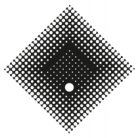

Comparison of Chinese and Western English language proficiency measures in transnational business degrees

Farley, Alan, Yang, Hong (helen), Min, L and Ma, S (2019) Comparison of Chinese and Western English language proficiency measures in transnational business degrees. Language, Culture and Curriculum. pp. 1-16. ISSN 0790-8318

Comparison of Chinese and Western English language proficiency measures in transnational business degrees

Alan Farley

Victoria University, Melbourne, Australia

La Trobe University, Melbourne, Australia

ORCID: 0000-0002-1835-3029

Helen Hong Yang*

La Trobe University, Melbourne, Australia

Email: helen.yang@latrobe.edu.au

ORCID: 0000-0003-1130-6825

* *Corresponding Author*

Lee Le Min

Liaoning University, China

Shuang (Diana) Ma

Liaoning University, China

Acknowledgement: This work was supported by the *Australian Government Office of Learning and Teaching under OLT grants (Project Number: SD15-5138)*

Note: This is the accepted version (post-print) manuscript published online on 17 June 2019 in Language, Culture and Curriculum. Doi: https://doi.org/10.1080/07908318.2019.1630423

Comparison of Chinese and Western English language proficiency measures in transnational business degrees

Abstract

This is the first empirical study to examine the comparative predictive power of Chinese national college English language program (CEP), a Western university's English for academic purpose (EAP) language program, a standardised English language proficiency (SELP) test, and business diploma for the academic performance of Chinese students in transnational business degree programs. The study analysed the academic results of 759 Chinese students who studied in the programs between 2006 and 2014. Crossover repeated measures design and multivariate regressions were used for analysis. Findings revealed that both CEP and EAP were good predictors of the performance in the first-year pathway diploma program. However, neither of them added any predictive power in later year degree subjects beyond the performance in the diploma program. There was a weak link between SELP and either CEP or EAP. This study also found that every student who failed the SELP test but had passed the diploma and the CEP and EAP programs successfully completed the degree. The study raises questions about English speaking universities using SELP alone as an admission criteria. The study has implications for the admission policy of English speaking universities.

Keywords: Admission policy; Chinese; Cross-over repeated measures; English language proficiency measures; transnational business education

1. Introduction

Transnational education programs (TNE) are programs which are conducted with a formal agreement between a Western university and an overseas institution. Western universities that offer their academic programs have a responsibility for overseeing the academic standards and completion (Yang, 2012). In China, TNE programs are officially referred to as Chinese-Foreign Cooperation in Running Schools (中外合作办学). The growth of Western universities' TNE programs since the 1990s was driven by the Chinese government's call for internationalisation of Chinese higher education to support the country's economic development and expansion in the international market (Regulations of the People's Republic of China on Chinese-Foreign Cooperation in Running Schools, 2003). Australia is among the top three countries (along with the United Kingdom and the United States) that offer TNE programs in China (China's Ministry of Education, 2018).

To be admitted to the TNE programs, Chinese students must submit evidence of their English language proficiency (ELP), in addition to the evidence of their academic qualifications. Similar to most of the Western English speaking universities, the commonly accepted evidence of ELP by Australian universities includes an Australian university's English for academic purpose (EAP) program, or a specified score in a standardised ELP test (SELP) such as the International English Language Testing System (IELTS academic), the Test of English as a Foreign Language (TOEFL). An alternative ELP evidence normally accepted by universities also includes the completion of a minimum of one year full time study in a discipline specific tertiary program instructed in English language[1].

The compulsory ELP requirement of Western universities is often justified by the argument that ELP is critical for the success of international students in English-speaking programs (see Oliver, Vanderford, & Grote, 2012), and their future career prospects (Cummin, 2000). Hence, a better understanding of the comparability of different ELP measures will assist Western universities to review their admission policies. It will also benefit the teaching and learning of English language programs and discipline-specific studies in English-medium universities. However, no prior study has examined the comparability of the predictive power of different forms of ELP evidence for students' performance in the context of the TNE programs based in China. The sheer volume of Chinese students studying in TNE programs warrants more research in this area.

[1] This is based on the data manually collected by the author from the English language requirement information published on universities' websites.

The objective of this study, as part of a larger project, is to investigate how different Western and Chinese ELP measures compare in relation to predicting the academic performance of EFL students in English-medium degree study. Researchers had a unique opportunity to draw on data sourced from a total of 759 students that studied in the TNE program between 2006 and 2014. Each student *concurrently* undertook a Chinese ELP program – the national college English program (CEP); a Western EAP program; a SELP test; and a post-secondary English-medium business diploma program. The CEP refers to the teaching of English language to undergraduate students who specialise in any discipline area other than English language in China. It is a compulsory program for the first two years in all Chinese universities (Jin & Yang, 2006). The Western EAP program in this study refers to the traditional English language support programs (see Fox, Cheng, & Zumbo, 2014).

The research design of this study differs from previous comparative studies (e.g. Hill, Storch, & Lynch, 1999; Oliver et al., 2012; Tweedie & Chu, 2017) in that most prior research has been limited by a research design that involves a different cohort of EFL students for each ELP measure compared. While recognising the difficulty of collecting data for a single cohort of students concurrently having multiple ELP evidence (Davies, Brown, Elder, Lumley, & McMamara, 1999), the design in most prior research cannot effectively control for the cohort variation, and risks results being inconclusive or misleading, without the inclusion of a large number of student characteristics as control variables (for example, individual differences, level of commitment, and social and demographic characteristics). To address this limitation, this study adopts a comparative design methodology – a crossover repeated measures (CRM) design (Jones & Kenward, 2014; Yang & Farley, 2019), where each student has results for all four ELP measures (CEP, EAP, SELP, and diploma) analysed, to control for student specific factors other than ELP measures in predicting students' academic performance. This is the first study that uses a CRM design to undertake the comparative analysis of Chinese and Western ELP measures, or even just three Western ELP measures, to predict academic performance in English-medium TNE programs. Findings of this will have implications for the admission policy and the curriculum development of Western universities' TNE programs.

