

The Model of Poetry Writing Learning Based on Authentic-Collaborative Assessment in High School

Suci Sundusiah

Department of Indonesian Language and Literature
Universitas Pendidikan Indonesia
Bandung, Indonesia
suci.sundusiah@upi.edu

Abstract—The insistence on the need for poetry writing instruments in school encourages efforts to compile authentic instruments for poetry assessment. The assessment did not only focus on the final work of poetry, but also on the process of creating poetry. This study examined the effect of learning to write poetry based on authentic assessment instruments combined collaboratively to assess the process of creating student poetry. The learning model in this study hereinafter referred to as the PMP-PAK Model (Learning to Write Poetry - Collaborative Authentic Assessment). Quantitative methods were employed with two quasi-experimental research designs namely single group pre-experimental that is testing instruments in three experimental classes with different school clusters; (2) quasi-experimental in the experimental and control classes by comparing the use of instruments in the experimental class and the use of conventional assessment instruments, as well as partial authentic assessments such as the class of self-assessment, peer evaluation, and conference evaluation. Subjects were high school students in class XI in three schools representing high, medium and low clusters in the city of Bandung, West Java, Indonesia. The results revealed that the PMP-PAK model influenced the learning of poetry significantly. This influence proved that the PMP-PAK model (1) enhances students' imagination and creativity in writing poetry; (2) fosters the ability to cooperate with students in small groups; (3) fosters the initiative for the positive character of students, for example, mutual respect for peer work as well as own work.

Keywords: *Collaborative Authentic Assessment, learning to write poetry*

I. INTRODUCTION

Authentic assessment and learning to write poetry are two things that support each other (Behizadeh, 2015; Ward, 2013; Winch, Johnston, March, Ljungdahl, & Holliday, 2006). As teaching material in a language class, poetry writing requires the authenticity of the task in the form of the final product of student poetry (Walker, 1997). In addition, poetry requires students to process the creation of ideas and express ideas that need guidance from the teacher in class. The process of creating this poem is a high-level cognitive activity that involves a variety of potential literacy students (Hughes, 2007). However, the process of writing the poem overlooked by the teacher in class in a way that the task of writing poetry is only focused on the final work. Therefore, an authentic assessment model that recognizes the learning process,

including the process of writing poetry, is very appropriate in assessing students' poetry authorship.

The teachers' need for authentic assessment instruments in poetry writing classes, especially in Indonesia is incredibly urgent. Previous research in writing poetry classes in Indonesia shows that teachers do not yet have poetry assessment instruments. The poetry instrument is the teacher itself, consequently, students do not engage in assessment activities (Sundusiah, Rofiuddin, Suwignyo, Basuki, 2019). Other studies note that one of the factors causing the failure of teaching poetry in the classroom is the teacher's competence in poetry (Dymoke, 2003, 2012ab; Xerri, 2013; 2016). Poetry is considered a difficult subject therefore it is characterized in language learning (Okonkwo, 2016; Ofsted, 2012; Cremin, 2011). Thus, teachers need instruments for poetry writing assessment that supports them in teaching poetry in class. The existence of authentic assessment instruments that can be used by teachers in relevant conditions and situations is one of the convenient solutions for teachers.

This study examines a model of learning to write poetry based on authentic assessment activities, called Learning of Collaborative-Authentic Poetry Writing (PMP-PAK). This model consists of a poetry study; 4 steps of authentic assessment technique to write poetry collaboratively, and 5 authentic assessment rubrics integrated into a student workbook. Activities to learn writing poetry include (1) the 1st Poetry Creation, and (2) The Xth Poetry Creation. Authentic assessment activities include (1) Poetry Expression, (2) Poetry Distribution, (3) Poetry Discussion, and (4) Poetry Reflection.

