

第卅七屆國際化學奧林匹亞競賽實作 有機合成試題評量之研究

吳佳儒、鄭嘉華、陳穩仁、方泰山*

國立臺灣師範大學化學系 chetsf@scc.ntnu.edu.tw

摘要

本研究簡介 2005 年第 37 屆國際化學奧林匹亞實作有機合成試題之編製、施測、評分過程，以及賽後的試題性質與測驗結果分析。

學術委員會編製實作試題的過程如下：選定主題，設計實作評量的內容，量化測驗結果；舉行預試，修訂試題與施測流程，以及訂定評分標準。

實作考試施測對象為 225 名來自 59 個國家的高中化學資優生，施測時間五小時。實作有機合成試題鑑別力指數 .46，試題難易度指數 .73，具有良好的評分者信度與效標關聯效度；選手之實作有機合成試題成績與另一實作分析試題、及理論有機試題成績均呈低度相關，顯示各試題所考驗的選手能力重複性低，符合競賽目的與學術委員會之預期。

關鍵詞：國際化學奧林匹亞，預試，實作有機合成評量。

一、緒論

(一) 研究背景與目的

國際化學奧林匹亞 (International Chemistry Olympiad, IChO) 是為世界各國高中資優生代表所舉辦的化學競賽，競賽內容分為實作考試 (40%) 與理論考試 (60%)，兩項測驗的舉行至少間隔一日，比賽時間各為五小時；參賽選手使用經由該國教練翻譯後的試題，並可於競賽時要求主辦國提供經國際評審團 (the International Jury, IJ) 修訂通過後原版英文試題。

我國於 1992 年開始組團參加第 24 屆 IChO，經由代表團選手與教練的努力，迄今屢戰皆捷、成績斐然；在持續提升代表團成員的參賽實力之際，我國也積極爭取 2005 年主辦權並獲得回應，隨即展開各項競賽籌辦事務，為在臺北舉行的第 37 屆 IChO 作周全的準備 (方泰山，民 94)。

本研究之目的為呈現學術委員會編製實作有機合成試題的命題過程，記錄實作考試施測、評分流程，並在賽後進行試題性質與測驗結果分析，檢視試題與結果是否符合 IChO 競賽目的與規章，以及我國學術指導委員會所預設目標。

(二) 名詞釋義

1. IChO 實作課程綱要

IChO 競賽規章中明定參賽選手應具備的知識和技能，在實作方面列有實作課程綱要，將 66 個實驗操作技巧歸納為六大類型，供教練與選手參考。

2. 準備題

IChO 主辦國應於競賽前半年公布以英文書寫的「準備題」，其中至少包含 25 道理論準備題和 5 道實作的準備試題，供參賽選手從中了解該屆主辦國競賽試題的形式與難度。

3. 國際評審團 (the International Jury, IJ)

IJ 由 IChO 指導委員會 (the Steering Committee, SC) 和參賽國教練組成，負責在競賽前審查試題，並監督競賽後的試題批改，以維護競賽的公正與公平性。

4. 成績仲裁 (Arbitration)

各國教練可依據主辦國提供的評分標準及選手答案卷影印本進行閱卷，並於成績仲裁會上核對主辦國的評分結果，取得雙方的共識，以確認每位選手的成績無誤且合理。

(三) 研究限制

1. 本研究所編製的試題，適用於 IChO 競賽，參賽選手為經過訓練的高中資優生，不宜作為一般高中生的測驗題。
2. 本研究設定對象為第 37 屆 IChO 參賽選手，其分析的結果不宜推論至其他學生。

二、文獻探討

(一) 測驗與評量

1. 測驗的特徵

效度 (validity) 和信度 (reliability) 是傳統標準測驗最重要的兩個特徵，效度以信度為基礎，衡量一個測驗能夠測量欲測特質的程度，而信度則檢驗測驗分數的一致性或穩定性 (陳英豪、吳裕益，民 92；郭生玉，民 93)。

French & Michael (1966) 提出三種效度類型：內容效度、效標關聯效度和建構效度，可視測驗性質與目的選擇使用之；常用的信度類型有重測信度、複本信度、折半信度、庫李信度、 α 係數以及評分者信度。

2. 預試

歐滄和 (民 75、民 82) 提出預試的兩個層次為非正式預試和正式預試：在非正式預試中，可確定測驗的指導語清晰度、事先得知預試時可能遇到的困難，並能粗略估計施測所需時間；此外，可由正式預試獲得實證資料，了解試題的難易度或鑑別度。