The remainder of this paper is structured as follows: Section 2 describes the context of this study. Section 3 reviews relevant literature. Section 4 presents research questions and explains the research methodology. Section 5 reports the results, and Section 6 discusses the results and concludes the study.

2. Context

The Australian university in this study (hereafter referred to as AU) is a large public university. AU offers business degrees through its TNE program in conjunction with a Chinese university (hereafter CU). The TNE program consists of CEP and EAP programs (taught and assessed by CU and Australian English language teachers respectively), Chinese first-year undergraduate degree core subjects (taught in Chinese by CU academics), and the Australian business diploma program (i.e. diploma). The diploma program is a tertiary discipline pathway program that gives students who have successfully completed the program full credit for first year undergraduate degree study. The diploma program covers Accounting, Business Law, Economics, Information Systems, Management, Marketing, Statistics and Business Communication. All AU diploma and degree subjects are assessed by a mix of continuous assessment and examinations in English. AU academics are responsible for the curriculum and all assessments including marking. To ensure the consistency and validity of grading in each subject, AU's learning and teaching policy requires (AU) peer moderation and validation of all assessments and marking in diploma and degree subjects.

Chinese students admitted to AU's TNE program are selected based on the Chinese national college entrance exam. The students simultaneously complete both an Australian degree and a Chinese degree. At the end of the pathway program successful students who have met both academic requirements of the diploma program and satisfied the English language requirement (i.e. pass AU's SELP test equivalent to IELTS academic model score 6.0) are allowed to proceed to the second year of the degree with full credit for the first year. Online Appendix 1 provides further details about CEP, EAP and SELP. For the remaining two years, students can complete AU's degree either in China or in Australia from Semester 5 onwards. The structure of AU's pathway program is illustrated in Table 1 (see online additional material).

Insert online Table 1 here

Unlike most Western English speaking universities, AU's internal admission policy specifically requires Chinese students to sit a SELP test in addition to the completion of the embedded CEP and EAP programs, and business diploma course. The SELP test score is used as sole English language evidence for admitting students to the later year degree study. Table 2 (see online additional material) provides a summary of equivalence levels of the commonly accepted standardized English language proficiency tests at AU.

Insert online Table 2 here

AU's admission policy, albeit controversial, offers a rare opportunity to compare the predictive power of Chinese and Western ELP measures for the academic performance of Chinese students in Western English-medium transnational business programs.

3. Chinese and Western ELP measures as predictors of academic performance

3.1 CEP

Since its inception, the fundamental principles underlying CEP have consistently been on developing the English language proficiency (in a broad sense) of non-English major students to support the country's economic activities. However, there has been sparse research into CEP, partly due to the discouraging institutional environment of academic research in this area in many Chinese universities (Borg & Liu, 2013). Current literature on CEP is dominated by opinion papers on what CEP should (or should not) be (see Cai, 2017). An exception is the study by Teng (2009) who adopted a CRM design to compare the predictive power of CEP and the one-off Chinese College English Test Band-4 (CET-4) test for academic performance in a Chinese undergraduate program. Teng's (2009) study found CEP had a better predictive power than the standardised CET-4 test score. However, Teng (2009) was unable to offer any explanation for his results. Given no other relevant empirical studies in the literature (English or Chinese), it is difficult to confirm Teng's (2009) findings.

3.2 EAP

Similar to Chinese CEP, traditional EAP programs are concerned with general academic skills and strategies for tertiary study (Floyd, 2015). However, an increasing number of researchers (e.g. Fox et al., 2014; Murray, 2016) question the value of traditional EAP approaches. They argue that language is embedded in discipline-specific contexts. EFL students need to develop cognitive academic language proficiency (CALP), i.e. both the knowledge of academic language and the knowledge of specialized subject matter (see Cummins, 2000). Therefore, the general nature of traditional EAP programs may be of little value to EFL students.

A number of Australian studies (e.g. Anderson, Reberger, & Doube, 2004; Dyson, 2014; Floyd, 2015; Leask, Ciccarelli, & Benzie, 2003; Oliver, et al., 2012) and an American study (Tweedie & Chu, 2017) consistently report the relatively weaker performance of students who were admitted to tertiary studies through general EAP

programs when compared to those who were accepted by SELP tests (e.g. IELTS). However, findings of the above mentioned studies cannot be generalised due to some inherent limitations of their data analysis (discussed later).

3.3 Standardised English language proficiency tests

IELTS (academic) and TOEFL (iBT) are the two most popular SELP tests used by Western university admission officers (see IELTS, 2018; ETS, 2018). Prior research (see Arkoudis, Baik, & Richardson, 2012; ETS, 2010; Gardiner& Howlett, 2016; Green, 2018) have found SELP tests are correlated (although recognition is also given to the differences among the SELP tests), and hence justifies the use of those tests as ELP evidence by university admission officers. However, prior studies report a generally weak relationship between SELP test scores and the academic performance of international students. For example, several studies found either weak relationships between IELTS scores and academic performance (Bayliss & Ingram, 2006; Feast, 2002; Kerstjens & Nery, 2000) or no relationship (e.g. Dooey & Oliver, 2002; Fiocco, 1992; Garinger & Schoepp, 2013), similar to the findings in the studies using TOEFL score as a predictor of academic performance (e.g. Cho & Bridgeman, 2012; Ginther & Yan, 2018; Wongtrirat, 2010).

The weak relationship between SELP and academic performance reported in prior research has led to a growing number of researchers (e.g. Bachman, 1990; Murray, 2016) questioning the content validity of SELP tests and their ability to predict the future performance of EFL international students in English-medium discipline studies. For example, Murray (2016) critiques that SELP tests currently 'used by English-medium universities as screening mechanisms lack authenticity for they do not sufficiently reflect the actual language requirements of students' future degree programs' (p. 107). SELP tests (also applicable to Chinese CET-4 test) focus on the general academic English ability of test takers, not testing the discipline specific CALP.

