

討論子題七:認證評量 (標準設定)

2008 年參考書目：

謝進昌、余民寧 (2007)。不同測驗難度對精熟標準設定與分數轉換效果之影響。*測驗學刊*, 54 (1), 頁 1-29。

杜佳真、林世華 (2007)。九年一貫課程數學領域能力指標「數與量」、「代數」主題軸第一、二階段表現標準適切性評估之研究。*師大學報：教育類*, 52 (1), 頁 63-85。

吳淑鶯、劉欣怡 (2007)。ISO 認證之服務業績評估構面之關係模式——平衡計分卡之應用。*商管科技季刊*, 8 (1), 頁 69-89。

周惠民、林陳涌、任宗浩、李哲迪 (2007)。科學教學與學習評量標準芻議。*科學教育*, 298, 頁 2-18。

謝進昌、黃秋華 (2006)。運用定錨點技術診斷學生學習表現。*國教學報*, 18, 頁 75-101。

謝進昌 (2006)。精熟標準設定方法的歷史演進與詮釋的新概念。*國民教育研究學報*, 16, 頁 157-193。

巫義政 (2005)。國內各項證照考試由考選部辦理之妥適性探討。*國家菁英*, 1 (4), 頁 81-108。

陳清溪、勞賢賢 (2005)。「高級中等以下學校及幼稚園教師資格檢定考試」之現況與展望。*研習資訊*, 22 (6), 頁 8-13。

黃財尉 (2004)。美國教師證照考試——Praxis 系列評量的探究與評論。*教育研究*, 127, 頁 100-117。

周春美、沈健華、張雅婷、賴銘娟 (2003)。我國推展商管類職業證照檢定有效作法之探討——以日本辦理中小企業診斷士的經驗為借鏡。*技術及職業教育*, 77, 頁 20-23。

楊江益、吳柏青、練建志、葛珠蘭 (2002)。技專校院推動 ISO 9000 品質管理系統之方法與問題探討。*宜蘭技術學報*, 9(人文及社會專輯), 頁 239-256。

李允傑、李天祥 (2002)。ISO9000 與組織績效之迷思——以我國大學院校實施 ISO9000 系列驗證為例。*空大行政學報*, 12, 頁 1-36。

涂金堂 (2000)。淺談設定最低能力測驗通過標準的方法。*測驗與輔導*, 158, 頁 3304-3308。

魏惠貞 (2001)。幼稚園老師證照教學複試認知基本能力評量內容參照標準之發展。*臺北市立師範學院學報*, 32, 頁 485-517。

2006 年參考書目：

謝進昌、余民寧(2005)。以最大測驗訊息量決定通過分數之研究。*測驗學刊*，52(2)，頁 149-176。

涂金堂(2000)。淺談設定最低能力測驗通過標準的方法。*測驗與輔導*，158，頁 3304-3308。

鄭明長、余民寧(1994)。各種通過分數設定方法之比較。*測驗年刊*，41，頁 19-40。

林惠芬(1993)。通過分數設定方法在護理人員檢覈筆試測驗之研究。*測驗年刊*，40，頁 253-262。

吳裕益(1988)。九種通過分數設定方法之比較研究。*初等教育學報(臺南師範)*，1，頁 47-120。

吳裕益(1988)。標準參照測驗通過分數設定方法之研究。*測驗年刊*，35，頁 159-165。

吳幼吾(1985)。標準參照測驗通過分數與測驗長度之探討。*高雄師院學報*，13，頁 404-350。

吳幼吾(1984)。標準參照測驗通過分數之探討。*臺灣教育*，402，頁 9-16。

Giraud, G., Impara, J. C., & Plake, B. S. (2005). Teachers' conceptions of the target examinee in Angoff standard setting. *Applied Measurement in Education*, 18(3), 223-232.

McKinley, D. W., Boulet, J. R., & Hambleton, R. K. (2005). A work-centered approach for setting passing scores on performance-based assessments. *Evaluation and the Health Professions*, 28(3), 349-369.

Plake, B. S. (2005). Doesn't everybody know that 70% is passing? In R. P. Phelps (Ed). (2005). *Defending standardized testing*. (pp. 175-185). Mahwah, NJ: Lawrence Erlbaum Associates.

Beretvas, S. N. (2004). Comparison of bookmark difficulty locations under different item response models. *Applied Psychological Measurement*, 28(1), 25-47.

Brandon, P. R. (2004). Conclusions about frequently studied modified Angoff standard-setting topics. *Applied Measurement in Education*, 17(1), 59-88.

Miller, G. E., Yoes, M. E., & Twing, J. S. (2004). Estimation of the all tests pass rate when no examinee took all of the tests. *Applied Measurement in Education*, 17(4), 393-406.

Reckase, M. D. (2004). What if there were a "True Standard Theory" for standard setting like the "True Score Theory" for tests? *Measurement: Interdisciplinary Research and Perspectives*, 2(2), 114-119.

Cizek, G. J., Bunch, M. B., & Koons, H. (2004). An NCME instructional module on setting performance standards: Contemporary methods. *Educational*

Measurement: Issues and Practice, 23(4), 31-50.

Wang, N. (2003). Use of the Rasch IRT model in standard setting: An item-mapping method. *Journal of Educational Measurement*, 40 (3), 231-253.

Wang, N. (2003). "Use of the Rasch IRT model in standard setting: An item-mapping method." Erratum. *Journal of Educational Measurement*, 40 (4), 373.

Downing, S. M., Lieska, N. G., & Raible, M. D. (2003). Establishing passing standards for classroom achievement tests in medical education: A comparative study of four methods. *Academic Medicine*, 78 (Suppl10), S85-S87.

Hurtz, G. M., & Auerbach, M. A. (2003). A meta-analysis of the effects of modifications to the Angoff method on cutoff scores and judgment consensus. *Educational and Psychological Measurement*, 63(4), 584-601.

Chinn, R. N., & Hertz, N. R. (2002). Alternative approaches to standard setting for licensing and certification examinations. *Applied Measurement in Education*, 15(1), 1-14.

Stephenson, A. S., Elmore, P. B., & Evans, J. A. (2000). Standard-setting techniques: An application for counseling programs. *Measurement and Evaluation in Counseling and Development*, 32(4), 229-244.