used for assessment. It encourages students, informs instructions, helps evaluate and grade students' works, and enables teachers to evaluate programs (van de Walle, 1998).

Group assessment can be perceived as a teacher assessing the collaborative effort of a group of students or members of a group of students assessing the contribution of each other using a mutually set and agreed criteria. But, the latter is frequently referred to as peer group assessment (MoE, 2006). Teacher assessment of groups instead of individuals has multiple benefits to teachers, including: easing their jobs to assess more groups, enabling them to provide evidence-based and realistic feedbacks on performance, and helping them address students' understandings or misunderstandings based on the summarized works of the group. Teacher assessment of groups can only be effective if and when teachers have established enabling instructional context. Specifically, teachers need to: design lesson activities and train students on how to implement group assessment, design mechanisms of assessing the processes and products of learning, and supervise and monitor the groups in their learning endeavors while acting as group members (Marin-Garcia & Lloret, 2008; Willcoxson, 2006). Admissible group products of learning, e.g. reports, proposals, and oral presentations, are the outcomes of reliable group process of committed members (Marin-Garcia & Lloret, 2008).

Peer group assessment in mathematics learning would involve assessing the selection, reflection, rationalization, and evaluation of a group's activities by another group. It requires groups to assess or cross-moderate and to grade other groups' works with constructive and positive feedbacks based on an established tool. Then, both the assessing and the assessed groups come together to discuss and agree on the results of the assessment before feedbacks and comments are presented in whole-class deliberations (MoE, 2006).

RESEARCH METHODS

Participants

Forty freshman pre-service teacher education students, enrolled in one Ethiopian teacher education college, were participated in the study. After the students were listed alphabetically according to their first names, the odd-numbered students (n = 20) were assigned into Experimental Class and the even-numbered (n = 20) were assigned into Comparison Class.

Intervention

As indicated above, the purpose of this study was to investigate the effectiveness of PE as an authentic assessment tool in teaching college geometry, presented in five chapters. Hence, the intervention has involved: a. integrating PE as assessment tool in teaching the Experimental Class, and b. using traditional paper-and-pencil assessment techniques in the Comparison Class. Students in the Experimental Class have completed the following seven activities, as on-going assessment tools of learning and progress, throughout the study. The first six activities were used as formative assessment tools while the last one, chapter-end assessment, is used as summative assessment tool by the students and the teacher. Students were required to undertake all of the activities throughout the intervention.

a. Summary Making: Involves students in summarizing big ideas of learned materials, and solving good representative problems showing all steps. Summary making and problem solving also involve the use of various tools (graphs, diagrams, representations, etc).
b. Reflective Activities: Involve students in expressing their learning and progress explicitly including: likes and dislikes, areas of strength, and areas needing improvement with the greatest of their ability and as detailed as possible (George & Cowan, 2004).
c. Journal Writing and Communication: Allow students to reflect on what they are learning, extend ideas, discuss solutions and strategies, and create meaning to themselves and to their peers. Students were also expected to communicate in mathematical terms and language (Santiago & Spanos, 1993).

d. Contribution to Group Work: Requires peer-assessment of the contributions of individual students to peer-learning using objective evidences. Every group member's contribution to group learning is assessed by her/his peers. It requires students to record in their portfolios the comments and suggestions provided by peers.
e. Portfolio Assignment: This activity requires students to produce problems (five multiple choice questions) that are good representatives of a chapter and solve them showing all steps and using diagrams, graphs, and/or representations. It also requires student to choose problems from assignments, class works, home works, and textbooks enjoyed solving and/or did especially good job on to represent the important concepts and skills learned in a given unit. Students are also required to provide justifications for their selection (Hoey & Watoon, 1994a, b).
f. Exposition: This activity requires students to explore, expand, and extend further the materials presented in the classroom to promote understanding (Posamentier & Stepelman, 1996). In this study, the students were involved in explaining concepts, algorithms, and theorems, describing and interpreting graphs, discussing and writing problem solutions.
g. Punctuality and Attendance: This activity is included to help students develop time management skills and prioritizing activities according their degree of importance. The students were required to put their reflections about these behaviors in their portfolios.
h. Chapter-end Assessment: This activity involves chapter-end assessment of the learning and teaching processes by the students (self-assessment) and the teacher (teacher assessment). It assesses the six activities, namely: summary making, reflective activities, journal writing and communication, portfolio assignment, exposition, and punctuality and attendance using a four-item rubric.

These on-going and continuous authentic assessment processes of students' learning and progress were accompanied with timely feedback and scoring and grading of students' works. Discussions were held with the students whenever they disagree with the teacher. Both classes were taught by the first author. The intervention was run for nine months.

Data Collection and Analyses

Three instruments were employed in the study, namely: a. Self-/Teacher Assessment Rubric (STAR), b. Mathematical test measuring Skills, Properties, Uses, and Representations (SPUR), and c. Learning and Study Strategies Inventory-High School Version (LASSI-HS). STAR was used to assess the implementation of PE in the Experimental Class through end-of-chapter evaluation of students' reflective activity, summary making and journal writing, contributions to groups, portfolio assignments, exposition, and attendance and punctuality. This instrument was developed by consulting College Preparatory Mathematics (Dietiker, 1997a, b; Hoey & Wotton, 1994a, b) and the Higher Diploma Program (MoE, 2006). A fourpoint scale scoring rubric adapted from the MoE (2006) was used with some modifications to assess student learning (Appendix 1). Students in the Experimental Class were allowed to assess their own learning progress using the scoring rubric (students' self-assessment) out of 20 at the end of each chapter. The students were made aware through training and advice to be free of biases in assessing their own learning activities and progress. Likewise, the teacher assessed the students using the same rubric (teacher assessment). Teacher assessment was accompanied with constructive feedbacks and comments on students' classroom learning behaviors. Students and teacher assessment scores were compared using independent sample t-test. Seventeen students (out of 20) have completed the instrument.