4. Research design

4.1 Research questions

While prior research provides useful information about the predictive power of different ELP measures for the academic performance of EFL international students, they are limited by the small data size (with the exception of Oliver, et al., 2012). Prior studies are also limited by the analysis of academic performance of the EAP cohort and a separate test-taker cohort

without controlling for other factors (e.g. academic ability) that also affect academic performance. For example, a finding about an EAP cohort who performed weaker than a SELP test taker cohort could possibly be explained by the difference in academic ability between the EAP cohort and the SELP test taker cohort. Another issue is the use of raw academic grades to measure students' performance across different disciplines with different proportions of EAP and SELP students, without controlling for the assessment and grading differences in different discipline-specific areas of study (see Yang & Farley, 2019). Results would be more rigorous if the comparison is made through comparing the results of the same cohort of international students who concurrently have multiple different forms of ELP evidence. The research team of this study is fortunate to be granted the access to a unique dataset which makes it possible to adopt the CRM design to investigate the following research questions.

Research Question 1 (RQ1): How do CEP, EAP and SELP compare in predictive power of performance in Year 1 (business diploma) and later years of an Australian University's degree program?

Research Question 2 (RQ2): How do CEP, EAP and SELP perform as incremental predictors of students' performance in later years of an Australian university degree program beyond the predictive power of the business diploma results?

RQ1 and RQ2 are closely related but each has its own focus. RQ1 aims to compare the three *general* English language proficiency (ELP) measures (i.e. CEP, EAP, and SELP). RQ2 builds upon RQ1 and investigates if the three general ELP measures add any incremental predictive power beyond the predictive power of the performance in the business diploma (which includes a *discipline-specific* ELP measure) in predicting later year degree performance.

4.2 Data collection

Consistent with previous studies (Garinger & Schoepp, 2013; Tweedie & Chu, 2017), this study will use the academic results (grades) to measure academic performance in the TNE program. Prior studies in the secondary school sector (see Allen, 2005; Thorsen & Cliffordson, 2012) have raised the issues of the predicative validity of using grades as a measure of students' academic achievements due to teachers' bias through taking into account non-academic factors. However, this is not a major concern in this current study due to the CRM design. Further, the systematic assessment and grading moderation processes imposed by AU and CU (i.e.

moderators and/or markers do not know the students) support the reliability and validity of student results as measures of student academic performance.

The primary data set (hereafter referred to as the Primary Data) was collected for 307 Chinese students who were admitted to all business streams of the AU pathway program in 2006 and completed the TNE degree in 2010. Data for each student covered gender, results in CEP, EAP, and SELP, results in business diploma subjects, and results in second and third year for all business subjects.

A supplementary sample (hereafter referred to as the Supplementary Data) was collected for 452 Chinese international students who were admitted to the Accounting stream of the AU pathway program in China in 2008, 2009 and 2010 and completed the program in 2012, 2013, and 2014 respectively. Data for each student covered gender, results in the CEP, results in the EAP program, results for the diploma subjects, and results in second and third year degree subjects. While these students completed a major in accounting their other degree subjects overlapped heavily with students completing other business majors in the same degree. This data set did not include the SELP results because the SELP program was amended in 2009 to record only a Pass/Fail result. Hence, unlike the Primary Data this dataset could not be used to compare the predictive power of all three programs; only to compare CEP and EAP. The students in this cohort also undertook an updated version of the EAP program compared to those in the Primary Data cohort. The main changes in the EAP program were using prescribed textbooks, and an integrated team teaching of Chinese CEP teachers and Australian English teachers. The above changes in the EAP program were in recognition of issues similar to those found in this article.

This study also collected the data of a cohort of students (who studied between 2006 and 2010) who failed the SELP test in 2008 but were still allowed to progress to the degree program (hereafter referred to as the Failed Cohort data). All of these students passed the CEP and EAP programs as well as all subjects in the diploma. Data for each student covered results in second and third year for all business subjects.

4.3 Empirical analysis

The unique aspect of the empirical analysis in this study is the use of a CRM design where matching results for the same student are available for all four forms of ELP evidence (i.e. CEP, EAP, SELP and diploma). This is a significant advance on previous studies because it controls for most student specific characteristics when comparing the different ELP evidence, including the key ones of academic ability and motivations.

Descriptive statistics and correlation analysis are used to show the relationship between results in the four different forms of ELP evidence. Regression analysis is used to determine the predictive power of CEP, EAP and SELP with respect to performance in the diploma and later year subjects.

To assess how each ELP evidence performs as a predictor of student performance, as per RQ1, a series of regressions were run as follows. Gender is used as control variable because previous studies (see Voyer & Voyer, 2014) found it influences the relationship between academic performance of Chinese students at different levels of study.

(1) CEP model

$$Av.Diploma\ Result\ or\ Later\ Year\ Av.Result = \alpha + \beta_1 CEP\ Result + \beta_2 Female + \varepsilon$$

(2) EAP model

$$Av.Diploma\ Result\ or\ Later\ Year\ Av.Result = \alpha + \beta_1 EAP\ Result + \beta_2 Female + \varepsilon$$

(3) SELP model

$$Av.Diploma\ Result\ or\ Later\ Year\ Av.Result = \alpha + \beta_1 SELP\ Category1 + \beta_2 SELP\ Category2 + \beta_3 SELP\ Category3 + \beta_4 SELP\ Category4 + \beta_5 Female + \varepsilon$$

The comparison of the predictive power of each ELP evidence is achieved by comparing the R^2 values from the regressions using the same dependent variable.

To assess how CEP, EAP and SELP perform as incremental predictors of student success beyond the predictive power of the diploma, as per RQ2, a series of regressions were run as follows.

(4) Dip CEP model

$$Later\ Year\ Av.Result = \alpha + \beta_0 Av.Dip\ Result + \beta_1 CEP\ Result + \beta_2 Female + \varepsilon$$

(5) Dip EAP model

$$Later\ Year\ Av.Result = \alpha + \beta_0 Av.Dip\ Result + \beta_1 EAP\ Result + \beta_2 Female + \varepsilon$$

(6) Dip SELP model

$$Later\ Year\ Av.Result = \alpha + \beta_0 Av.Dip\ Result + \beta_1 SELP\ Category1 + \beta_2 SELP\ Category2 + \beta_3 SELP\ Category3 + \beta_4 SELP\ Category4 + \beta_5 Female + \varepsilon$$

The incremental predictive power of each ELP evidence is assessed by looking at the significance of the relevant English language variable/s in each of these models. The significance level of 5% is represented by *, and the significance level of 1% is represented by ** in all relevant tables reported in Section 5.

