Language Testing in Asia
a SpringerOpen Journal

CrossMark

Chinese university students' perceptions of assessment tasks and classroom assessment environment

Liying Cheng, Yongfei Wu[*] and Xiaoqian Liu

* Correspondence:
yongfei.wu@queensu.ca
Faculty of Education, Queen's
University, 511 Union Street,
Kingston, Ontario K7M 5R7, Canada

Abstract

Classroom assessment tasks and environment are central to supporting student learning, yet are under-studied at the tertiary level, especially in China's test-driven culture. This study explores the relationship between students' perceptions of assessment tasks and classroom assessment environment, within the university context of teaching English as a foreign language (EFL) in China. A questionnaire was designed and administered, based on Dorman & Knightley's (2006) Perceptions of Assessment Tasks Inventory (PATI) and Alkharusi's (2011) scale, in order to measure students' perceptions of the classroom assessment environment. PATI includes five subscales: congruence with planned learning, authenticity, student consultation, transparency, and diversity. Alkharusi's scale comprises two subscales: learning-oriented classroom assessment environment and performance-oriented classroom assessment environment. Participants were 620 university students from three universities in China. The factor analysis findings identified the original five-factor PATI, and Alkharusi's two-factor scales within this Chinese research context. Multiple regression analyses exploring the interrelationship showed that *congruence with planned learning, authenticity, student consultation*, and *transparency* significantly predicted the learning-oriented classroom assessment environment, explaining 48 % of the variance. *Congruence with planned learning* and *student consultation* negatively, and *diversity* positively, predicted the performance-oriented classroom assessment environment, explaining 12 % of the variance. The findings highlight the two core values in classroom assessment tasks: *congruence with planned learning* and *student consultation* in mediating the classroom environment. This study addresses the research gap in our limited understanding of the relationship between classroom assessment tasks and assessment environment, and aids teachers in structuring their day-to-day classroom assessment practices in support of their students' learning.

Keywords: Assessment tasks; Classroom assessment environment; Perceptions

Background

Assessment plays a central role in teaching and learning. It has long been used in education for various high stakes decision-making purposes, e.g., selection and placement, guiding and correcting learning, and grading achievement (Crooks 1988). The prevalence of large-scale high-stakes testing, and its impact on stakeholders, has been well documented in education. There are a set of relationships, intended and unintended, between testing, teaching and learning. However, not until recently has research on

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 2 of 17

classroom assessment begun to explore the role that assessment tasks that teachers use, and that the classroom environment plays in supporting student learning.

The classroom assessment environment (Stiggins & Conklin 1992) is an important part of the classroom atmosphere. The ways teachers communicate their expectations to students, and the ways they provide feedback on how well these expectations are being met, help students form concepts of what is important to learn and how good they are at learning it. The importance of the classroom assessment environment and the particular classroom assessment tasks within that environment signify a salient area of research to study. These include (a) student perceptions of the tasks, their interest and importance; (b) student self-efficacy to accomplish them and their reasons for doing so; and (c) the goal orientations at the task level. Without knowing the relationship between classroom assessment tasks and classroom assessment environment, especially from the students' perspective, it is not possible for us to know how teacher assessment practices are supporting student learning.

As part of the worldwide movement to combine assessment *of* learning with assessment *for* learning, in order to promote student learning, the Chinese Ministry of Education developed the College English Curriculum Requirements (CECR) (CMoE 2004; 2007) to introduce this balanced assessment concept. The stipulated College English assessment framework is taking a two-pronged approach: Firstly, the original external examination system, the College English Test (CET) testing system, which almost all undergraduate students in China are required to take, is being reformed in a substantial way (Jin 2005). Secondly, formative assessment elements are to be incorporated into the existing summative assessment framework (Wang 2007).

Teachers and students are now caught between a long examination history for selection purposes and continued large-scale national English language testing, and this recent curriculum reform promoting formative classroom assessment practices. Recent research also demonstrates that translating the CECR formative assessment initiative into classroom practices is a complex issue that involves many factors such as teachers' beliefs, students' perceptions, institutional differences, and educational tradition (Chen et al. 2013). Concern remains as to whether this worldwide ('foreign') assessment movement works within the local Chinese university context, which is highlighted as test-centered, textbook-centered, and teacher-centered (Cheng 2010). China has a long history of using large-scale testing, which still enjoys widespread acceptance and recognition as a fair measurement for selecting the best talents. Within such a highly summative assessment context, the success of the balanced assessment framework will foremost involve an informed understanding of assessment tasks and classroom assessment environment. Within this context, we conducted an empirical study to explore Chinese university students' perception of assessment tasks and their classroom assessment environment.

Nature of assessment tasks and assessment environment

A substantial proportion of classroom time involves exposing students to a variety of assessment tasks (Stiggins & Conklin 1992). As students process these tasks, they develop beliefs about their importance, utility, value, and difficulty. The characteristics of the assessment tasks, that is, how these tasks are conducted in the classroom as perceived by the students, are central to the understanding of the quality of student learning (Dorman et al. 2006; Watering et al. 2008). Since students are direct participants in

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 3 of 17

the assessment process, their perceptions of the assessment tasks are the foundation of successful formative assessment (Brookhart 2013).

Dorman and Knightley (2006) developed and validated an instrument to investigate secondary school students' perceptions of assessment tasks along five dimensions: congruence with planned learning, authenticity, student consultation in the assessment process, transparency about the purposes and forms of the assessment, and diversity. This study argues for more research identifying the perceived characteristics of assessment tasks supportive of a classroom environment that is conducive to enhanced student learning. The nature of the tasks, the environment in which they are undertaken, and how these tasks and their environment are perceived significantly affect the depth of student engagement (Fox et al. 2001; Lizzio & Wilson 2013).

