

Submission No. 156
(Overseas Trained Doctors)
Date: 1/08/2011



**Pearson Australia Group Submission to the House
of Representatives Standing Committee on Health
and Ageing on the Pearson Test of English
Academic (PTE A)**

Prepared by Pearson Australia Group 29 July 2011

Contact Details:

Name: Fraser Cargill/Sonia Wilson

Email: Fraser.cargill@pearson.com / Sonia.wilson@pearson.com.au

Tel: 02 9454-2200

Contents:

1.0	Background	3
2.0	Pearson Test of English – Overview	4
3.0	The Registration Process	5
4.0	PTE Academic Scores	5
4.1	Automated Scoring and Standardisation	6
4.2	The PTE Academic Scale and the Common European Framework (CEF)	7
4.3	Concordance with Other Tests	7
4.4	Recommended Score Concordance with IELTS	8
4.5	Standard Setting Entry Scores with the Medical Profession	8
5.0	Test Security	9

Appendices:

Appendix 1 *Interpreting the PTE Academic Score Report*

Appendix 2 *PTEA_Automated Scoring*

Appendix 3 *Preliminary Estimates of Concordance between Pearson Test of English Academic and other Measures of English Language Competencies*

Appendix 4 *NCLEX Technical Brief_PTE 2010*

1.0 Background

Pearson understands that the House of Representatives Standing Committee on Health and Ageing has adopted a reference from the Hon Nicola Roxon MP, Minister for Health and Ageing to inquire into and report on the registration processes and support available to Overseas Trained Doctors (OTD's) in Australia.

The Committee has also taken on board feedback from overseas doctors with regards to the English language proficiency testing component of the process, and with that in mind is seeking further clarification from Pearson on the following:

- An overview of the Pearson Test of English Academic (PTE Academic)
- An outline of the way in which the test scores should be interpreted.

Pearson is very pleased to be able to respond to the committee and would like to submit the following document to the committee to assist the members with the overall inquiry. There are also a set of appendices with more details on specific areas of the test for those members that require additional information.

Summary of key Features of PTE Academic:

- Computer-based exam utilising state-of-the-art technologies
- Comprehensively tested English language proficiency test
- Measurement of English language 'Enabling' Skills (grammar, oral fluency, pronunciation, spelling, vocabulary and written discourse) in addition to the four basic English language skills of reading, writing, listening and speaking
- Inclusion of a candidate's personal introduction 'speech sample'
- Automated scoring, with high correlation to human raters' scores, ensuring standardised scoring across countries
- Globally recognised candidate, test and score security systems
- No fixed dates and quick turn-around of results - within five days
- Institution score verification site including searchable databases
- Global availability of test Centres
- Developed by leading international English language test experts

Pearson is the world's leading learning company. Our education business combines 150 years of experience in publishing with the latest learning technology and online support. We are also part of the wider Pearson family which includes Penguin, Dorling Kindersley and the Financial Times.

We provide education and assessment services in more than 65 countries. Our courses and resources are available in print, online and through multi-lingual packages, helping people learn whatever, wherever and however they choose. Products such as MyLabs are changing education practice worldwide.

At Pearson, we take learning personally. Every day our work helps learning flourish, and wherever learning flourishes, so do people.

2.0 Pearson Test of English Academic

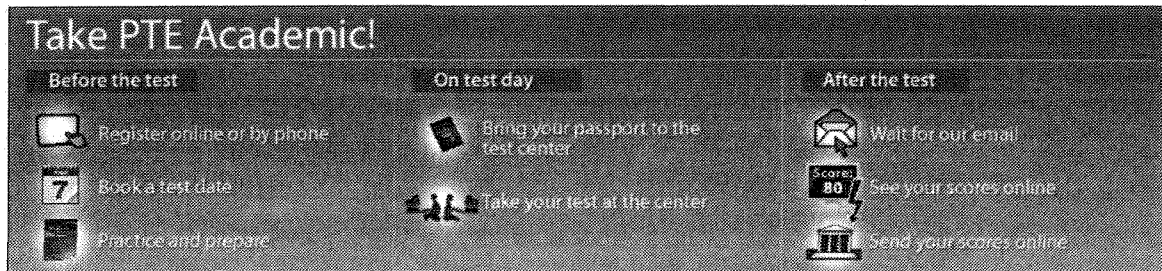
Overview:

Figure 1: Overview of PTE Academic

Format	<ul style="list-style-type: none"> • Computer-based test • Measures all four language areas: Reading, Writing, Listening, Speaking + Enabling Skills of; <ul style="list-style-type: none"> ○ Grammar ○ Oral fluency ○ Pronunciation ○ Spelling ○ Vocabulary ○ Written discourse
Structure	<ul style="list-style-type: none"> • Authentic content – measures real world skills • 20 innovative and integrated question types
Content	<ul style="list-style-type: none"> • Authentic academic test content from a variety of contexts and scenarios • A test of international English
Duration	<ul style="list-style-type: none"> • The test will normally be completed in one 3-hour session
Delivery	<ul style="list-style-type: none"> • Delivered in Pearson Test Centres through our secure global Pearson VUE delivery channel network <p>http://pearsonpte.com/TestMe/Taking/Pages/TestCentersandFees.aspx</p>
Cost	<ul style="list-style-type: none"> • The test fee is market priced and varies from US\$160 - 250
Registration	<ul style="list-style-type: none"> • The test is available on demand which means test takers are able to register online and schedule their test whenever a seat is available at a Pearson test centre. The test is not restricted to a fixed date schedule.
Access to Scores	<ul style="list-style-type: none"> • Test takers will access scores online, generally within 5 business days after the test • Institutions will access scores online and have access to a secure database of results • Secure online results service

3.0 The Registration Process:

Figure 2: PTE Academic Registration process



1) Registration: Is done online at www.pearsonpte.com where test takers can visit the booking website to create a user account. Once they have their web account details, they can search for a test centre and schedule and pay for the test online. Alternatively, test takers can call Pearson's regional call centres to register and book.

2) Results: Once the test taker has taken the test, results are generally available within 5 business days. Test takers are informed via email that their test results are ready and can be accessed online via their secure account. Once in their account, test takers can send their score report to up to seven institutions of their choice at a time. If they have taken the test on more than one occasion, all reports will be available for up to two years from the test date.

3) Institutions: Receiving institutions log on to the secure site and check results. They can download a PDF of a test taker's score report to print or store offline. The test taker must have given permission for the institution to access their results by assigning the scores from within their online account. Institutions can also receive score information via a data download.

4.0 PTE Academic Scores

Scores for PTE Academic are reported as a profile of your level of ability in listening, reading, speaking and writing, and in six Enabling Skills. Test takers can access this through their secure on-line user account.

Figure 3: Sample PTE Academic Score Report

LASTNAME FIRSTNAME Score: 56 090-08-97

PEARSON

Example Test Taker

Example Test Taker Score Report

Example, Test Taker

Address: 123 Example Lane
Example City, CA 92345
USA

Email Address: example@pearson.com

Telephone Number: 019 555 0100

Date of Birth: 12 August 1985

Country of Citizenship: USA

Gender: Male

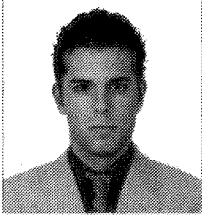
Registration ID: 123456789

Report Issue Date: 04 October 2009

Test Date: 30 September 2009

First-Time Test Taker: No

Score Valid Until: 29 September 2011

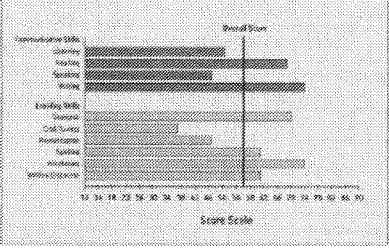


Overall Score: 56

The Overall Score for PTE Academic is based on the test taker's performance on all items in the test. The scores for Communicative Skills and Enabling Skills are based on the test taker's performance on all items. Score Range: 10-90. Score 56 is above the minimum score of 10. To view more details about the test taker's performance, click on the Overall Score. The Overall Score is based on the test taker's performance on all items in the test. The scores for Communicative Skills and Enabling Skills are based on the test taker's performance on all items. Score Range: 10-90. Score 56 is above the minimum score of 10. To view more details about the test taker's performance, click on the Overall Score. The Overall Score is based on the test taker's performance on all items in the test. The scores for Communicative Skills and Enabling Skills are based on the test taker's performance on all items. Score Range: 10-90. Score 56 is above the minimum score of 10. To view more details about the test taker's performance, click on the Overall Score.