This study aims to (1) test the PMP-PAK model in the experimental class with different student qualifications based on school clusters namely high-school, medium-school, and low-school; (2) test the PMP-PAK model compared to the poetry learning model with conventional (non-authentic) assessment and partial authentic assessment, namely self-assessment, peer evaluation, and conference evaluation.

II. METHOD

A. Design

Two quasi-experimental research designs were used namely: (1) a single group pre-experimental design (Creswel, 1994) to test instruments in three experimental classes with different school clusters; (2) quasi-experimental design of the experimental and control classes (Creswel, 1994) in the middle-school cluster by comparing the use of instruments in the experimental class and the use of conventional assessment instruments, as well as partial authentic assessments, namely self-assessment, peer evaluation, and conference evaluation.

B. Participants

This study involved 7 classes of students with a total of 84 students in the experimental class and 109 students in the control class. Class samples were randomly selected from the class population in 3 high schools with the characteristics of high-school clusters, middle-school clusters, and low-school clusters. The selection of schools is done randomly from the school population based on each of these clusters in the city of Bandung, West Java, Indonesia. The characteristics of school clusters are distinguished by passing grades for new students. The details information about the research design and research sample is as seen in Table I.

TABLE I. DESIGN AND RESEARCH SAMPLE

Design	Treatment	Sample code	School clusters
Quasi pre-experiment pre-test post-test single group	Authentic-collaborative assessment instrument in 3 school clusters	K-EKS01	High
		K-EKS02	Middle
		K-EKS03	Low
Quasi-experiment pre-test post-test Experimental and control group	Authentic-collaborative assessment instrument	K-EKS02	Middle
	Without instrument	K-EKS02 K-KON01	
	Self-assessment signs	K-EKS02 K-KON02	
	Peer assessment signs	K-EKS02 K-KON03	
	Conference assessment signs	K-EKS02 K-KON04	

Information:

- O_1 : Pre-test ;
- O_2 : Post-test;
- K_1 : Experimental Group;
- K_{2a} : Conventional assessment Control Group;
- K_{2b} : Self-assessment Control Group;
- K_{2c} : Peer assessment Control Group;
- K_{2d} : Conference assessment Control Group;
- $K-EKS01,02,03$: Experiment Class 01, 02, and 03;
- $K-KON01,02,03,04$: Conventional Class 01,02,03,04

C. Research Hypothesis

Research hypotheses to test the PMP-PAK model in the pre-experimental class are:

H_0 there is no difference between the ability (written/cognitive/affective expression) for Poetry Creation students who are treated with the PMP-PAK model in high, medium, and low clusters.

H_a there is a difference between the ability (written/cognitive/affective expression) for Poetry Creation students who are treated with the PMP-PAK model in high, medium, and low clusters.

Research hypotheses to test the PMP-PAK model compared to other poetry writing learning models based on other assessment models are:

H_0 there is no difference between the ability to write poetry for students who are treated with the PMP-PAK model and the ability of students who are given the conventional/Spear/self/conference assessment.

H_a There is a difference between the ability to write poetry for students who are treated with the PMP-PAK model and the ability of students who are treated with the conventional/peer/self/conference assessment.

D. Research Instruments

The instruments in PMP-PAK are compiled based on a review of (1) curriculum achievements that have been and are currently in force in the field of poetry writing in Indonesian language subject material in high school; (2) the concept of authorship of poetry that proceeds and requires practice (Withworth, 2003; Mock 1998; Beach & Marshal, 1991); (3) the concept of the construct of poetry consisting of language and forms that are centered on meaning (Altenbernd & Lewis, 1966; Kennedy & Gioia, 1994); (4) authentic assessment studies on writing poetry (Griswold, 2006; LeNoir, 2002; Walker, 1997).