SPUR was prepared to measure the four dimensions of understanding, namely skills, properties, uses, and representations. SPUR is an acronym for Skills, Properties, Uses, and Representations. 'Skills' questions require knowledge of procedures to get answers whereas 'Properties' require the understanding of the principles behind the questions. 'Uses' questions are designed to link mathematics to real life situations. 'Representation' questions require students to use pictures, graphs, or objects to illustrate concepts (Coxford et al., 1991). A 50-item test was adapted from the University of Chicago School Project Mathematics: Geometry (Coxford et al., 1991) to compare the performances of the Experimental and Comparison classes. The test was comprised of four parts that respectively deal with Skills, Properties, Uses, and Representations. Test items were carefully selected from each of the fifteen chapters. Students in the Experimental and Comparison classes took the test at the beginning (pre-test) and the end (post-test) of the intervention. Students were advised to exert utmost efforts to score maximum marks and to complete all the problems in the test paper (Appendix 2). Out of the 20 students in each class, 17 in the Experimental and 19 in the Comparison class have completed the tests. Learning gain (post-test score minus pre-test score) was calculated for each student. Mean learning gain of both classes were compared using independent sample t-test.

LASSI-HS is a five-point scale questionnaire designed to measure student self-regulating skills and strategies, such as how they study, how they learn, and how they feel about learning and studying. The instrument is described in Weistein and Palmer (1990). Ten subscales make up the LASSI-HS, namely: attitude, motivation, time management, anxiety, concentration, information processing, selecting main idea, study aid, self-testing, and test strategies. Descriptions of each scale with examples are available in the work of Olaussen and Braten (1999). In this study, only 40 items – four items from each of the ten scales – were selected from the original 76 items (Appendix 3). The students in both classes have completed the questionnaire at the beginning (pre-test) and the end (post-test) of the study. In the Experimental Class, 20 and 17 students have completed the pre-test and the post-test, respectively. In the Comparison Class, 18 students have completed the pre-test and post-test. Mean pre-test and post-test scores of both classes were compared using independent sample t-test. In all cases, comparisons were made at a priori significance level of $p \leq 0.05$

Validity and Reliability of the Tools

Studies have shown that validities of peer rating are difficult to attain (Dancer & Dancer, 1992 quoted in Sluijsmans et al., 1998). Peers are prone to produce poor ratings because they prefer to focus only on positive contributions to group projects. Jonsson and Svingby (2007) have studied the validity of judgments in regard to performance and promotion of student learning and the quality of instruction. They have established the types of validity and reliability researchers should use; and have suggested that the validities of judgments need be generalized. Moreover, they have recommended that content, external, structural, substantive, and consequential validities are used without exception. Nonetheless, myriads of researchers rely only on content validity or construct validity or both.

The present study has ensured content validity by including relevant and representative knowledge and skills in the assessment by subjecting the SPUR and the Rubric to critical proofread and evaluation by experienced experts in education and mathematics. Moreover, the Rubric of the PE is an adaptation from an established national document with slight modifications to fit our purpose. Likewise, the LASSI-HI version is a well-established tool having high consistency of its items. The coefficients of alpha for the ten categories in the original work of Weinstein and Palmer (1990) ranged from 0.68 for 'study aids' to 0.82 for 'anxiety and concentration'. The tool was also pilot-tested to check if its internal consistency matches with the original findings. Consequently, the Cronbach's Alpha for our instrument ranged from 0.73 for 'attitudes' to 0.87 for 'test strategies'. The drops in alpha values ranged between 0.004 and 0.072, quite insignificant guaranteeing the maintenance of the selected items. Interestingly, no raise in alpha values was observed as given task item was removed.

RESULTS

Self-/Teacher Assessment Rubric

Portfolio of Evidence was employed as an authentic formative assessment tool in teaching a five-chapter college geometry course in the Treatment Class. It was used to help students learn with understanding and foster their self-regulation abilities. It is well established that proper implementation of PE provides students and teachers with valuable information about their learning and progress over time (National Research Council [NRC], 2000). Thus, each student has assessed her/his own learning progress at the end of each chapter using STAR. Likewise, the teacher has assessed every student at the end of each chapter using the same rubric. Comparisons of mean students self-assessment and teacher assessment of all chapters lack statistically significant difference ($t = 0.26$ to 1.83, $p \geq 0.05$), implying that students' self-assessment and teacher assessment are congruent (**Table 1**).

Table 1. Comparing self and teacher assessment

Chapters of Assessment	Mean (SD) Assessment Score		t-test	
	Self [n = 17]	Teacher [n = 17]	$t_{(1,32)}$	*p*
Chapter 1 (20 Points)	12.94 (1.89)	12.53 (1.12)	0.77	.445
Chapter 2 (20 Points)	14.65 (1.06)	14.00 (1.00)	1.83	.076
Chapter 3 (20 Points)	15.76 (1.52)	15.06 (1.48)	1.37	.180
Chapter 4 (20 Points)	16.35 (1.66)	16.23 (0.83)	0.26	.795
Chapter 5 (20 Points)	17.35 (1.69)	17.18 (1.23)	0.35	.731

NB: p is 2-tailed.

Mathematical Test Measuring SPUR

SPUR measures four dimensions of understanding in learning geometry education, namely Skills, Properties, Uses, and Representations. The study assumed that using PE as authentic assessment procedure in teaching geometry fosters students' abilities and understanding to tackle geometry problems related to skills, properties, uses, and representations. Hence, it was predicted that the mean learning gain of the Treatment Class will be greater than that of the Comparison Class. Comparisons of results are given in Table 2. The Treatment Class achieved significantly greater mean learning gain in Properties, Uses, and Total SPUR ($t = 2.42$ to 4.23, $p \leq 0.05$).

Table 2. Comparing students' mathematics test measuring SPUR

Learning Gain	Mean (SD) Test Scores		t-test	
	Treatment [n = 17]	Comparison [n = 19]	$t_{(1,34)}$	*p*
Skills	7.76 (2.17)	6.63 (2.17)	1.57	.126
Properties	8.12 (4.06)	3.79 (3.51)	3.43	.002
Uses	6.76 (1.60)	5.53 (1.47)	2.42	.021
Reprensetations	9.23 (2.61)	7.58 (2.83)	1.81	.078
Total	**31.82 (6.77)**	**23.26 (5.36)**	**4.23**	**.000**

NB: Learning gain = Post-test – Pre-test; *p* = 2-tailed.