Skills Profile

Category	Sub-Skill	Score
Communicative Skills	Listening	51
	Reading	69
	Speaking	47
	Writing	74
Enabling Skills	Grammar	70
	Oral Fluency	37
	Pronunciation	47
	Spelling	61
	Vocabulary	74
	Written Discourse	61



Sample score report

The score report has a number of different sections:

- The first section contains personal details and the photograph taken on the day of the test.
- In the middle of the report, the Overall Score is shown.
- The final section of the score report shows more detail of performance on the test.

The individual scores for listening, reading, speaking, and writing are shown, along with the scores for the six Enabling Skills.

Note: the test taker score report is for personal use only, and should not be accepted by recognizing institutions. Institutions should only accept scores that have been allocated to them by test takers via the website.

More information can be found in Appendix 1 Interpreting the PTE Academic Score Report

4.1 Automated scoring and standardisation

With PTE Academic, reliable and standardized grading will assist institutions with high-quality recruitment decisions. The test uses state-of-the-art technology for test scoring.

Test takers' verbal and written responses are scored using Pearson's proprietary, patented automated scoring technologies. Research shows that the automated scoring technology underlying PTE Academic produces scores comparable to human raters but with the precision, consistency and objectivity of a machine, ensuring maximum reliability and validity.

Hundreds of academic institutions, corporations and government agencies around the world rely on Pearson's proven automated scoring technologies to measure the abilities of students, staff or applicants.

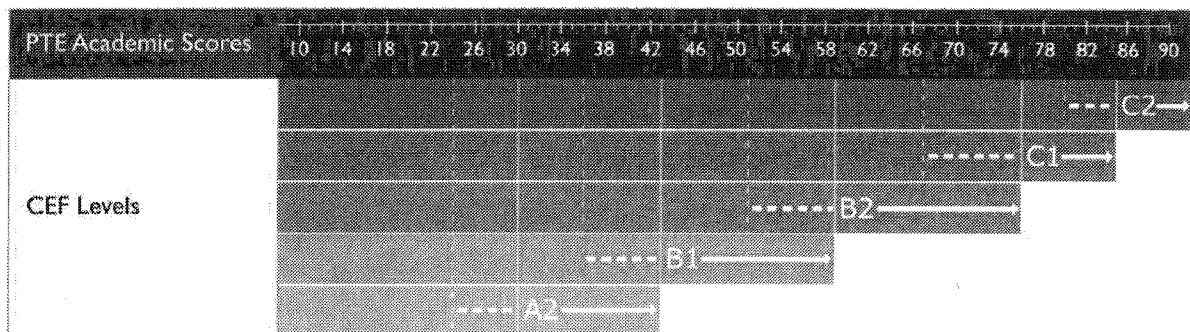
For more information see Appendix 2 PTEA_Automated Scoring

4.2 The PTE Academic Score Scale and the CEF

The test is aligned to the Common European Framework of Reference for Languages (CEF or CEFR; Council of Europe, 2001). This is a widely recognized benchmark for language ability.

The explanation of the alignment of PTE Academic to the CEF is that to stand a reasonable chance at successfully performing any of the tasks defined at a particular CEF level; learners must be able to demonstrate that they can do the average tasks at that level.

Figure 4: PTE Academic Score Scale and the CEF



The above diagram shows PTE Academic scores aligned to the CEF levels A2 to C2. The dotted lines on the scale show the PTE Academic score ranges that predict that test takers are likely to perform successfully on the easiest tasks at the next higher level. For example, if a test taker scores 51 on PTE Academic, this places them in the CEF B1 level. At the same time, it predicts that a PTE A test taker is likely to perform successfully on the easiest tasks at B2.

More information on the concordance with CEF and the level descriptors can be found in **Appendix 3 Preliminary Estimates of Concordance between Pearson Test of English Academic and other Measures of English Language Competencies**.

4.3 Concordance with other tests

Based on research, Pearson has produced concordance tables for PTE Academic with other major English language tests. These details can also be found in **Appendix 3** but a ready reference score comparison widget can be found at:

<http://www.pearsonpte.com/PTEAcademic/Pages/TestScores.aspx>.

4.4 Recommended score concordance with the International English Language Testing System (IELTS)

Table 1 demonstrates the relationship between scores on IELTS with that of PTE Academic across the most common score range. More comprehensive details are included in **Appendix 3**

Table 1: PTE Academic Concordance with IELTS

	IELTS score							
	4.0	4.5	5.0	5.5	6.0	6.5	7.0	7.5
PTE scores recommended by Pearson	Less than 29	30	36	42	50	58	65	73

4.5 Standard setting entry scores with the medical profession

In November 2009 the National Council of State Boards of Nursing (NCSBN®) conducted a standard setting study to establish a PTE Academic minimal passing standard, as related to entry-level nursing practice. NCSBN's recommendation regarding the English proficiency passing standard of an overall score of 55, with no sub-scores lower than 50, is consistent with the previous English proficiency passing standards recommended for other English language tests. Additionally, this recommendation is consistent with the NCSBN policy position on international nurse immigration, which states that "domestic and international nurses need to [be] proficient in written and spoken English skills" (NCSBN, 2003). Boards of nursing can use this information to make decisions regarding the level of English proficiency needed for entry-level RNs and LPN/VNs in their jurisdictions.

The full report, including the range of medical tasks assessed, can be found in **Appendix 4 NCLEX Technical Brief_PTE 2010**

5.0 Test Security

High stakes assessments are repeatedly under scrutiny from the end users and certain less scrupulous test takers. This means that all test developers should put the strictest security measures in place to ensure that the users of the scores can rely upon the validity of the test results.

Pearson uses a variety of cutting edge assessment security measures to ensure that all elements of test security are covered.


Randomized tests: Multiple test forms are distributed during any given test, making it highly unlikely that more than one test taker in the same room is using the same test form.

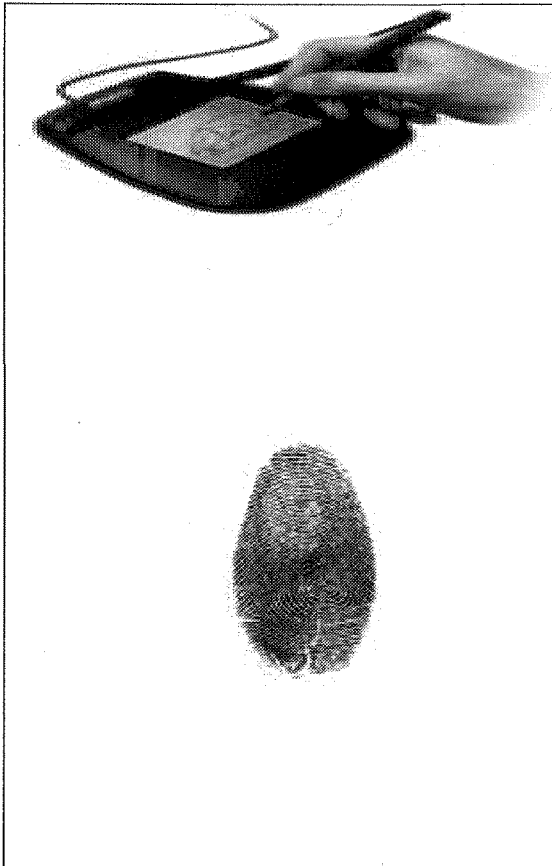
Secure item banking software: The item banking software used for PTE Academic is on a secure platform that stores all information relevant to every item in the bank.

No access to test material by centre staff: Prior to a test taker's appointment, the files and data required to admit the test taker and deliver the test are automatically downloaded to the test centre server from the Pearson VUE hub. This is done via a protected communications link and stored securely on the test centre file server for subsequent access by the admissions and test delivery software. Test centre staff cannot access the data held on the server.

No paper versions: PTE Academic is a computer-based test which eliminates the possibility of obtaining a copy of a test prior to taking it. It also means that even if a test taker memorizes the tasks they do in a test and passes on the information to other test takers globally, the likelihood of meeting those same items in another test is minimal. Memorization is a factor all test developers must take steps to combat, and the design of PTE Academic does this very effectively.

Figure 4: PTE Academic Biometric Security

	<p>Three types of digital biometrics: photograph, signature & palm-vein capture: Pearson VUE test centres are equipped with the most sophisticated identity verification capabilities used in the testing industry today. Test takers are asked to provide a digital signature and have their photograph taken, after which their palm vein pattern is captured and verified using state-of-the-art PalmSecure™ biometric technology. Advanced biometric matching is performed to verify test taker identity upon entry to and exit from the testing room.</p> <p>Video and proctor monitored test taking: Each individual testing workstation is monitored by digital audio and video surveillance. Test takers are also monitored by proctors seated in a station</p>
---	---



with a direct line of sight to each workstation.