The classroom environment is the overall sense or meaning that students develop from the various assessment tasks (Brookhart & DeVoge 1999). Brookhart (2004) rightly points out that each classroom has its own assessment environment perceived by the students while their teacher establishes assessment purposes, assigns assessment tasks, sets performance criteria and standards, gives feedback, and monitors students' progress. Alkharusi (2011) investigated students' perception of the classroom environment and established two facets: learning-oriented, focusing on assessment practices to enhance student learning; and performance-oriented, focusing on grading and comparing students' learning. However, despite the increasing literature on classroom assessment practices (MacMillan 2013), research on the relationship between assessment tasks and the classroom assessment environment perceived by students is still lacking.

Among the limited existing empirical studies, Alkharusi et al. (2014a) examined the inter-correlations between students' perceptions of the assessment tasks and the classroom assessment environment. Their focus, however, was on the gender difference in a research context where male and female students were educated separately, rather than coeducationally, in the Middle East. They found statistically significant gender differences concerning the perceptions of the assessment tasks and classroom assessment environment. For both male and female students, learning-oriented classroom assessment environment was congruently correlated with planned learning, authenticity, student consultation, transparency, and diversity. Performance-oriented classroom assessment environment, however, was only correlated with student consultation among females while it was associated with all assessment task variables among males.

In the same research context, Alkharusi et al. (2014b) investigated how teachers' classroom assessment practices and students' perceptions of classroom assessment tasks related to student academic self-efficacy by collecting data from 1, 457 secondary school students and 99 teachers. They found that students' perceptions of classroom assessment tasks all had significant positive influences on their academic self-efficacy. This outcome indicates that there exists a relationship between the assessment tasks and the assessment environment that supports student learning.

Wang & Cheng, 2010 explored the relationship between students' perceptions of the classroom assessment environment and their goal orientations among 503 first-year Chinese EFL university students. She found that they perceived their classroom assessment environment to be learning-oriented, test-oriented, and praise-oriented. Students' perception of the assessment environment as being learning-oriented positively predicted their adoption of mastery goals, whereas perceptions of the assessment

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 4 of 17

Table 1 Demographic information

University	Anhui	46.8 %	College Year	1		44.4 %
	Chongqing	11.3 %		2		45.9 %
	Guangdong	41.9 %		3		9.7 %
Gender	Male	48.4 %	English Level	Low		23.0 %
	Female	51.6 %		Medium		74.0 %
				High		3.0 %
Age	23–24	2.6 %	Program	Humanities and Social Sciences		14.9 %
	21–22	40.8 %		Sciences		17.8 %
	19–20	40.7 %		Engineering		31.5 %
	17–18	5.9 %		Business		35.8 %
Province	Anhui	36.3 %	Start Grade	1–2		20.6 %
	Guangdong	34.5 %		3–4		35.7 %
	Chongqing	5.5 %		5–6		13.6 %
	Other	31.7 %		7		20.1 %

environment as being test-oriented or praise-oriented positively predicted their adoption of performance goals.

Despite the fact that China has the largest number of students learning English, studies of classroom assessment practices within this context are limited. Cheng et al. (2004) conducted two of the first studies of classroom assessment practices within the Chinese university context (see also Cheng et al. 2008), but these studies were undertaken to examine these practices from the perspectives of teachers. Furthermore, such studies should be conducted within the macro-societal context, as well as the micro-instructional assessment context, so as to better understand the factors that influence students' perceptions of classroom assessment practices. Based on the previous research studies and the Chinese university context, this study answered the following research questions:

1. How do students perceive their classroom assessment tasks and environment?
2. What is the relationship, if any, between the assessment tasks and the classroom assessment environment?
3. Are there significant differences in students' perceptions of assessment tasks and classroom assessment environment by university, subject major and self-perceived language proficiency?

Method

Participants

Participants were students from three Chinese universities respectively located in the provinces of Anhui (46.8 % or n = 296), Chongqing (11.3 % or n = 67), and Guangdong (41.9 % or n = 257), as outlined in Table 1. Female students (51.6 % or n = 315) slightly outnumbered male students (48.4 % or n = 295) and the majority of the participants were between 19 and 22 years of age. Most of them came from the provinces of Anhui (36.3 % or n = 221), Guangdong (34.5 % or n = 210) and Chongqing (5.5 % or n = 34). The majority of the participants (79.9 %) started learning English in primary school, typically in Grade 3–4, while a fifth of them began their English studies in

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 5 of 17

Grade 7 in junior high school. Respectively, they were enrolled in humanities and social sciences (14.9 % or n = 91), sciences (17.8 % or n = 109), engineering (31.5 % or n = 192), and business (35.8 % or n = 218). Close to half (44.4 % or n = 271) were in their first year of university with another 45.9 % (or n = 280) in their second year, and 9.7 % (or n = 59) in their third. The vast majority of the participants rated their own English proficiency as medium level (74 % or n = 451), with about a quarter of them rating themselves as low level (23 % or n = 140), and a very small remainder as high level (3 % or n = 18).

Instrument

The questionnaire used in this study is designed based on Dorman & Knightley's (2006) perceptions of assessment tasks inventory (PATI) (35 items), and Alkharusi's (2011) assessment environment (16 items). We chose these two instruments based on their theoretical grounding and the psychometric quality. The Cronbach alpha of PATI is ranging from 0.85 to 0.63 (Dorman & Knightley 2006, p. 54), and .82 and .75 respectively for learning-oriented and performance-oriented classroom assessment environment (Alkharusi 2011). These 51 items constitute the two major sections of the questionnaire. Section 1 on PATI consists of 5 scales: congruence with planned learning, i.e., the extent to which assessment tasks align with the goals, objectives, and activities of the learning program (items 1–7); authenticity, i.e., the extent to which assessment tasks feature real-life situations (items 8–14); student consultation, i.e., the extent to which students are consulted and informed about the forms of assessment tasks being employed (items 15–21); transparency, i.e., the extent to which the purposes and forms of assessment tasks are well-defined and clear to the learner (items 22–28); and diversity, i.e., the extent to which all students have an equal chance at completing assessment tasks (items 29–35).