LASSI-High School Version

The purpose of using LASSI-HS version is to evaluate the contributions of PE in promoting student behaviors that support instructional processes in which students regulate their own learning and reflect on their own progresses and limitations. This instrument was completed by both groups of students at the beginning and the end of the intervention. Comparisons of mean pre-test and post-test scores (of LASSI-HS) of the Treatment Class revealed that the students made: a. 22 significant shifts towards demonstrating supportive behaviors out of the possible 22 shifts, and b. 15 significant shifts towards abandoning limiting behaviors out of the possible 18 shifts ($p \leq 0.05$). On the other hand, comparisons of mean pre-test and post-test scores (of LASSI-HS) of the Comparison Class showed that the students could only make: a. 17 significant shifts towards demonstrating supportive behaviors out of the possible 22 shifts, and b. seven significant shifts towards abandoning limiting behaviors out of the possible 18 shifts ($p \leq 0.05$) (**Table 3**).

DISCUSSION

This section provides discussions of the findings of the study on the effects of using PE on pre-service teachers' mathematics achievement and performance. The study has attempted to look into the contributions of PE on the pre-service teacher trainees': reflective thinking capacity, understanding of geometry, and self-regulation skills.

Promoting Students Reflective Thinking

Portfolios employed for assessment and evaluation purposes engage learners in higher order thinking through reflection and inquiry (Johnson et al., 2006). The present study assumed that effective implementation of PE as assessment component of instruction fosters students' abilities to assess and evaluate their own learning activities including: reflective activity, summary making and journal writing, contributions to group activities, portfolio assignments, exposition, and attendance and punctuality as measured using STAR. STAR was employed as an authentic teacher and student summative assessment tool in teaching college geometry to the Treatment Class. The lack of any statistically significant difference between the mean students' self-assessment and teacher assessment in each chapter ($t = 0.26$ to 1.83, $p \geq 0.05$) (**Table 1**) implies that the students were able to self-evaluate their own learning and progress as accurately as their teacher. Moreover, despite open-ended questions, like STAR, are less liable to agreement among different assessors (e.g. Brown & Knight, 1994; Mahalski, 1992; Orsmond & Merry, 1997), mean students and teacher assessment scores were consistently increasing from the first to the last chapter. Lack of agreement between students' self-assessment and teacher assessment score, as exemplified above, is a usual phenomenon. As this problem did not occur in this study, we argue, the PE was employed correctly to the extent that helped the students assess their own learning and progress as objectively as possible. This is because, we further argue, the PE has helped the students develop reflective thinking skills. Several studies have shown that learners who have been engaged in building portfolio of evidence favorably improve their self-reflection, critical thinking, learning responsibility, and multiple intelligences (e.g., Cole & Struyk, 1997; Davies & Le Mahieu, 2003; Gardner, 1999). For example, Davies and Le Mahieu (2003, p.13) attested that metacognitive skills are residues left over during portfolio development – where students are engaged in reflecting on their own learning, in selecting, organizing and sequencing work samples in their portfolios, and in systematizing self-assessment activities that show successes and the significance of each piece of work towards the successes.

Table 3. Mean (SD) student teachers' responses to LASSI-HS statements

LASSI-HS Statements		Mean (SD) Responses							
		Treatment [n = 20/17]				Comparison [n = 18/18]			
		Pre-test	Post-test	$t_{(1,35)}$	*p*	Pre-test	Post-test	$t_{(1,34)}$	*p*
Anxiety	1–	3.8 (1.2)	1.9 (0.8)	5.55	0.00	3.8 (1.2)	2.6 (1.1)	3.17	0.00
	2–	3.7 (1.3)	2.3 (1.1)	3.58	0.00	3.8 (1.7)	3.7 (1.4)	0.13	0.89
	3–	3.8 (1.1)	2.5 (1.4)	3.26	0.00	4.1 (0.9)	3.2 (1.5)	2.17	0.04
	4–	4.0 (1.2)	2.1 (1.1)	4.91	0.00	3.7 (1.5)	3.3 (1.6)	0.87	0.39
Attitude	1–	4.1 (1.3)	1.5 (1.2)	6.42	0.00	4.1 (1.1)	1.7 (0.9)	7.15	0.00
	2–	3.5 (1.2)	1.5 (0.9)	5.64	0.00	3.6 (1.3)	2.2 (1.4)	3.05	0.00
	3–	3.8 (1.3)	2.1 (1.1)	4.16	0.00	3.7 (1.4)	1.6 (1.0)	5.20	0.00
	4–	3.6 (1.1)	1.4 (0.9)	6.68	0.00	3.7 (1.3)	1.3 (0.5)	7.43	0.00
Concentration	1–	2.3 (1.4)	2.1 (1.1)	0.46	0.65	2.3 (1.2)	2.6 (1.2)	-0.83	0.42
	2–	3.8 (1.2)	2.3 (1.1)	4.04	0.00	3.4 (1.4)	3.0 (1.1)	1.05	0.30
	3–	3.7 (1.5)	1.7 (1.0)	4.83	0.00	3.6 (1.4)	3.1 (1.3)	1.11	0.28
	4–	3.8 (1.2)	1.9 (1.0)	5.18	0.00	3.7 (1.4)	3.1 (1.2)	1.42	0.17
Information Processing	1+	2.5 (1.1)	3.9 (1.1)	-4.03	0.00	2.3 (1.3)	3.3 (1.4)	-2.26	0.03
	2+	1.8 (1.1)	4.1 (0.5)	-8.00	0.00	1.9 (1.1)	3.6 (1.0)	-4.66	0.00
	3+	2.3 (1.1)	4.1 (0.9)	-5.26	0.00	2.2 (1.1)	2.9 (1.3)	-1.78	0.08
	4+	1.9 (1.1)	4.2 (0.7)	-7.86	0.00	1.9 (1.1)	3.3 (0.8)	-4.23	0.00
Motivation	1+	2.2 (0.8)	3.5 (0.7)	-4.92	0.00	1.9 (1.2)	3.6 (1.0)	-4.66	0.00
	2+	2.5 (1.2)	4.3 (0.9)	-5.32	0.00	2.2 (1.2)	2.9 (1.5)	-1.77	0.09
	3+	1.9 (1.1)	3.2 (1.1)	-3.52	0.00	1.9 (1.1)	3.2 (1.1)	-3.53	0.00
	4–	2.5 (1.2)	2.1 (1.1)	1.03	0.31	2.2 (1.2)	2.8 (1.3)	-1.63	0.11
Selecting Main Idea	1+	1.8 (1.1)	4.2 (0.4)	-8.50	0.00	1.9 (1.1)	3.7 (0.9)	-5.29	0.00
	2+	2.1 (1.2)	4.2 (0.7)	-6.19	0.00	2.1 (1.2)	3.9 (0.8)	-5.29	0.00
	3–	1.8 (1.1)	2.3 (1.1)	-1.52	0.14	1.9 (1.2)	2.2 (1.2)	-0.83	0.41
	4–	3.8 (1.1)	2.1 (0.9)	4.78	0.00	3.9 (1.2)	3.4 (1.4)	1.17	0.25
Self-Testing	1+	2.3 (1.3)	3.9 (0.8)	-4.70	0.00	2.2 (1.2)	4.1 (0.8)	-5.50	0.00
	2+	2.2 (1.1)	3.7 (1.1)	-3.95	0.00	2.2 (1.2)	2.8 (1.3)	-1.49	0.15
	3+	2.2 (1.1)	3.9 (1.2)	-4.50	0.00	2.0 (1.0)	3.9 (1.1)	-5.38	0.00
	4+	2.4 (1.1)	3.8 (0.9)	-4.18	0.00	2.3 (1.3)	2.7 (0.8)	-0.92	0.36
Study Aid	1+	2.4 (1.1)	4.2(0.8)	-5.66	0.00	2.4 (1.3)	3.2 (1.0)	-2.16	0.04
	2+	2.3 (1.2)	3.6 (0.8)	-4.08	0.00	2.1 (1.1)	2.5 (0.7)	-1.24	0.23
	3+	2.1 (1.0)	4.1 (0.7)	-7.08	0.00	2.0 (1.0)	3.6 (1.0)	-4.81	0.00
	4+	2.3 (1.3)	3.9 (0.7)	-4.61	0.00	2.2 (1.2)	3.4 (1.2)	-2.97	0.00
Test Strategies	1+	1.8 (1.3)	4.2 (0.9)	-6.70	0.00	1.9 (1.1)	3.4 (1.4)	-3.40	0.00
	2+	2.4 (1.3)	3.7 (1.1)	-3.21	0.00	2.0 (1.0)	2.9 (1.4)	-2.25	0.03
	3+	2.3 (1.3)	3.7 (0.9)	-3.64	0.00	2.1 (1.2)	3.2 (1.1)	-2.97	0.00
	4+	1.9 (1.1)	4.1 (1.1)	-6.09	0.00	2.0 (1.1)	3.3 (1.2)	-3.44	0.00
Time Management	1–	4.1 (0.8)	2.4 (1.1)	5.36	0.00	3.6 (1.3)	3.2 (1.2)	0.94	0.36
	2–	3.9 (1.2)	3.0 (1.4)	2.29	0.03	3.9 (1.3)	3.7 (1.4)	0.61	0.54
	3–	3.4 (1.4)	1.8 (1.1)	3.63	0.00	4.1 (1.2)	2.5 (1.4)	3.59	0.00
	4+	2.3 (1.3)	3.9 (1.0)	-4.10	0.00	2.2 (1.2)	3.4 (1.2)	-3.26	0.00