Testing room arrangement: The testing rooms are arranged so that test takers cannot see each others' workstations, and those checking in cannot see into the testing room.

No personal belongings allowed into the testing room: Test takers are provided with secure lockers to store their personal items. They can take notes during the test using an erasable note board booklet and pen provided by the test centre.

Lockdown of computer workstation: The system locks the workstation desktop to prevent all activity except taking the test. Only one program on the workstation can run at any given time and short cut functions are disabled, so a test taker cannot attempt to download information or screen shots."

Interpreting the PTE Academic Score Report

This document provides key information for test takers about understanding the PTE Academic Score Report. It provides simple guidance on accessing and understanding scores.

More detailed information about scoring is available within the documents *Using PTE Academic Scores, Skills Tested and Scored in PTE Academic and Preliminary Estimates of Concordance Between PTE Academic and Other Measures of English Language Competencies*, available at pearsonpte.com/PTEAcademic/Pages/TestScores.

1. Get your scores

After taking the test, you will be notified by email when your PTE Academic scores are available (typically within five business days from your test date). Using the online account you set up when you registered for the test, you can view your results and then request to make them available to an unlimited number of institutions of your choice.

The PTE Academic Score Report includes your overall score, communicative skills scores and enabling skills scores.

Example Test Taker Score Report

Example Test Taker

Example Test Taker Score Report

Example Test Taker

Address: 123 Example Lane
Example City, 10 12345
USA

Email Address: example@pearson.com

Telephone Number: 312 333 0100

Date of Birth: 12 August 1988

Country of Citizenship: USA

Gender: Male

Registration ID: 123456789

Report Issue Date: 04 October 2009

Test Date: 30 September 2009

First-Time Test Taker: No

Scores Valid Until: 29 September 2011

Overall Score: 56

Skills Profile

Communicative Skills

- Listening: 51
- Reading: 69
- Speaking: 47
- Writing: 74

Enabling Skills

- Grammar: 70
- Oral Fluency: 37
- Pronunciation: 57
- Spelling: 93
- Vocabulary: 74
- Written Discourse: 63

Example Test Taker Score Report

1. The **overall score** reflects your overall English language ability. The score is based on performance on all items in the test.

The range for the overall score is 10-90 points.

2. **Scores for communicative skills** (listening, reading, speaking and writing) are based on all test items that assess these skills, either as a single skill or together with other skills.

The range for each communicative skill score is 10-90 points.

3. **Scores for enabling skills** (grammar, oral fluency, pronunciation, spelling, vocabulary and written discourse) are based on all test items assessing one or more of these skills.

The range for each enabling skill score is 10-90 points.

4. The display of the scores in a graph allows you to quickly see your strengths and weaknesses, and how each skill relates to your overall performance.

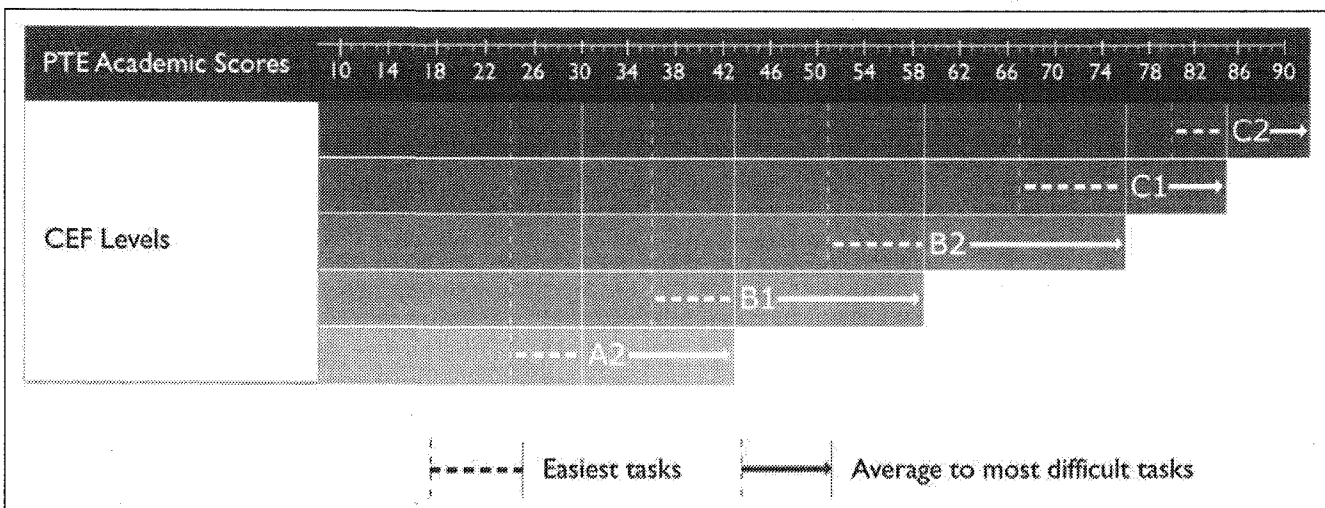
2. Understand the PTE Academic Score Scale and the CEF

PTE Academic is aligned to the Common European Framework (CEF or CEFR). This is a widely recognized benchmark for language ability developed by the Council of Europe (2001).

The explanation of this alignment is that to stand a reasonable chance at successfully performing any of the tasks defined at a particular CEF level, you must be able to demonstrate that you can do the average tasks at that level.

As you grow in ability, for example within the B1 level, you will become successful at doing even the most difficult tasks at that level and will also find you can cope with the easiest tasks at the next level. In other words, you are entering into the B2 level.

The table below shows PTE Academic scores aligned to the CEF levels A2 to C2. The dotted lines on the scale show the PTE Academic score ranges that predict that you are likely to perform successfully on the easiest tasks at the next higher level. For example, if you score 51 on PTE Academic, this means that you are likely to be able to cope with the more difficult tasks in the CEF B1 level. At the same time, according to your PTE Academic score, it predicts that you are likely to perform successfully on the easiest tasks at B2.



Preliminary alignment of PTE Academic scores to CEF levels

3. Find out what PTE Academic scores mean

PTE Academic alignment with the CEF can only be fully understood if it is supported with information showing what it really means to be 'at a level'. In other words, are you likely to be successful with tasks at the lower boundary of a level; do you stand a fair chance of doing well on any task, or will you be able to do almost all the tasks, even the most difficult ones, at a particular level? The table below shows for each of the CEF levels A2 to C2 which PTE Academic scores predict the likelihood of you performing successfully on the easiest, average and most difficult tasks within each of the CEF levels.

PTE Academic scores predicting the likelihood of successful performance on CEF level tasks			
CEF Level	Easiest	Average	Most Difficult
C2	80	85	NA
C1	67	76	84
B2	51	59	75
B1	36	43	58
A2	24	30	42

For example, if your PTE Academic score is 36, this predicts that you will perform successfully on the easiest tasks at B1. From 36 to 43, the likelihood of successfully performing the easiest tasks develops into doing well on the average tasks at B1. Finally, reaching 58 predicts that you will perform well at the most difficult B1 level tasks.

Use the table below to find out what PTE Academic scores in the range from 10 to 84 (A1 to C1) mean. The table shows the score ranges that predict some degree of performance at the next level, and it explains what you are likely to be able to do within those score ranges.

PTE Academic Score	Common European Framework Level	Level Descriptor © Council of Europe	What does this mean for a score user?
76 - 84	C1	Can understand a wide range of demanding, longer texts and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.	C1 is a level at which a student can comfortably participate in all post-graduate activities including teaching. It is not required for students entering university at undergraduate level. Most international students who enter university at a B2 level would acquire a level close to or at C1 after living in the country for several years, and actively participating in all language activities encountered at university.
59 - 75	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	B2 was designed as the level required to participate independently in higher level language interaction. It is typically the level required to be able to follow academic level instruction and to participate in academic education, including both coursework and student life.
51 - 58	Predicts success on easiest tasks at B2	Has sufficient command of the language to deal with most familiar situations, but will often require repetition and make many mistakes. Can deal with standard spoken language, but will have problems in noisy circumstances. Can exchange factual information on familiar routine and non-routine matters within his/her field with some confidence. Can pass on a detailed piece of information reliably. Can understand the information content of the majority of recorded or broadcast material on topics of personal interest delivered in clear standard speech.	
43 - 58	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst in an area where the language is spoken. Can produce simple connected text on topics, which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.	B1 is insufficient for full academic level participation in language activities. A student at this level could 'get by' in everyday situations independently. To be successful in communication in university settings, additional English language courses are required.
36 - 42	Predicts success on easiest tasks at B1	Has limited command of language, but it is sufficient in most familiar situations provided language is simple and clear. May be able to deal with less routine situations on public transport e.g., asking another passenger where to get off for an unfamiliar destination. Can re-tell short written passages in a simple fashion using the wording and ordering of the original text. Can use simple techniques to start, maintain or end a short conversation. Can tell a story or describe something in a simple list of points.	
30 - 42	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.	A2 is an insufficient level for academic level participation.
10 - 29	A1 or below	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.	A1 is an insufficient level for academic level participation.