Before being tested, the instruments in PMP-PAK have gone through the stages of content validity and reliability testing. Validity test was conducted on 3 experts, namely a professor in the field of language learning assessment, a professor of poetry learning and a practitioner of poetry writing trainers and renowned poets. The validity test is focused on the content validity test, namely testing the correctness of the contents of the device with the achievement of poetry writing in the curriculum applied in Indonesia and the determination of the contents of the device with poetic writing composition both theoretically and practically (Gulikers, Bastiaens, & Kirschner, 2005; O'Malley & Pierce, 1995; Darling-Hammond, 1994; Shepard, 2001).

Meanwhile, the instrument reliability test on PMP-PAK which focuses on 5 rubric assessments was accompanied by an inter-rater reliability test (Herman, 1992) which is conducting rater training as many as 8 meetings between teachers and experts. Rater-training activities were carried out to minimize differences in interpretation and ambiguity between assessors (Ross & LeGrand, 2017; Attali, 2016; Rezaei & Lovorn, 2010; Knoch, 2009; Weigle, 2002). The rubric was tested on 2 teacher evaluators and 35 students with each writing 2 poems in two stages of assessment. The test results showed that the

rubric had content validity with a percentage of agreement between experts of 100% (compatibility of the device with curriculum achievements); 94.79% (accuracy of the device with poetic writing theory); and 94.79% (compatibility of the device with the practice of writing poetry). Likewise, the rubric was stated to be reliable with the percentage of agreements between the appraisers on the Poetry Creation Rubric (84.72%), Poetry Expression (81.67%), Poetry Distribution (90%), Poetry Discussion (83.34%), and Poetry Reflection (96.67 %).

III. FINDINGS AND DISCUSSION

A. Effectiveness of the PMP-PAK Model in the Pre-Experimental Group with the School Cluster Differences

After the normality test using the Kolmogorov-Smirnov and Shapiro Wilk tests were assisted by SPSS 22 for Windows, obtained each experimental class namely K-EKS01, K-EKS-02, and K-EKS03 in the sample had a confidence level in the percentage of 95% or level of the significance of each class is above 0.05 so it can be said that the sample class is normally distributed. The homogeneity test showed that the significance level of each class was greater than the significance level of 0.05. Thus, the variance data is homogeneous. Therefore, the sample group data can be tested with parametric statistical tests.

Hypothesis test results showed that the F count on the ability of Poetry Creation 17.442 was greater than the F table, or F count > F table (17.442 > 3.11). Thus, H₀ was rejected and H_a was accepted, meaning that there were differences in the ability to write poetry (Poetry Creation) between students who were treated with the PMP-PAK model in the high, medium, and low clusters. On the other hand, the results of the hypothesis test on cognitive abilities in understanding the process of assessing poetry showed that the F count of 20,235 was greater than the F table, then H₀ was rejected and H_a was accepted.

TABLE II. PRE-EXPERIMENT GROUP TEST RESULT

Rubric Testing	Competence	Class	Assumption Testing Sig. > 0,05		Hypothesis class F _{count} > F _{table}	Conclusion	
			Normality				
			K-S	S-W			
Poetry Creation	Cognitive	K-EKS01	0.200	0.897	0.073	17,429 > 3,11	Ho rejected Ha accepted
		K-EKS02	0.200	0.388			

(1) Poetry Creation; (2) Poetry Distribution; (3) Poetry Discussion (4) Poetry Reflection	Cognitive	K-EKS03	0.200	0.425	0.431	20.235 > 3,11	Ho rejected Ha accepted
		K-EKS01	0.200	0.881			
		K-EKS02	0.200	0.307			
	Affective	K-EKS03	0.119	0.161	0.065	21.905 > 3,11	Ho rejected Ha accepted
		K-EKS01	0.200	0.712			
		K-EKS02	0.009	0.007			
		K-EKS03	0.178	0.627			

As shown in Table II, Pre-experimental group hypothesis test results namely K-EKS-01, K-EKS-02, and K-EKS-03 concluded that H_a was accepted. This means that there was a significant difference between students' ability to write poetry in high, medium and low school clusters after being treated with the PMP-PAK model. This hypothesis test data showed that the PMP-PAK model can be applied in schools with various levels or different clusters. The application in various schools with diverse clusters influenced the ability to write poetry for students in a variety of ways.

