

財團法人語言訓練測驗中心



英語能力測驗

成績統計報告

(民國91~92年)

測驗研發組

中華民國九十三年十一月

目次

一、前言	2
二、測驗方式與說明	2
三、各項測驗成績統計	4
四、各項測驗成績分布	5
五、成績百分等級	7
六、測驗目的與成績分析	8
七、應試者背景資料及成績分析	10
八、與「全民英檢」中級、中高級成績之比較	15
九、信度與測驗標準誤	16
十、效度與相關性	18
附錄	20
參考文獻	26

一、前言

本中心外語能力測驗 (Foreign Language Proficiency Test, FLPT) 自民國五十四年研發完成，至今已逾四十年。本測驗廣為公、民營機構採用，作為人才遴選、陞遷、海外受訓及開會時，鑑定外語能力的工具及標準。目前 FLPT 提供英、日、法、德、西班牙語五種外語能力測驗，每兩個月施測三次，測驗人口平均每年三千人左右。目前採用 FLPT 測驗的政府機構包括行政院人事行政局、經濟部等各級公家單位，台灣電力公司、中華電信研究所、中國石油公司等國營事業機構，及工研院、各銀行等民營機構。自八十八年度起，教育部公費留學考試¹、台灣大學教育學程及數所大學研究所及醫學院亦採用本測驗作為申請或報考資格。

為能提供各送考機構充分的測驗訊息，作為選才或政策制定的參考，本中心自民國八十五年起，根據參加 FLPT 英語能力測驗應試者的性別、年齡、主修科系及學歷等背景，及聽力、用法、字彙與閱讀及口試四項測驗的表現進行各項分析。本報告將就民國九十一年一月至民國九十二年十二月參加 FLPT 英語能力測驗的 4,106 人次考生²，根據其報名表及問卷上所填各項背景資料與其各項成績表現進行統計分析，並與本中心於八十九年開辦之「全民英檢」中級及中高級筆試成績比較，期能提供各送考單位機構在訂定標準時更詳盡的資訊。

二、測驗方式與說明

FLPT 外語能力測驗為一標準化³常模參照測驗⁴ (standardized, norm-referenced test)，測驗項目包含筆試及口試，不分階段，一次施測完畢。筆試含聽力測驗、用法測驗及字彙與閱讀測驗三項，均為四選一之選擇題，採電腦閱卷。各項測驗分別計分，採「標準分數」，以 60 分為平均數，每一標準差加或減 20 分。每項測驗分數範圍為 0 至 120 分。口試部分為開放式題型，採錄音方式，於語言實習室中進行。評分採用整體式評分量表，由兩位評分人員獨立評分，差距在一級分內 (含) 者，取兩道平均；大於一級分者，交付資深評分人員複評，決定最後成績。級分範圍為 0 至 5 級分。詳細筆試、口試評分方式及成績說明請參見附錄一、二。

¹ 根據教育部八十八年之公費留學考試留學國語文公告，外語能力測驗合格標準為聽力、用法、字彙與閱讀平均六十分以上、口試成績 S-2 以上。

² 以四項測驗均有成績之應試考生作為統計基礎。

³ Borg & Gall 指出標準化測驗應具有以下特徵：客觀性、施測條件一致、有根據百分等級編製而成的常模及有信度、效度 (Borg & Gall, 1989)。

⁴ 依分數詮釋方式，測驗可分為兩種：一為常模參照測驗 (norm-referenced test)，如 TOEFL；一為效標參照測驗 (criterion-referenced test)，如全民英檢。常模參照測驗中，應試者的得分並非原始分數，而是一與事先定義好的常模比較、轉換後所產生的標準分數，其分數高低代表應試者在此常模的相對位置 (王文科，1999)。

聽力測驗共有 50 題，分問答、釋義及談話三部分，測驗時間約 30 分鐘，試題均由錄音機播出，美式發音，讀題速度為常速⁵。應試者在聽完問題後，由試題冊上 A、B、C、D 四個備選答案中，選出最適合的作答，答錯不倒扣。每題塗答時間十五秒。聽力測驗旨在評量應試者是否具備日常生活所需的一般英語聽解能力，包括聽解問句並了解其回應模式的能力，聽解直述句並了解其間接或直接傳達的意涵及了解連續對話與短文之主旨、大意及細部資訊的能力。

用法測驗因試題別不同，有 80 題及 100 題兩種，皆為單項式⁶試題 (discrete-point item)，測驗時間分別為 40 與 50 分鐘。試題冊上，每題含有一空格的句子及四個選項，應試者須由四個選項中選出最適合題意的字或詞填入，使其成為一完整的單句。此部分測驗旨在評量應試者對於一般常用或慣用的英語文法及句構的了解。

字彙與閱讀測驗共 60 題，分字彙與閱讀兩部分，測驗時間為 60 分鐘。字彙部分有填充題及同義詞兩種題型，為單項式試題。閱讀部分有六至八篇文章，每篇文章長度在 80 至 250 字之間，為題組型式，每篇有二至四個問題。此部分測驗旨在評量應試者是否具備利用上下文推測字詞意義的能力，以及閱讀時能否有效運用閱讀技巧，掌握文章主旨、大意及細部資訊的能力。

口試分唸句、中翻英、問答、申論及看圖說話等五部分，測驗時間約 15 分鐘。除第三部分題目由錄音機播出外，其餘則印在試卷上，回答經由麥克風錄在錄音帶上。第一部分唸句共有五題，作答時間 45 秒，旨在測驗應試者是否具備清晰可解的英語發音。第二部分中翻英亦有五題，作答時間一分鐘，旨在評量應試者能否以英語表達一般生活中常用的社交用語或句構。第三部分問答共有 10 題，第 1 至 5 題每題有 15 秒回應時間，第 6 至 10 題每題有 30 秒回應時間，旨在評量應試者是否能以英語應答有關個人經歷、工作及時事等問題。第四部分申論題及第五部分看圖說話，試題冊上分別印有一討論主題及一張圖片，應試者各有 1.5 分鐘思考及 1.5 分鐘回答時間。前者旨在評量應試者是否能針對一抽象話題進行較深入的探討及意見表達；後者旨在評量應試者能否就圖中的人、事、物做具像描述。

⁵ Pimsleur, et al. (1977) 將英語說話速度劃分成以下五類：fast 為每分鐘 220 字以上，moderately fast 為每分鐘 190 至 220 字，average 為每分鐘 160 至 190 字，moderately slow 為每分鐘 130 至 160 字，slow 為每分鐘 130 字以下。

⁶ 單項式試題 (discrete-point item) 係針對單一、特定的語言項目進行檢測，譬如時態、詞類、單複數等。其優點為可採樣的語言點較廣，且受試生在任一單題的作答結果，不會連帶影響其在其他題的表現，造成連續扣分 (double penalty) 的情形。

四項測驗時間、題型與題數如表一，各題型試題樣例請參閱附錄三。

表一、FLPT 英語能力測驗 測驗時間、題型與題數表

測驗項目	施測方式	部分	題型	題數	總題數	測驗時間
聽力測驗	紙筆	第一部分	問答	15 題	50 題	約 30 分鐘
		第二部分	釋義	15 題		
		第三部分	談話	20 題		
用法測驗	紙筆	文法、句型結構		80/100 題	80/100 題	40/50 分鐘
字彙與閱讀測驗	紙筆	第一部分	字彙：填充題	20 題	60 題	60 分鐘
			字彙：同義詞	20 題		
		第二部分	閱讀：6~8 篇	20 題		
口試	錄音	第一部分	唸句	5 題	22 題	約 15 分鐘
		第二部分	翻譯	5 題		
		第三部分	問答題	10 題		
		第四部分	申述題	1 主題		
		第五部分	看圖說話	1 主題		