Section 2 Classroom Assessment Environment consists of two scales: learning-oriented assessment environment, focusing on classroom assessment practices that improve student learning and mastery of content materials (items 1–9); and performance-oriented assessment environment, focusing on harshness of assessment, grading, public evaluation and recognition practices (items 10–16). The third section of the questionnaire consists of seven demographic items to collect information on the participants' gender, age, home province, years of learning English, years at university, major area of study, and perceived level of English proficiency. These items provide the participants' contextual information and their English learning background.

Data collection

The questionnaire was pre-tested on five Chinese students studying in a Canadian university. As a result, the directions were made clearer and ambiguous items were revised. The questionnaire was then translated into Chinese by an experienced translator and double-checked by a researcher proficient in both English and Chinese. In addition, pilot tests were conducted with a small group of Chinese students (n = 4) whose characteristics were similar to the survey participants. Consequently, three items were revised to improve clarity and idiomatic expressions to Chinese participants. Questionnaires were then administered, in the summer of 2013, to 652 students from a comprehensive university in Anhui (East China), a foreign language studies university in Guangdong (South China), and a polytechnic university in Chongqing (Southwest China). Altogether, 637 students

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 6 of 17

Table 2 Perceptions of classroom assessment tasks

Item	N	M	SD	Item	N	M	SD
1	610	3.50	.94	19	610	2.11	1.02
2	610	3.34	1.02	20	610	3.25	1.08
3	610	4.17	.89	21	610	3.95	.94
4	610	3.44	.94	22	610	3.40	1.02
5	610	3.51	.84	23	610	3.43	1.00
6	610	3.76	.89	24	610	3.22	1.13
7	610	3.93	.88	25	610	3.41	1.06
8	610	3.54	1.01	26	610	3.65	1.04
9	610	3.49	.96	27	610	3.55	1.04
10	610	3.38	.92	28	610	3.69	.91
11	610	3.25	.98	29	610	3.44	.98
12	610	3.72	.86	30	610	3.26	1.02
13	610	3.78	.85	31	610	2.83	1.03
14	610	3.78	.85	32	610	2.33	1.04
15	610	3.28	.98	33	610	3.23	.92
16	610	3.26	1.01	34	610	2.45	1.03
17	610	2.90	1.04	35	610	2.31	1.07
18	610	2.62	1.06				

responded to the questionnaire – a return rate of 97.7 %. The data were entered into SPSS 20 independently for each university, and then merged into one data set.

Data analysis

The data were first checked for missing values. Cases were deleted if more than 10 % of the questionnaire items were not completed, which resulted in 10 cases being removed, and 610 cases retained for subsequent analysis. The remaining missing values were construed to have been omitted at random and were replaced by means. Data were then analyzed in four phases: First, means and standard deviations were calculated with regard to participants' perceptions of the classroom assessment tasks and the classroom assessment environment, in order to answer the first research question.

Second, exploratory factor analysis was run to detect the latent factor structures from Dorman & Knightley's (2006) PATI and Alkharusi's (2011) assessment environment scale. Principal component analysis was employed to extract the factors primarily because this method was used in the original validation processes for both instruments, and is also widely used in the language assessment literature (Ockey 2014). To improve the interpretability of the factor extraction results, direct oblimin was selected since oblique rotation allows the factors to be correlated, which is most often the case in social sciences (Bentler 2008). This choice was also supported by previous studies showing that all subscales of the PATI are moderately or highly correlated (Dorman & Knightley 2006), and learning-oriented assessment environment and performance-oriented assessment environment are moderately correlated among both male and female students (Alkharusi et al. 2014a). To determine the items retained for further analysis, this study adopted Tabachnick and Fidell's (2001) suggestion that .32 be a good rule of thumb for the lowest factor loadings.

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 7 of 17

Table 3 Perceptions of classroom assessment environment

| Item	N	M	SD	Item	N	M	SD
1	610	3.56	.90	9	610	3.66	1.04
2	610	3.77	.90	10	610	3.36	.94
3	610	3.88	.92	11	610	2.96	1.03
4	610	3.38	.99	12	610	2.48	1.04
5	610	3.84	.86	13	610	2.38	1.01
6	610	3.41	1.03	14	610	2.94	1.13
7	610	4.20	.76	15	610	2.89	1.13
8	610	4.14	.82	16	610	2.95	1.04

Third, stepwise multiple regressions were conducted to examine the relationship between the perceived classroom assessment tasks and the perceived classroom assessment environment. This analysis, together with the above factor analysis, was undertaken to answer the second research question. Students' perceptions of the classroom assessment tasks were used as independent variables (predictors) and their perceptions of the classroom assessment environment were treated as dependent variables.

Finally, one way ANOVAs and t-tests were conducted to investigate the differences of students' perceptions of assessment tasks and classroom assessment environment among the three participating universities, the four subject majors included, and students with differing levels of self-perceived language proficiency – the third research question.

Results

Descriptive

Students' perceptions of the 35 items on assessment tasks and the 16 items on classroom assessment environment were put into descriptive analysis. In Table 2, the greatest mean is item 3 ("My assignments are related to what I am learning in English") ($M = 4.17$), followed by the next two largest means: items 21 ($M = 3.95$) and 7 ($M = 3.93$). Item 21 states "I ask my teacher about English assessment", and item 7 asserts "I have answered English questions on topics that have been covered in class". The smallest mean among