5. Results

5.1 RQ1: Comparison of predictive power of CEP, EAP and SELP

To address RQ1, the link between the embedded CEP, EAP and the SELP was done by examining the link between results in SELP, EAP, CEP, and results in the business diploma and degree programs. The link between SELP and EAP and Chinese CEP is shown in Table 3 (see online additional material).

Insert online Table 3 here

Table 3 shows the pattern of the average, maximum and minimum AU EAP and Chinese CEP results across the categories of SELP. It indicates that there is a relationship between SELP and both AU EAP and Chinese CEP performance. However, the extreme overlap in the range of scores within each SELP category suggests that this relationship is not a strong one. The average of 62.28 in EAP and 54.19 in Chinese CEP in the SELP failed category (Category 5) suggests that even a failed result in SELP corresponds to a relatively high level of English competence in the EAP program. This finding is reinforced by the fact that the correlation (See Table 4 below) between SELP category and average EAP result is -0.43, which means that either measure explains only 18.4% of the variation in the other measure and with CEP the equivalent results are -0.52 and 27.4%.

Table 4: R Squared (% of variation explained) comparison

Measure	**CEP**	**EAP**
EAP	57.14%	
SELP	27.41%	18.43%

Table 5 shows the detailed results of estimating RQ1 regression equations (Models 1, 2 and 3).

Table 5: Detailed results of CEP, EAP and SELP in predicting diploma and degree results using Primary Data

	Dependent Variable					
Explanatory Variable	**Av. Diploma Result**			**Av. Degree Result**		
	CEP Model	EAP Model	SELP Model	CEP Model	EAP Model	SELP Model
Constant	25.394**	27.727**	67.331**	35.670**	35.964**	62.621**
CEP Result	0.812**			0.518**		
EAP Result		0.712**			0.462**	
SELP Category 1			9.044**			8.101**
SELP Category 2			6.997**			7.081**
SELP Category 3			3.608			3.940*
SELP Category 4			1.903			2.838
Female	0.234	-0.584	7.845**	1.556	1.977*	3.397**
Adjusted R Squared	0.470	0.305	0.309	0.296	0.223	0.151
F Statistic	136.359**	67.769**	28.699**	64.724**	44.427**	11.912**

The Adjusted R Squared row in Table 5 shows that when comparing the power to predict diploma results CEP is better than EAP and SELP (which are very similar), and for degree results CEP is stronger than EAP which in turn is stronger than SELP. However, all three ELP measures have reduced predictive power for the later year degree performance than for the diploma results.

The comparability of CEP and EAP can be further tested by reference to the Supplementary Data set. This allows a comparison of the two after the EAP program was updated. The CEP and EAP regression models are re-estimated using this new data and the results shown in Table 6. Given this data set covers intakes over three years year variables have been added to the model to allow for the relationship to alter over time.

Table 6: Regressions predicting the average mark in the diploma and degree subjects using the Supplementary Data set

	Dependent Variable			
Explanatory Variables	**Av. Diploma Result**		**Av. Degree Result**	
	CEP Model	**EAP Model**	**CEP Model**	**EAP Model**
Constant	27.586**	26.294**	25.876**	24.797**
Average Mark in CEP	0.718**		0.585**	
Average Mark in EAP		0.668**		0.542**
Female	3.233**	3.086**	2.448**	2.470**
2009 Dummy	-2.559**	2.335**	-2.558**	1.291
2010 Dummy	0.513	3.317**	2.983**	5.255**
Adjusted R squared	0.569	0.512	0.373	0.335
F statistic	149.917**	117.794**	65.754**	55.602**

Table 6 shows that the updating of the EAP program led to the predictive power of the EAP program moving much closer to that of the CEP program. Similar to the Primary Data, both CEP and EAP had reduced predictive power for the later year degree program compared to the diploma program. Further supporting evidence of the weakness of the SELP program as a predictor of success in either the diploma or degree can be obtained from tables 7 and 8.

Table 7: Regressions predicting the average mark in the diploma and degree subjects based upon SELP categories using different reference categories

	Dependent Variables					
	Av. Diploma Result			**Av. Degree Result**		
	Reference Category			**Reference Category**		
Explanatory Variable	**1**	**3**	**5**	**1**	**3**	**5**
Constant	76.376**	70.939**	67.331**	70.722**	66.561**	62.621**
SELP Category 1		5.436**	9.044**		4.161**	8.101**
SELP Category 2	-2.048	3.389**	6.997**	-1.021	3.141**	7.081**
SELP Category 3	-5.436**		3.608	-4.161**		3.940*
SELP Category 4	-7.141**	-1.705	1.903	-5.263**	-1.102	2.838
SELP Category 5	-9.044**	-3.608		-8.101**	-3.940*	
Female	7.845**	7.845**	7.845**	3.397**	3.397**	3.397**
Adjusted R Squared	0.309	0.309	0.309	0.151	0.151	0.151
F Statistic	28.699**	28.699**	28.699**	11.912**	11.912**	11.912**

Table 8: Failed SELP students' performance in degree

Grade	Failed SELP	All Business Students
HD	7%	9%
D	23%	17%
C	25%	22%
P	25%	20%
N	5%	13%
WD	15%	20%

Key: HD= High Distinction (80 marks or above); D= Distinction (70-79); C=Credit (60-69); P= Pass (50-59); N= Fail (0-49); WD (withdrawal)

Tables 7 highlights that it is common for adjacent categories of SELP to not have significantly different predictive power to each other, e.g. 1 vs 2, 3 vs 4 and 4 vs 5 for both the diploma and the degree and also 3 vs 5 for the diploma. Table 8 is based upon the Failed Cohort data and shows their grade distribution compared to that of all business students. All 22 students in the cohort successfully completed the degree and on average did better than the full business cohort.