© The copyright of the level descriptors reproduced in this document belongs to the Council of Europe.

Pearson Test of English Academic: Automated Scoring

Introduction

Universities, higher education institutions, government departments and other organizations are increasingly faced with the need for an English language proficiency test that will accurately measure the communication skills of international students in an academic environment. In response to this need, Pearson Test of English Academic (PTE Academic) has been developed. The new test from Pearson will reliably measure the reading, writing, listening and speaking abilities of test takers who are non-native speakers of English and who want to study at institutions where English is the principal language of instruction.

Launching globally in 2009, PTE Academic will be offered in collaboration with the Graduate Management Admission Council® (GMAC®). GMAC is well known worldwide as the owner of the Graduate Management Admission Test® (GMAT®). In addition, PTE Academic will be delivered in a phased approach through Pearson VUE's test centers in a variety of countries. Pearson VUE is the global leader in electronic testing for regulatory and certification boards, providing a full suite of services from test development to test delivery to data management.

As the worldwide leader in publishing and assessment for education, Pearson is using several of its proprietary, patented technologies to automatically score test takers' performance on PTE Academic. Academic institutions, corporations and government agencies around the world have selected Pearson's automated scoring technologies to measure the abilities of students, staff or applicants. Pearson customers using automated spoken and written assessments include eight of the 2008 Fortune Top 20 companies; 11 of the 2008 Top 15 Indian BPO companies; the U.S., German and Dutch governments; world sports organizations, such as the FIFA (organizers of the World Cup) and the Asian Games; major airlines and aviation schools; and leading universities and language schools.

An extensive field test program was conducted to test PTE Academic's test items and evaluate their effectiveness as well as to obtain the data necessary to train the automated scoring engines to evaluate PTE Academic items. Over the past 18 months, test data were collected from more than 10,000 test takers from 38 cities in 21 countries who participated in PTE Academic's field test. These test takers came from 158 different countries and spoke 126 different native languages, including (but not limited to) Cantonese, French, Gujarati, Hebrew, Hindi, Indonesian, Japanese, Korean, Mandarin, Marathi, Polish, Spanish, Urdu, Vietnamese, Tamil, Telugu, Thai and Turkish. The data from the field test were used to train the automated scoring engines for both the written and spoken PTE Academic items.

This paper provides a description of the automated scoring engines used to score PTE Academic and information about how and why they are accurate measures of written and spoken test taker performance.

Why automated scoring?

Research supports that, in many ways, automated scoring gives more analytical, objective results than humans do. Unlike human judgment, which is prone to be influenced by a variety of factors, an automated scoring system is impartial. This means that the system is not “distracted” by language-irrelevant factors such as a test taker’s appearance, personality or body language (as can happen in spoken interview tests). Such impartiality means that test takers can be confident that they are being judged solely on their language performance, and stakeholders can be confident that a test taker’s scores are “generalizable” – that they would have earned the same score if the test had been administered in Beijing, Brussels or Bermuda.

Also, automated scoring allows individual features of a language sample (spoken or written) to be analyzed independently, so that weakness in one area of language does not affect the scoring of other areas. Human raters often exhibit “transfer of judgment” from one area of language to another. For example, test takers who speak smoothly may be marked as proficient even though their grammar is very poor. Automated scoring, on the other hand, assesses the different language skills objectively.

When developing its automated scoring technologies, Pearson conducts “validation studies” to make sure that the machine’s scores are comparable to scores given by skilled human raters. In a validation study, a new set of test taker responses (never seen by the machine) is scored by both human raters and by the automated scoring system. During Pearson’s validation studies, when the human scores are compared with the machine scores, they are found to be similar. In fact, the difference between the human score and the machine score is so small that it is usually less than the difference between one human score and another human score. This is true for both written and spoken assessments.

Research shows that the automated scoring technology underlying PTE Academic produces scores comparable to those obtained from careful human experts who are trained to consider only relevant language skills. This means that the automated system “acts” like a human rater when assessing test takers’ language skills, but does so with the precision, consistency and objectivity of a machine.

Scoring written English skills

The written portion of PTE Academic is scored using the Intelligent Essay Assessor™ (IEA), an automated scoring tool that is powered by Pearson’s state-of-the-art Knowledge Analysis Technologies™ (KAT™) engine. Based on more than 20 years of research and development, the KAT engine automatically evaluates the meaning of text by examining whole passages. The KAT engine evaluates writing as accurately as skilled human raters using a proprietary application of the mathematical approach known as Latent Semantic Analysis (LSA). Using LSA (an approach that generates semantic similarity of words and passages by analyzing large bodies of relevant text) the KAT engine “understands” the meaning of text much the same as a human.

IEA can be tuned to understand and evaluate text in any subject area, and includes built-in detectors for off-topic responses or other situations that may need to be referred to human readers. Research conducted by independent researchers as well as Pearson supports IEA’s reliability for assessing knowledge and knowledge-based reasoning. IEA was developed more than a decade ago and has been used to evaluate millions of essays, from scoring student writing at elementary, secondary and university level, to assessing military leadership skills.

Intelligent Essay Assessor and PTE Academic

IEA automatically evaluates a test taker's writing skills and knowledge and can be trained to score any writing traits that humans can reliably score. It assesses the total content of a test taker's response, using responses that were previously scored by expert human readers as a guide.

When taking PTE Academic, test takers will be asked to write 200 – 300 word essays and 50 – 70 word summaries. When a response is submitted for scoring, the system will evaluate the meaning of the response, as well as mechanical aspects of the writing. The system compares the response with the large set of training responses, computes similarities, and assigns a score based on content, in part by placing the response in a category with the most similar training responses. Scoring the mechanical aspects of the writing occurs in much the same way. The system assesses each trait (grammar, structure and coherence, etc.) in the test taker response, compares it with the large set of training responses, and then ranks the response according to that trait.

For the training of IEA, more than 50,000 written responses (essays and summaries) were collected in the field test. These written responses were scored on a number of traits including content, formal requirements, grammar, vocabulary, general linguistic range, spelling, development, structure and coherence. All test takers' responses in the field test were first scored by two human raters, and then by a third human rater when the first two did not agree. The scores from these human raters served as input for training IEA.

Because test takers' written responses were assigned randomly to raters drawn from a pool of more than 200 from Australia, the United Kingdom and the United States, the machine is trained on a rich set of international human judgments. The result is a person-independent rating. Based on the scores for all the traits mentioned above, an overall measure of writing performance can be formed by summing the trait scores for each test taker across all of the written items. This measure can be formed for the human raters and for the machine-generated scores. The correlation of these overall scores on this measure between pairs of human raters was 0.87. The correlation between the human score and the machine-generated score was 0.88. The reliability of the measure of writing in PTE Academic is 0.89.

Scoring spoken English skills

The spoken portion of PTE Academic is automatically scored using Pearson's Ordinate technology. Ordinate technology is the result of years of research in speech recognition, statistical modeling, linguistics and testing theory. The technology uses a proprietary speech processing system that is specifically designed to analyze and automatically score speech from native and non-native speakers of English. In addition to recognizing words, the system locates and evaluates relevant segments, syllables and phrases in speech and then uses statistical modeling technologies to assess spoken performance.

To understand the way that the Ordinate technology is "taught" to score spoken language, think about a person being trained by an expert rater to score speech samples during interviews. First, the expert rater gives the trainee rater a list of things to listen for in the test taker's speech during the interview. Then the trainee observes the expert testing numerous test takers, and, after each interview, the expert shares with the trainee the score he or she gave the test taker and the characteristics of the performance that led to that score. Over several dozen interviews, the trainee's scores begin to look very similar to the expert rater's scores. Ultimately, one could predict the score the trainee would give a particular test taker based on the score that the expert gave.