Other data (see Table III) showed that the cognitive and affective abilities of K-EKS02 students were higher than K-EKS01 and K-EKS03. This means the ability of students in schools with middle clusters to understand the process of poetry writing and appreciation in poetry assessment activities was better than students in high and low cluster schools. The data also showed that the average cognitive and affective abilities of high and low-grade students were not too far away.

TABLE III. DESCRIPTION OF COGNITIVE AND AFFECTIVE ABILITY IN EXPERIMENTAL GROUP

Class	Mean	
	<i>Cognitive</i>	<i>Affective</i>
K-EKS01	3.2009	3.2235
K-EKS02	3.7360	3.6980
K-EKS03	3.1655	3.2158

Based on the description table of the ability of Poetry Creation, it was found that the mean of K-EKS-02 was in the first place. The second place was K-EKS-03 and the last place was K-EKS-01. K-EKS-02 was a sample group of students taken at random in medium or middle cluster schools. K-EKS-03 was a sample group of students taken at random in low cluster schools. While K-EKS-01 was a sample group of students taken at random in high school clusters. With the data above, there were unique things found in this study. The ability of students to write poetry in middle and low cluster school groups can outperform students in high cluster school groups. This gave rise to several interpretations, namely: (1) the PMP-PAK model can be easily applied to students with the character of middle and low-school clusters, 'periphery' schools and rural schools; (2) The PMP-PAK books can be applied to students with middle and lower levels if the cognitive level was determined by the passing grade of the school; (3) furthermore, poetry as the main content in the PMP-PAK model turned out to be easily accepted by students in the middle and low-school cluster levels if approached with an authentic learning model and a collaborative authentic assessment model.

This is supported by respondent data in the discussion of previous findings, that students with the character of the 'periphery' school can receive poetry far better than students in high cluster schools (high graduation standard schools). Other studies prove that students' attitudes toward learning, including literature, were based on the location of schools and classrooms (Ghazali, 2008). In this study, the response of high-class students to poetry was lower than students in the middle and low clusters.

The data on the ability to write poetry in Poetry Creation was not in line with the cognitive abilities of students towards the PMP-PAK model. Data from the hypothesis test results showed that the PMP-PAK model had a diverse effect on students in various clusters. The PMP-PAK model was able to influence the cognitive abilities of medium cluster students better than high and low cluster students. Table description of cognitive abilities of pre-experimental class students showed K-EKS-02 students had the highest cognitive abilities, K-EKS-01 students had second cognitive abilities, and K-EKS-03 students had third cognitive abilities. That means the PMP-PAK model can be more accepted in groups of students with intermediate cognitive abilities. High cluster students can certainly understand books better than low cluster students. However, due to their responses to PMP-PAK books were lower than those of moderate cluster students, their cognitive abilities were not higher than those of moderate cluster students.

Hypothesis test data for students' affective abilities in the high, medium and low clusters after being treated with the PMP-PAK model also varied. That means the PMP-PAK model can distinguish the affective abilities of students in high, medium and low clusters. The description of affective abilities showed that the mean of K-EKS-02 students was in the highest position, while the affective abilities of K-EKS-01 students were in second, and only 0.0077 numbers of students of K-EKS-03 in third. These findings indicated that in intermediate cluster students (K-EKS-02) the ability to write poetry was parallel to the cognitive ability to understand the content of poetry material and the affective ability to respond to poetry. While in high and low cluster students, the ability to write poetry was not directly proportional to the cognitive and affective abilities of students towards poetry. High cluster students were cognitively able to understand the content of poetry material and were affectively able to respond to poetry better than low cluster students who had better poetry writing skills but lacked cognitive and affective abilities towards poetry.