(二) 實作評量的編製

綜合 Airasian (1994) 和 Stiggins (1987, 1991, 1994) 兩位學者的研究，實作評量編製過程主要可分為明確界定評量的目的、設計實作評量的任務、確定實作評量的評分標準，以及實作評量的評分方式四個步驟；而常見的評分工具的類型則有檢核表、評定量表、軼事記錄和主觀紀錄四種，測驗編製者可依測驗目的選擇適用的評分紀錄方式。

三、研究方法

(一) 研究對象

1. 預試對象共 94 人，為行政方便皆為臺師大化學系學士班與碩、博士班學生，具有穩定而純熟的實驗操作技巧，其實驗操作能力和受過訓練的競賽對象相近，依預試目的不同而分為五組。
2. 競賽施測對象為第 37 屆 IChO 參賽選手共 225 人，分別來自 59 個國家，是經過選拔和訓練的高中化學資優生。
3. 分析對象為第 37 屆 IChO 競賽之理論和實作考試的競賽結果。

(二) 研究流程

本研究配合 IChO 試題編製與競賽時程，依序確定研究問題與範圍、界定評量的目標、設計實作評量的任務、尋找量化競賽測驗結果的方式、舉行預試、修訂試題與施測流程、確定試題內容及評分標準、試題審議、正式施測、結果測量與成績仲裁，資料分析

(三) 研究工具

1. 儀器設備：核磁共振儀 (Nuclear Magnetic Resonance Spectroscopy, NMR)，以及 Japan JASCO Co. DIP-1010 Digital polarimeter 自動旋光光度計。
2. 試題分析方法：參考學者整理之分析公式與方法 (林清山，民 81；簡茂發，民 91；郭生玉，民 93)，採用非選擇題的鑑別力指數和難易度指數計算方式、皮爾遜積差相關 (Pearson's product-moment correlation)、Cronbach's α 係數，以及斯布公式 (Spearman-Brown formula) 進行競賽試題分析。

四、研究結果與討論

第 37 屆 IChO 學術委員會篩選出「有機合成」與「無機化合物之定性分析」兩道實作試題，依序編為實驗一和實驗二；本研究配合學術委員會運作，只進行實驗一的實作試題編製工作，茲將命題相關程序描述如下。

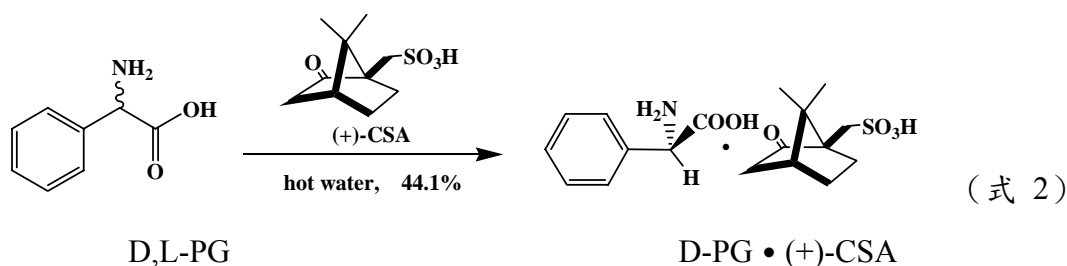
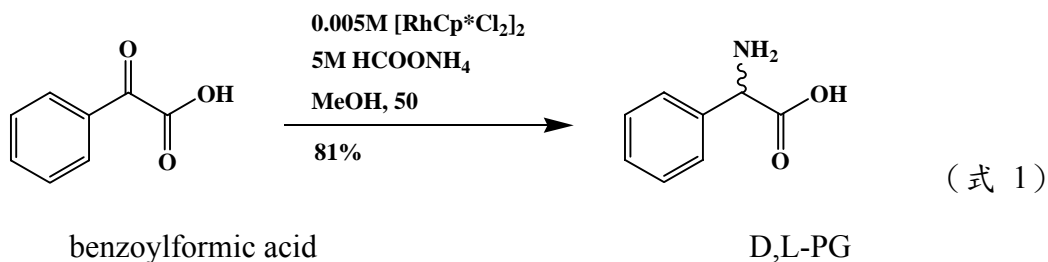
(一) 實作試題編製

1. 確定實作試題之主題

學術委員會選擇具有「樟腦」架構衍生物為實驗一的主題，即是著眼於民國五、六十年代，臺灣中樟出口的樟腦粉和樟腦油稱霸全球 (張嘉琪，民 92)，足以代表我國特有的產物與技術。由樟腦製品提煉出的樟腦磺酸 (camphorsulfonic acid, [CSA])，經處理後可做為製作抗生素時的光學分離劑，極具生技製藥方面的應用價值。