5.2 RQ2: Incremental power of CEP, EAP and SELP to predict performance in degree above the predictive power of diploma

To address RQ2 the incremental ability of CEP, EAP and ELP to predict student performance in the later year degree beyond the explanatory power of the results in the diploma was also examined. This is done by using the average mark in the degree subjects as the dependent variable and the average mark in the diploma program combined with the average mark in CEP, EAP, and category in SELP as the explanatory variables (i.e. Models 4, 5, and 6). The results of this are shown in Table 9. Gender is again used as a control variable.

Table 9: Regressions predicting the average mark in the later year degree subjects using average mark in pathway program

Explanatory Variables	**Dip CEP Model**	**Dip EAP Model**	**Dip SELP Model**
Constant	19.214	17.234	20.864**
Average Mark in Business Subjects in Diploma Program	0.590**	0.593**	0.614**
Average Mark in CEP	0.074		
Average Mark in EAP		0.094	
SELP Category 1			1.704
SELP Category 2			1.955
SELP Category 3			1.324
SELP Category 4			1.128
Female	0.224	0.115	0.096
Adjusted R squared	0.537	0.539	0.538
F statistic	116.078**	117.111**	57.579**

Table 9 shows that when performance in the diploma is added to the model predicting performance in the later year degree the explanatory power of the model increases significantly and the incremental explanatory power of all the other three ELP measures (CEP, EAP, and SELP) is accepted as being zero. This finding shows that although each of the other three ELP measures alone has predictive power (albeit differing) for academic performance at the later year degree level (as analysed in Subsection 5.1), when considered concurrently with the diploma, they have no additional/incremental value beyond the predictive power of the diploma (see Section 6.2 for further discussion).

6. Discussion and Concluding Remarks

6.1 Comparison of CEP, EAP and SELP as predictors of the academic performance in the pathway diploma and later year degree programs

The weak correlation found between CEP (Chinese College English Program), EAP (Western English for Academic programs) and SELP (Western standardised English language proficiency test) in this study (Table 4) suggests they may be measuring different aspects of English language skills, consistent with prior studies on the comparison of different Western ELP measures (Arkoudis et al., 2012; Green, 2018; Murray, 2016).

Results from Table 5 confirm that CEP, EAP and SELP are all capable of predicting performance in different year levels of a degree program. However, as found in other studies (Teng, 2009; Oliver et al., 2012; Garinger & Schoepp, 2013) they are all better at predicting performance in first year (in this study the equivalent is the diploma) than in later years of the degree. This is consistent with English language development being an on-going process and

hence evolving, and thus English language proficiency (ELP) changing, over the life of the degree (Arkoudis et al., 2012). As a result, ELP measured through first year becomes less appropriate as a measure of the student's ELP as they progress through later years of the degree.

Table 5 findings also suggests that CEP is a substantially better predictor than EAP or SELP. However, results from Table 6 suggest that the initial result for EAP could be a consequence of the poor design of the initial EAP program. Once the EAP program was redesigned in 2007 results suggest a well-structured EAP program has very similar predictive power to CEP. This leaves SELP as the sole poor predictor of diploma and degree results, which is consistent with prior research based in English-speaking countries (see Section 3.3).

One possible explanation for SELP's weaker predictive power can be that it uses only five categories to distinguish students' ELP versus the 100 point scale used by CEP and EAP. However, a deeper insight into SELP's poor predictive power can be gleaned from the results in Table 7. The results confirm that adjacent categories of SELP (e.g. category 4 versus category 5) generally offer no discriminating power. It is only when there is a difference of at least two categories (e.g. category 4 versus category 2) does SELP discriminate diploma or degree performance (although for the diploma even category 5 versus category 3 does not discriminate). This suggests poor correlation between the type of ELP required for success in diploma and degree study and the type of ELP measured by SELP. The findings support the view that it is a highly arbitrary judgement to use an IELTS score of 6.0 to determine that a student can be admitted to a particular discipline study, while a student who gains an IELTS score of 5.5 cannot (see Arkoudis et al., 2012). This view is further supported by the results in Table 8 which show that the Failed Cohort of students, who all failed to meet the SELP hurdle requirement of category 4 for admission, went on to not only complete the degree but on average outperform the average degree student. This finding is consistent with the results in Deygers, den Branden, & Gorp (2018). Given these students all passed the diploma and the CEP and EAP programs it supports all these being more appropriate hurdle requirements than a single category difference in SELP. The finding also points out the unfairness of using SELP alone to determine the ELP of Chinese international students in AU's TNE degree programs.

The findings about Chinese CEP in both the Primary Data set (Table 5) and the Supplementary Data set (Table 6) lend some support to Teng's (2009) findings about a stronger predictive power of CEP than the standardised ELP test (Chinese CET-4 Test). The findings suggest Chinese CEP can be accepted as an alternative English language proficiency evidence for English-medium TNE programs (at least in the AU's TNE programs in China).

Findings about the equivalent good predictive power of CEP and Australian university's EAP program (but only after its revision as shown in the Supplementary Data set analysis in Table 6) suggest the benefits of collaboration between English language teachers of both CEP and EAP in TNE programs in China. The findings point to the potential of having CEP and Western EAP programs being integrated as one language support program embedded in the TNE program offered to Chinese students.

The findings about the comparative predictive power of the (revised) EAP program when compared with SELP for the academic performance in diploma/degree studies differ from the previous studies based in English speaking countries (Anderson et al., 2004; Dyson, 2014; Leask et al., 2003; Oliver et al., 2012; Tweedi & Chu, 2017) which found SELP test takers performed better than the EAP cohort. This difference can be explained by the data used for this study. As mentioned earlier, comparison of different forms of ELP of the same cohort of students is rare due to the resource constraints. The current study has addressed this limitation by comparing the same cohort of students' across different ELP measures by using the crossover repeated measures design which has effectively controlled for most other factors influencing the academic performance. This was not done in the other referenced studies.