Based on the explanation above, this study found several facts that (1) the PMP-PAK model had a better influence on the ability to write poetry, the cognitive abilities of poetry and the affective abilities of poetry on students in secondary cluster schools; (2) the PMP-PAK model gave an effect on cognitive and affective abilities about poetry that were better for students in high clusters than students in low clusters; but (3) the PMP-PAK model had a better influence on the ability to write poetry to students in the lower cluster than students in the high cluster. Thus, it can be concluded that the PMP-PAK model can be accepted and applied across diverse school clusters.

B. Effectiveness of the PMP-PAK Model in the experimental group and other instruments in the control group

Normality test using the Kolmogorov-Smirnov and Shapiro Wilk tests, obtained samples (K-EKS02 and K-KON01) had a confidence level in the percentage of 95% or the significance level of each class is above 0.05, therefore, it can be said that the sample class was distributed normally. Meanwhile, the homogeneity test showed that the significance level of 0.855 was greater than 0.05 ($0.855 > 0.05$), meaning that the variance of the compared sample group was stated to be insignificant which means homogeneous variance. Thus, hypothesis testing can be done with parametric statistics.

The research hypothesis test was conducted based on the data of the increase in pre-test and post-test scores of students' Poetry which had been processed in the form of N gain (normalized gain). Based on the results of the independent sample t-test, it appeared that the t-test was 5,216 while the t-table was 2007 so that the t-count > t-table was $5,216 > 2,007$ then H_0 was rejected and H_a was accepted. Thus, there were differences in the learning outcomes of students who were treated with the PMP-PAK model with conventional assessments.

TABLE IV. EXPERIMENTAL AND CONTROL GROUP TEST RESULT

Comparison	Assumption Testing Sig. > 0,05		Homo geneity	Hypothesis Test $T_{count} >$ table	Conclusio n
	Normality				
	K-S	S-W			
K-EKS02	0.200	0.687	0.855	5.216 > 2.007	Ho ditolak Ha diterima
K-KON01	0.200	0.155			
K-EKS02	0.200	0.687	0.084	4.777 > 2.001	Ho ditolak Ha diterima
K-KON02	0.200	0.826			
K-EKS02	0.200	0.687	0.269	2.667 > 2.001	Ho ditolak Ha diterima
K-KON03	0.054	0.058			
K-EKS02	0.200	0.678	0.060	3.434 > 2.008	Ho ditolak Ha diterima
K-KON03	0.058	0.069			

As seen in Table IV, four control classes served as a comparison for the application of the PMP-PAK model in the experimental class (K-EKS-02). First, the control class which was given the conventional assessment treatment (K-KON-01); Second, the control class that was given a self-assessment treatment (partial) (K-KON-02); and third, the control class which was given partial peer assessment (K-KON-03); and fourth, the control class given the conference evaluation treatment (K-KON-04). The results of the hypothesis test concluded that there were significant differences in learning outcomes between the application of the PMP-PAK model in the experimental class when compared to the application of conventional assessment in all control classes. That is, the poetry work of students who were treated with the PMP-PAK model had a higher score than the poetry work of students in the control class. Thus, the PMP-PAK model has a role in improving the ability to write poetry of students compared to learning poetry with conventional assessments, including authentic assessment models (self-assessment, peers, and conferences) that were applied partially.

Interesting findings can be traced through the following Table V. The data showed that t arithmetic on the comparison of experimental class (K-EKS-02) and K-KON-03 was closer to t table compared to the comparison of t arithmetic with t table of the experimental class and the control class K-KON-01, K-KON-02, and K-KON-04. K-KON-03 was given peer assessment treatment. This can be interpreted that the application of peer assessment had a character that is closer to the application of the PMP-PAK model. While the application of conventional assessments as well as self-assessment models and conferences conducted partially cannot match the application of the PMP-PAK model. This finding was corroborated by the findings of students' responses to peer assessment activities on K-KON-03 higher than conventional assessments, self-assessments, and conference evaluations.