三、各項測驗成績統計

聽力測驗 4,106 人次考生平均成績為 62.57 分，標準差 21.86 分，中間值⁷ 63 分，眾數⁸ 76 分，峰度 (kurtosis) -0.78，即聽力成績分布曲線較常態分布平坦，考生間的程度差異稍大。偏態值 (skewness) -0.07，趨近常態分布。

用法測驗考生平均成績 61.79 分，標準差 19.29 分，中間值為 62 分，眾數為 60 分，峰度 -0.53，偏態值 -0.15，用法成績分布曲線屬常態分布。

字彙與閱讀測驗考生平均成績 61.38 分，標準差 21.24 分，中間值為 62 分，眾數為 69 分，峰度 -0.75，偏態值 -0.10，成績分布曲線與聽力測驗相似。

口試測驗，為便於分析統計，各級分成績轉換為以下數值：

- ◆ S-0 級 = 0
- ◆ S-0+ 級 = 5
- ◆ S-1 級 = 10
- ◆ S-1+ 級 = 15
- ◆ S-2 級 = 20
- ◆ S-2+ 級 = 25
- ◆ S-3 級 = 30
- ◆ S-3+ 級 = 35
- ◆ S-4 級 = 40
- ◆ S-4+ 級 = 45
- ◆ S-5 級 = 50

⁷ 所有考生成績排序中，位於第百分之五十位置者的成績。

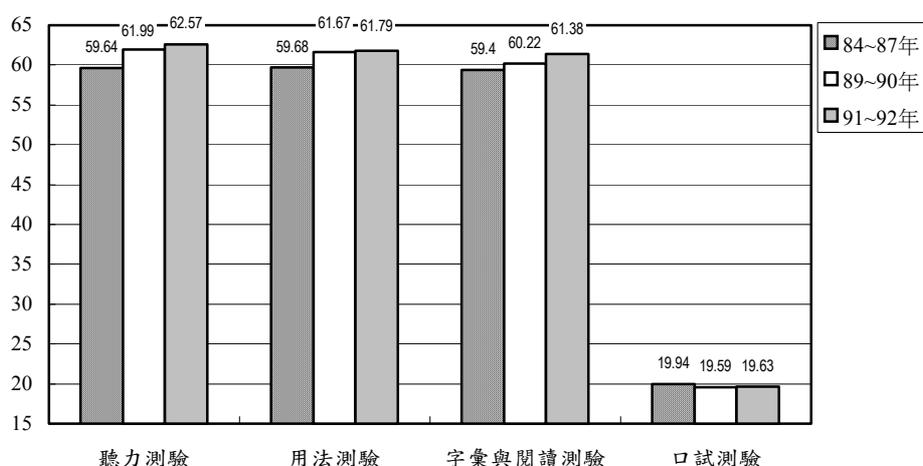
⁸ 所有考生成績中，最多人得到的分數。

以上述定義，91~92 年間所有口試考生之平均分數、中間值及眾數均為 20（相當 S-2），標準差為 4.90（約為半個級分），峰度 0.13、偏態-0.03，口說成績分布曲線呈常態分布。四項成績統計表請參考表二。

表二、各項測驗成績統計表

項目	聽力測驗	用法測驗	字彙與閱讀測驗	口試測驗
平均分數	62.57	61.79	61.38	S-2 (19.63)
標準差	21.86	19.29	21.24	4.90
中間值	63	62	62	20
眾數	76	60	69	20
峰度	-0.78	-0.53	-0.75	0.13
偏態	-0.07	-0.15	-0.10	-0.03

與民國 84~87 及 89~90 年的統計比較，聽力、用法及字彙與閱讀測驗三項平均分數雖有 1 至 3 分的小幅成長，但與本次四項測驗平均成績相較，並無顯著差異 ($p>0.05$)。84~92 年間各項測驗平均成績比較請參閱圖一。



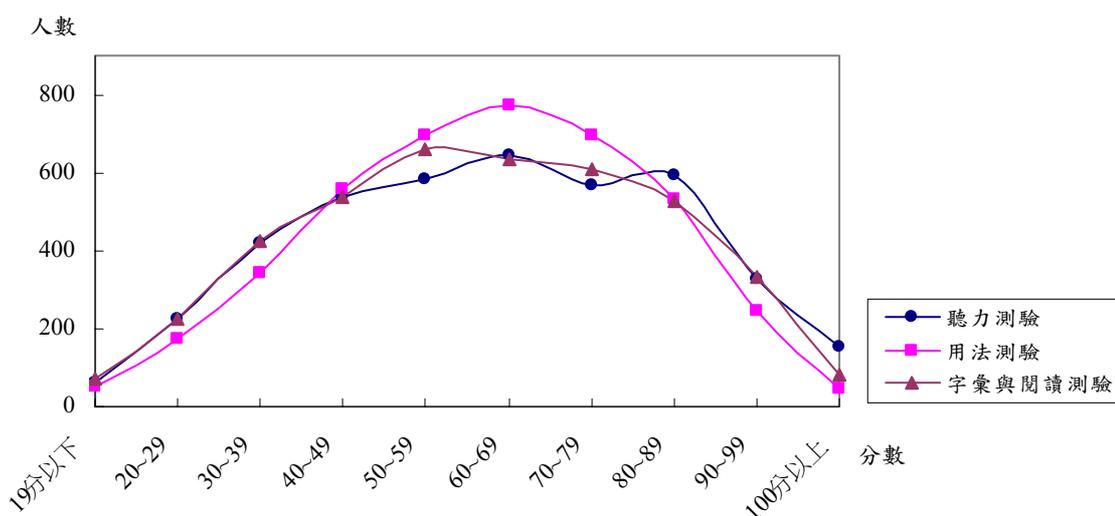
圖一、84~92 年間各項測驗平均成績比較圖

四、各項測驗成績分布

依各項筆試成績分布分析，聽力測驗及用法測驗 60 分以上考生人次均佔全部考生人次之 56%，字彙與閱讀測驗略低，佔 53%。請參見表三及圖二筆試各項測驗成績分布表、分布圖。

表三、筆試各項測驗成績分布表

分數範圍	聽力測驗			用法測驗			字彙與閱讀測驗		
	人數	人數	累計%	人數	人數	累計%	人數	人數	累計%
100 分以上	151	151	3.7%	47	47	1.1%	82	82	2.0%
90~99	326	477	11.6%	246	293	7.1%	331	413	10.1%
80~89	592	1069	26.0%	530	823	20.0%	527	940	22.9%
70~79	569	1638	39.9%	694	1517	36.9%	609	1549	37.7%
60~69	642	2280	55.5%	771	2288	55.7%	635	2184	53.2%
50~59	585	2865	69.8%	697	2985	72.7%	662	2846	69.3%
40~49	537	3402	82.9%	555	3540	86.2%	537	3383	82.4%
30~39	419	3821	93.1%	344	3884	94.6%	426	3809	92.8%
20~29	223	4044	98.5%	172	4056	98.8%	223	4032	98.2%
19 分以下	62	4106	100.0%	50	4106	100.0%	74	4106	100.0%