Fig. 1 Scree Plot of Principal Component Analysis (Perceptions of Assessment Task Inventory)
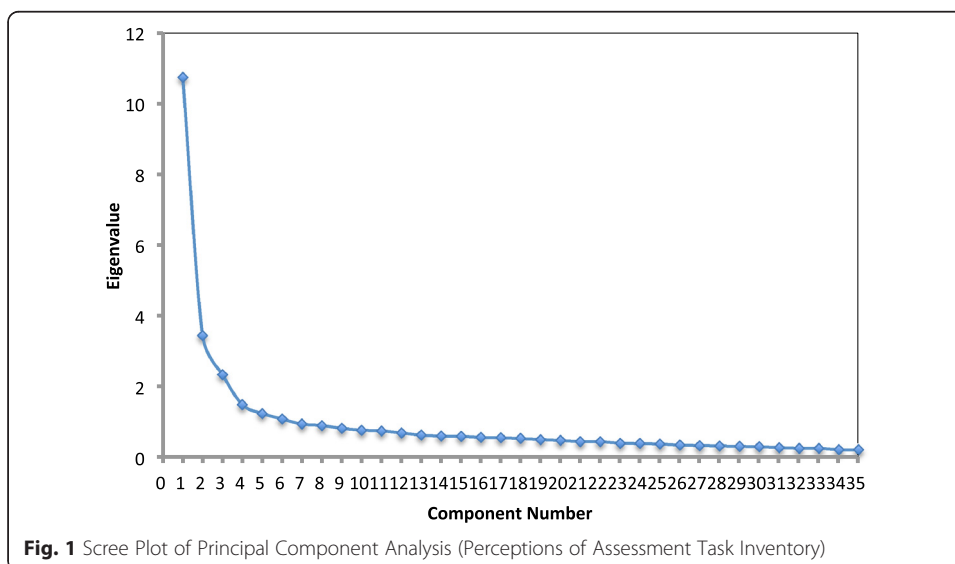
the assessment tasks is item 19 ("I have helped the class develop rules for assessment in English") ($M = 2.11$), followed by the next two smallest means: items 35 ($M = 2.31$) and 32 ($M = 2.33$). Item 35 says "I do work that is different from other students' work", while item 32 declares "I am set assessment tasks that are different from other students' tasks". The means of seven items (17, 18, 19, 31, 32, 34, 35) are below 3. These fall into Dorman and Knightley's (2006) original scales of student consultation (17, 18, 19) and diversity (31, 32, 34, 35).

The range of the standard deviation (SD) of all the items is between .84 and 1.13. The greatest SD is item 24 ("I know in advance how I will be assessed") ($SD = 1.13$), followed by another two relatively large SDs: items 20 ($SD = 1.08$) and 35 ($SD = 1.07$). Item 20 states "My teacher has explained to me how each form of assessment is used", and item 35 asserts "I do work that is different from other students' work". The smallest SD is item 5 ("I am assessed in similar ways to the tasks I do in class") ($SD = .84$), followed by the next three smallest SDs: items 13 ($SD = .85$), 14 ($SD = .85$) and 12 ($SD = .86$). Item 13 says "Assessment in English tests my ability to apply learning", item 14 states "Assessment in English examines my ability to answer important questions", and item 12 affirms "English assessment tasks check my understanding of topics". These items (12, 13, 14) are in the scale of authenticity. In addition, many items' SDs in the scale of student consultation (5 out of 7), transparency (6 out of 7), and diversity (5 out of 7) are above 1.00, which indicates that students' perspectives on these items are more varied.

In Table 3, generally the means of items in the scale of learning-oriented assessment environment are greater ($M > 3.00$) than those in performance-oriented assessment environment. The greatest mean is item 7 ("Our teacher holds us the responsibility to learn") ($M = 4.20$), followed by item 8 ("Our teacher uses a variety of ways to assess our mastery of English") ($M = 4.14$). The smallest mean is item 13 ("There is a mismatch between the learned subject materials and the assigned homework and tests") ($M = 2.38$), followed by item 12 ("Our teacher gives more importance to the grades than to the learning") ($M = 2.48$). This indicates that students showed more agreement on the items in the scale of learning-oriented assessment environment than that of the performance-oriented assessment environment. In addition, there are greater SDs of

Fig. 2 Scree Plot of Principal Component Analysis (Perceptions of Classroom Assessment Environment)
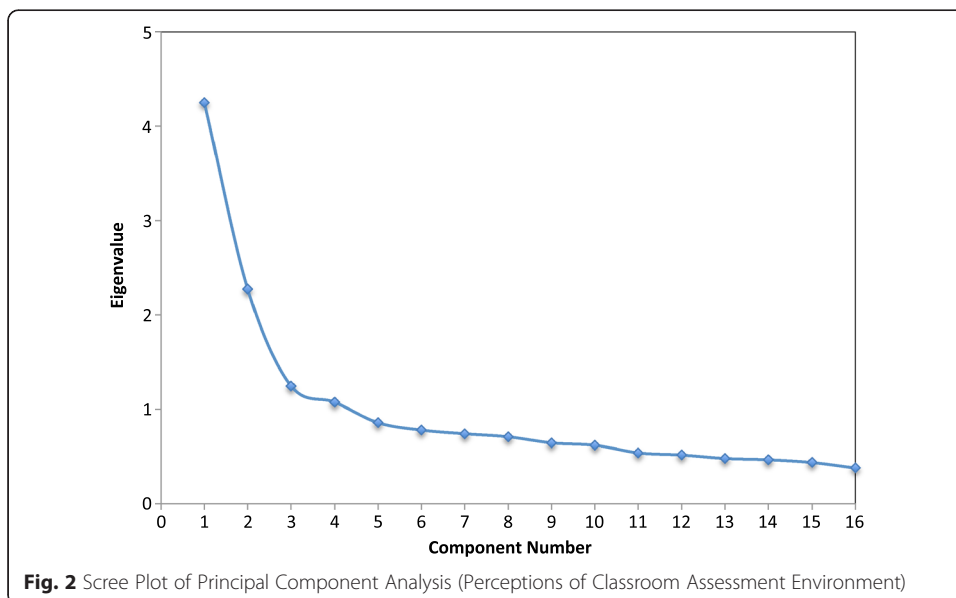
items in the scale of performance-oriented assessment environment (e.g., 6 out 7 items' $SD > 1.00$) than those in learning-oriented assessment environment (7 out of 9 items' $SD < 1.00$). This implies that students' perceptions were more varied on items in the scale of performance-oriented assessment environment than those in learning-oriented assessment environment.