NB: "+" Positive behavior statement, "–" Negative behavior statement; p = 2-tailed

Development of a portfolio requires one to: be aware of the audience of the portfolio, be aware of personal learning needs, understand the criteria of quality portfolio and the way quality is promoted, and be skillful in compiling and completing one. The task helps students: develop their self-reflecting capacity, develop sense of belongingness in and ownership of the instruction, monitor their learning progress and selection of learning strategies, and build deep-learning and critical thinking skills (Julius, 2000; Sambell et al., 1997; Segers & Dochy, 2001; Slater, 1996; Struyven et al., 2005).

Improving Students Understanding in Geometry

Learning environments need to be designed in such a way that instructions become learner-centered, knowledge-centered, assessment-centered, as well as community-centered (NRC, 2000) to ensure that learning results in connection and understanding. Learner-centeredness is ensured by considering the knowledge, skills, attitudes, and beliefs students bring to the class. Likewise, knowledge-centered instruction strives "to help students become knowledgeable by learning in ways that lead to understanding and subsequent transfer" (NRC, 2000, p. 136). Moreover, as instruction is designed in such a way that assessment becomes an integral part, the later should be designed to "provide opportunities for feedback and revision" to support "one's learning goals" (NRC, 2000; p. 139–140). Community-centeredness refers to how the social norms of the classroom as well as the norms a larger social entity (community, state, nation, and the world) affect and connect to student learning in one way or another.

One of the purposes of the present study was to explore the contributions of integrating instruction and portfolio-based assessment towards students' understanding in teaching college geometry course. It looked into the contributions of building PE as integral part of instruction in enhancing students' understanding of college geometry. The PE was integrated in the instructional process to engage the students in the process; and their activities and results in reflection, summary making and journal writing, contributions to groups, portfolio assignments, exposition, and attendance and punctuality were assessed. Comparisons of student self-assessment and teacher assessment of the Treatment Class revealed the lack of any significant differences. Moreover, the increments in mean students' self-assessment scores between subsequent chapters are fairly comparable to mean teacher assessment scores. These findings imply that the students were able to assess their own learning and progress as authentically as the teacher (**Table 1**).

As students are able to assess their own learning and progress authentically, supported with teacher's feedbacks, it is apparent that they are engaged in higher order thinking and reflection (e.g. Julius, 2000). The purpose of employing SPUR in this study was to measure four dimensions of understanding in learning geometry, namely skills, properties, uses, and representations (Coxford et al., 1991). Solving geometry problems in properties, uses, and representations require understanding for application and transfer. Data analyses showed that the learning gains among the students in the Treatment Class are significantly greater than that of the Comparison Class as revealed in the tests that demand high order thinking, i.e. properties and uses, thus, Total mean SPUR score ($p \leq 0.05$) (**Table 2**). Skill questions, which merely involve the use of procedures (Coxford et al., 1991) yielded no significant difference between the classes as is usually the case (e.g. Cobb et al., 1991). Hence, the present study shows that integrating PE as an authentic assessment tool in teaching college-level geometry promotes learning with understanding for application, connection, and transfer. This finding is consistent with previous findings (e.g. Coxford et al., 1991; van den Berg, 2004). Studies on relations between assessment practices and students' learning in mathematics consistently indicated that whereas traditional paper-and-pencil assessments offer limited contributions, authentic assessment tools, including portfolio-based assessment, help students build deeper understanding with connections, commit towards their own learning, and retain and articulate what they have learned better (Frey et al., 212; NCTM, 2000; Sambell et al., 1997; Struyven et al., 2005).