6.2 Incremental power of CEP, EAP and SELP to predict performance in degree above predictive power of diploma

The findings about the lack of any incremental predictive power for the later year degree performance of any of the three general ELP measures (CEP, EAP and SELP) beyond the strong predictive power of the diploma (Table 9) does not mean that English proficiency does not influence performance in the degree but rather that the English proficiency of each student is already sufficiently represented by their performance in the diploma. This important finding supports researchers' views (Evans & Morrison, 2011; Murray, 2016) that discipline specific language proficiency better reflects the actual language requirements of students' future degree study than generic English language programs (CEP or EAP) or SELP tests. The findings of this study confirm that evaluation of English language proficiency is better placed in the context of specific discipline study. The findings support the suitability of using a discipline specific pathway program as acceptable ELP evidence to meet the admission requirement. What is not supported is the frequent claim (within AU and as expressed in media and public forums) that adding a SELP test after students have successfully completed an English language pathway program will *protect* the degree program from students with an unacceptable level of English. The findings of this study support the reverse argument (see Deygers et al.,

2018) since it shows students below the SELP hurdle but with alternative evidence of ELP can be highly successful when allowed to proceed. A belief that performance in a discipline specific English-medium pathway program is not a sufficient measure of English language ability seems to be one of the greatest myths in transnational education and is directly rebutted by the findings in relation to RQ2. Blindly imposing a higher level of SELP score (a simplistic approach of assessing students' English language proficiency) while ignoring discipline-specific English language proficiency measures risks denying the tertiary education opportunity to those students 'who are at an academic level where their peers are achieving success academically' in an English as foreign language country (Garinger & Schoepp, 2013, p.12).

It should be noted that while CEP, EAP and SELP in isolation are predictors of performance in the diploma and the degree, the analysis in Table 9 does not allow a statement about how well the discipline specific English language proficiency embodied in the diploma result would predict the performance in the degree, only that the CEP, EAP and SELP results add nothing beyond the predictive power of business diploma. A separate discipline specific ELP measure is not directly assessed and cannot be calculated because the diploma result is a composite of discipline specific ELP and other student characteristics, such as academic ability, which are not available in this study. However, this is not a limitation because of the cross-over repeated measure research design used in this study.

6.3 Implications, limitations and direction for future study

This interdisciplinary study involves researchers from Western and Chinese universities in the fields of business and applied linguistics. It serves as an ice-breaker for more collaborations between Western and Chinese scholars (and university administrators) to continuously improve the curriculum design of transnational degree programs in China (and other EFL countries). The study offers the first empirical evidence of the comparison of the predictive power between Chinese (CEP) and Western (EAP, SELP, business diploma) ELP measures for the academic success of Chinese students in English-medium TNE programs. Even though the CEP and EAP program results add no additional predictive power to the diploma result it does not imply these programs could be eliminated. It is the completion of these programs that develops the foundation for the discipline specific ELP further developed throughout the diploma. This does not apply to the SELP since it is not a program to develop an ELP foundation but rather just a test of ELP.

The use of cross-over repeated measures design in this study has addressed the limitations in prior literature by effectively controlling for factors other than the ELP. Thus, it enables a more accurate comparative analysis of Chinese and Western ELP measures. It is recommended future researchers consider adopting CRM to undertake comparative studies of student performance if such data is available, or control for other factors that can affect student performance if it is not.

Findings of this study raise serious questions for Western universities to reflect upon when it comes to admission decisions. It also challenges the dominant practice in some English speaking countries of using SELP alone for purposes such as student visa applications and professional registration. Authors of this study believe its results firmly support a principled approach (see Green, 2018) to admission decisions which involves weighing the attributes of applicants, the academic language requirements of the applicant's future discipline specific course of study, and the learning support available for students.

The study has several limitations that may be addressed in future studies. First, this study is limited to one Australian university's TNE program. Future study can include more universities from other English-speaking countries. Second, due to resource constraints, data for the tertiary entry scores as well as Chinese CET-4 results were not available for the study period. Future research may be extended to include the analysis of the Chinese CET-4 test and other forms of standardised English language proficiency tests.

The variety of ELP measures currently employed by Western English-speaking universities means a consistent framework for English language proficiency is imperative for EFL learners and their future employers. Recent development in China's convergence of its newly issued *Standards of English Language Ability* with the *Common European Framework of Reference for Languages* (see British Council, 2018; ETS, 2018) is expected to produce more consistent Chinese and international ELP measures (albeit with challenges, see Jin, Wu, Alderson, & Song, 2017). Future research can monitor the comparability of the predictive power of Chinese and Western ELP measures in the context of a converged Chinese and international ELP framework.

References

Allen, J.D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearinghouse: A Journal of Educational Strategies, Issues and Ideas, 78*(5), 218-223.

Arkoudis, S., C. Baik, & S. Richardson. (2012). English Language Standards in Higher Education: From Entry to Exit. Camberwell: ACER.

Anderson, P., Reberger, H., & Doube, L. (2004). Language proficiency and academic outcomes for international postgraduate students at the University of Adelaide. Adelaide: University of Adelaide

Bachman, L. F. (1990). Fundamental considerations in language testing. Oxford University Press

Bayliss, A. and Ingram, D. E. (2006) IELTS as a Predictor of Academic Language Performance, Australian International Education Conference, 2006, 1-12.

Borg, S., & Liu, Y. (2013). Chinese College English Teachers' Research Engagement. *TESOL Quarterly, 47*(2), 270-299.

British Council. (2018). APTIS and CSE, https://www.britishcouncil.cn/en/exams/aptis/cse, accessed 28 December 2018

Cai J. (2017). Debates Around the Orientation of TEFL in Chinese Tertiary Education. In: Reinders H., Nunan D., Zou B. (eds, 115-153) Innovation in Language Learning and Teaching, New Language Learning and Teaching Environments. London: Palgrave Macmillan.

Cho, Y., & Bridgeman, B. (2012). Relationship of TOEFL iBT® scores to academic performance: Some evidence from American universities. *Language Testing, 29*(3), 421-442.

