The Development of Assessment Literacy in Chinese Pre-Service Primary Teachers

Bing Yan[1,a], Rachelle Kisst Hackett[2,b] and Juan Xu[1,c]

[1]1661 North Shengxin Rd, Jiading, SH, China, 201815

[2]3601 Pacific Avenue, Stockton, California, America, 95211

[a]bingfeng0506@163.com, [b]rhackett@pacific.edu, [c]millexu@aliyun.com

Keywords: Assessment literacy; pre-service primary teacher; training course.

Abstract. This quasi-experiment using two classes, each of these 40 participants, examined the effectiveness of a backward-designed training course for improving Chinese pre-service primary teachers' assessment literacy. Framed by the Understanding by Design approach, the 12-week course was taken by those in the treatment group. The average level of assessment literacy among the Chinese pre-service teachers was initially low, with only 42% of the items on the modified Assessment Literacy Inventory correctly answered. ANCOVA results suggest the assessment literacy training course improved pre-service teachers' assessment literacy, $F(1,77)=135.911$, $p<.001$, with an adjusted posttest mean difference of 7 points (on a 35-item test; partial eta squared= .64). Implications for teacher training and policy are discussed.

1. Introduction

In the past decades, since educators have come to realize the enormous positive impacts of a balanced assessment system on students' learning, teachers' assessment literacy has attracted more and more attention from researchers and educators. It follows, then, that growing attention has been paid to the improvement of current classroom assessment training courses. According to Poth (2013), current assessment education courses are incompatible with properly preparing teacher candidates for scoring, interpreting, and communicating assessment outcomes in classroom settings and there is a need to pay more attention to assessment for learning [1]. Likewise, DeLuca and Bellara (2013) assert that most of the current assessment courses put their emphasis on "assessment processes," "assessment fairness," and "measurement theory" (p. 366) and thus these courses promote pre-service teachers' "definitional understanding" (p. 363), rather than "higher order thinking" (p. 363) [2]. Jin (2010) carries this conversation into the Chinese teacher education context. Her study highlights the fact that, in order to support pre-service language teachers' development of assessment competency, the current college-based language testing and assessment courses in China need to be improved, not only in course content, but also in teaching methodologies [3].

The purpose of this study was to examine the effectiveness of a backward-designed assessment training course for improving the assessment literacy levels of pre-service primary teachers who participate in college-level teacher preparation programs in Shanghai, China.

2. Theoretical Framework

In the current study, the theoretical framework Understanding by Design (UbD) (also called "Backward Design") was used for integrating the assessment of student work into the lesson planning process. In contrast to previous curriculum design theories, this approach begins with the end of the process (the intended learning outcome) and designs curriculum content and lesson planning backward through the sequence of the learning process. According to Wiggins & McTighe (2005), UbD includes three important stages. In the first stage, the teacher considers where he/she needs to take students in terms of learning outcomes before specifying any learning activities. The next stage moves to determining the assessment methods. Different assessment methods designed to gather a variety of evidence need to be developed so as to function in support of the intended

learning outcomes. After determining learning goals and assessment methods, the curriculum designer moves to the final stage wherein he/she addresses lesson planning and instruction [4].

3. Methods

This study employed a nonequivalent groups pretest-posttest control group quasi-experimental design, where data collected with the Assessment Literacy Inventory was analyzed via an ANCOVA (Analysis of Covariance) to evaluate the effectiveness of a one-semester backward-designed assessment training course.

3.1 Instrumentation

The Assessment Literacy Inventory (ALI) contains 35 contextualized multiple choice items, which are scored dichotomously (0 = incorrect response, 1 = correct response). As a modified version of the Teacher Assessment Literacy Questionnaire [5], this evaluation tool was developed to examine the level of assessment literacy of pre-service and in-service teachers. The Assessment Literacy Inventory displays a relatively high reliability coefficient of .74 [6]. Minor modifications were made to the ALI in adapting it for use in China.

3.2 Participants

The sample was composed of 80 students who were enrolled in pre-service primary teacher training programs offered by a private pre-service teacher education institution "XT" (this pseudonym has been adopted to protect the identity of the institution) in Shanghai, China. The control group consisted of 40 juniors taking the Introductory Foundations of Education course whereas the treatment group consisted of 40 seniors taking the Assessment Training course.

3.3 Procedures

IRB approval was obtained and informed consent was gathered to ensure, in part, that students in the two classes realized they were not required to be research participants. The pretest was given to all, then Seniors in the treatment group participated in the Assessment Training course (see Table 1 for course outline) for 12 weeks while Juniors in the control group completed a course (Introductory Foundations of Education) already taken by the treatment group earlier in their course of study. At the semester's end, both classes took the posttest.