Enhancing Students Self-Regulating Skills

Students' self-assessment activities are integral parts of any instructional processes supported with metacognitive strategies (e.g. NRC, 2000; Sbhatu, 2006). One of the purposes of integrating PE in the instructional process was to enhance students' use of self-regulating strategies in learning. Students' self-regulating skills were studied using LASSI-HS version. The instrument was designed to measure student self-regulating strategies and skills, such as how they study, how they learn, and how they feel about learning and studying (Weldeana, 2008). The study predicted that students assessed using PE will exhibit supportive student learning behaviors, and employ self-regulating strategies. Thus, students in the Treatment Class were expected to make significant shifts towards demonstrating more of the supporting behaviors and activities while abandoning the hindering behaviors and activities. Students in the Treatment Class have made more significant shifts towards showing the supportive learning behaviors; and towards abandoning inhibitive learning behaviors compared the Comparison Class (**Table 3**).

The chain of events revealed in this study is that the use of authentic assessment strategy using PE promoted students' self-regulation skills. Students with well-developed self-regulation skills have managed their own learning and learning progress so as to learn with understanding; and subsequently solve application and transfer problems much better than students exposed to traditional instruction. Previous studies have indicated that students encouraged be engaged in self-regulating and metacognitive reflection activities learn with understanding and are capable of tackling application and transfer questions (e.g. Allen & Armour-Thomas, 1991; Georghiades, 2000; Sbhatu, 2006). Assessment strategies similar to those integrated in this study are known to foster reflection capacity and self-regulation skills (e.g. Brown & Knight, 1994; Pollard, 2002; Posamentier & Stepelman; 1996; Topping, 1998; van de Walle, 1998; White & Mitchell, 1994). Several studies have long shown that authentic assessment tools boost students' (the assessors and the assessed) cognitive and metacognitive skills by motivating the students to stay on task longer and by engaging them in thinking, comparing and contrasting, reviewing, clarifying, correcting mistakes, and communicating with their peers (Purwanti, 2015; Sluijsmans et al., 1998; Struyven et al., 2005;

Topping, 1998; van den Berg et al., 2006; VanLehn et al., 1995). For instance, college students who were involved in authentic assessment tasks, including self-assessment, earned better scores on tests of academic courses than their peer who were not exposed to similar tasks (Hassmen et al., 1997; Sluijsmans et al., 1998).

CONCLUSIONS

When assessment is integrated into instructional processes, with the intension of assessing students' learning and progress, we call it formative assessment. In the present study, PE was employed for exactly the same purpose. It has opened important opportunities for the students and the classroom teacher. It has provided students with opportunities to assess and regulate their own learning and progress. Students' self-assessment activities, in turn, have enhanced the development of an important metacognitive skill called reflection. Students capable of reflecting on what they know/don't know as well as on their own learning progress learn with understanding; thus, can connect what they have learned with broader body of knowledge. They become capable of linking school problems with real-life problems. PE also has enabled the teacher to deliver timely and appropriate feedback on students' learning activities. Timely feedbacks and reflections are helpful as instruction allows students to use them to revise their thinking as they are working on a problem (e.g. NRC, 2000; Sbhatu, 2006). Furthermore, PE and other formative assessment tools increase students' learning and transfer.

Nowadays, the integration of assessment practices with instructional processes is highly sought in the Ethiopian higher education institutes. Ethiopian higher education students (aged 19 to 24) need to learn how to learn and how to assess their learning progress. Thus, authentic assessment, like portfolio of evidence, would become instrumental in helping the students grow towards independent learners with limited teacher support. However, it is unfortunate that the assessment practices pursued in the institutions are neither authentic nor aim at promoting understanding with connection and for transfer. They are procedurally continuous and characteristically summative. Portfolio of evidence can be implemented successfully to bring about better youth development and learning. Thus, college instructors and educators are compelled to design assessment strategies like PE to promote student learning with understanding and subsequent transfer. The use of PE with higher education students can serve as an action research to generate additional data and some more insights for enhancing and expanding the tool.

REFERENCES

Allen B. A. & Armour-Thomas, E. (1991). Constract validation of metacognition. *The Journal of Psychology*, *127*(2), 203−211.

Birgin, O. & Baki, A. (2007). The use of portfolio to assess student's performance. *Journal of Turkish Science Education*, *4*(2): 75−90.

Birgin, O. (2011). Pre-service mathematics teachers' views on the use of portfolios in their education as an alternative assessment method. *Educational Research and Reviews*, *6*(11), 710−721.

Black, P. (1999). Assessment, learning theories, and learning systems. In P. Murphy (Ed.), *Learners, Learning, and Assessment* (pp. 118–134). London, UK: Paul Chapman.

Brown, S. & Knight, P. (1994). *Assessing Learners in Higher Education*. London, UK: Routledge Falmer.

Burns, N. (1995). *Writing in Math Class*. White Plains, NY: Cuisenaire.

Cathcart, W. G., Pothier, Y. M, Vance, J. H., & Bezuk, N. S. (2001). *Learning Mathematics in Elementary and Middle Schools*. 2nd Ed. Upper Saddle River, New Jersey: Merril and Prentice-Hall.

Chang, C. C. & Chou, P. P. (2011). Effects of reflection category and reflection quality on learning outcomes during web-based portfolio assessment process: a Aase study of high school students in computer application course. *The Turkish Online Journal of Educational Technology*, *10*(3), 101–114.

Clements, D. H., Sarama, J., & Batista, M. J. (1998). Develeopment of concept of geometric figures in a specially designed Logo computer environment. *Focus on Learning Problems in Mathematics*, *20*(2&3), 47−64.

Cobb, P., Wood, T., Yackel, E., Nicholls, J., Wheatly, G., Trigatti, B., & Perlwitz, M. (1991). Assessment of problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education*, *22*(1), 3–29.

Cole, K. B. & Struyk, L. R. (1997). Portfolio assessment: Challenges in secondary education. *High School Journal*, *80*(4), 261–273.

Coxford, A., Usiskin, Z., & Hirshchhorn, D. (1991). *The University of Chicago School Mathematics Project: Geometry*. Sunnyvale, CA: Scott, Foresman and Co.