圖二、筆試各項測驗成績分布圖

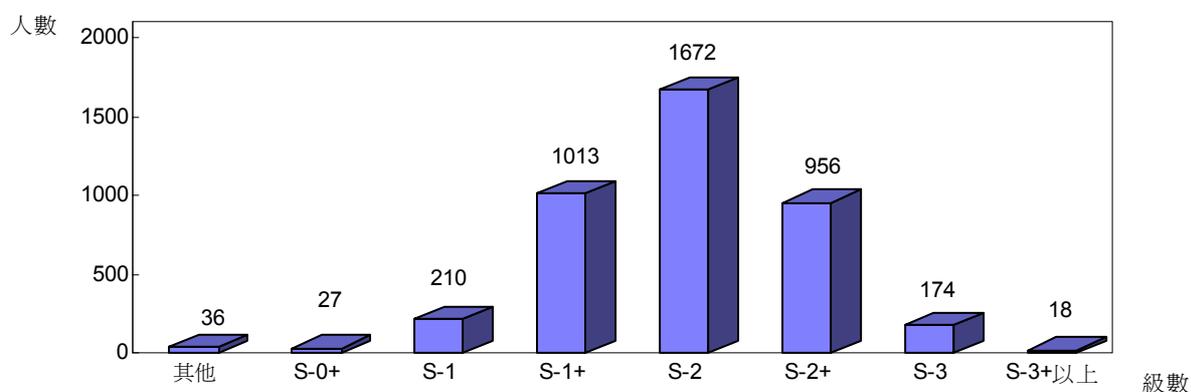
口試各級分人數統計顯示，以 S-2 級分者所佔比例最高，佔總人次 41%；其次為 S-1+及 S-2+，分別佔總人次 25%及 23%。詳細口試成績人數分布表、分布圖請參見表四、圖三。

進一步分析三項筆試與口試成績分布情形，三個人數分布最多的口試級數(1) S-2 的三項筆試平均成績為聽力 63.89 分、用法 62.26 分、字彙與閱讀 62.37 分；(2) S-1+ 為 42.19、47.09 及 44.39；(3) S-2+ 為 82.79、77.53 及 78.81。三個口說級數之筆試平均成績差距落在 15 至 22 分之間，將近一個標準差。請參見表四口試成績與筆試平均成績統計表。

表四、口試成績與筆試平均成績統計表

口試成績	人數	累計人數	累計%	聽力測驗	用法測驗	字彙與閱讀測驗
S-3+以上	18	18	0.4%	101.50	96.56	98.33
S-3	174	192	4.7%	95.74	89.41	91.13
S-2+	956	1148	28.0%	82.79	77.53	78.81
S-2	1672	2820	68.7%	63.89	62.26	62.37
S-1+	1013	3833	93.4%	42.19	47.09	44.39
S-1	210	4043	98.5%	29.74	35.16	31.86
S-0+	27	4070	99.1%	23.89	22.89	23.63
其他	36	4106	100.0%	77.92	69.25	68.81

註：其他類為未採用 FLPT 標準口試者。



圖三、口試成績人數分布直條圖

五、成績百分等級

百分等級(Percentile Rank, PR)是將所有考生的標準分數，轉化為一百級均等分數，應試者可藉由百分等級表評估個人在此群落中之排名。以表五為例，如某一考生三項筆試成績均落在 60~69 分成績範圍間，口試成績為 S-2，則可推知該考生聽力、用法、字彙與閱讀測驗及口試的 PR 值分別為 52、54、55 及 52，各數值代表該生各項測驗表現高過曾於 91~92 兩年參加 FLPT 英語測驗應試者之百分比。民國 91~92 年 FLPT 英語筆試及口試測驗成績百分等級表請參閱表五。

與 89~90 年的百分等級表比較，本次 PR 值的浮動不大，皆在 3% 以內。

表五、民國 91~92 年 FLPT 英語筆試及口試測驗成績百分等級表

筆試成績	聽力測驗 百分等級	用法測驗 百分等級	字彙與閱讀測驗 百分等級	口試成績	口試測驗 百分等級
100 分以上	98	99	99		
90~99	92	96	94	S-3+以上	99
80~89	81	86	84	S-3	97
70~79	67	72	70	S-2+	84
60~69	52	54	55	S-2	52
50~59	37	36	39	S-1+	19
40~49	24	21	24	S-1	4
30~39	12	10	12	S-0+	1
20~29	4	3	5		
10~19	1	1	1		
0~9	0	0	0		

六、測驗目的與成績分析

依應試者「參加測驗目的」分析，4,106 人次考生中，以 A 類「一般機構送考」最多，有 3,751 位，佔總人次 91% 以上；其餘依序為 B 類「申請公費留考」共 268 位，佔 7%；及 C 類「申請台大教育學程」共 87 位，佔 2%。

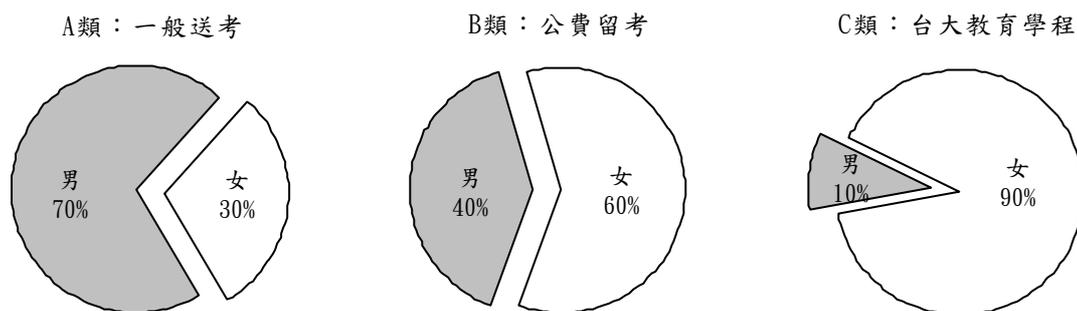
三組成績表現，以 C 類考生平均成績最高，各項平均成績分別為聽力 84.68 分、用法 82.78 分、字彙與閱讀 78.85 分、口試 S-2+ (24.89)。此類考生的平均年齡為 25.09 歲；女性居多，近九成；科系背景以「英語相關主修」最多，近七成，請參見表六測驗目的與成績統計表。

B 類平均成績居次，分別為聽力 69.06 分、用法 68.57 分、字彙與閱讀 68.21 分、口試約 S-2 (20.97)。此類考生較 C 組年長，平均年齡為 29.25 歲；男女比例為四比六；各類別主修科系均有，其中以主修文、史、哲學、教育相關科系最多，80 人，約佔此類考生人數三成。

A 類考生包含 306 所公、民營機構送考之新進員工、陞遷、受訓或外派考察人員，為 FLPT 主要對象考生。其各項平均成績分別為聽力 61.59 分、用法 60.82 分、字彙與閱讀 60.49 分、口試接近 S-2 (19.42)。此類考生平均年齡最長，為 35.41 歲；男性比例亦較 B、C 類高出甚多，男女比例約為七比三；主修科系則以商業及工程相關科系最多。

表六、測驗目的與成績統計表

類別	測驗目的	人數	人數%	聽力	用法	字彙與閱讀	口試
A	一般送考	3751	91.4%	61.59	60.82	60.49	S-2 (19.42)
B	公費留考	268	6.5%	69.06	68.57	68.21	S-2 (20.97)
C	台大教育學程	87	2.1%	84.68	82.78	78.85	S-2+(24.89)