Table 4 Principal component analysis with direct oblimin rotation of perceptions of assessment task inventory

	Item	CPL	D	T	SC	A
T1	My assessment in English is a fair indicator of my work	.69				
T2	My English tests are a fair indicator of what my class is trying to learn.	.68				
T4	My assessment is a fair indication of what I do in English	.66				
T5	I am assessed in similar ways to the tasks I do in class.	.65				
T6	I am assessed on what the teacher has taught me.	.62				
T7	I have answered English questions on topics that have been covered in class.	.53				
T3	My assignments are related to what I am learning in English.	.46				
T32	I am set assessment tasks that are different from other students tasks.		.82			
T34	I use different assessment methods from other students.		.82			
T35	I do work that is different from other students work.		.82			
T19	I have helped the class develop rules for assessment in English.		.68			
T31	I am given a choice of assessment tasks.		.64			
T18	I select how I will be assessed in English.		.59			
T33	I am given assessment tasks that suit my ability.		.34			
T26	I am told in advance when I am being assessed.			.79		
T27	I am told in advance on what I am being assessed.			.75		
T25	I am told in advance why I am being assessed.			.72		
T24	I know in advance how I will be assessed.			.65		
T28	I understand the purpose of English assessment.			.47		
T30	When I am faster than others, I move on to new assessment tasks.			.43		
T29	I complete assessment tasks at my own speed.			.42		
T16	I am clear about the forms of assessment being used.				-.76	
T15	I am aware of the types of assessment in English.				-.69	
T20	My teacher has explained to me how each form of assessment is used.				-.63	
T22	I understand what is needed in all English assessment tasks.				-.58	
T23	I know what is needed to successfully accomplish an English assessment task.				-.49	
T17	I am asked about the types of assessment I would like to have in English.				-.48	
T21	I ask my teacher about English assessment.				-.32	
T8	I am asked to apply my learning to real-life situations.					-.79
T9	My English assessment tasks are meaningful.					-.68
T11	I find English assessment tasks relevant to the real world.					-.65
T13	Assessment in English tests my ability to apply learning.					-.62
T10	My English assessment tasks are useful.					-.60
T12	English assessment tasks check my understanding of topics.					-.58
T14	Assessment in English examines my ability to answer important questions.					-.45
	Cronbach's Alpha	.82	.86	.82	.84	.86

CPL Congruence with planned learning, *D* Diversity, *T* Transparency, *SC* Student Consultation, *A* Authenticity

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 10 of 17

Exploratory factor analysis

The scree plots in Figs. 1 and 2 show that the exploratory factor analyses identified the original five-factor PATI and Alkharusi's two-factor scales. The five factors of PATI, namely, congruence with planned learning, authenticity, diversity, transparency, and student consultation, cumulatively explained 55 % of the total variance of classroom assessment tasks. The two factors of Alkharusi's (2011) scale, learning-oriented assessment environment and performance-oriented assessment environment, cumulatively explained 41 % of the total variance for the classroom assessment environment. As shown in Table 4, the five subscales respectively consist of seven items in this study as of the original PATI. Items 1–7 loaded on congruence with planned learning and items 8–14 loaded on authenticity. But item 18 (selecting English assessment method) and item 19 (helping develop the English assessment rules) loaded on diversity rather than student consultation. Likewise, item 22 (understanding what English assessment tasks entail) and item 23 (understanding how to accomplish English assessment tasks successfully) loaded on student consultation instead of transparency. Item 29 (completing assessment tasks at one's own speed) and item 30 (moving onto new assessment tasks when completing earlier than others) loaded on transparency, not diversity. Table 5 indicates that the subscale learning-oriented assessment environment also comprised items 1–9, and the subscale performance-oriented assessment environment comprised items 10–16, as of Alkharusi's (2011) assessment environment scale.

The internal consistency of all PATI subscales reached acceptable level, ranging from .82 to .86 (see Table 4), according to George and Mallery (2003) who anchored the internal consistency in the following way: excellent ($\alpha \geq .9$), acceptable ($.9 > \alpha \geq .8$), good ($.8 > \alpha \geq .7$), questionable ($.7 > \alpha \geq .6$), poor ($.6 > \alpha \geq .5$), and

Table 5 Principal component analysis with direct oblimin rotation of perceived classroom assessment environment

Item		LAE	PAE
E5	I am given a chance to correct my mistakes.	.74	
E4	I receive continuous feedback from my teacher about my performance in English.	.73	
E2	Our teacher helps us identify the places where we need more effort in future.	.71	
E3	The assignments and tests encourage thinking.	.67	
E8	Our teacher uses a variety of ways (e.g., tests, in-class tasks, homework assignments) to assess our mastery of English.	.65	
E6	The assignments and activities are related to our every day lives.	.65	
E7	Our teacher holds us the responsibility to learn.	.61	
E1	I can find out my strengths in English.	.54	
E9	The assignments and tests are returned in a way that keeps our scores private.	.43	
E12	Our teacher gives more importance to the grades than to the learning.		.71
E13	There is a mismatch between the learned subject materials and the assigned homework and tests.		.68
E15	Our teacher's grading system is not clear.		.62
E16	The assessment results do not fairly reflect the effort put in us studying the subject.		.59
E11	Our teacher compares our performances to each other.		.58
E14	The in-class and homework assignments are not interesting.		.56
E10	The tests and assignments are difficult to us.		.36
Cronbach's Alpha		.82	.70

LAE Learning-oriented Assessment Environment, *PAE* Performance-oriented Assessment Environment

unacceptable (.5 > α). The reliability coefficients of the two assessment environment subscales were respectively .82 for learning-oriented assessment environment and .70 for performance-oriented assessment environment (see Table 5). These exploratory factor analysis results were used for the following regression analysis.