Craine, T. V. & Rubenstein, R. (1993). A quadrilateral hierarchy to facilitate learning in geometry. *The Mathematics Teacher*, *86*(1), 30−36.

Dandis, M. A. (2013). The assessment methods that are used in a secondary mathematics class. *Journal for Educators, Teachers and Trainers*, *4*(2), 133–143.

Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic Assessment in Action*. New York, NY: Teachers College.

Davies, A. & Le Mahieu, P. (2003). Assessment for learning: reconsidering portfolios and research evidence. In M. Segers, F. Dochy, & E. Cascallar (Eds.), *Innovation and Change in Professional Education: Optimizing New Modes of Assessment. In Search of Qualities and Standards* (p. 141–169). Dordrecht: Kluwer Academic Publishers.

De Fina, A. A. (1992). *Portfolio Assessment: Getting Started*. Scholastic Inc.: New York, NY.

Dietiker, L. (Ed.) (1997a). *College Preparatory Mathematics 1.* Vol. 1, 2nd Ed. CPM Educational Program.

Dietiker, L. (Ed.) (1997b). *College Preparatory Mathematics 1.* Vol. 2, 2nd Ed. CPM Educational Program.

MoE (Ethiopian Ministry of Education) (2006). *Higher Diploma Programme for Teacher Educators*. Unpublished Handbook, Addis Ababa, Ethiopia.

Frey, B. B., Schmitt, V. L., & Justin, P. A. (2012). Defining authentic classroom assessment. *Practical Assessment, Research and Evaluation*, *17*(2), 1–18.

Gardner, H. (1999). Assessment in context. In P. Murphy (Ed.), *Learners, Learning, and Assessment* (pp. 90–117). London: Paul Chapman.

George, J. & Cowan, J. (2004). *A Handbook of Techniques for Formative Evaluation.* London, UK: Routledge Falmer.

Georghiades, P. (2000). Beyond conceptual change learning: Focusing on transfer, durability and metacognition. *Educational Research*, *42*(2), 119–139.

Gutierrez, A. & Jaime, A. (1998). On the assessment of the van Hiele levels of reasoning. *Focus on Learning Problems in Mathematics*, *20*(2&3), 27−46.

Hassmen, P., Sams, M. R., & Hunt, D. P. (1996). Self-assessment responding and testing methods: Effects on performers and observers. *Perceptual and Motor Skills*, *83*, 1091–1104.

Hoey, B. & Wotton, K. (1994a). *College Preparatory Mathematics 2: Geometry* (Unit 1–6). Lodi, California: CPM Educational Program.

Hoey, B. & Wotton, K. (1994b). *College Preparatory Mathematics 2: Geometry* (Unit 7–12). Lodi, California: CPM Educational Program.

Huetinck, L. & Munshin, S. N. (2000). *Teaching Mathematics in the 21st Century: Methods and Activities for Grades 6–12.* Upper Saddle River, NJ: Merrill and Prentice Hall.

ICMI (International Commission on Mathematical Instruction) (1995). *Perspectives on the Teaching of Geometry for the 21st Century.* Kluwer.

Johnson, R. S., Mims-Cox, J. S., & Doyle-Nichols, A. (2006). *Developing Portfolio in Education: a Guide to Reflection, Inquiry, and Assessment.* Thousand Oaks, CA: SAGE.

Jones, M. G. (1994). *Assessment Potpourri: Science and Children.* University of North Carolina at Chapel-Hill, Chapel-Hill, North Carolina.

Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*, 130–144.

Julius, T. M. (2000). *Third Grade Students' Perceptions of Portfolios*. University of Massachusetts Amherst. Unpublished Doctoral Dissertation.

Kantrov, I. (2000). *Issues in Mathematics Education: Assessing Students' Mathematics Learning*. Educational Development Center, Inc.

Kuhs, T. M. (1994). Portfolio assessment: Making it work the first time. *Mathematics Teacher*, *87*(50), 332–335.

Kyriacou, C. (1998). *Essential Teaching Skills*. 2nd Ed. Cheltenham, UK: Nelson Thrones.

Lambdin, D. V. & Walker, V. L. (1994). Planning for classroom portfolio assessment. *Arithmetic Teacher*, *41*(6), 318–324.

Lee, F. & Silverman, F. L. (2001). Using portfolio assessment to assess children of diverse language and cultural background. *Focus on Learning Problems in Mathematics*, 32(2&3), 102–117.

Lockhart, C. & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions and content. *Language Learning*, *45*, 605–655.

Mahalski, P. A. (1992). Essay-writing: Do study manuals give relevant advice? *Higher Education 24*(1), 113–132.

Marin-Garcia. J. A., & Lloret., J. (2008). Improving teamwork with university engineering students. The effect of an assessment method to prevent shirking. *WSEAS Transactions on Advances in Engineering Education*, *1*(5), 1–11.

Marsh, C. J. (2004). *Key Concepts for Understanding Curriculum*. 3rd Ed. London, UK: Routledge Falmer.

Melograno, V. J. (2000). Designing a portfolio system for K-12 physical education: a Step-by-step process. *Measurement in Physical Education and Exercise Science*, *4*(2), 97–116.

Mires, G. J., Friedman Ben-David, M., Preece, P. E., & Smith, B. (2001) Educational benefits of student self-marking of short-answer questions. *Medical Teacher*, *23*(5), 462–466.

Movshoritz-Hadar, N., Inbar, S., & Zaslasky, O. (1986). Students' distortions of theorems. *Focus on Learning Problems in Mathematics*, *8*(1), 49−57.

NCTM (National Councils of Teachers of Mathematics). (1989). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: National Councils of Teachers of Mathematics.

NCTM. (1995). *Assessment Standards for School Mathematics*. Reston, VA: National Councils of Teachers of Mathematics.

NCTM. (2000). *Principles and Standards for School Mathematics*. Reston, VA: National Councils of Teachers of Mathematics.

NRC (National Reserach Council). (2000). *How People Learn: Brain, Mind, Experience, and School.* Expanded Edition. Washington, DC: National Academy Press.

Olaussen, B. S. & Braten, I. (1999). Students' use of strategies for self-regulated learning: Cross-cultural perspectives. *Scand. Journal of Educational Research*, *43*(4), 409–432.