Table 1. Assessment Training Course Outline.

Weeks	Intended learning outcomes	Topics	Assessment evidence
1	Pre-service teachers will be able to: 1) choose assessment methods appropriate for instructional decisions; 2) recognize unethical uses of assessment information	Keys to quality classroom assessment	KWL chart: students list what they already know about the topic (K), what they want to know (W), and, toward the end of the unit, what they learned (L)
2		Roles of assessment in the education process; Assessment for and of learning; Multiple users and uses of assessment data.	Classroom observation Log: Students observe a teacher assessing in his/her class and finish an observation log. Interview: students conduct a 15-minute interview of the teacher to get more information on his/her understanding of assessment. Oral report: Based on the observation log and interview, students make an oral report to present the assessment being used around us.
3		Reliability and validity of assessment; Assessment bias; Fairness of assessment.	
4	Pre-service teachers will be able to: Integrate assessment into teaching and learning process.	Types of learning targets; Benefits of clear learning targets.	Plan for matching assessment methods with achievement targets
5		Assessment options; Matching Assessment Methods to learning targets; Assessment blueprints & development cycle	Developing test blueprint
Cont. Table 1			

6	Pre-service teachers will be able to: 1) Develop assessment methods; 2) Develop valid pupil grading procedures.	Selected response assessment.	Assessments critiques: Students review the assessments being used within teaching and learning context. Student-Generated Assessments: Students refine the assessments they reviewed and make them more reliable and valid.
7		Essay assessment.	
8		Performance assessment.	
9		Portfolios	
10	Pre-service teachers will be able to: Administer, score, interpret, and communicate assessment results.	The meaning of scores; Standardized testing; common uses of test scores;	Develop student record-keeping forms. In-class Quiz for interpreting standardized test scores. Final reflection: Student will provide a written reflection for the whole learning process of this course covering important points and noting what s/he wants to eventually understand better.
11		Guidelines for effective grading.	
12		Feedback based on assessment results; Reports and communication.	

4. Findings

4.1 Entry Level of Assessment Literacy in Chinese Pre-service Primary Teachers

Table 2 presents the means and standard deviations of the pre-service primary teachers' assessment literacy as a whole and by subgroups. Before the intervention, the average assessment literacy for the pre-service primary teachers being tested was 14.79 points with a standard deviation of 3.01 points. Because there are 35 points possible, this represents an average of 42% of the items being answered correctly.

Table 2. Assessment literacy pretest scores by experimental condition.

	N	Min	Max	M	SD
All participants	80	8	23	14.79	3.013
Control group	40	8	18	13.63	2.853
Treatment group	40	8	23	15.95	2.736

Note: The maximum point possible is 35.

4.2 Development of Assessment Literacy in Chinese Pre-service Primary Teachers Taking a One-Semester Backward-Designed Assessment Training Course

To determine the degree of similarity between the control and treatment groups before the intervention began, independent-samples *t*-tests and chi square tests of association on the demographic variables were employed. No statistically significant differences between the two groups were found, except for age (see Table 3).

Table 3. Categorical demographic information of the participants by experimental condition (treatment, n=40; control; n=40) with Chi Square Tests of Association results.

		Control Group		Treatment Group		
Variable	Response Categories	N	%	n	%	χ^2
Gender	Male	6	15	3	7.5	1.127
	Female	34	85	37	92.5	
Age	20	22	55	1	2.5	41.674*
	21	14	35	8	20	
	22	3	7.5	21	52.5	
	23	1	2.5	10	25	
Ethnic Identity	Han	39	97.5	40	100	1.013
	Minority	1	2.5	0	0	
Residence	Rural	20	50	23	57.5	1.935
	Town	6	15	4	10	
	Prefecture-level city	5	12.5	2	5	
	Big cities	9	22.5	10	25	
Career Plan	Education career	33	82.5	33	82.5	1.077
	Not education	0	0	1	2.5	
	Undecided	7	17.5	6	15	

* p< .05

As show in Table 4, the groups did not differ in terms of college entrance exam scores or college GPA. As expected, and shown by the independent samples t-test results, the groups did differ in age, since the treatment group consisted of college seniors while the control group consisted of college juniors. However, age was not controlled and put into the model as a covariate variable because, within each group, it was not significantly correlated with the dependent variable. Most important to note is that a difference existed between the two groups' pre-test assessment literacy levels, t (78) = 3.72, p<.05. Therefore, it is appropriate and statistically powerful to use ANCOVA in the data analysis procedure where the pretest scores serve as a covariate in order to make statistical adjustments between the groups.