圖四、ABC 三類應考生男女比例圖

與民國 89~90 年間同類應考生的成績相較，公、民營機構送考人數增加六百名以上，其四項測驗平均成績亦小幅提升 1 至 3 分；公費留考與台大教育學程兩組報考人數及成績皆呈下降趨勢。此或與公、民營機構近年來積極提升國際競爭力，強調員工英語能力，國內大學、研究所增加，選擇出國深造人數下滑，以及中小學英語教職對大學生的吸引力降低等因素有關。各類考生四項成績與民國 89~90 年應考生的比較詳參表七~九。

表七、民國 89~92 年「一般送考生」四項測驗成績比較

民國年	人數	聽力	用法	字彙與閱讀	口試
91~92 年	3751	61.59	60.82	60.49	S-2 (19.42)
89~90 年	3146	59.28	58.96	57.64	S-2 (18.92)
進/退步情形	605	2.31	1.86	2.85	

表八、民國 89~92 年「公費留考生」四項測驗成績比較

民國年	人數	聽力	用法	字彙與閱讀	口試
91~92 年	268	69.06	68.57	68.21	S-2 (20.97)
89~90 年	555	70.95	70.41	69.2	S-2 (21.85)
進/退步情形	-287	-1.89	-1.84	-0.99	

表九、民國 89~92 年「台大教育學程考生」四項測驗成績比較

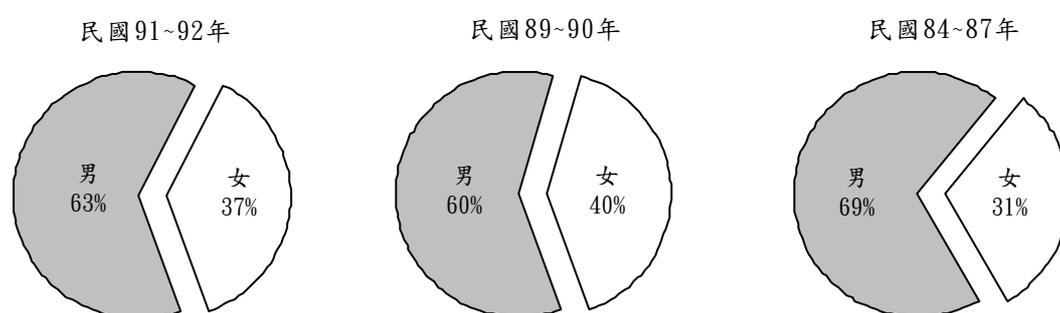
民國年	人數	聽力	用法	字彙與閱讀	口試
91~92 年	87	84.68	82.78	78.85	S-2+ (24.89)
89~90 年	158	84.44	84.97	80.07	S-2+ (24.84)
進/退步情形	-71	0.24	-2.19	-1.22	

七、應試者背景資料及成績分析

以下針對應試者之性別、年齡、主修、學歷等四項主要人口特徵 (demographic features) 與其語言能力表現做交叉統計分析比較。

7.1 性別

4,106 人次應試者中，男性考生 2,603 人次，佔全部考生人次六成三，女性考生 1,503 人次，佔三成七。本次女性考生人口比例較民國 89~90 年略減，但仍較民國 84~87 年高。詳參圖五、歷年應考生男女比例。



圖五、歷年應考生男女比例圖

成績表現方面，男性考生各項平均成績分別為聽力 59.94 分、用法 59.90 分、字彙與閱讀 60.77 分，口試約為 S-2 (19.20)，三項筆試平均成績相當。女性考生聽力 67.11 分、用法 65.07 分、字彙與閱讀 62.45 分，口試 S-2 (20.39)，三項筆試平均成績以聽力成績表現最好，其次為用法及字彙與閱讀。

表十、男女考生平均成績統計表

性別	人數	人數%	聽力	用法	字彙與閱讀	口試
男性	2603	63.4%	59.94*	59.90*	60.77	S-2* (19.20)
女性	1503	36.6%	67.11*	65.07*	62.45	S-2* (20.39)

* $p < 0.05$

此外，女性之聽力、用法及口試顯著高於男性考生，字彙與閱讀平均成績則無顯著差異。請參考表十男女考生平均成績統計表。此現象與 84~87 及 89~90 年之統計結果一致。

7.2 年齡

依考生年齡分析，全部考生平均年齡為 34.79 歲，女性平均年齡較輕，33.06 歲；男性平均年齡較長，35.78 歲。以 5 歲為一級距統計，則 30~34 歲組考生人次最多，佔全部考生人次 25.2%，其次為 25~29 歲組 22% 及 35~39 歲組 19%。請參見下表。

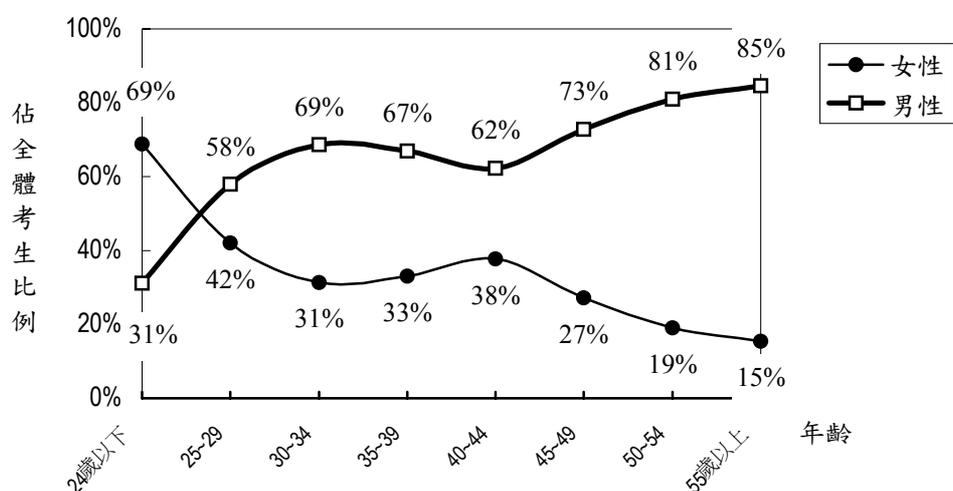
表十一、各年齡層平均成績統計表

年齡	人數	人數%	聽力	用法	字彙與閱讀	口試
24 歲以下	302	7.4%	80.76	77.46	74.97	S-2 (21.92)
25~29	910	22.2%	62.10	60.17	57.88	S-2 (19.66)
30~34	1034	25.2%	61.34	61.17	60.65	S-2 (19.69)
35~39	781	19.0%	59.04	60.09	60.18	S-2 (18.37)
40~44	538	13.1%	60.24	60.35	60.69	S-2 (18.91)
45~49	327	8.0%	64.19	62.19	64.74	S-2 (19.06)
50~54	162	4.0%	62.48	60.50	63.00	S-2 (19.69)
55 歲以上	52	1.3%	56.33	53.37	57.35	S-2 (18.56)

成績表現方面，以 24 歲以下考生表現最好，各項平均成績分別為聽力 80.76 分、用法 77.46 分、字彙與閱讀 74.97 分、口說 S-2 (21.92)，顯著高於其他年齡層考生 ($p < 0.05$)。另一值得注意的現象為 30~34 歲 (含) 以前各組的聽力平均成績皆高於用法、字彙與閱讀；30~34 歲以後，各組原居末的字彙與閱讀表現躍居三項之首。請參見表十一各年齡層平均成績統計表。