D,L-苯甘胺酸 (D,L-phenylglycine, [D,L-PG]) 的其中一個對掌體是製備 β -內醯胺 (β -lactam) 抗生素的重要原料，學術委員會擬以此相關的反應來做為實作試題；根據日本學者 Kitamura (2002) 等人之研究，將苯甲醯甲酸 (benzoylformic acid) 在銻

催化劑 (Rh Catalyst) 的作用下，進行還原胺化反應 (reductive amination) 來合成消旋的 D,L-PG (式 1)，再採用學者 Yoshioka 等人 (2000) 所發表的方法，以(+)-CSA 與 D,L-PG 反應進行光學分離 (式 2)。



2. 設計實作評量的任務

確定實作主題後，要將實驗內容試題化，使之符合選手安全、競賽時間與場地設備的考量；學術委員會參考式 1 與式 2，並找尋適當實驗條件，將其修訂為實驗一的步驟一與步驟二，修訂過程中進行變更溶劑系統、縮短反應時間、去除產物中主要雜質，以及尋找最佳反應物比例等各項實驗。

實作考試時間為五個小時，試題內容的設計應可讓大多數選手在考試時間內完成實作任務，其中包含實驗一的兩個步驟，以及需時約 1~1.5 小時的實驗二。

3. 尋找量化測驗結果的方式

對於進行實驗的實作比賽而言，可由數字呈現的結果較為客觀且容易相互比較；學術委員會擬以兩步驟的產物來評分；其中，產物的產率是第一個評分標準，而產率多寡對應分數高低的轉換方式，需經由預試所得之產率分佈來劃定評分標準。

除了以產率來評分，產物的純度也是另一項重要指標，產物純度越高，可對應於選手具有較好的實驗技巧；實驗一的兩步驟產物純度鑑定方式依序為測定產物的 ^1H NMR 光譜和比旋光度，學術委員會於試題編製期間測試相關因素對測定結果的影響，包含 D,L-PG 與雜質溶於氘溶劑的穩定度、儀器穩定度或人為操作的影響、雜質的種類與影響、溶劑或溶液濃度對比旋光度之影響，以及環境溫度對儀器穩定度之影響，嚴密掌控上述因素，以期測量結果可客觀且真實呈現選手成果。

4. 舉行預試

學術委員會將 94 名化學系學生分為五組，依不同目的進行預試，依序確認試題內

容與施測流程，以及建立對照團體效標；綜合上述兩部分的五次預試，學術委員會將預試中發現的問題逐一修正，並以預試三～五的實驗結果建立特殊團體效標。唯有透過預試找出施測流程、實驗條件與器材的不足，並設法改善問題，方能減少測驗編製與實施對選手的影響，測驗出選手真正的待測能力。

5. 修訂試題與施測流程

預試時，研究者與監考人員共同紀錄施測流程中待改進的問題，並於實驗結束後晤談參與預試的學生，了解預試對象在實作過程中的疑問與建議；學術委員會參考預試實驗結果、監考人員紀錄、軼事紀錄，以及晤談內容，對於試題內容及施測流程進行修訂。

當多數預試對象可依照試題提供之條件，在時限內完成實驗，並獲得合理範圍內之產物產率與純度，即表示此實驗條件具有適當性，可作為 IChO 實作考試的試題。

6. 確定試題內容及評分標準

經過數次「預試-修訂試題」過程，學術委員會可確定試題內容具適當性與清晰度，並參考預試結果分布來訂定評分標準，試題於 IChO 競賽期間通過試題審議後，即可成為正式試題。

(二) 實作考試之施測流程

第 37 屆 IChO 競賽於 2005 年 7 月 16 ~ 25 日於臺灣-臺北舉行，闈場設於三峽國家教育學院、實作考試假臺灣大學思亮館舉行，選手村、測量所使用的儀器與理論試場則安排在臺灣師範大學分部校區。與實作考試相關的流程包含試場檢視與試題審議、正式施測，以及結果測量與成績仲裁。

其中，在試題審議時，IJ 降低部分評分標準，其主要考量在於 IChO 為具有教育意義和引發選手興趣的競賽，應避免整體成績分布偏低，另有 IJ 成員認為預試對象的實驗操作能力優於參賽選手過多，應降低根據預試結果制定的評分標準。此外，工作人員測量選手成果時，利用專業且具高難度的 NMR 及旋光儀進行定性與定量分析，並將外在因素的影響降至最低，使測量結果足以作為選手得分的依據。