TABLE V. COMPARISON OF T COUNT AND T TABLE OF EXPERIMENTAL AND CONTROL GROUP

CLASS COMPARISON	T count: T table	Conclusion
K-EKS-02 and K-KON-01	5.216 > 2.007	Ho rejected; Ha accepted
K-EKS-02 and K-KON-02	4.777 > 2.001	Ho rejected; Ha accepted
K-EKS-02 and K-KON-03	2.667 > 2.001	Ho rejected; Ha accepted
K-EKS-02 and K-KON-04	3.434 > 2.008	Ho rejected; Ha accepted

The PMP-PAK model was designed with the predominance of group assessment activities, both peer group and community assessment. Peer and communal assessment models were collaborated with self-assessment so that the PMP-PAK model was superior in improving students' poetry writing skills. The PMP-PAK model that involved students collaboratively was close to the peer assessment model, although the application of partial peer assessment cannot yet be maximized to improve the ability to write student poetry. Peer assessment in K-KON-03 was carried out only by applying general rating signs. This can also affect the results of the assessment. Peer or self-assessment that was not equipped with a clear assessment rubric would cause high subjectivity in the assessment. Students would feel insecure because their peer rate is subjective as opposed to be based on an objective device (Duers, 2017).

On the other hand, this finding corroborates previous research that peer evaluation will be maximized if they collaborate with self-assessment (Basuki, 2007; Shin, 2017). Peer assessment is more effective if it is applied reciprocally (Tsivitanidou, Constantinou, Labudde, Rönnebeck, & Ropohl, 2018) which in this study is called conference evaluation. In addition, peer assessment and self-assessment will run optimally if it is equipped with a set of scoring rubrics that are measurable and optimally collaborated with complementary activities, for example, activities are providing sufficient feedback. All of these activities are summarized in the PMP-PAK model.

The finding that poetry turned out to be better responded to by students in the low academic competency group was corroborated by several studies that poetry was closer to the marginalized group (Lipenga, 2015). This marginalized group used to express social criticism, moral criticism, and other criticisms delivered through poetry. Indonesian poetry knows Chairil Anwar, Rendra, Taufiq Ismail, K.H. Mustofa Bisri, Acep Zam Zam Noor, etc. as a voice of social and moral criticism. Students in schools with low academic qualifications were determined by the passing grade scores of the school entrance selection. The passing grade determined the qualifications of a group of students. Students with high grades will be free to enter schools with high qualifications. Students with moderate grades will enter schools with moderate qualifications. While students with low grades can only be accepted at schools with low passing grade acceptance qualifications. Characteristics of students who are in this low cluster including those who (1) come from lower-middle economic families since family economic conditions affect student academic competence (Dietrichson, Bøg, Filges,

& Klint, 2017); (2) vulnerable to criminalization and harassment problems involving excessive emotional and empathetic problems (Solomontos-Kountouri, Tsagkaridis, Grading, & Strohmeier, 2017); (3) lack of academic achievement (Hardré, Sullivan, & Crowson, 2009). However, research in Chicago public schools showed that academic achievement was not determined by the school cluster but by the level of school performance. In low cluster schools that were not selected based on the highest test scores, graduation rates increased both academically and non-academically (Allensworth, Moore, Sartain, & de la Torre, 2017).

This study corroborates the data that students in K-EKS-03, with the characteristics of schools with low passing grade grades, with a lower-middle economic background, who are vulnerable to emotional problems and crime (Dietrichson et al., 2017) can easily associate with poetry. The researcher found that students in this school were more expressive in expressing ideas, more spontaneous and open in thinking, not confined to the image that they should look better in writing poetry. Students in this marginalized group voiced their emotions and thoughts more relaxed, flexible, weightless, and free. In K-EKS-03, the researchers found the impression that the PMP-PAK model was new to them. Students at K-EKS-03 were enthusiastic because the model invites students to express and imagine. Poetry for students becomes lighter and more relaxed. Although on the contents of the diction assessment material, students in K-EKS-03 must practice more. However, an authentic emotional approach to the PMP-PAK book was able to have a positive effect on students in low-cluster schools.