根據性別與年齡交叉分析顯示，除「24 歲以下」組女性應考比例較男性高外 (7:3)，其餘各組皆以男性比例較高，且比例隨年齡增長呈正比，「50 歲以上」各組男性比例高達八成以上，請參見圖六各年齡層男女應考生比例圖。

此比例與民國 89~90 年的統計結果類似。此現象可能解釋有二：一為各送考機構 25 歲以上的青壯年成員以男性為主，故而送考比例較女性高；或者，各送考機構或因職能、家庭因素等考量，較傾向送考男性職員。



圖六、各年齡層男女應考生比例圖

7.3 主修科系

根據考生資料卡所填寫之背景資料分析，4,106 名考生中，以主修管理、財經等商業相關科系者最多，佔全部考生人數 34.8%；其次為工科相關科系，佔 28.2%。該二組應試者即佔去所有考生六成以上的比例，此現象應與現今國家經濟發展政策及各公、民營機構經營理念有密不可分的關係。

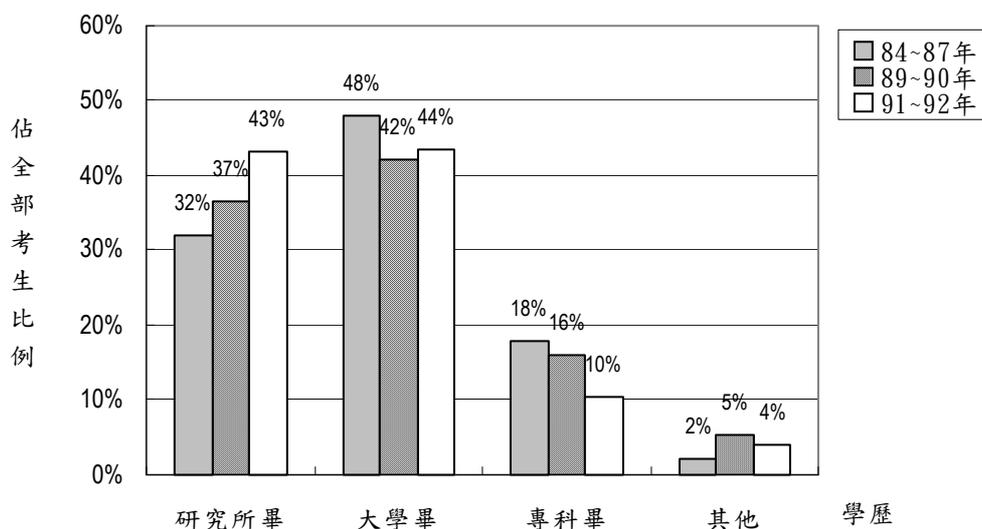
成績表現方面，以主修醫學相關科系的表現最好，各項平均成績分別為聽力 78.76 分、用法 77.63 分、字彙與閱讀 79.54 分、口試接近 S-2+ (22.81)；其次為文科（文、史、哲、教育等），四項平均成績依序為 75.32、72.76、71.05 及 S-2+ (22.96)。其他學科背景應試者成績表現請參見表十二各主修科系平均成績統計表。

7.4 最高學歷

依最高學歷分析，4,106 名考生中，以「研究所畢」及「大學畢」者最多，各佔全部考生人次 43%左右。與民國 82~90 年的應考人數比較，「研究所畢」者有大幅增加的趨勢，由民國 82~84 年的 25%、84~87 年的 32%、89~90 年的 37%至 91~92 年的 43%，研究所人口增加達 18%以上。此現象應與政府近十年的教育改革，大學與研究所的增設，國人教育程度大幅提升有關（圖七）。

表十二、各主修科系平均成績統計表

科系所	人數	人數%	聽力	用法	字彙與閱讀	口試
文	384	9.4%	75.32	72.76	71.05	S-2+ (22.96)
法	362	8.8%	72.89	69.74	71.70	S-2 (21.70)
商	1428	34.8%	60.10	59.89	58.64	S-2 (18.85)
理	288	7.0%	58.60	59.11	58.51	S-2 (19.05)
工	1158	28.2%	55.50	55.29	55.24	S-2 (18.26)
醫	219	5.3%	78.76	77.63	79.54	S-2+ (22.81)
農	92	2.2%	64.72	64.75	65.79	S-2 (19.46)
藝術	25	0.6%	75.16	70.44	68.72	S-2 (22.00)
體育	24	0.6%	65.04	61.25	57.54	S-2 (19.79)
其他	126	3.1%	63.20	61.61	60.81	S-2 (20.56)



圖七、民國 84~92 年應試者學歷背景分析圖

成績方面，以填寫「高中畢」考生表現最好，各項平均成績分別為聽力 75.89 分、用法 73.79 分、字彙與閱讀 72.24 分、口試 S-2 (20.50)。進一步分析發現，填寫「高中畢」者多為大學肄業生，而非已就業人士。「研究所畢」各項平均成績分別為聽力 64.61 分、用法 63.91 分、字彙與閱讀 64.87 分、口試 S-2 (20.26)。其中，字彙與閱讀測驗及口試兩項，「研究所畢」表現顯著高於「大學畢」者。其他組群成績表現請參考表十三最高學歷各項平均成績統計表。

表十三、最高學歷各項平均成績統計表

學歷	人數	人數%	聽力	用法	字彙與閱讀	口試
研究所畢	1777	43.3%	64.61	63.91	64.87	S-2 (20.26)
大學畢	1790	43.6%	63.74	62.81	61.69	S-2 (19.91)
專科畢	425	10.4%	45.85	45.74	42.89	S-1+ (15.72)
高職畢	3	0.1%	26.67	24.67	24.67	S-1 (8.33)
高中畢*	111	2.7%	75.89	73.79	72.24	S-2 (20.50)

* 主要為大學肄業者。

上表顯示在學者的英語能力最佳，此似意涵應試者在就學期間較能保持較佳的英語學習狀態。就業後，或因工作、家庭、興趣等因素，疏於練習，英語能力有下滑趨勢。此現象與民國 84~87 及 89~90 年的統計結果一致。

另發現本次的大學肄業生，即填寫「高中畢」者，筆、口試四項成績表現較民國 89~90 年同類應試者平均成績明顯下降，將近 4~7 分左右(表十七)；其次為「專科畢」，筆試三項平均成績下降 2 分左右。此現象或與國內大學法修改，部分原被歸類於技職體系的應試者，因改制而於 91~92 年統計中被歸類於大學組有關。除上述兩組在 89~92 年間成績變化較大外，其餘各類組四項成績表現變動幅度不大，詳參表十四~十七。

表十四、民國 84~92 年「研究所畢考生」四項測驗成績比較

民國年	聽力	用法	字彙與閱讀	口試
91~92 年	64.61	63.91	64.87	S-2 (20.26)
89~90 年	64.83	64.72	64.55	S-2 (20.14)
84~87 年	64.24	65.23	66.51	S-2 (20.40)
89~92 年進/退步情形	-0.22	-0.81	0.32	
84~92 年進/退步情形	0.37	-1.32	-1.64	

表十五、民國 84~92 年「大學畢考生」四項測驗成績比較

民國年	聽力	用法	字彙與閱讀	口試
91~92 年	63.74	62.81	61.69	S-2 (19.91)
89~90 年	62.48	61.83	60.46	S-2 (19.71)
84~87 年	61.01	61.11	60.22	S-2 (19.70)
89~92 年進/退步情形	1.26	0.98	1.23	
84~92 年進/退步情形	2.73	1.7	1.47	