Regressions

Based on the results of factor analysis, two separate stepwise regression analyses were conducted to investigate the relationship between students' perceptions of assessment tasks and classroom assessment environment. The five assessment task factors were used as independent variables and the two classroom assessment environment factors as dependent variables. Stepwise regression was employed because, in the current context, there was no strong theory or empirical evidence that could be used to decide the order in which the variables were to be entered into the equation. Tests of normality were first applied to all the independent and dependent variables, to check if it was appropriate to employ regression analysis. The test results showed that all the variables' skewness and kurtosis were less than +3 and greater than −3, which meant that they were fairly normally distributed (Tabachnick & Fidell 1989). Correlation analyses were then conducted on the five independent and two dependent variables (see Table 6). The five independent variables of assessment tasks were moderately correlated with each other (.68 ≥ r ≥ .32) and the two dependent variables were weakly yet negatively correlated with each other (r = −.19). All independent variables were moderately correlated with the dependent variable learning-oriented assessment environment (.59 ≥ r ≥ .34). All independent variables of assessment tasks were weakly correlated with performance-oriented assessment environment. However, the correlation between diversity and performance-oriented assessment environment was positive (r = .10) whereas the correlations between the other four independent variables and performance-oriented assessment environment was negative (−.11 ≥ r ≥ −.28).

The regression results are presented in Tables 7 and 8. Table 7 used learning-oriented assessment environment as the dependent variable, with diversity removed from the equation model. The other four factors were positive significant predictors of the learning-oriented assessment environment, explaining 48 % of the variance with their beta weights and significance being authenticity (β = .39, p < .001), congruence with planned learning (β = .23, p < .001), transparency (β = .10, p < .05), and student consultation (β = .09, p < .05). Table 8 used performance-oriented assessment environment as the dependent variable, and authenticity and transparency were removed from the equation model. The other three factors were significant predictors of the performance-oriented assessment environment, explaining 12 % of the variance with their beta weights and significance being

Table 6 Correlations between classroom assessment tasks and environment

Variables	1	2	3	4	5	6	7
1. Congruence with Planned Learning	1						
2. Authenticity	.68[**]	1					
3. Diversity	.32[**]	.41[**]	1				
4. Transparency	.50[**]	.50[**]	.33[**]	1			
5. Student Consultation	.54[**]	.56[**]	.34[**]	.64[**]	1		
6. Learning-oriented Assessment Environment	.59[**]	.65[**]	.34[**]	.47[**]	.50[**]	1	
7. Performance-oriented Assessment Environment	-.28[**]	-.20[**]	.10[*]	-.11[**]	-.18[**]	-.19[**]	1

Table 7 Stepwise regression analysis results (learning-oriented assessment environment as dependent variable)

Variables	β	t	Sig.	Adj. R^2	ΔR^2	95.0 % CI
(Constant)		11.28	.000			[1.03, 1.46]
Authenticity	.39	9.28	.000			[.27, .41]
Congruence with planned learning	.23	5.47	.000			[.14, .29]
Transparency	.10	2.57	.010			[.02, .15]
Student consultation	.09	2.22	.027			[.01, .14]
				.48	.00	

$P < .05$

congruence with planned learning ($\beta = -.30$, $p < .001$), diversity ($\beta = .23$, $p < .001$), and student consultation ($\beta = -.10$, $p < .05$). This result implied that congruence with planned learning and student consultation in conducting assessment tasks were negatively contributing to the formation of the performance-oriented assessment environment. The diversity of assessment tasks, on the other hand, contributed positively to the formation of the performance-oriented assessment environment.

Not surprisingly, while diversity was a significant predictor of performance-oriented assessment environment, it was removed from the equation when learning-oriented assessment environment was used as the dependent variable. Similarly, while authenticity and transparency were significant in predicting learning-oriented assessment environment, they were removed when using performance-oriented assessment environment as the dependent variable. In addition, the co-occurrence of congruence with planned learning and student consultation showed that these two factors positively predicted learning-oriented assessment environment, and negatively predicted performance-oriented assessment environment.

One-way ANOVA and T-tests

In order to examine whether there was significant difference among students between the three universities in this study, in terms of the five assessment tasks scales and the two assessment environment scales, seven one-way ANOVA tests were to be applied to each (five for PATI and one each for the learning- and performance- assessment environment). Before the ANOVA tests were conducted, we put the variables into Levene tests. The results showed that two of the seven variables did not meet the assumptions required for further ANOVA tests: *transparency* and *diversity*. Therefore, they were excluded from further analysis. The ANOVA results showed that significant differences existed between the university in Anhui and the university in

Table 8 Stepwise regression analysis results (performance-oriented assessment environment as dependent variable)

Variables	β	t	Sig.	Adj. R^2	ΔR^2	95.0 % CI
(Constant)		25.04	.000			[3.42, 4.01]
Congruence with planned learning	-.30	−6.42	.000			[−.38, −.20]
Diversity	.23	5.67	.000			[.13, .26]
Student consultation	-.10	−2.22	.027			[−.17, −.01]
				.12	.01	

$P < .05$

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 13 of 17

Table 9 Post hoc investigation on the differences between universities

Variables	University		Mean diff.	Sig.	95 % CI
Congruence with planned learning	Anhui	Guangdong	.21	.00	[.08, .34]
Authenticity			.16	.02	[.02, .31]
Learning-oriented environment			.18	.00	[.05, .30]

P < .05

Guangdong in terms of *planned learning* (F(2607) = 7.76, *p* < .01), *authenticity* (F(2607) = 5.08, *p* < .01), and *learning-oriented environment* (F(2607) = 6.50, *p* < .01). The Scheffe post hoc tests showed that each of these three aspects was higher for the university in Anhui than for the university in Guangdong (see Table 9).

Next, in order to examine whether there was significant difference between majors, in terms of the five assessment tasks scales and the two assessment environment scales, seven one-way ANOVA tests were again performed. Levene tests showed that the diversity variable violated the assumption for ANOVA tests, and therefore we excluded it in the further analysis. The ANOVA results showed that significant differences existed between the four different majors in terms of *student consultation* ($F(3593) = 4.80$, $p < 0.01$) and *transparency* ($F(3593) = 6.23$, $p < 0.01$). The Scheffe post hoc tests showed that each of the two variables for humanities/social sciences, engineering, and business was greater than those for sciences (see Table 10).

Finally, we compared students with different levels of self-perceived language proficiency to determine if there was significant variation in their perceptions of classroom assessment tasks and environment. We compared two groups – students with low and medium language proficiency – through a *T*-test. We excluded the group of students with high language proficiency as the sample size was very small (3 %). Levene tests showed that all the variables met the assumptions. *T*-test results showed that students with medium language proficiency perceived *transparency* in the classroom assessment tasks significantly higher than students with low language proficiency did [$t(579) = 2.79$, $p < .01$].