IV. CONCLUSION

The PMP-PAK model enhanced students' ability to work in small groups, which was to highlight the role of peer educators in understanding the content of the material and writing poetry; teamwork to help each other understand and understand the assessment sections together. The PMP-PAK model fostered the student's positive character, which was (1) mutual appreciation of the work of peers, in the form of appreciation of praise and appreciation both verbal and written; (2) an attitude towards one another and to be able to accept criticism and advice on the shortage of his work; (3) the courage to present opinions with appropriate arguments and to respect one's opinions; (4) Respect for self-esteem. The PMP-PAK model enhanced students' imagination in writing poetry; teamwork and end-of-life poetic celebrations provide a non-serious and monotonous assessment process experience. This reinforced the impression that authentic-collaborative assessment activities in PMP-PAK were fun for students.

REFERENCES

Allensworth, E. M., Moore, P. T., Sartain, L., & de la Torre, M. (2017). The Educational Benefits of Attending Higher Performing Schools: Evidence From Chicago High Schools. *Educational Evaluation and Policy Analysis*, 39(2), 175-197.

Altenbernd, L & Lewis, L.L. (1966). *A handbook for the study of poetry*. New York: MacMillan Publishing Co., Inc.

Attali, Y. (2015). A comparison of newly-trained and experienced raters on a standardized writing assessment. *Language Testing*, 33(1), 99–115.

Basuki, I.A. (2007). *Pengembangan model penilaian sebaya untuk meningkatkan hasil pembelajaran menulis di SMP* (Unpublished doctoral dissertation). Universitas Negeri Malang.

Beach, R. & Marshall, J. (1991). *Teaching literature in the secondary school*. New York: Harcourt Brace Jovanovich.

Behizadeh, N. (2015). Xavier's take on authentic writing: Structuring choices for expression and impact. *Journal of Adolescent and Adult Literacy*, 58(4), 289–297.

Cremin, T. (2011). "Reading teachers/teaching readers." *English Drama Media*, 19, 11–18.

Creswel, J. (1994). *Research design: Qualitative and quantitative approaches*. London: SAGE Publications.

Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review*, 53(1), 5-30.

Dietrichson, J., Bøg, M., Filges, F., & Klint Jørgensen, A. M. (2017). Academic interventions for elementary and middle school students with low socioeconomic status. *A Systematic Review and Meta-Analysis*, 87(2), 243-282.

Duers, L.E. (2017). The learner as co-creator: A new peer review and self-assessment feedback form created by student nurses. *Nurse Education Today*, 58, 47-52.

Dymoke, S. (2003). *Drafting and assessing poetry*. London: SAGE Publications Company.

Dymoke, S. (2012a). Poetry is an unfamiliar text: locating poetry in secondary english classrooms in new zealand and england during a period of curriculum change. *Changing English*, 19(4), 395–410.

Dymoke, S. (2012b). Opportunities or constraints? Where is the space for culturally responsive poetry teaching within high-stakes testing regimes at 16+ in Aotearoa New Zealand and England? *English Teaching: Practice and Critique*, 11(4), 19-35.

Ghazali, S.N. (2008). Learner Background and their Attitudes towards Studying Literature. *Malaysian Journal of ELT Research*, 4, 1-17.

Griswold, A. (2006). Assessment List: One solution for evaluating student poetry. *English Journal Arts & Humanities*, 96(1): 70-75.

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2005). A Five-Dimensional Framework for Authentic Assessment. *Educational Technology Research and Development*, 52(4), 67-86.