表十六、民國 84~92 年「專科畢考生」四項測驗成績比較

民國年	聽力	用法	字彙與閱讀	口試
91~92 年	45.85	45.74	42.89	S-1+(15.72)
89~90 年	48.02	48.31	44.43	S-1+(16.55)
84~87 年	48.11	46.59	45.42	S-1+(17.10)
89~92 年進/退步情形	-2.17	-2.57	-1.54	
84~92 年進/退步情形	-2.26	-0.85	-2.53	

表十七、民國 84~92 年「高中/職畢考生」四項測驗成績比較

民國年	聽力	用法	字彙與閱讀	口試
91~92 年高中畢	75.89	73.79	72.24	S-2 (20.50)
89~90 年大學肆	81.95	80.87	77.19	S-2+(24.26)
84~87 年高中/高職畢	57.52	55.72	54.32	S-2 (18.30)
89~92 年進/退步情形	-6.06	-7.08	-4.95	-0.5 級分

註：民國 84~87 年高中/職考生（含大學肆業者）成績合併計算，其分類基礎與 89~90 年及 91~92 年差異過大，故不列入比較。

八、與「全民英檢」中級、中高級成績之比較

民國 91 至 92 年間的 FLPT 考生中，曾報考本中心另一項測驗—全民英檢（GEPT）中級及中高級初試，且與 FLPT 受測時間相隔在一年之內者，共有 451 名。其中「FLPT-GEPT 中級初試⁹」有 207 位，110 位達通過標準，通過率為 53%；「FLPT-GEPT 中高級初試」有 244 位，83 位達通過標準，通過率 34%。

451 名考生測驗成績的迴歸分析結果顯示，FLPT 英語能力測驗三項平均 60~69 分約相當於中級初試兩項平均 90~99 分，中高級初試兩項平均 70~79 分。請參見表十八 FLPT 英語能力測驗與 GEPT 中級、中高級成績比較。

另，讀者在使用表十八時，宜將下列因素納入考量，裨以做出對送考單位及應考者最適當的決定：

- (1) FLPT 測驗的主要對象考生為成年、在職的英語學習者，其心理成熟度及生活經歷與 GEPT 測驗的對象考生，即高中、大學在校生並不完全相同。

⁹ 「全民英檢」中高級及中級初試包括聽力及閱讀能力測驗兩項，均為選擇題，滿分 120 分，初試通過標準為兩項測驗均達 80 分。

- (2) FLPT 為常模參照測驗， GEPT 為效標參照測驗，兩測驗在命題及分數詮釋上不盡相同，使用者不宜將兩項測驗結果或能力值直接等化，做相同詮釋。
- (3) 兩項測驗的建構 (construct) 並不完全相同，例如，FLPT 含用法測驗，GEPT 則無；GEPT 有 cloze 題型，FLPT 則無。因此用法能力可以充分反映在 FLPT 的成績中，但並無法充分反映在 GEPT 的聽力或閱讀成績中。反之亦然，考生在 cloze 部分所展現的綜合能力，可以直接反映在 GEPT 的閱讀成績中，但無法由 FLPT 聽力、用法或字彙與閱讀成績中直接推估。
- (4) 451 名考生雖皆於一年內接受兩項測驗，但並非同期施測或隨機抽樣。因此比較表中數值尚包含兩項測驗施測中間可能產生的考生能力變化與評量誤差，決策者在採用時，宜將這些變數一併納入考量。

表十八、FLPT 英語能力測驗與 GEPT 中級、中高級成績比較

FLPT 英語能力測驗 (三項平均)	GEPT 中級初試 (兩項平均)	GEPT 中高級初試 (兩項平均)
100 分以上		107 分以上
90~99	116 分以上	99~106
80~89	110~115	90~98
70~79	100~109	80~89
60~69	90~99	70~79
50~59	80~89	60~69
40~49	70~79	50~59
30~39	60~69	40~49

九、信度與測驗標準誤

信度(reliability)係指考生在不同的時間，接受相同或難度、內容相當之試卷(parallel forms)，所得到的測驗分數的一致性 (Anastasi, 1988, p.109)。如果兩次施測所得到測驗分數的一致性高，表示該測驗的分數具穩定性、可靠性與可預測性 (郭生玉，1997，p.44)。

估計信度的方法有重測法 (test-retest method)、複本法 (equivalent-forms method) 與內部一致性法 (internal consistency method)。重測法與複本法需在短時間內施測兩次，易受應試者動機、疲勞與學習因素等影響。因此一般多採只施測一次的「內部一致性信度」，此係根據應試者答題情況分析試題內容的一致性。當試題內容同質性較高時，其所得到的信度亦較高。當試題內容異質性較高時，即測驗內容包含多個以上特質，例如不同領域 (如數學、歷史或文學等)，其內部一致性信度則愈低。

內部一致性信度常用 KR20、KR21 或 Alpha 值來表示。本中心 FLPT 英語能力測驗聽力測驗 50 題，內部一致性信度介於 0.85~0.92 之間，平均信度為 0.88；用法測驗題數 80 或 100 題，信度介於 0.89~0.93 之間，平均信度為 0.91；字彙與閱讀能力測驗 60 題，信度介於 0.83~0.91 之間，平均信度為 0.88。

測量標準誤 (Standard Error of Measurement, SEM) 係指應試者因受到非測驗相關因素影響，譬如，受測時身體突然不適或手機鈴響等，而導致實際分數 (observed score) 與真實分數 (true score) 之間產生的差異值；或指應試者在接受兩套平行試卷時，其成績所產生變動的範圍。誤差的大小取決於信度與測驗分數標準差。

FLPT 英語聽力測驗的平均標準誤為 7.21 分，用法測驗 5.76 分，字彙與閱讀測驗 7.09 分。換言之，在相同情境下，如應試者重複接受同一測驗，其聽力的真實分數有 68% 的機率落在實際分數加減 7.21 分之間，用法真實分數落在實際分數加減 5.76 分之間，字彙與閱讀測驗真實分數落在實際分數加減 7.09 分之間。詮釋分數時，如能將測量標準誤加入考量，可更客觀、正確地評量個人真正的能力。

以下以實例說明 true score、observed score 與 SEM 間的相互關係及應用方式。

實際分數 (observed score) = 真實分數 (true score) + 測量標準誤 (SEM)

- ◆ FLPT 聽力測驗平均 SEM=7.21
- ◆ FLPT 用法測驗平均 SEM=5.76
- ◆ FLPT 字彙與閱讀測驗平均 SEM=7.09

假設考生甲在某一次 FLPT 英語測驗施測中得到聽力 x 分、用法 y 分及字彙與閱讀 z 分，則甲生的真實分數有 68% 的機會落在：

- ◆ 聽力真實分數 = $x \pm 7.21$
- ◆ 用法真實分數 = $y \pm 5.76$
- ◆ 字彙與閱讀真實分數 = $z \pm 7.09$

FLPT 口試因採主觀評分方式，故其信度取決於評分人員間的評分標準是否一致。如一致性 (consistency) 愈高，則評分人員間評分的相關係數 (inter-rater reliability) 亦愈高。評分者間信度愈高，則成績誤差愈小，愈接近考生的真實分數。為維持評分人員評分的穩定，本中心除定期舉辦評分說明外，並於評分過程中做人工監控及評分後統計分析。評分時，如兩道差異大，即安排資深評分人員做複評，以確保評分的穩定性；另亦要求評分不穩人員重新檢視評分標準，進行評分校準。FLPT 英語口試測驗的評分者間信度為 0.87。