Discussions and conclusions

This study addresses the research gap resulting from the lack of empirical research on student perceptions of their classroom assessment tasks and environment, and the relationship between these tasks and environment as it relates to student learning within the context of Chinese EFL tertiary education. A number of findings, as reported above,

Table 10 Post hoc investigation on the differences between majors

Variable	Majors		Mean diff.	Sig.	95 % CI
Student consultation	Humanities and social sciences	Sciences	.32	.02	[.03, .61]
	Engineering		.25	.04	[.01, .49]
	Business and management		.29	.01	[.06, .53]
Transparency	Humanities and social sciences	Sciences	.32	.02	[.04, .61]
	Engineering		.27	.02	[.03, .51]
	Business and management		.35	.00	[.12, .58]

P < .05

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 14 of 17

have enhanced our understanding of the nature of classroom assessment within this context, and identified areas for further research.

Nature of assessment tasks and environment

Our results from the descriptive analysis showed that the participants of this study perceived their assessment tasks as being highly related to what they were learning in their English classes. This demonstrates a match between assessment and learning from these students' viewpoint. However, these students were seldom involved in developing the criteria for assessment, which has been empirically supported by previous studies (Cheng et al. 2004; Wang et al. 2013). Involving students in developing the criteria for assessment is a step further in supporting their learning. The process of setting goals can clarify the process in reaching the learning goals. In terms of the classroom assessment environment, students strongly perceived that their teachers held them responsible for learning. This result has also been supported through the discussion of the role that teachers play in student learning in and around the Asian context (Carless 2011). Although these students did not identify a mismatch between what they learned and what was assessed, they stated that their assessment results did not fairly reflect the effort they put in. This outcome echoed what we found regarding the assessment tasks, and pointed to an aspect of assessment environment that needs further research. We need to know how efforts are included in the assessment results within this context and listen to students' voices regarding assessment of their learning. After all, the involvement of student learning including perceptions of their learning is an indicator of quality classroom assessment.

Relationship between assessment tasks and environment

In order to explore the relationship between assessment tasks and classroom environment, exploratory factor analyses were conducted and identified the original five-factor of PATI and Alkharusi's two-factor scales on learning- and performance-oriented environment. This demonstrates the robustness of the instruments in another research context and at different level of learning (secondary school vs. higher education), especially for the PATI scale on assessment tasks.

Multiple regression analyses showed that congruence with planned learning, authenticity, student consultation, and transparency significantly predicted the learning-oriented classroom assessment environment, explaining 48 % of the variance. Congruence with planned learning and student consultation negatively, while diversity positively, predicted the performance-oriented classroom assessment environment, explaining 12 % of the variance. This analysis highlights the two core values in classroom assessment tasks: congruence with planned learning and student consultation in mediating the classroom environment. These two assessment task variables had a medium correlation with each other. These two core values have been discussed in many educational assessment literatures (e.g., Barksdale-Ladd & Thomas 2000; Stiggins & Conklin 1992).

Our findings show that congruence with planned learning and student consultation are positive predictors of the learning-oriented classroom assessment environment and negative predictors of performance-oriented classroom assessment environment at the same time. This implies that aligning assessment tasks with the goals and objectives of the learning programs and effectively informing students regarding how

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 15 of 17

they will be assessed potentially have twofold benefits. On the one hand, they may contribute to an environment where students focus on learning and mastery. On the other hand, they may potentially prevent the development of an environment where students compare themselves against each other. Furthermore, our findings also show that assigning authentic assessment tasks and clarifying the purposes and forms of these tasks may also help foster a learning-oriented assessment environment in the classroom.

Interestingly, diversity is the only predictor (among the above three statistically significant predictors) that had a positive relationship with performance-orientated assessment environment. Diversity had a small correlation with the other four assessment tasks (from .32 to .41), while each of these had medium correlations with one another (from .50 to .68). This relationship might have something to do with the results from the second regression model. This finding challenges the existing research literature, which concludes that the extent to which all students have an equal chance at completing assessment tasks, helps create a fair environment (Taylor & Nolen 2008), yet it is not clear as to whether this supports students' attention on mastery or performance.

This finding may be explained by the context in which this research was conducted. Due to the nature of the Chinese examination-centered educational system, students in China tend to put an emphasis on scores and rankings (Guo 2012; Kirkpatrick & Zang 2011). It is possible that diversity in classroom assessment, such as the extent to which all students have an equal chance at completing assessment tasks, may have potentially intensified their performance-oriented assessment environment. When students are given an equal chance at completing assessment tasks, they might tend to be more competitive, which would contribute to the performance-oriented assessment environment. Future research is needed to arrive at a better interpretation on how and why diversity is associated with performance-oriented classroom assessment environment in the Chinese educational context.

The results derived from students' perceptions of assessment tasks and classroom assessment environment among the three participating universities, the four subject majors, and students with differing levels of self-perceived language proficiency demonstrated that significant differences exist in terms of the universities, the subject majors and how students perceived their own English language proficiency. Again congruence with planning learning and student consultation (along with transparency) were the variables signifying the differences. We are not able to explain the reasons between the differences, but it seems clear to us that we need to conduct further research into the macro-societal context in addition to the micro-instructional/university assessment context. We recognize that this study is an initial exploratory effort into a restricted sample of Chinese tertiary students from three universities. Many of the findings provide new insights into the nature of classroom assessment, yet at the same time point out aspects of this environment that cannot be tapped into through a survey study of this nature. A follow-up study using a qualitative research approach could enhance our understanding of classroom assessment within the Chinese EFL tertiary context.