Hardré, P. L., Sullivan, D. W. & Crowson, H. M. (2009). Student characteristics and motivation in rural high schools. *Journal of Research in Rural Education*, 24(16), 1-19.

Herman, J. L. (1992). what research tell us about good assessment. *Educational Leadership*, May, 72-78.

Hughes, J. (2007). *Poetry: A powerful medium for literacy and technology development*. Ontario: Literacy and Numeracy Secretariat.

Kennedy, X.J. & Gioia, D. (1994). *An introduction to poetry eight edition*. New York: Harper Collins College Publisher.

Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, 26, 275-304.

LeNoir, W. D. (2002). Grading student poetry: A few words from the devil's advocate. *English Journal Arts & Humanities*. 91(3), 59-63.

Lipenga, K. J. (2015). Voicing marginality: Disability in Leila Aboulela's *Lyrics Alley*. *Journal of African Cultural Studies*, 30(1), 93-104.

Mock, J. (1998). *You Can Write Poetry*. Ohio USA: Writer's Digest Books.

O' Malley, J. M. & Pierce, L.V. (1996). *Authentic assessment for english language learner practical approach for teacher*. USA: Longman.

Okonkwo, A. F. P. (2016). Effective Teaching of Poetry : Impediments and sustainability option in senior secondary schools in onueke education zone. *International Journal of English Language, Literature and Humanities*, 4(10), 23–36.

Ofsted (Office for Standards in Education). (2012). *Moving English Forward: Action to Raise Standards in English*. London: Ofsted.

Rezaei, A.R. & Lovorn, M. (2010). Reliability and Validity of Rubrics for Assessment through Writing. *Assessing Writing*. 15 (1), 18-39.

Ross, V. & LeGrand, R. (2017). Assessing writing constructs: toward an expanded view of inter-reader reliability. *Journal of Writing Analytics*, 1, 227-275.

Shepard, L. A. (2001). *The Role of Classroom Assessment in Teaching and Learning*. In *Handbook of Research on Teaching*, ed. V. Richardson, 4th ed. Washington, D.C.: American Educational Research Association.

- Shin, Dh. (2017). The Reliability of Self- and Peer Assessment and Students' Perceptions toward Peer and Teacher Assessment. *English Education Curriculum, 16*(4), 19- 43.
- Solomontos-Kountouri, O., Tsagkaridis, K. , Gradinger, P., & Strohmeier, D. (2017). Academic, socio-emotional and demographic characteristics of adolescents involved in traditional bullying, cyberbullying, *International Journal of Developmental Science, 11*(1-2), 19-30.
- Sundusiah, S., Rofiuddin, Ah., Suwignyo, H., Basuki, I.A. (2019) Indonesian Senior High School student's perspective of the poetry teacher's role in authentic assessment of poetry writing. *Journal of Poetry Therapy, 32*(1), 37-52.
- Tsivitanidou, O.E., Constantinou, C.P., Labudde, P., Rönnebeck, S., & Ropohl, M. (2018). Reciprocal peer assessment as a learning tool for secondary school students in modeling-based learning. *European Journal of Psychology of Education, 33* (1) : 51–73.
- Walker, M. (1997). Authentic Assessment in the Literature Classroom. *The English Journal, 86*(1), 69-73.
- Ward, L. (2013). Integrating poetry and writing: ways in which to engage literacy learners. *Practically Primary, 18* (2), 17-20.
- Weigle, S. C. (2002). *Assessing writing*. New York: Cambridge University Press.
- Winch, G., Johnston, R. R., March, P., Ljungdahl, L., & Holliday, M. (2006). *Literacy: Reading, writing and children's literature* (3rd Edn.). South Melbourne, VIC: Oxford University Press.
- Withworth, J. (2006). *Writing Poetry Second Edition*. London: A & C Black.
- Xerri, D. (2013). Poetry writing in the post-16 English curriculum. *English Teaching: Practice and Critique, 12*(2), 140-155.