十、效度與相關性

效度 (validity) 是指測驗分數的正確性。簡言之，一個測驗能否測量到它所想要測量的特質 (trait) 或能力 (郭生玉, 1997)。譬如，一英語能力測驗所要測量的是英語能力，如果測量結果顯示，不同英語能力的考生，在此測驗中的得分有顯著不同，且英語能力愈強者，得分愈高，能力愈弱者，得分愈低。則表示該英語測驗具有效度，它確實可以測量出受試者的英語能力。反之，不同英語能力的考生，在此測驗中的得分無顯著不同，反而是數理能力不同者，在此測驗產生顯著不同，且數理越強者，分數越高，數理愈弱者，分數愈低。則表示該英語測驗不具效度，因為它的分數未反映出受試者的英語能力，反而是數理能力。

檢測效度的方式大致可分為質化與量化兩種，質化研究著重在測驗理論、對象考生特質、測驗建構及製卷藍圖 (blueprint) 的分析；量化研究方面則有內部相關 (internal correlation)、因素分析 (factor analysis) 及多特質多方法分析 (MTMM, multi-trait and multi-method analysis) 等。

根據 Campbell & Fiske 提出之 MTMM，以兩種以上的方法 (multi-method) 來評量多種的特質 (multi-trait)，藉由方法與特質間的組合來判斷測驗的建構效度 (construct validity)，透過各種方法與特質間的相關係數，檢視不同測驗項目所檢測特質差異程度。如果一個測驗與不同特質測驗的相關係數低於其與相同特質測驗的相關係數，則可顯示該測驗具區別效度。

FLPT 四項測驗間的相關係數以「用法—字彙與閱讀」相關係數最高 0.84，顯示兩項測驗所測得的特質較相近；而「口試—用法」及「口試—字彙與閱讀」的相關係數較低，分別為 0.66 及 0.67，顯示口試所測得的建構與用法或字彙與閱讀測驗較不同。其他測驗項目的相關係數落在 0.70~0.80 範圍內。顯示 FLPT 四項測驗除該具有的信度外，亦有適度的建構效度。請參見表十九各項測驗相關係數表。

表十九、各項測驗相關係數表

測驗項目	聽力測驗	用法測驗	字彙與閱讀測驗
聽力測驗	1		
用法測驗	0.77	1	
字彙與閱讀測驗	0.79	0.84	1
口試	0.73	0.66	0.67

另就 FLPT 四項筆試成績與「全民英檢」中級及中高級初試成績的相關係數分析，「FLPT-GEPT 中級」聽力相關係數為 0.71，「FLPT 字彙與閱讀-GEPT 中級閱讀」相關係數為 0.65，而「FLPT 用法-GEPT 中級閱讀」相關係數為 0.61。「FLPT-GEPT 中高級」聽力相關係數為 0.70，「FLPT 字彙與閱讀-GEPT 中高級閱讀」相關係數為 0.75，「FLPT 用法-GEPT 中高級閱讀」相關係數為 0.65。詳參表二十。

上列數據顯示（1）性質較相近的測驗項目，其相關係數高於不同性質測驗項目的相關係數，符合 Campbell & Fiske 區別效度的立論；（2）FLPT 英語測驗所測量的能力範圍較大，幾乎涵蓋「全民英檢」中級及中高級兩個級數。

表二十、FLPT 與 GEPT 中級、中高級各項測驗相關係數表

全民英檢	測驗項目	FLPT 英語測驗			
		聽力	用法	字彙與閱讀	口試
中級 (N=207)	聽力	0.71	0.48	0.51	0.61
	閱讀	0.58	0.61	0.65	0.53
中高級 (N=244)	聽力	0.70	0.57	0.63	0.32
	閱讀	0.66	0.65	0.75	0.30

附錄一 評分方式說明

英語能力測驗四項成績分別計分。聽力、用法及字彙與閱讀測驗成績採「標準分數」計分，以 60 分為平均分數，每一標準差距加或減 20 分調整。每項測驗分數介於 0 至 120 分之間，每套試題因其難易度不完全相同，最高、最低分略有差異，一般應考人成績分布大致如下：

標準分數	20	40	60	80	100	
分數分布	2%	14%	34%	34%	14%	2%

採「標準分數」計分的優點是應試者之成績較不會因當次應試者程度偏高、偏低及所考試題難易不同的影響。換言之，排除學習效果與測驗誤差等因素外，應試者在相近的兩次測驗，其成績不會因試卷不同而有太大差異。

口說測驗時，考生的回答經由麥克風錄下。考後，口試錄音帶交由兩位評分人員就應試者之發音、文法、字彙、流利度及溝通表達能力作整體式評分。給分範圍為 0~5 級。如兩道評分結果差異大或口試成績與筆試成績相差過大時，則安排資深評分人員進行複評。

錄音帶評分方式不同於面試評分方式，所有應試者於同一時間進行口試測驗，所接受測試的題目相同，較一對一面試方式公平，且評分人員較不易受到應試者舉止、容貌、穿著或態度影響，評分可較為客觀公正。

FLPT 四項測驗成績採分項報分，樣例如下。各送考機構可依據所需人才之專業學科或業務需求（例如派赴國外留學、研習、考察等）自訂通過標準。一般機構常用之通過標準請參考附錄二。

准考證號碼	姓名	語言	聽力	用法	字彙與閱讀	口試	備註
41XXX	王大同	英	70	80	65	S-2XXX	

訂定門檻標準常用方法可分為「互補性」與「無互補性」分數兩種(Bachman & Palmer, 1996)。採互補性分數 (compensatory composite score) 時，受試者只需在總分上達到門檻，即可通過測驗。換言之，如受試者在某一項測驗得分較低時，可由另一項得分較高之測驗彌補，如托福測驗。採無互補性分數 (non-compensatory composite score) 時，受試者必須在各分項測驗上皆達門檻，始能通過測驗。換言之，無論受試者測驗總分多高，只要有一項成績未達標準，就無法通過測驗，如「全民英檢」。其他常用的門檻設定機制尚有針對某一特定科目，例如國文、英文或數學等，進行加權計分 (weighting) 或設定最低門檻，如國內各高中、大學的入學考試。

附錄二 外語能力測驗 (FLPT) 成績說明

1、筆試

單項分數	說明		備註
80 以上	聽力	可自在聽懂大部分常速說話。	可考慮出國參加學術性或技術性訓練計畫，得免語言訓練。
	用法	熟悉語法結構。	
	字彙與閱讀	能閱讀較深入或專業的讀物，例如大學用書。	
70~79	聽力	能聽懂一般常速說話，但偶須重複。	可考慮出國參加學術性或技術性訓練，但如能在國內先受全日語言訓練再出國，更佳。
	用法	頗知語法結構，但可能仍有語法錯誤。	
	字彙與閱讀	能閱讀一般讀物，例如報章雜誌。	
60~69	聽力	經過重複或解釋後，大致能聽懂一般會話。	(1) 65~69 分：可考慮出國參加非學術性訓練。 (2) 60~64 分：可考慮出國參加非學術性訓練，但如能在國內先接受全日語言訓練再出國，更佳。如為學術性訓練，則必須先接受全日語言訓練，始能出國。
	用法	大致瞭解語法結構，可應用基本句型。	
	字彙與閱讀	對常用字在 3000~4000 字的讀本，閱讀大致無障礙。	
50~59	聽力	說話必須重複且減慢速度，方能聽懂。	無論參與學術性或非學術性訓練，均須先接受全日語言訓練，始能出國。
	用法	語法結構瞭解不夠，常有語法錯誤。	
	字彙與閱讀	字彙有限，閱讀常有困難。	
49 以下	聽力	聽力程度不足，無法有效溝通。	須先接受基本語言訓練。
	用法	對語法結構認識粗淺。	
	字彙與閱讀	閱讀技能尚未發展。	