Table 4. Continuous demographic information of the participants by experimental condition (treatment, n=40; control; n=40) with Independent-Samples t-Test results.

Variable	Group	*M*	*SD*	*t*
Age (at Pretest)	Treatment	22.00	.751	8.507*
	Control	20.58	.747	
College Entrance Exam Score	Treatment	457.29	33.671	1.874
	Control	435.82	50.376	
Grade Point Average (GPA)	Treatment	2.917	.3563	1.157
	Control	2.797	.4620	
Assessment Literacy Pretest Score	Treatment	15.95	2.736	3.720*
	Control	13.63	2.853	

* p< .05

Based on the data reported in Table 5, there is evidence to suggest that the average level of assessment literacy test scores varies between students who received the assessment training course and those who did not (F (1, 77) = 135.91, p<.001). Participation in the assessment training course is a statistically significant predictor of pre-service teachers' assessment literacy, with their previous assessment literacy controlled. The partial eta squared of .638 suggests a strong relationship between treatment and post-test scores, controlling for the pre-test scores.

Table 5. ANCOVA summary table for comparing treatment versus control groups (coded 1 and 0, respectively) with pretest as the covariate.

Source	Sum of Squares	df	Mean Square	F	p	Partial Eta Squared
Pretest Score	74.709	1	74.709	12.607	.001	.141
Group	805.388	1	805.388	135.911	.000	.638
Error	456.291	77	5.926			

Table 6. Parameter estimates based on ANCOVA.

Parameter	*B*	SE_B	*T*	*p*	95% Confidence Interval Lower Bound	Upper Bound	Partial Eta Squared
Intercept	16.465	1.619	10.167	.000	13.240	19.689	.573
Pretest Score	.350	.099	3.551	.001	.154	.547	.141
Group=0	-6.886	.591	-11.658	.000	-8.062	-5.710	.638

Note: The unstandardized regression coefficient, B, for the row labeled "group=0" represents the control group's (lower) performance relative to the treatment group.

Table 6 shows that we are 95% confident that training course participants score at least 5.71 points and at most 8.06 points higher on the ALI than those in the control group (who do not take any assessment-related course). It is worth noting that information regarding the extent to which assessment-related topics were covered during the intervention in other pre-service courses was gathered and determined to be minimal.

5. Conclusion

Teachers should acquire sufficient knowledge of and skills related to assessment in order to collect evidence that demonstrates students' understanding or lack thereof. However, there has been very little research conducted in China to explore how to improve assessment-related curricula and develop pre- and in-service teachers' assessment literacy. The current study expands our knowledge by reporting that the Chinese pre-service primary teachers involved in the study, before the intervention, were basically assessment-illiterate as they had not been exposed to assessment-related coursework in the current college-level teacher preparation program. Moreover, it is revealed from this study that a one-semester backward-designed assessment training course can have positive impacts on pre-service primary teachers' assessment literacy levels.

Limitations of the study include the fact that the sample is one of convenience, not fully representative of Shanghai pre-service primary teachers, much less those of China, as a whole. As a result, the ability to generalize the findings beyond the participant pool to a wider target population is certainly limited to populations demonstrating the same characteristics of and working in contexts similar to those experienced by this sample.

6. Acknowledgment

7. References

[1]. C. Poth, "What assessment knowledge and skills do initial teacher education programs address? A Western Canadian perspective," Alberta Journal of Educational Research, 58(4), pp. 634-656, 2013.

[2]. C. DeLuca, T. Chavez, A. Bellara, and C. Cao, "Pedagogies for Preservice Assessment Education: Supporting Teacher Candidates' Assessment Literacy Development," The Teacher Educator, 48(2), pp. 128-142, 2013. doi:10.1080/08878730.2012.760024

[3]. Y. Jin, "The place of language testing and assessment in the professional preparation of foreign language teachers in China. Language Testing," 27(4), pp. 555-584, 2010. doi:10.1177/0265532209351431

[4]. G. Wiggins, and J. McTighe, Understanding by design, 2nd ed., Alexandria: VA: Association of Supervision and Curriculum Development, 2005.

[5]. B. Plake, J. Impara, and J. Fager, "Assessment competencies of teachers: A national survey," Educational Measurement: Issues and Practice, 12(4), pp. 10-12, 39, 1993.

[6]. C. A. Mertler, and C. S. Campbell, "Measuring teachers' knowledge and application of classroom assessment concepts: Development of the Assessment Literacy Inventory," Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada, April 2005.