Orsmond, P., & Merry, S. (1997). A study in self-assessment: Tutor and students' perceptions of performance criteria. *Assessment and Evaluation in Higher Education*, *22*(4), 357–369.

Pollard, A. (2002). *Reading for Reflective Thinking*. London, UK: Continuum International Publishing Group Ltd.

Posamentier, A. S. & Stepelman, J. (1996). *Teaching Secondary School Mathematics: Techniques and Enrichment Units.* Englewood Cliffs, New Jersey: Prentice Hall.

Purwanti, T. T. (2015). The implementation of self-assessment in writing class: a Case study at Stba Lia, Jakarta. *TEFLIN Journal*, *26*(1), 97–115.

Race, P., Brown, S., & Smith, B. (2005). *{500 Tips} on Assessment*. 2nd Ed. London, UK: Routledge Falmer.

Reichel, A. G. (1994). Performance assessment: Five practical approaches. *Science and Children*, *32*(2), 21–25.

Robinson, G. E. & Bartlet, K. T. (1993). Assessment and evaluation of learning. In D.T. Owens (Ed.), *Research Ideas for The Classroom: Middle Grades of Mathematics* (pp. 329–348). New York, NY: McMillan.

Sambell, K., McDowell, L. & Brown, S. (1997). 'But is it fair?' An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, *23*(4), 349–371.

Santiago, F. & Spanos, G. (1993). Meeting the MCTM communication standards for all students. In G. Cuevas and M. Driscoll (Eds.), *Reaching All Students With Mathematics* (pp.133−145). Reston, VA: National Councils of Teachers of Mathematics.

Sbhatu, D. B. (2006). *Investigating the Effects of Metacognitive Instruction in Learning Primary School Science in Some Schools in Ethiopia.* Unpublished Doctoral Dissertation, North Carolina State University, Raleigh, NC.

Schmidt, M. E. & Bronsnan, P. A. (1996). *Mathematics Assessment Practices Reporting Methods*. Vol. 96, No. 1. Mansfield, OH: Ohio State University.

Segers, M. & Dochy, F. (2001). New assessment forms in problem-based learning: the Value-added of the students' perspective. *Studies in Higher Education*, *26*(3), 327–343.

Senk, S. L. (1989). van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education*, *20*, 309−321.

Slater, T.F. (1996). Portfolio assessment strategies for grading first-year university physics student in the USA. *Physics Education*, *31*, 82–86.

Slauijsmans, D., Dochy, F., and Moerkerke, G. (1998). The use of self-, peer and co-assessment in higher education: a Review. *Studies in Higher Education*, *24*(3), 331–350.

Somervell, H. (1996). Issues in assessment, enterprise and higher education: the Case for self, peer, and collaborative assessment. *Assessment and Evaluation in Higher Education*, *18*(3): 221 – 233.

Stallings-Roberts, V. (1994). Exploratory geometry: Let the students write the text. *The Mathematics Teacher*, *87*(6), 403–408.

Stenmark, J. K. (1989). *Assessment Alternatives in Mathematics: an Overview Assessment Technique that Promotes Learning–EQUALS*. Berkley, CA: University of California.

Struyven, K., Dochy, P., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a Review. *Assessment & Evaluation in Higher Education*, *30*(4), 331–347.

Taki, S. & Heidari, M. (2011). The effect of using portfolio-based writing assessment on language learning: the Case of young Iranian EFL learners. *English Language Teaching*, *4*(3), 192–199.

Tanner, H. & Jones, S. (1994). Using peer and self-assessment to develop modeling skills with students aged 11 to 16: Socio-constructivist view. *Educational Studies in Mathematics*, *27*, 413–431.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, *68*(3), 249–276.

van de Walle, J.A. (1998). *Elementary and Middle Schools Mathematics: Teaching Developmentally*. 3rd Ed. New York, NY: Longman.

van den Berg, G. (2004). Use of assessment in the development of higher order thinking skills. *African Education Review*, *1*(2), 279–294.

Van den Berg, I., Admiraal, W., & Pilot, A. (2006). Peer assessment in university teaching: Evaluating seven course designs. *Assessment & Evaluation in Higher Education*, *31*(1), 19–36.

VanLehn, K. A., Chi, M. T. H., Baggett, W., & Murray, R. C. (1995). *Progress Report: Towards a Theory of Learning during Tutoring*. Pittsburg: Learning Research and Development Center, University of Pittsburg.

Webb, N. L. & Welsch, C. (1993). Assessment and evaluation for middle grades. In D.T. Owens (Ed.), *Research Ideas for the Classroom: Middle Grades of Mathematics* (pp. 299–316). New York, NY: McMillan.

Weldeana, H. N. (2008). *The Effect of Problem-Based Teaching and Learning on the Achievement of High School Mathematics Learners.* MEd. Dissertation, UNISA, Pretoria.

Weistein, C. E. & Palmer, D. (1990). *LASSI-HS Manual.* Cleanwater, FL: H&H Publishing.

White, R. T. & Mitchell, I. J. (1994). Metacognition and the quality of learning. *Studies in Science Education, 23*, 21–37.

Willcoxson, L. E. (2006). "It's not fair!" Assessing the dynamics and resourcing of teamwork. *Journal of Management Education*, *30*(6), 798–808.

Williams, N. B. (2000). Journal writing in mathematics classroom: a Beginner's approach. *The Mathematics Teacher*, *93*(2), 132–135.