This, in effect, is how the machine is trained to score only instead of one expert “teaching” the trainee, there are many expert scorers feeding scores into the system for each response, and instead of a few dozen test takers, the system is trained on thousands of responses from hundreds of test takers. Furthermore, the machine does not need to be told what features of the speech are important; the relevant features and their relative contributions are statistically extracted from the massive set of data when the system is optimized to predict human scores.

While no human listener is likely to be accustomed to more than 100 different foreign accents, the speech processor for PTE Academic has been trained on more than 126 different accents and can deal with all of these accents equally. If the speaker has a very heavy accent and would be assigned a low score by typical human raters, then this test taker will receive a low pronunciation score from the machine. Importantly, the poor pronunciation would not influence the test taker’s grammar or vocabulary scores.

Ordinate technology powers the Versant™ line of language assessments, which are used by organizations such as the U.S. Department of Homeland Security, schools of aviation around the world, the Immigration and Naturalization Service in the Netherlands, and the U.S. Department of Education. Independent studies have demonstrated that Ordinate’s automated scoring system can be more objective and more reliable than many of today’s best human-rated tests, including one-on-one oral proficiency interviews.

Ordinate technology and PTE Academic

The Ordinate scoring system collects hundreds of pieces of information from the test takers’ spoken responses, such as their pace, timing and rhythm, as well as the power of their voice, emphasis, intonation and accuracy of pronunciation. It also recognizes the words that the speakers select (even if they are mispronounced) and evaluates the content, relevance and coherence of the response. Because the system is sensitive to many hundreds of linguistic and acoustic features in each response, it is able to provide a very precise estimate of how a skilled human rater would score each component of the response if paying specific attention to the component in question.

PTE Academic field testing provided data to create the automated scoring models for the spoken part of the test, just as it did for the written part. Nearly 400,000 spoken responses from more than 10,000 test takers were collected. These included test takers’ spoken performances when describing figures or graphs, and re-telling lectures or presentations. Test takers’ responses were recorded and sent to human raters to be scored. Human raters scored test takers’ responses on a number of traits. The traits included content, vocabulary, language use, pronunciation, fluency and intonation. Aspects of the test takers’ responses, which were objectively observable by the advanced speech processing system, such as rate of speech, rhythm and word choice, were then compared with the raters’ scores. Scoring models were then built, which are used to predict how trained human raters would score any “new” incoming response. The correlation between the human scores and the machine scores for an overall measure of speaking was 0.96 thus proving the reliability of the measure of speaking in PTE Academic.

When taking PTE Academic, test takers will be required to respond verbally to various kinds of questions. Their spoken responses will be captured as audio files and analyzed by the patented Ordinate scoring system. Some test questions require short spoken responses. In these cases, the Ordinate scoring system measures the accuracy of the test taker’s word identification, pronunciation, fluency and grammatical facility. Other questions are more complex, with test takers providing longer, more elaborate responses requiring many sentences or paragraph-level utterances. In addition to the traits listed above, the automated scoring system provides content and vocabulary scores on these responses.

Conclusion

By combining the power of a comprehensive field test, in-depth research and Pearson's proven, proprietary automated scoring technologies, PTE Academic fits a critical gap by providing a state-of-the-art test that accurately measures the English language speaking, listening, reading and writing abilities of non-native speakers. Colleges, universities, government organizations and other institutions interested in becoming a PTE Academic-recognizing institution can visit www.pearsonpte.com or send an e-mail to the appropriate email address below for more information.

North and South America
usreco@pearson.com

Canada
canreco@pearson.com

United Kingdom and Ireland
ukireco@pearson.com

Europe, Middle East, African and India
emaireco@pearson.com

Asia-Pacific, Australia and New Zealand
apacreco@pearson.com

References

About Knowledge Analysis Technologies (KAT) Engine, Latent Semantic Analysis (LSA), and Intelligent Essay Assessor (IEA)

Calfee, R. (2000). To grade or not to grade. *IEEE Intelligent Systems* 15(5), 35–37.
<http://www.pearsonkt.com/papers/IEEEdebate2000.pdf>

Landauer, T.K., D. Laham, & P.W. Foltz. (2003). Automatic essay assessment. *Assessment in Education: Principles, Policy & Practice*, 10(3), 295-308.

Landauer, T.K., D. Laham, & P.W. Foltz. (2000). The Intelligent Essay Assessor. *IEEE Intelligent Systems* 15(5), 27–31.

Landauer, T.K., P.W. Foltz, P. & D. Laham. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes*, 25, 259–284.
<http://www.pearsonkt.com/papers/IntroLSA1998.pdf>

Landauer, J.K., & S.T. Dumais. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240.
<http://www.pearsonkt.com/papers/plato/plato.annotate.html>

Pearson (2008). Reliability and Validity of the KAT Engine. <http://pearsonkt.com/researchVRSum.shtml>

About Ordinate technology and Versant tests

Bernstein, J., J. De Jong, D. Pisoni, & B. Townshend. (2000). Two experiments on automatic scoring of spoken language proficiency. In P. Delcloque (Ed.), *Proceedings in InSTIL2000*, pp. 57–61. Dundee, Scotland: University of Abertay.
<http://pearsonpte.com/research/automatedscoring>

Harcourt (2006). *Predicting ICAO levels from Versant™ for English*. Author.
<http://harcourtassessment.com/hai/images/dotcom/vaet/ICAOPredictionFromVersant.pdf>

Kerkhoff, A., P. Poelmans, J. de Jong, & M. Lennig (2005). *Verantwoording Toets Gesproken Nederlands*. [Account of the Test of Spoken Dutch] Den Bosch: CINOP.

Pearson (2008). *Versant English Test: Test Description and Validation Summary*
<http://pearsonpte.com/research/automatedscoring>

Pearson (2008). *Versant Aviation English Test: Test Description and Validation Summary*. Author.
<http://pearsonpte.com/research/automatedscoring>

Pearson (2008). *Versant Spanish Test: Test Description and Validation Summary*
<http://pearsonpte.com/research/automatedscoring>

Pearson (2004). *Versant English Test: Can do Guide; Ordinate® SET-10®*. Author.
<http://pearsonpte.com/research/automatedscoringw>

Preliminary Estimates of Concordance Between PTE Academic and Other Measures of English Language Competencies

This document outlines the research carried out to produce preliminary estimates of concordance between PTE Academic and other measures of English language, including the Common European Framework (CEF or CEFR, Council of Europe, 2001), TOEFL® iBT and IELTS™.

Test comparisons using field test data

PTE Academic has been field tested using over 10,400 test takers. Field testing took place in 2007 and 2008. Test takers were representative of the global population of students seeking admission to universities and other tertiary education institutions where English is the language of instruction. Test takers were born in 158 different countries and spoke 126 different languages.

During the field tests several sets of secondary data were collected. Among these were ratings for all test takers on descriptive scales published by the Council of Europe (2001). In addition, a number of test takers reported their scores on other tests of English, including TOEIC, TOEFL PBT, TOEFL CBT, TOEFL iBT and IELTS.

A limited number of the self-reported data were invalid as the reported scores were outside the possible score range for the particular test. A small number of the test takers also submitted copies of their official score reports on the tests, for which they had provided self-reported data. Table 1 shows the following for each test: the numbers of self-reported data, how many of these were valid, the mean self-reported scores, the number of official score reports sent in, the mean official scores and the correlations with the PTE Academic field test scores. All correlations are significant at $p < .01^1$.

Test	Self-Reported Data				Official Score Report		
	N Total	N Valid	Mean	Correlation	n	Mean	Correlation
TOEIC	328	327	831.5	0.76	No data	-	-
TOEFL PBT	96	92	572.3	0.64	No data	-	-
TOEFL CBT	110	107	240.5	0.46	No data	-	-
TOEFL iBT	144	140	92.9	0.75	19	92.1	0.95
IELTS	2436	2432	6.49	0.76	169	6.61	0.73

Table 1: PTE Academic field tests: test takers on other tests of English

From Table 1 it can be concluded that the self-reported scores are, in general, quite accurate. Indeed, the correlation between the self-reported results and the official score reports was .82 for TOEFL iBT and .89 for IELTS. This finding is in agreement with earlier research on self-reported data. For example, Cassady (2001) found students' self-reported Grade Point Average (GPA) scores to be 'remarkably similar' to official records. The data are also consistent. According to ETS (2005, p.7) the score range 75–95 on TOEFL iBT is comparable to the score range 213–240 on TOEFL CBT and to the score range 550–587 on TOEFL PBT. The mean self-reported scores in Table 1 for these three tests are therefore comparable.