(三) 實作考試之試題性質與測驗結果分析

1. 試題性質分析

實驗一試題鑑別力指數 .46，試題難易度指數 .73，屬於鑑別力優良、難易度適中之試題，符合競賽目的與學術委員會之預期目標。

評分時以實驗數據量化測驗結果，並經過成績仲裁會討論，具有良好的評分者信度；另以特殊團體效標檢驗效標關聯效度，亦得到高度相關 ($r = .997, p < .01$)，試題品質良好。試題之 Cronbach's α 係數為 .41，解讀時需同時考量試題題數較少、施測對象同質性高，以及設定試題難度經 IJ 大會調整為中等偏易等因素之影響。

本屆參賽選手實作考試有機合成試題成績，與其實作考試分析試題、理論考試有機試題成績均為低度相關（分別為 $r = .26, p < .01$ ； $r = .38, p < .01$ ），顯示兩道實作試

題考驗之實驗技能重複性低，且選手之有機實驗技能受其理論基礎影響低。

2. 測驗結果分析

本屆實作考試有機合成試題採用 IJ 修正後評分標準，計算得平均分數為 80.0 分，測驗分數集中於高分一端；如改用學術委員會原定評分標準計算得平均分數為 74.3 分，測驗分數較接近常態分配趨勢，分布於中高分區域。

其中，本試題中變更評分標準的兩個小題，均呈現以原標準較新標準有較趨近常態的成績分布曲線，成績分布範圍也較為合理並符合預期，顯示 IJ 應給予主辦國的學術委員會更多信任，採用參考預試結果所訂定的評分標準。

五、結論與建議

綜合上述研究成果，學術委員會依循實作試題編製過程，完成歷時三年的實驗一試題編製，並於第 37 屆 IChO 競賽期間時順利進行施測與實驗結果測量；由研究者試題性質與測驗結果分析得知，實驗一為符合 IChO 競賽目的和學術委員會預期之實作試題，其測驗編製與施測過程可供相關研究者作為參考。

實作試題的編製、施測與評量相當耗時費力，用於各種規模的競賽或日常教學，均有推廣上的困難。然而實作考試可提起學生對操作實驗的興趣與重視，進而督促教育者培養學生的各層次科學技能，是值得研究與推廣的評量方式；建議相關研究者可進行長期實作評量研究，並定期進行交流與討論，共同累積經驗，持續提升我國在科學實作評量方面的研究實力。

六、後續發展

第 37 屆 IChO 競賽落幕後，實驗一試題經細節修改成為教學實驗內容，編入臺灣師範大學化學系大二有機實驗課程，學生於兩週分別進行實驗與產物純度測量，並由實驗助教說明反應機構與純度鑑定原理，使實驗一試題的教育意義更加完備，延續其引發學生興趣之目的。

七、致謝

感謝教育部及國科會提供研究經費的補助，並感謝 37th IChO 2005 學術委員會、臺師大化學系與實作小組的學術與行政的全力支援。

八、參考文獻

- 方泰山 (民 94, 8 月)。台灣主辦 2005 年第 37 屆國際化學奧林匹亞 -回顧與展望-。
科學教育月刊, 281, 43-64。
- 林清山 (民 81)。心理與教育統計學 (初版)。台北市：東華書局。
- 張嘉琪 (民 92, 11 月 27 日)。吳正中走過的地方空氣中都有樟腦的香味。*新新聞周報*, 873, 93-95。
- 郭生玉 (民 93)。教育測驗與評量 (修訂版第一版)。台北市：精華書局。
- 陳英豪、吳裕益 (民 92)。測驗與評量 (六版)。高雄市：復文圖書出版社。
- 歐滄和 (民 75)。談測驗編製之預試。*測驗與輔導*, 74, 1382-1387。

- 歐滄和 (民 82)。標準化測驗的編製發展程序。 *測驗統計年刊*, 1, 33-42。
- 簡茂發 (民 91)。 *心理測驗與統計方法* (三版)。台北市：心理出版社。
- Airasian, P. D. (1994). *Classroom assessment*. New York: McGraw-Hill.
- French, J.W. & Michael, W.B. (1966). *Standard for Educational and psychological Test*. (pp. 12-24). Washington: American Psychological Association.
- Kitamura, M.; Lee, D.; Hayashi, S.; Tanaka, S.; Yoshimura, M. (2002). *J. Org. Chem*, 67, 8685.
- Stiggins, R. J. (1987). Design and development of performance assessment. *Educational Measurement : Issue and Practice*, 6(3), 33-42.
- Stiggins, R. J. (1991). Facing the challenges of a new era of education assessment. *Applied Measurement in Education*, 4(4), 263-273.
- Stiggins, R. J. (1994). *Student-centered classroom assessment*. NY: Macmillan College Publishing Company.
- Yoshioka, R.; Hiramatsu, H.; Okamura, K.; Tsujioka, I.; Yamada, S. (2000). *J. Chem. Soc., Perkin Trans. 2*, 10, 2121