The research findings have pedagogical implications for teachers structuring their day-to-day classroom assessment practices. To create a learning-oriented assessment environment where students focus on mastery learning, teachers are expected to align the classroom assessment tasks with the learning goals of the program, maximize the transferability

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 16 of 17

of the knowledge and skills assessed in the tasks to real-life situations, and clarify the assessment purposes and forms before the assessment tasks are assigned to students. In addition, teachers are encouraged to consult with students about what forms of assessment tasks will be used. However, teachers need be cautious when giving students equal opportunities to complete these tasks at various speeds. Efforts need to be made so as to minimize the possibility of creating an environment where students focus on mutual comparison and higher grades rather than self-improvement and meaningful learning.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors read and approved the final manuscript.

References
Alkharusi, H. (2011). Development and datametric properties of a scale measuring students' perceptions of the classroom assessment environment. *International Journal of Instruction, 4*(1), 105–120.
Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2014a). Modeling the Relationship between perceptions of assessment tasks and classroom assessment environment as a function of gender. *The Asia-Pacific Educational Researcher, 23*(1), 93–104.
Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2014b). Classroom assessment: Teaching practices, student perceptions, and academic self-efficacy beliefs. *Social Behavior and Personality, 42*(5), 835–856.
Barksdale-Ladd, M.A, & Thomas, K.F. (2000). What's at stake in high-stakes testing: teachers and parents speak out. *Journal of Teacher Education, 51*, 384–397.
Bentler, P.M. (2008). *EQS program manual*. Encino, CA: Multivariate Analysis.
Brookhart, S.M. (2004). Classroom assessment: Tensions and intersections in theory and practice. *Teachers College Record, 106*(3), 429–458.
Brookhart, S.M. (2013). Classroom assessment in the context of motivation theory and research. In JH MacMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 35–54). Los Angeles: Sage.
Brookhart, S.M., & DeVoge, J.G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education, 12*, 409–425.
Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. New York: Routledge.
Chen Q., May L., Klenowski V., Kettle M. (2013). The enactment of formative assessment in English language classrooms in two Chinese universities: Teacher and student responses. *Assessment in Education: Principles, Policy & Practice*. doi:10.1080/0969594X.2013.790308
Cheng, L. (2010). The history of examinations: Why, how, what, whom to select? In L Cheng & A Curtis (Eds.), *English language assessment and the Chinese learner* (pp. 13–26).). New York: Routledge: Taylor's & Francis Group.
Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods and procedures. *Language Testing, 21*(3), 360–389.
Cheng, L., Rogers, T., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education, 33*(1), 9–32.
CMoE document. (2004). College English Curriculum Requirements (trial), retrieved from http://www.edu.cn/20040120/3097997.shtml
CMoE document. (2007). College English Curriculum Requirements, retrieved from http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/moe_1846/201011/xxgk_110825.html
Crooks, T.J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438–481.
Dorman, J.P., & Knightley, W.M. (2006). Development and validation of an instrument to assess secondary school students' perceptions of assessment tasks. *Educational Studies, 32*, 47–58.
Dorman, J.P., Fisher, D.L., & Waldrip, B.G. (2006). Classroom environment, students' perceptions of assessment, academic efficacy and attitude to science: A lisrel analysis. In D Fisher & MS Khine (Eds.), *Contemporary approaches to research on learning environment: Worldviews* (pp. 1–28). Australia: World Scientific Publishing.
Fox, R.A., McManus, I.C., & Winder, B.C. (2001). The shortened study process questionnaire: An investigation of its structure and longitudinal stability using confirmatory factor analysis. *British Journal of Educational Psychology, 71*, 511–530.
George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference 11.0 update* (4th ed.). Boston: Allyn & Bacon.
Guo, L. (2012). New curriculum reform in China and its impact on teachers. *Canadian and International Education, 41*(2), 87–105.
Jin, Y. (2005). CET4/6 reform framework and prospect. *China College Teaching (Zhongguo Daxue Jiaoxue), 5*, 49–53.
Kirkpatrick, R., & Zang, Y. (2011). The negative influences of exam-oriented education on Chinese high school students: Backwash from classroom to child. *Language Testing in Asia, 1*(3), 37–45.
Lizzio, A., & Wilson, K. (2013). First-year students' appraisal of assessment tasks: implications for efficacy, engagement and performance. *Assessment & Evaluation in Higher Education, 38*(4), 389–406.

Cheng *et al. Language Testing in Asia* (2015) 5:13

Page 17 of 17

MacMillan, J.H. (2013). *SAGE handbook of research on classroom assessment*. Los Angeles: Sage.

Ockey, G.J. (2014). Exploratory factor analysis and structural equation modeling. In AJ Kunnan (Ed.), *The companion to language assessment* (pp. 1–21). Chichester, West Sussex: John Wiley & Sons, Inc. doi:10.1002/9781118411360wbcla114.

Stiggins, R.J., & Conklin, N.F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany: State University of New York Press.

Tabachnick, B.G., & Fidell, L.S. (1989). *Using multivariate statistics* (2nd ed.). New York, NY: HarperCollins College Publishers.

Tabachnick, B.G., & Fidell, L.S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.

Taylor, C.S, & Nolen, S.B. (2008). *Classroom assessment: Supporting teaching and learning in real classrooms*. Upper Saddle River, NJ: Pearson.

Wang, X., & Cheng, L. (2010). Chinese EFL students' perceptions of the classroom assessment environment and their goal orientations. In L. Cheng & A. Curtis (Eds.), English language assessment and the Chinese learner (pp. 202–218). New York, NY: Routledge.

Wang, Q. (2007). The national curriculum changes and their effects on English language teaching in the People's Republic of China. In J Cummins & C Davison (Eds.), *International handbook of English language teaching* (Vol. 1, pp. 87–105). New York: Springer.

Wang, W., Zeng, Y., He, H. (2013). Students' perceptions of the effects of rubricreferenced self-assessment on EFL writing: A developmental perspective. Paper presented at 35th Language Testing Research Colloquium. Seoul, South Korea, retrieved from http://www.ltrc2013.or.kr/download/LTRC2013Program0729.pdf

Watering, G., Gijbels, D., Dochy, F., & Rijt, J. (2008). Students' assessment preferences, perceptions of assessment and their relationships to study results. *Higher Education, 56*, 645–658.