July 2011 V2

¹ Significant at $p < .01$ means there is less than 1% chance to observe this correlation if the measures are not related.

In addition, according to ETS (2001, p.3) a score range of 800–850 on TOEIC corresponds to a score range of 569–588 on TOEFL PBT, which makes the self-reported TOEIC mean score of the test takers on the PTE Academic field test also fall in line with data published by ETS.

Based on the data presented in Table 1, concordance between PTE Academic and other tests of English can be estimated, taking into account a less than optimal effort of test takers during field testing where test results have no direct relevance to the test takers.

Relation to the Common European Framework

The preliminary relation of the PTE Academic score scale with the descriptive scale of the Common European Framework for Languages (CEF) is based on both an item-centered and a test taker-centered method. For the item-centered method, the CEF level of all items was estimated by item writers, reviewed and, if necessary, adapted in the item-reviewing process. For the test taker-centered method, three extended responses (one written and two spoken) per test taker were each rated by two independent, trained raters. If there was a disagreement between the two independent raters, a third rating was gathered and the two closest ratings were retained. A dataset of over 26,000 ratings (by test takers self-reporting, by items and by raters) on up to 100 different items was analyzed using the computer program FACETS (Linacre, 1988; 2005). Estimates of the lower boundaries of the CEF levels, based on the item-centered method, correlated at .996 with those based on the test taker-centered method, which effectively means that the two methods yielded the same results except for less than 1% of error variance.

Validity check using BETA testing data

In addition to the initial field testing of 10,400 students during 2007–08, a further 364 test takers participated in the 2009 BETA testing of PTE Academic. The concordance between the score scale of PTE Academic and the score scales of TOEFL iBT and IELTS (each estimated from the field test data) were used as predictors of TOEFL iBT and IELTS scores of test takers participating in BETA testing. Test takers provided self-reported scores and a smaller, partially overlapping, number of test takers sent in copies of their official score reports.

Table 2 shows the mean scores as self-reported and from the official score reports, the mean scores for the same test takers as predicted from their PTE Academic score and the correlations between the reported scores and the predictions from PTE Academic. All correlations are significant at $p < .01^2$.

It can be concluded that this preliminary concordance produces fairly accurate and coherent predictions.

Test	Self-Reported Data				Official Score Report			
	n	Mean	Predicted	Correlation	n	Mean	Predicted	Correlation
TOEFL iBT	42	98.9	97.3	0.75	13	92.2	98.2	0.77
IELTS	57	6.80	6.75	0.73	15	6.60	6.51	0.83

Table 2: PTE Academic BETA: test takers on other tests of English

² Significant at $p < .01$ means there is less than 1% chance to observe this correlation if the measures are not related.

Concordance of PTE Academic with other measures of English

Based on the research described, Pearson has produced preliminary concordance tables. Table 3 (p.4) shows Pearson's current best estimate of concordance between PTE Academic scores and the CEF. In addition, shaded score ranges indicate the PTE Academic scores that predict some degree of performance at the next CEF level. More information is available in the document *Using PTE Academic Scores*, available at pearsonpte.com/PTEAcademic/Pages/TestScores.aspx.

Table 4 (p.5) shows the relation between scores on TOEFL iBT and PTE Academic. Table 5 (p.6) shows the relation between scores on IELTS and PTE Academic.

It must be noted that any attempt to predict a score on a particular test, based on the score observed on another test, will contain measurement error. This is caused by the inherent error in each of the tests in the comparison and in the estimate of the concordance. Furthermore, tests in the comparison do not measure exactly the same construct.

References

Cassady, Jerrell C. (2001) Self-Reported GPA and SAT Scores. ERIC Digest. ERIC Identifier: ED458216

Council of Europe (2001) Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Cambridge: CUP

ETS (2001) TOEFL Institutional Testing Program (ITP) and TOEIC Institutional Program (IP): Two On-Site Testing Tools from ETS at a Glance. Handout Berlin Conference 2001. Princeton: Educational Testing Service

ETS (2005) TOEFL® Internet-based test: Score comparison tables. Princeton: Educational Testing Service

Linacre, J.M (1988; 2005) A Computer Program for the Analysis of Multi-Faceted Data. Chicago, IL: Mesa Press

Preliminary estimates of concordance between PTE Academic and the descriptive scale of the CEF

PTE Academic Score	Common European Framework Level	Level Descriptor © Council of Europe	What does this mean for a score user?
>85	C2	Can understand with ease virtually everything heard or read. Can summarize information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.	C2 is a highly proficient level and a student at this level would be extremely comfortable engaging in academic activities at all levels
76 - 84	C1	Can understand a wide range of demanding, longer texts and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.	C1 is a level at which a student can comfortably participate in all post-graduate activities including teaching. It is not required for students entering university at undergraduate level. Most international students who enter university at a B2 level would acquire a level close to or at C1 after living in the country for several years, and actively participating in all language activities encountered at university.
59 - 75	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	B2 was designed as the level required to participate independently in higher level language interaction. It is typically the level required to be able to follow academic level instruction and to participate in academic education, including both coursework and student life.
51 - 58	Scores in this range predict success on the easiest tasks at B2	Has sufficient command of the language to deal with most familiar situations, but will often require repetition and make many mistakes. Can deal with standard spoken language, but will have problems in noisy circumstances. Can exchange factual information on familiar routine and non-routine matters within his/her field with some confidence. Can pass on a detailed piece of information reliably. Can understand the information content of the majority of recorded or broadcast material on topics of personal interest delivered in clear standard speech.	
43 - 58	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst in an area where the language is spoken. Can produce simple connected text on topics, which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.	B1 is insufficient for full academic level participation in language activities. A student at this level could 'get by' in everyday situations independently. To be successful in communication in university settings, additional English language courses are required.
36 - 42	Scores in this range predict success on the easiest tasks at B1	Has limited command of language, but it is sufficient in most familiar situations provided language is simple and clear. May be able to deal with less routine situations on public transport e.g., asking another passenger where to get off for an unfamiliar destination. Can re-tell short written passages in a simple fashion using the wording and ordering of the original text. Can use simple techniques to start, maintain or end a short conversation. Can tell a story or describe something in a simple list of points.	
30 - 42	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.	A2 is an insufficient level for academic level participation.
10 - 29	A1 or below	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.	A1 is an insufficient level for academic level participation.

Table 3: Concordance estimate PTE Academic and CEF

© The copyright of the level descriptors reproduced in this document belongs to the Council of Europe.

Preliminary estimates of concordance between PTE Academic and TOEFL iBT

TOEFL iBT Score	PTE A Score	TOEFL iBT Score	PTE A Score
No data	85 - 90	90	61
120	84	89	60
119	83	87-88	59
118	82	86	58
117	81	85	57
115-116	80	83-84	56
114	79	82	55
113	78	81	54
112	77	79-80	53
110-111	76	78	52
109	75	76-77	51
107-108	74	74-75	50
106	73	72-73	49
105	72	70-71	48
103-104	71	67-69	47
102	70	65-66	46
101	69	63-64	45
99-100	68	60-62	44
98	67	57-59	43
97	66	54-56	42
95-96	65	52-53	41
94	64	48-51	40
93	63	45-47	39
91-92	62	40-44	38
		No data	10 - 37

Table 4: Concordance estimate PTE Academic and TOEFL iBT

Preliminary estimates of concordance between PTE Academic and IELTS

IELTS Score	PTE A Score
9.0	86 - 90
8.5	83 - 85
8.0	79 - 82
7.5	73 - 78
7.0	65 - 72
6.5	58 - 64
6.0	50 - 57
5.5	42 - 49
5.0	36 - 41
4.5	29 - 35
No data	10 - 28

Table 5: Concordance estimate PTE Academic and IELTS

Setting an English Language Proficiency Passing Standard for Entry-Level Nursing Practice Using the Pearson Test of English Academic

Ada Woo, PhD
Senior Psychometrician, NCLEX® Examinations

Philip Dickison, PhD, RN
Director, NCLEX® Examinations

John de Jong, PhD
Vice President, Test Development, Pearson Language Tests



Pearson Test of English (PTE) Academic is a new computer-based English language proficiency test launched in October 2009 that will be available in 35 countries and more than 200 locations in 2010. Anticipating widespread use of PTE Academic, the National Council of State Boards of Nursing (NCSBN®) conducted a standard-setting study in November 2009 to establish a PTE Academic minimal passing standard, as related to entry-level nursing practice.