China Ministry of Education. (2007). College English Program Syllabus Requirement (大学英语课程教学要求). Policy No.: Jiaogaoting (教高厅) 2007(3), http://www.moe.gov.cn/srcsite/A08/s7056/200707/t20070710_110825.html, accessed 20 April 2018.

China Ministry of Education (2018). 中外合作办学监管工作信息平台. List of transnational education programs, http://www.crs.jsj.edu.cn/index.php/default/approval/orglists, accessed 23 April 2018.

China Ministry of Education (2003). Regulations of the People's Republic of China on Chinese-Foreign Cooperation in Running Schools, http://www.crs.jsj.edu.cn/news/index/3, accessed 23 November 2018.

Cummins J. (2000). *Language, power and pedagogy: Bilingual children in the cross-fire*. Clevedon, UK: Multilingual Matters.

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T, & McNamara, T. (1999). Dictionary of Language Testing. Cambridge: Cambridge University Press.

Deygers, B., den Branden, K. V., & Gorp, K. V. (2018). University entrance language tests: A matter of justice. *Language Testing*, *35*(4), 449–476.

Dooey, P., & Oliver, R. (2002). An Investigation into the predictive validity of the IELTS Test as an indicator of future academic success: AMEP Research Centre. *Prospect: an Australian journal of TESOL*, *17*(1), 36-54

Dyson, B. (2014). Are onshore pathway students prepared for effective university participation? A case study of an international postgraduate cohort. *Journal of Academic Language & Learning, 8(2),* A28-A42.

Educational Testing Service (2010). Linking TOEFL iBT ™ Scores to IELTS® Scores – A Research Report, http://www.ets.org/toefl/institutions/scores/compare, accessed 20 December 2018

Educational Testing Service (2018). ETS and NEEA Start to Map TOEFL® Test Scores to CSE Levels, http://www.ets.org, accessed 20 December 2018

Feast, V. (2002). The Impact of IELTS Scores on Performance at University. *International Education Journal*, *3*(4), 70-85.

Fiocco, M. (1992). English proficiency levels of students from a non-English speaking background: a study of IELTS as an indicator of tertiary success, Unpublished Research Report, Curtin University of Technology, Perth.

Floyd, C. B. (2015). Closing the gap: International student pathways, academic performance and academic acculturation. *Journal of Academic Language & Learning*, *9*(2), A1-A18.

Fox, J., Cheng, L., & Zumbo, B. D. (2014). Do They Make a Difference? The Impact of English Language Programs on Second Language Students in Canadian Universities. *TESOL Quarterly*, *48*(1), 57-85.

Gardiner, J., & Howlett, S. (2016). Student perceptions of four university gateway tests. *University of Sydney Papers in TESOL*, 11, 67-96.

Ginther, A., & Yan, X. (2018). Interpreting the relationships between TOEFL iBT scores and GPA: Language proficiency, policy, and profiles. *Language Testing*, *35*(2), 271-295.

Garinger, D., & Schoepp, K. (2013). IELTS and Academic Achievement: A UAE Case Study. Perspectives (TESOL Arabia), *20*(3), 7-13.

Green, A. (2018). Linking Tests of English for Academic Purposes to the CEFR: The Score User's Perspective. *Language Assessment Quarterly, 15*(1), 59-74.

Hill, K., Storch, N., & Lynch, B. (1999). A Comparison of IELTS and TOEFL as Predictors of Academic Success. IELTS Research Reports 1999 Volume 2, https://www.ielts.org/-/media/research-reports/ielts_rr_volume02_report3.ashx , accessed 20 December 2018

IELTS. (2018). https://www.ielts.org/what-is-ielts/ielts-for-study

Jin, Y., & Yang, H. (2006). The English Proficiency of College and University Students in China: As Reflected in the CET. *Language, Culture and Curriculum, 19*(1), 21-36.

Jin, Y., Wu, Z., Alderson, C., & Song, W. (2017). Developing the China Standards of English: challenges at macropolitical and micropolitical levels. *Language Testing in Asia, 7*(1), 1-19

Jones, B., Kenward, M. G. (2014). *Design and Analysis of Cross-Over Trials, Third Edition*: Taylor & Francis, New York: Chapman and Hall/CRC.

Kerstjens, M & Nery, C. (2000). Predictive validity in the IELTS test: A study of the relationship between IELTS scores and students' subsequent academic performance [online]. In: Tulloh, Robyn (Editor). International English Language Testing System (IELTS) Research Reports 2000: Volume 3. [Melbourne]: IELTS Australia, 2000: 85-108.

Leask, B., Ciccarelli, A., & Benzie, H. (2003). Pathways to tertiary Learning: A framework for evaluating English language programs for undergraduate entry. *EA Journal, 21*(2), 17-30.

Murray, N. (2016). Standards of English in Higher Education: Issues, Challenges and Strategies. Cambridge: Cambridge University Press.

Oliver, R., Vanderford, S., & Grote, E. (2012). Evidence of ELP and academic achievement of non-English-speaking background students. *Higher Education Research and Development, 31*(4), 541-555.

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes, 30*(3), 198-208.

Thorsen, C., & Cliffordson, (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. *Educational Research and Evaluation, 18*(2), 153-172.

Tweedie, M. G., & Chu, M.-W. (2017). Challenging equivalency in measures of ELP for university admission: data from an undergraduate engineering programme. *Studies in Higher Education*, 1-13.

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174-1204.

Wongtrirat, R. (2010). *ELP and Academic Achievement of International Students: A Meta-analysis*: Old Dominion University.

Yang, H.H., (2012). Western concepts, Chinese context: A note on teaching accounting offshore. *International journal of pedagogies and learning, 7*(1), 20-30.

Yang, H. H., & Farley, A. (2019). Quantifying the impact of language on the performance of international accounting students: A cognitive load theory perspective. *English for Specific Purposes*, 55 (July), 12-24

Comparison of Chinese and Western English language proficiency measures in transnational business degrees

Additional online materials

The structure of AU's pathway program is illustrated in Table 1 below.