APPENDICES

Appendix 1

STAR

STAR (Circle One) [1 = Poor; 2 = Fair; 3 = Good; 4 = Excellent]	Score
1. Reflective Activity	
Makes very limited reflections often incomplete; makes very brief and incomplete comments; completes very few reflective activities.	1
Makes some reflections; makes some extended comments with little relevance to topics of discussion; completes some reflective activities.	2
Makes good reflections and critics; completes most reflective activities with good quality; makes discussions that are directly related to lesson topics and tasks.	3
Makes well-developed reflections; exhibits self-critics and self-analysis ability; makes relevant and related discussions to topics; completes all reflective tasks.	4
2. Summary Making and Journal Writing	
Makes little thought and spends little time to summary making; exhibits little understanding of summary making tasks; improper use of mathematical notions and terms; makes little explanation of tasks; does not use multiple representation; makes little communication of concepts, ideas, and definitions.	1
Makes some thoughts and spends some time to summarize; exhibits some ability to be self-critical; uses mathematical notation and terms; gives some quality explanations of tasks; communicates concepts, ideas, and definitions clearly to some extent; provides some multiple mathematical representations.	2
Has good grasp of summary making skills with good improvements; uses proper mathematical notations and terms; provides quality explanation for most of the tasks; communicates concepts, ideas, and definitions clearly oftentimes; provides most of the multiple mathematical representations of tasks.	3
Demonstrates high degree of summary making skills; selects big ideas accurately during summary making; shows clear use of mathematical notation and terms; gives quality explanations of tasks; communicates concepts, ideas, and definitions clearly; provides multiple mathematical representations of tasks exhaustively.	4
3. Contributions to Group Work	
Let others get on tasks often; rarely participates constructively in discussions; often works in isolation within the group.	1
Takes a reasonable share of group tasks, usually when asked to do so; listens to others and makes occasional contributions to discussions.	2
Takes on works within group willingly, occasionally helping to organize; is often a good active listener and contributes readily to discussions; often helps the group to move forward.	3
Shows clear ability to help the group get on tasks and to involve group members; encourages others to speak and makes them feel involved; good leadership skills; gets on tasks readily and reliably.	4
4. Portfolio Assignment	
Fails to select good representative problems; words, drawings, and diagrams are not relevant to problems, but are evidence of efforts to cope up with the idea.	1
Selects problems that represent lessons slightly; submits incomplete works with no clear solution processes.	2

STAR (Circle One) [1 = Poor; 2 = Fair; 3 = Good; 4 = Excellent]	Score
Shows solid understanding in selecting problems representing the lessons; uses proper mathematical language with figures describing terms, but missing minor elements; understands features of problems, explores them and selects appropriate strategies; and reviews, revises, extends.	3
Shows understanding in selecting problems with complete understanding of their features; explores problems with detailed understanding; selects appropriate and workable strategies; reviews, revises, and extends problems with proper figures.	4
5. Exposition	
Makes little attempt to explain concepts, algorithms, and theorems; makes very little attempt to interpret graph and figures; makes almost no attempt in writing problems; makes invalid generalizations.	1
Makes some attempt to explain concepts, algorithms, and theorems; makes little attempt to interpret graphs and figures; makes incomplete attempts in writing problems; states some elements of generalizations.	2
Makes good attempt to explain concepts, algorithms, and theorems; makes some sensible interpretations of graphs and figures; writes representative problems of lessons; makes valid generalizations.	3
Makes quality explanation of concepts, algorithms, and theorems; makes sensible interpretation of graphs and figures; writes representative problems of lessons; makes valid and sensible generalizations.	4
6. Attendance and Punctuality	
Attends less than 80% of lessons; completes lesson tasks late; participates rarely in class discussions with the teacher.	1
Attends 80% of lessons with excused absence; completes lesson tasks sometimes late; participates sometimes in class discussion with the teacher.	2
Attends 90% of lessons with excused absence but usually punctual; completes lesson tasks on time; participates in class discussion with the teacher.	3
Attends all the lessons; always punctual for lessons; completes all tasks on time; actively participates in class discussion with the teacher.	4
Total (Maximum 24 Points)	

Sources: Dietiker (1997a, b), Hoey and Wotton (1994a, b), and MoE (2006)

Appendix 2

SPUR Test Questions

Nature of Test Items							
Skills		Properties		Uses		Representation	
Items*	Reference*	Items	Reference	Items	Reference	Items	Reference
1	150, 20	14	151, 32	30	152, 42	43	153, 54
2	150, 23	15	151, 32	31	152, 43	44	153, 55
3	150, 22	16	251, 21	32	251, 46	45	251, 50
4	249, 12	17	251, 45	33	251, 47	46	413, 64
5	249, 5	18	352, 29	34	353, 35	47	57, 60-1
6	411, 33	19	56, 48	35	412, 47	48	516, 5
7	411, 24	20	412, 44	36	412, 50	49a	519, 45
8	463, 22	21	412, 43	37	464, 34-7	49b	519, 46
9	463, 26	22	518, 32	38	518, 34	50	561, 33
10	517, 15	23	464, 32	39	518, 36		
11	624, 6	24	518, 28	40	627, 37		
12	728, 11	25	518, 33	41	731, 56		
13	788, 4	26	626, 26	42	791, 44		
		27	679, 15				
		28	790, 36				
		29	791, 40				

* "Items": Refers to test items in this study. "References": Refers to page and question number (e.g. 150, 20) in Coxford et al. (1991).

Appendix 3

LASSI-HS Version

LASSI-HS Subscales	Statement Nos. (Nos. in the Original 76-Item LASSI-HS, +/−)*			
	1	2	3	4
1. Anxiety	1 (1, −)	15 (25, −)	31 (51, −)	17 (31, −)
2. Attitude	5 (5, −)	20 (37, −)	30 (50, −)	38 (68, −)
3. Concentration	6 (6, −)	21 (38, −)	23 (42, −)	25 (45, +)
4. Information processing	22 (39, +)	26 (46, +)	37 (66, +)	40 (75, +)
5. Motivation	9 (13, +)	16 (28, +)	18 (35, +)	28 (48, −)
6. Selecting main idea	2 (2, +)	8 (8, +)	24 (43, +)	34 (59, −)
7. Self testing	4 (4, +)	14 (24, +)	19 (36, +)	39 (72, +)
8. Study aid	7 (7, +)	29 (49, +)	32 (52, +)	35 (62, +)
9. Test strategies	10 (17, +)	12 (21, +)	33 (58, +)	36 (64, +)
10. Time management	3 (3, −)	11 (20; −)	13 (22, −)	27 (47, +)

* "Statement Nos.": Represented by the numbers outside brackets refer to "statement numbers" in the 40-item LASSI-HS of the present study. "Nos. in the Original 76-Item LASSI-HS": Represented by the numbers inside the brackets refer to the "statement numbers" in the original LASSI-HS. "+": Refers to statements that imply supportive learning behavior. "−": Refers to statements that imply limiting learning behaviors. The numbers "1 to 4" below the row of "Statement Nos." are those indicated vertically in **Table 3** next to each of the 10 items of the LASSI-HS subscales (Source: Weistein & Palmer, 1990).

http://iserjournals.com/journals/eurasia