Similar standard-setting exercises targeting entry-level nursing practice have been performed on other English language proficiency tests. In 2003, NCSBN conducted its first English language proficiency workshop with the Test of English as a Foreign Language Computer-based Test (TOEFL® CBT). An updated TOEFL passing standard was set as a new version of the test, TOEFL Internet-based Test (TOEFL iBT™), became the primary choice for TOEFL candidates in a standard-setting workshop in 2008. In 2004, a similar standard-setting workshop was conducted with the International English Language Testing System (IELTS™) examination.

NCSBN is providing its member boards with another option for evaluating English language proficiency of licensure applicants by setting a minimum English proficiency requirement with PTE Academic. Given the number of internationally educated nurses who seek licensure in the U.S., defining a legally defensible passing standard, documenting the procedures used to identify that standard, and making the standard available to boards of nursing and other interested parties would be a benefit to member jurisdictions. An additional benefit to using the NCSBN recommended English proficiency standard is that it would make PTE Academic pass-fail decisions portable across member jurisdictions. This technical brief describes the PTE Academic standard-setting process and the NCSBN recommended passing standard.

PTE Academic

PTE Academic was developed to measure the English communication skills of nonnative English-speaking students in an academic environment. Owned and developed by Pearson, the test was cross-validated in several countries, including the United Kingdom, Australia and the U.S., to ensure that the test is representative of and measures international academic English.

PTE Academic measures four English language skills in four corresponding sections: Writing, Speaking, Reading and Listening. The test's format is shown in Table 1.

Part	Content	Time Allowed
Intro	Introduction	Not timed
Part 1	Speaking and Writing	77-93 minutes
Part 2	Reading	32-41 minutes
Optional Scheduled Break		10 minutes
Part 3	Listening	45-57 minutes

PTE Academic is aligned with the Common European Framework (CEF) (Council of Europe, 2001), a set of widely used benchmarks for language ability, made up of six consecutive levels of language ability defined by descriptors of language competencies (Appendix A). In terms of score reporting, PTE Academic results are expressed in three types of scores. The Overall Score reflects the overall English language ability of a test taker. Scores for Communicative Skills include listening, reading, speaking and writing. Scores for Enabling Skills include grammar, oral fluency, pronunciation, spelling, vocabulary and written discourse. The PTE Academic score scale ranges from 10 to 90.

Selection for the PTE Academic Standard-Setting Panel

The composition of a standard-setting panel is a key element in determining the validity of the standard-setting process. For the PTE Academic Standard-Setting Workshop, 19 subject matter experts were recruited from the four NCSBN geographic areas in order to provide a representative sample of the nursing profession and consumers of nursing services. The panel also includes a diverse range of demographic and practice characteristics following these

selection criteria: (1) nurses who speak a primary language other than English and have taken an English proficiency test; (2) nurses who work with clients who speak a primary language other than English; (3) clinical nurse supervisors of nurses who speak a primary language other than English; (4) nursing regulators; (5) nursing educators; and (6) consumers of nursing services. In all, the standard-setting panel represented 13 NCSBN boards of nursing, six languages and a variety of nursing specialties. Details of panelist composition are listed in Table 2.

Panelist Characteristic	Category	%	(N)
Gender	Female	89	17
	Male	11	2
Nursing license ¹	Registered nurse (RN)	82	14
	Licensed practical/vocational nurse (LPN/VN)	18	3
NCSBN geographic region	I	21	4
	II	37	7
	III	21	4
	IV	21	4
Years of post licensure experience ²	1 to 10 years	24	4
	11 to 20 years	29	5
	21 to 30 years	18	3
	31 to 40 years	24	4
	41 or more years	6	1
Primary language	English	47	9
	Spanish	26	5
	Tagalog	11	2
	Arabic	5	1
	Chinese	5	1
	French	5	1
Has taken an English proficiency test	Yes	42	8
	No	58	11
Job title	Staff nurse	32	6
	Supervisor	21	4
	Board of nursing regulator	21	4
	Nursing educator	21	4
	Director	5	1
Nursing specialty ³	Nursing education	24	4
	Emergency medicine	18	3
	Medical/Surgical	12	2
	Mental health	12	2
	Geriatric	6	1
	Nursing regulation	6	1
	Obstetrics	6	1
	Pediatrics	6	1
	Preoperative care	6	1
	Rehabilitation	6	1
	Rheumatology	6	1

¹ Two panelists, representing consumers of nursing services, do not hold nursing licenses.

² Among the 17 panelists who are licensed nurses, average years of post-licensure experience is 22.8 years (SD=12 years).

³ Several panelists listed more than one specialty area; therefore sum of percentages exceeds 100%.

Standard-Setting Panelist Orientation and Training

A set of preparatory materials were sent to panelists prior to the workshop. The materials included a brief description of PTE Academic, the NCLEX-RN® and NCLEX-PN® test plans, and a preworkshop assignment that asked panelists to consider entry-level nursing activities that require English skills in each of the four areas measured by PTE Academic. Panelists were also given access to a PTE Academic online practice test prior to the standard-setting workshop. They were encouraged to complete the practice test in order to familiarize themselves with different item types on PTE Academic.

When panelists arrived at the workshop, they were given an overview of PTE Academic and the test development process. In addition to receiving training on standard-setting methodology, panelists also received extensive education on CEF, which covered basic principles underlying CEF, the hierarchy of CEF scales and the interpretation of CEF levels.

Standard-Setting Procedures

PTE Academic consists of four sections: Writing, Speaking, Listening and Reading. The standard-setting panel recommended a separate passing standard for each. Prior to the start of the standard-setting exercise, workshop facilitators surveyed panel members on whether language requirements for entry-level registered nurses (RNs) and licensed practical/vocational nurses (LPN/VNs) are comparable. This was an important discussion because if the panel judged that language requirements for entry-level RNs and LPN/VNs are distinct, two separate passing standards would need to be set in the workshop. After the discussion, the panel concluded that the English competency level required to practice entry-level nursing tasks safely and effectively should be the same across all licensure types.

At the beginning of each section, panelists participated in a discussion of how much English proficiency in each language skill was required to safely and effectively perform entry-level nursing activities. From the discussion, panelists generated a list of tasks related to that specific language skill deemed necessary for safe and effective entry-level nursing practice (Appendix B). After reaching a consensus, panelists reviewed actual PTE Academic items and made recommendations for a passing standard.

There were two standard-setting methods used to set passing standards for the four sections. For the productive skills (Writing and Speaking sections), panelists independently reviewed response profiles given by test-takers at various score levels (ranging from basic English users to proficient English users). The responses reviewed were obtained from actual PTE candidates on pretest items. Panelists were asked to consider whether an entry-level nurse with minimally acceptable English language skills would be able to answer

the items similar to one of the sample candidates. For the receptive skills (Listening and Reading sections), panelists reviewed PTE Academic items and made judgments on each. The panelists considered whether an entry-level nurse with minimally acceptable language skills would be able to answer the item under review correctly.

Within both standard-setting methods, panelists were asked to provide two rounds of ratings. In the first round, panel members estimated the absolute ability of candidates or the absolute difficulty of items in relation to the coherent system of level descriptors defined in CEF. In the second round, the panel members estimated the ability of candidates or the difficulty of items in relation to the hypothesized linguistic requirements for entry-level nurses. Panel members were asked to indicate whether they thought the ability displayed in a response (Speaking and Writing) or the ability required to respond (Listening and Reading) represented a minimal level of required English proficiency for an entry-level nurse.

Standard-Setting Panel Recommendations

In all four sections, panelists were able to reach consensus on the recommended standard through group discussion. A summary of panel recommendations on section scaled scores is listed in Table 3. According to conversion data provided by Pearson, the PTE Academic passing standards recommended by the standard-setting panel are comparable to the NCSBN endorsed TOEFL and IELTS passing standards. For TOEFL iBT, a minimal total score of 84 and a minimum speaking score of 26 has been advised. A score of 84 on TOEFL iBT corresponds with a score of 56 on PTE Academic. For IELTS, the minimum requirement has been set at 6.5 (corresponding to 58 on PTE Academic) with no communicative skill below 6.0 (which corresponds with 50 on PTE Academic).

Table 3. Summary of Panel Recommendations

Reading	Writing	Listening	Speaking	Overall
50	50	50	50	55

NCSBN Board of Directors (BOD) Deliberation

The NCSBN BOD reviewed and discussed results from the PTE Academic Standard-Setting Workshop. They reviewed background on PTE Academic, along with English proficiency benchmarks on CEF and procedures used in the standard-setting workshop. The BOD then considered passing standards recommended by the PTE Academic panel, as well as the potential pass/fail impact of these recommendations. After weighing all available evidence, including comparability to other NCSBN endorsed English proficiency standards and the BOD's own knowledge about the nursing profession, they concurred with the panel's recommendation of an overall passing standard of 55 with no individual section below 50.