Table 1: AU's pathway program with integrated English language programs and SELP

Semester (S)*	English Language Programs	Discipline Program
S1-S2	CEP + EAP	Chinese first year Degree core subjects
SELP Year 1 (minimum grade equivalent to IELTS 5.5)		
S3- S4	CEP + EAP	AU English medium Diploma Course
SELP Year 2 (minimum grade equivalent to IELTS 6.0)		
S5-S6 Year 3	N/A	Second Year AU's Degree (either in China or in Australia)
S7-S8 Year 4	N/A	Third Year AU's Degree (either in China or in Australia)

*Chinese university's academic year

Table 2 provides a brief summary of the equivalence of different forms of standardised English language proficiency tests used widely for meeting English language entry requirements in English-speaking universities.

Table 2: Equivalence table of standardized English language proficiency tests[2]

SELP	**IELTS**	**TOEFL iBT**	**Pearson PET**
Category 1	7.5 and above	103 and above	72 and above
Category 2	7	95	67
Category 3	6.5	88	61
Category 4	6	80	56
Category 5	5.5 and below	72 and below	51 and below

Table 3: Average mark comparison in EAP, CEP and SELP

	EAP			**CEP**		
SELP Category	**Av**	**Max**	**Min**	**Av**	**Max**	**Min**
Category 1	75.51	88.00	62.75	71.76	86.00	58.75
Category 2	72.69	84.75	49.50	67.08	79.00	50.00
Category 3	68.53	83.00	58.67	61.97	75.50	50.00
Category 4	67.42	80.50	46.33	60.04	75.00	50.00
Category 5 (Failed)	62.28	80.00	40.00	54.19	65.75	50.00

[2] The equivalence table is adapted from AU's admission requirements for English language proficiency

Appendix 1: Context Information about CEP, EAP, and SELP

- *Chinese national College English language program (CEP)*

The curriculum of CEP for the current study was set by MoE's (2007) "*College English Curriculum Requirement*" (hereafter *CECR 2007*). The *CECR 2007* explicitly specifies the purpose of CEP is to develop students' ability in using English for academic purposes, professional communication and international engagement. The *CECR 2007* prescribes three levels of English language proficiency to be achieved through CEP, i.e. Standard, Intermediate, and Advanced (see Appendix 2). The *CECR 2007* allows Chinese universities to customise the English proficiency levels to their individual institution's circumstances.

The assessment of CEP in the TNE program includes continuous formative and summative assessments in four components, being reading, listening, writing, and speaking. The program uses prescribed textbooks authored by Chinese academics of English language and linguistics. The program is taught and assessed by the Chinese English language teachers. The English language proficiency (as per CECR 2007) of students at CU in the current study is expected to reach *Advanced level*. The total face-to-face class time of CEP is 384 hours, with an average class size of 30 students.

- *AU EAP program*

AU's EAP program is a general English language program, not specifically related to the intended discipline study of Chinese students. The program is made up of five components: Academic Culture, Academic Writing, Academic Reading, Academic Listening, and Academic Speaking. Similar to CEP, the assessment of EAP includes continuous formative and summative assessments. However, unlike CEP, initial version of AU's EAP program did not have prescribed textbooks, and the content of each component is managed by AU. The revised version of EAP added prescribed textbooks. CEP and EAP are studied simultaneously, and their instruction is complementary. However, compared to the CEP program, the EAP program has more flexibility in terms of the subject content and assessment. Unlike CEP (which carries credit in CU's degree), the EAP does not form a component of AU's degree despite the program being mandatory. The total teaching hours of EAP is 520 hours with a similar class size to CEP.

- *Standardised English language proficiency test*

A standardised ELP test (i.e. SELP) – The Diagnostic English Language Assessment (DELA) was used for this purpose. The DELA test is a two-hour assessment made up of three components: reading, writing, and listening. In addition, speaking was assessed separately modelled on the IELTS format. The final grade of DELA results (which includes the IELTS based speaking assessment) was classified into five categories with the highest being Category 1, followed by Categories 2, 3, 4, and Category 5, which is considered as failing the test. The test is conducted at CU but assessed by AU English language teachers. AU's controversial international admission policy has resulted in some students who failed the SELP test (even though they had passed the diploma, CEP and EAP) being suspended from progressing into the second year degree program. This has triggered a high level of debate between academics (both AU and CU) and the admission administrators on the justice of using SELP alone as ELP evidence for admission in TNE programs.

Appendix 2: Chinese College English Proficiency Levels

Element	Level		
	Standard	**Intermediate**	**Advanced**
Listening	130–150 general English words per minute, conversational English	140–180 general English words per minute; be able to basically understand discipline courses in English	Be able to understand normal speed of conversation between native English speakers; be able to fully understand discipline courses and seminars in English
Speaking	Conversational English	Fluently communicate on general topics	Fluently communicate on discipline topics; be able to synthesise long and difficult readings; be able to present academic papers at international conferences in English
Reading	70 words per minute for general purpose English articles; 100 words per minute for easy reading	Be able to: - read news and articles in native English-speaking countries - read at a speed of 70–90 words per minute for general topics and 120 words for easy topics	Be able to read more difficult discipline-related academic papers
Writing	Be able to write a short essay (general purpose) of no less than 120 words in 30 minutes	Be able to: - write academic purpose essays of no less than 160 words in 30 minutes - clearly express opinions and develop persuasive arguments	Be able to write: - discipline-specific academic papers - an academic essay of no less than 200 words in 30 minutes
Translation	With relative accuracy, be able to translate 300 English words into Chinese and 250 Chinese words into English per hour	Be able to: - translate familiar topics in discipline-specific English articles into Chinese - translate 350 English words into Chinese and 300 Chinese words into English per hour	Be able to translate: - moderate discipline-specific articles into Chinese - Chinese culture into English - 400 English words into Chinese and 350 Chinese words into English per hour
Vocabulary	4795 words and 700 phrases, including 2000 active words	6395 words and 1200 phrases, including 2200 active words	7675 words and 1870 phrases, including 2360 active words

Source: College English Curriculum Requirement (2007), MOE, translated by the first author