

摘要

◆ 研究團隊與研究目的

本研究由英國貝德福德大學 (University of Bedfordshire) Anthony Green 教授與 Chihiro Inoue 博士主持，依照 *Relating Language Examinations to the CEFR: A Manual* (Council of Europe, 2009) 建議的程序—包含 familiarization (熟悉 CEFR 分級)、specification (審試測驗品質與內容和 CEFR 級數的關聯)、standard setting (標準設定，即判斷試題對應的 CEFR 級數)，與 empirical validation (實證研究) 等四階段—由台北與英國兩地的兩個英語教學與評量專家小組 (twin-panel)，判斷全民英檢初級至高級口說能力測驗對應 CEFR 的級數，研究結果為全民英檢口說能力測驗提供更多的效度證據，且提供增進測驗品質的建議。

◆ 研究問題

1. 參照全民英檢各級口說能力測驗與 CEFR 級數。
2. 改進審視測驗品質與內容和 CEFR 級數參照時所使用的表格 (boilerplate)。

◆ 研究方法摘要

1. 測驗內容分析 (specification) 由 LTTC 研究團隊與貝德福德大學研究人員協力完成，過程中使用 CEFR 手冊所提供的表格與研究團隊研發的 CEFR 表格 (boilerplate)，分析全民英檢口說能力測驗各級試題內容，並根據分析結果判定全民英檢各級所對應的 CEFR 級數。
2. 標準設定 (standard setting) 由 18 位具英語教學與評量背景的老師與研究人員所組成的兩個專家小組分別在英國與台北兩地進行。英國組有 6 位成員，皆具有英語口說測驗的專業背景，其中一位曾進行過全民英檢相關研究；台北組有 12 位成員，對於全民英檢較熟悉，且都瞭解 CEFR 或曾參與相關研究。標準設定會議前，每位成員皆透過 CEFRtrain 網站 (www.helsinki.fi/project/ceftrain)、Cambridge English Language Assessment (UCLES, 2009) 與 Centre international d'études pédagogiques (CiEP, 2008) 提供的範例熟悉 CEFR 口說能力說明與練習判斷 CEFR 級數。標準設定採面對面小組討論的方式進行，每位成員依試題內容與考生作答，根據 CEFR 口說能力說明，判斷 CEFR 級數。

◆ 研究結果摘要

1. 本研究證實研究團隊研發的 CEFR 表格比 CEFR 手冊所提供的表格能夠更適切的將測驗內容傳達給使用者，且方便測驗單位進行不同測驗間的比較。
2. 測驗內容分析與標準設定的結果顯示全民英檢初級至高級口說能力測驗與所預期的 CEFR 級數大致相符 (A2-C1⁺)。
3. 惟中高級與高級口說能力測驗部分參照結果與預期略有落差，研判原因包括 CEFR 能力說明沒有涵蓋部分中高級能力指標與高級評分量表說明，例如切題度與作答量；而標準設定會議時由於時間有限，僅足夠聆聽數量有限的考生作答，且每位考生僅呈現局部作答，可能會影響專家小組對於考生作答與 CEFR 能力說明的對應結果，本結果尚需進一步的研究佐證。