Conclusions

NCSBN's recommendation regarding the English proficiency passing standard of an overall score of 55 with no subscores lower than 50 is consistent with the previous English proficiency passing standards recommended for other English language tests and the level of English proficiency needed to perform entry-level nursing safely and effectively. Additionally, this recommendation is consistent with the NCSBN policy position on international nurse immigration, which states that "domestic and international nurses need to [be] proficient in written and spoken English skills" (NCSBN, 2003). Boards of nursing can use this information to make decisions regarding the level of English proficiency needed for entry-level RNs and LPN/VNs in their jurisdictions.

References

- Angoff, W. H. (1971). Scales, norms and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 508-600). Washington, DC: American Council on Education.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: CUP.
- National Council of State Boards of Nursing. (2003). National Council of State Boards of Nursing (NCSBN) Position Statement International Nurse Immigration. Retrieved July 20, 2009, from <https://www.ncsbn.org/1272.htm>
- Taube, K. T. (1997). The incorporation of empirical item difficulty data in the Angoff standard-setting procedure. *Evaluation and the Health Professions, 20*, 479-498.

Appendix A: Overview of the Common European Framework (CEF)

The *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (CEFR) is a guideline used to describe achievements of learners of foreign languages across Europe. It was established by the Council of Europe between 1986 and 1996 as the main part of the project "Language Learning for European Citizenship." Its main purpose is to provide educators, students and employers a common scale for comparing and contrasting language ability across all languages in Europe. In November 2001, a European Union Council Resolution recommended using CEFR to set up systems of validation of language ability. The six reference levels (see below) are becoming widely accepted as the European standard for grading an individual's language proficiency.

PTE Academic is aligned (preliminary estimate) with the Common European Framework (CEF), which is a widely recognized benchmark for language ability developed by the Council of Europe (2001) and other international English language tests. CEF includes a set of consecutive language levels defined by descriptors of language competencies. These commonly defined levels enable language learners, teachers, universities and employers to compare and relate language qualifications by level.

CEF describes language proficiency along a number of scales. Each scale refers to a particular aspect of language use. The scales of descriptors have been drawn up on the basis of empirical research.

- A1 – Makes simple statements on personal details and very familiar topics
- A2 – Relates basic information on work, background, family, free time, etc.
- B1 – Relates comprehensibly the main point he/she wants to make on familiar matters
- B2 – Relates information and points of view clearly and without noticeable strain
- C1 – Shows fluent, spontaneous expression in clear, well-structured speech
- C2 – Conveys finer shades of meaning precisely and naturally

A person at the border between level A2 and B1 or B2 will probably do 80% of the tasks at level A2, but can also do 50% of the tasks at level B2, because the underlying scale is continuous.

The relation between PTE Academic scores and the CEF scale is shown in the figure below.

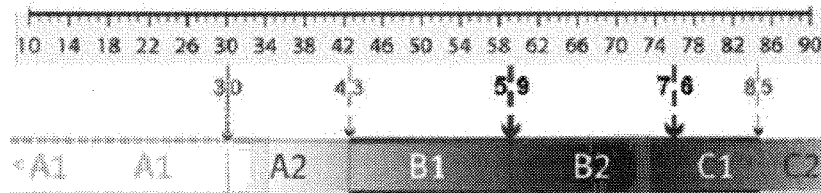


Table 1. Common Reference Levels: Global Scale		
Proficient user	C2	Can understand with ease virtually everything heard or read. Can summarize information from different spoken and written sources, and reconstruct arguments and accounts in a coherent presentation. Can express himself/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts and recognize implicit meaning. Can express himself/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organizational patterns, connectors and cohesive devices.
Independent user	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can handle most situations likely to arise while traveling in an area where the language is spoken. Can produce simple connected text on topics that are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions, and briefly give reasons and explanations for opinions and plans.
Basic user	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g., very basic personal and family information, shopping, local geography, employment). Can perform simple and routine tasks following a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar, everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce himself/herself and others, and can ask and answer questions about personal details, such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly, and is prepared to help.

Appendix B: Comprehensive Entry-Level English Task List Generated by Panelists

WRITING TASKS BY CATEGORY				
Documentation	Evaluation	Communicating with Health Care Workers	Nursing Interventions	Continuing Education and Training
Patient's status in health record	Taking notes on patient assessment	Communicating with health care workers	Instructions to patient/family	Test-taking skills (writing)
Progress note	Explaining in clear detail what the nurse observes	Notes to the charge nurse/supervisor	Developing a care plan and treatment plan	
Patient's change in condition		Agenda for staff meetings (nursing supervisor)	Contributing to care plan or patient specific instructions	
Vital signs/lab results		Transcribing a physician's order	Discharge instructions	
Notes related to postprocedure result		Line staff assignment		
Receiving or giving report over the phone		Unit report		
Narrative notes on a patient's chart		Documentation of quality assurance for blood glucose strips or urine dipsticks		
Medication administration record		Communicate accurately, clearly by e-mail		
Completion of ID bands for mom/baby in obstetrics		Taking health care providers' orders		
Completion of blood/blood products forms before and after infusion		Incident reports		
Computer documentation (electronic and paper records)		Provides instructions to unlicensed assistive personnel (UAPs) in writing		
		Notes (formal or informal) to other health care providers		
		An evaluation		

SPEAKING TASKS BY CATEGORY			
Assessment	Teaching	Interdisciplinary Education	Therapeutic Communication
Ask client questions when status change is noted	Any patient teaching	Give oral report to next shift	Offer support, hope, compassion to a grieving family
Asking patient about pain levels	Reinforce discharge teaching	Give report/updates to other health care professionals	Phrase questions to patients to elicit more than just a yes/no response
Obtain a health history	Explain new medications or disease process	General communication with colleagues regarding patients or work flow	
	Orient patient to room upon admission	Taking assessment	
	Provide patient/family with instruction/directions	Interacting with staff	
	Explain procedures	Able to articulate pertinent information	
	Instruction on diet	Explain accurately medical jargon into common language	
		Taking telephone orders/communicate via telephone with physician	

LISTENING TASKS BY CATEGORY				
Interacting with Health Care Team	Obtaining Information Over the Phone	Interacting with Clients	Participating in Continuing Education (CE) or Professional Development	Therapeutic Communication
Obtaining information from a supervisor	Obtaining a verbal or telephone physician's order	Identifying a patient/family's most urgent concerns	Attends CE session	Uses therapeutic conversations/counseling
Identifying pertinent information from verbal clinical report	Uses telephone properly	Understands questions posed by patient/family regarding care or instructions	Learns new procedures	Understands description of symptoms
Listens to shift report (verbal or dictated)	Obtains lab reports by phone		Educational inservices or programs	Understands pain level is what patient says it is
Verbal instructions			Participates in and contributes to team conferences	Recognize physical therapy needs
Listen to what other nurses are saying regarding treatment options				
Confirm physician orders				
Giving report to oncoming shift				

READING TASKS BY CATEGORY		
Interdisciplinary Communication	Education/Professional Development	Patient Safety
Read notes of other health care professionals/interdisciplinary notes	Reads nursing journal for ongoing professional development	Allergy lists
Able to read client chart	Understand reference materials as necessary	Dosage measurements (weights/heights)
Interoffice communication	Read equipment instructions	Read the medication/treatment administration record
Patient care plans	Policy/procedures manual for the organization	Check ID bands
Reads shift assignment and unit schedule	Drug inserts/Physicians' Desk Reference (PDR)/labels for drug information	Understand abbreviations used within the organization
History and physical	Information regarding diagnostic test or diagnosis which they are unfamiliar with	Understands signs found within hospital
Physician orders	Patient education materials	

©2010 The National Council of State Boards of Nursing (NCSBN®) is a not-for-profit organization whose members include the boards of nursing in the 50 states, the District of Columbia and four U.S. territories — American Samoa, Guam, Northern Mariana Islands and the Virgin Islands. There are four associate members.

Mission: The National Council of State Boards of Nursing (NCSBN), composed of Member Boards, provides leadership to advance regulatory excellence for public protection.



National Council of State Boards of Nursing, Inc.
111 E. Wacker Drive, Suite 2900, Chicago, IL 60601-4277
312.525.3600 | International Calls: +1.312.525.3688
Testing Services: 312.525.3750 | www.ncsbn.org