2、口試

級分	說明
5	<u>口語表達能力相當於以該語為母語，且受過教育人士所用的語言程度</u> 在各種場合能流暢使用該語，足為母語人士所接受，包括字彙、成語、俗語以及有關文化背景的知識在內。
4	<u>在職務上各種場合，均能正確並流利使用該語</u> 能使用精確的字彙，以流暢的語言參與職務上的談話，雖仍與母語人士有別，但即使在不熟悉的話題上，也能適切交談。鮮有發音或語法錯誤，並能從事非正式的口譯。
3	<u>能以正確的語法與足夠的字彙在正式或非正式的談話中討論關於實用、社交或職務上的話題、自己的興趣與專長</u> 常速的講話全懂。字彙足供表達，不必為用字而感到困擾。本國腔調仍明顯，但語法相當好，即使有錯亦不致使聽者不了解或困惑。
2	<u>能應付日常社交需要與工作上有限的要求</u> 說話雖然不太流利，但尚可在社交場合中自由應對：包括介紹或談及時事、自己的工作、家庭、經歷等。可以應付有限的工作需要。至於比較複雜或困難的，則需要他人幫助。對非專門性之話題，大多能了解主旨，而其字彙只能迂迴表達意思。音調雖常有錯，仍能讓人了解。基本構句雖謂正確，但語法尚不夠自在應用。
1	<u>能應付旅外時基本需要和禮儀上最低要求</u> 對十分熟悉的話題能發問並回答。因對該語的經驗有限，對方必須緩慢地說，或重複或重新解釋，才能了解簡單的問答。字彙除了表達一些最基本的需要外，很難表達任何意見。發音語法常有錯誤，但母語人士，若常接觸如此程度的人，尚可了解。雖說個人熟悉的話題和基本需要因人而異，但應該能叫簡單的菜、找旅館、問路或指路、購物與報時。
0	<u>無使用該語言之能力</u>

註：兩道評分結果顯示某一考生口說表現介於兩級分間時，以「+」表示。例如，S-2+表示考生的表現介於 S-2 及 S-3 之間。

附錄三 英語能力測驗試題樣例

一、聽力測驗：共 50 題，試題由錄音機播出，每題只播出一次。聽了之後，在試題冊上四個備選答案中選一個最適合者作答。

第 1~15 題是問句或敘述句。聽了之後，選一個最適合的回答或回應。

例：(聽) Where is your brother now?

(看) A. He's 20 years old.

B. He's a doctor.

C. He's in New York.

D. He's thirsty.

(正確答案：C)

第 16~30 題是敘述句。聽了之後，選一個說明最恰當的句子。

例：(聽) Fred spent his childhood in France.

(看) A. Fred used to work in France.

B. Fred lived in France as a boy.

C. Fred now has his home in France.

D. Fred has spoken French since childhood. (正確答案：B)

第 31~50 題是一段對話或短文，接著有一問句。聽了之後，選一個最適合的回答。

例：(聽) (M): What did Tom do with the money he won in the contest?

(W): He donated some to charity, and he invested the rest in his brother's business.

(Q): How did Tom get the money?

(看) A. He won a contest.

B. He worked for his brother.

C. He made money on his investment.

D. He got his money from a charity.

(正確答案：A)

二、用法測驗：共 80 或 100 題，限 40 或 50 分鐘作答，每題有四個備選答案，選出其中一個最適合一般情況下的用詞作答。

例：Who _____ your English teacher?

A. were B. am C. is D. are (正確答案：C)

三、字彙與閱讀測驗：共 60 題，限 60 分鐘作答。

字彙測驗共 40 題，包括單字、片語及慣用語測驗。

第 1~20 題：每題中有一空格，請由四個備選答案中，選出一個可使該句意思完整的字或詞作答。

例：A _____ is used to write with.

A. match B. spoon C. pen D. ruler (正確答案：C)

第 21~40 題：每題中有一劃線部份，請由四個備選答案中，選出一個最符合原字或詞意義者作答。

例：Uncle John likes to make up stories to tell the children.

A. invent B. remember C. read D. describe (正確答案：A)

閱讀測驗有 6~8 篇短文，每篇之後有幾個問題，請由每題的四個備選答案中，選出一個最適合者作答，共 20 題。

四、口試試題樣例

分唸句、翻譯、問答題、申述題與看圖說話五部分，時間約 15 分鐘。口試開始時先自我介紹，唸：My name is (姓 名) , my registration number is (准考證號碼) , and my seat number is (座位號碼) 。口試終了時，再將自我介紹的句子重覆一遍。

第一部分唸句：共 5 句。在 45 秒鐘之內依序將此五句連同題號唸一遍。

例：One：The weather was nice yesterday.

第二部分翻譯：共 5 句。在一分鐘之內將五個中文句子依序連同題號翻譯成英文。
(中文不唸)

例：1.我知道你的意思。

One：I know _____.

(答：One：I know what you mean.)

第三部分問答題：共 10 題。題目經由耳機播出，不在試題紙上。第 1~5 題，每題回答時間 15 秒，第 6~10 題，每題 30 秒，聽完問題後應立即儘量回答。題目只播出一遍，聽到下一題號應立刻停止回答，以免錯過下一題。

例：How long have you been working at your present job?

(答：I have been working at my present job for eight years...)

第四部分申述題：看題目思考 1 分 30 秒(備有稿紙供擬大綱)，聽到"Now record your answer for one and a half minutes."時，開始回答 1 分 30 秒。內容應儘量完整有條理。

第五部分看圖說話：看圖片及試題紙上的問題思考 1 分 30 秒(備有稿紙供擬大綱)，聽到 "Now record your answers for one and a half minutes."時，開始回答 1 分 30 秒。內容應儘量詳細有條理。

參考文獻

- 語言訓練測驗中心：〈語言訓練測驗中心英語能力測驗成績統計報告，民國 84~87 年〉，台北：語言訓練測驗中心，1999。
- 語言訓練測驗中心：〈語言訓練測驗中心英語能力測驗成績統計報告，民國 89~90 年〉，台北：語言訓練測驗中心，2002。
- 王文科. 1999. 教育研究法，五南圖書出版公司，1999 年：頁 111-113，129-130。
- 林清山：〈心理與教育統計學〉，台北：東華書局，1997。
- 郭生玉：〈心理與教育測驗〉，台北：精華書局，1997。
- Algina, J., Crocker, L. (1986). *Introduction to Classical & Modern Test Theory*. Orlando, Florida : Holt, Rinehart and Winston, Inc.
- Anatasi, A. (1988). *Psychological Testing*. New York: Macmillan.
- Bachman, L.F. & Plamer, A. S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- Borg, W.R., & Gall, M.D. 1989. *Educational Research: An Introduction (5th ed.)*. New York: Longman.
- Henning, G. (1987). *A Guide to Language Testing—Development, Evaluation, Research*. Boston: Heinle & Heinle Publishers
- Pimsleur, P. et al. (1977). "Speech Rate and Listening Comprehension." In M. K. Burt et al. (eds) *Viewpoints on English as a Second Language*. In honor of James E. Alatis. NY: Regents Publishing.