

編製桌球反手擰球技術效標參照測驗

國立體育大學體育研究所、明志科技大學體育室 陳建彰

國立體育大學競技與教練科學研究所 姚漢禱

國立陽明交通大學體育室 游鳳芸*

摘 要

目的：利用多層面 Rasch 模式編製桌球反手擰球技術效標參照測驗。**方法：**研究對象以 40 位競技選手，組別分為精熟組 23 位、非精熟組 17 位，實施桌球反手擰球技術效標參照測驗，以 FACETS 程式估計選手、評分者、試做次數和試題等四個層面的測驗。**結果：**測驗資料適合 Rasch 模式；效標參照評量的分界分數在能力 5.69 邏輯子 (logit)，其介於原始分數 422 分和 421 分之間，以及效度係數 0.75，信度的一致性比率 0.88、柯恆 kappa 係數 0.75 和修正 kappa 係數 0.75。**結論：**桌球反手擰球技術效標參照測驗具有良好的測驗品質，能鑑別競技選手反手擰球技術的精熟。

關鍵詞：一致性比率、分界分數、效度係數、桌球反手擰球技術、多層面 Rasch 模式

壹、緒論

效標參照 (criterion-referenced) 評量的看法,最早是 Glaser (1963) 提出根據效標標準 (criterion standard) 來評估學生成績,可提供學生所能學習達成程度的信息,且與其他人的表現無關。意思是指出教學的重點在於學生學會課程 (精熟),而不在於班級團體的排名比較。因此,每個教學單元都要設定效標標準,才能使教學順利的依進度推進,最終讓學生具備有教育設定的各項能力,與九年一貫課程培養學生帶得走的能力相吻合。張宏亮 (1995) 也指出:「改進的方法是採效標參照測驗,……同時效標參照測驗,亦可做為教師教學策略之參考。」如果學生未能精熟,則會造成後續學習的困難;因此,這種效標參照測驗關係著學生學習結果,應該普遍的使用在教學上。吳璧純與詹志禹 (2018) 論述從能力本位到素養導向教育,1990 年代能力本位教育,21 世紀隨環境和教育的發展,走向素養導向的評量策略。所謂素養導向的評量,注重形成性評量、促進教學的評量、多元整合評量,這個趨勢與前述的效標參照測驗吻合,偏重學習過程、促進教學正是效標參照的特點。林清山與程炳林 (1996) 也認同:「從訊息處理與後設認知的觀點而言,學習策略

是促進有效學習。」

在體育運動領域的應用上,魏豐閔與施登堯 (2014) 提出評分作為教育評價的一環,必須能進一步對體育課程實施與教學實踐做出貢獻。這項看法和 Glaser (1963) 的效標參照測驗一致。但從魏豐閔與施登堯 (2014) 對體育成績考查標準的選擇與設定所提出常模與效標參照標準之概念內涵、操作方式以及優缺點而論,可能混淆了常模參照測驗、效標參照測驗和體育成績給分三者的目的和功能,在教學現場中不易找到合適的測驗和給分;首先關於常模參照和效標參照兩種測驗是基於目的而分類,姚漢禱 (1995) 效標參照測驗目的是精熟組和非精熟組分類,常模參照測驗目的在區別受試者能力,使分數產生最大的變異。至於國內體育教師需要的是體育成績給分,我們的評分系統是百分制,設定 60 分及格、研究所 70 分及格,而體育技能測驗成績是各種紀錄,有時間、距離、次數……等各種情況,需要透過體育測驗成績給分量表,才能轉換為教師所要的分數。給分量表與兩種測驗目的關係密切,適當的使用相得益彰,用百分來區隔學生的表現屬於常模參照,但設定及格分數屬於效標參照,顯然一般人執

著於分數愈高愈好，但只要及格就是精熟該項能力，也就是說，教育目的希望學生能擁有精熟的能力素養。

回到效標參照的實務上，標準是最難設定的概念，以游泳為例：標準設定游完 50 或 100 公尺，這真的學會游泳嗎？關鍵在於未能真正抓住標準的設定，要設定游泳的標準，不在距離、而是會換氣，學會換氣就可輕鬆的享受游泳；不換氣游完 50 或 100 公尺者判定非精熟，與距離、速度無關，能夠輕鬆自在的換氣游泳，哪怕你是抬頭蛙泳、或是狗爬式，皆判定游泳精熟。另外，效標參照測驗的分級標準，是依不同目的而設定，並不一定是依能力高低的區別，主要著眼與使用工具、環境、服務對象、從事工作等因素考慮；例如駕照檢定依駕駛工具而定，分為機車、重機、小汽車、公車、貨櫃車、聯結車…等分類。托福 TOEFL 測驗為評量母語非英語者的英語能力，為外籍生申請美國大學及研究所之入學時語言能力證明。GRE[®]測驗成績為美國各大學研究所的申請入學參考條件之一，證明具有研究的能力。而 TOEIC 測驗和 TOEFL 考試的對象不同，TOEIC 測驗主要針對在英語社會求職的需求。

體育運動技能評量觀察或測驗對象是成績表現 (performance)，主

體不是認知而是技能，此運動技能評量又分成兩個部分，一是技術的表現結果，通常可以量化（丈量、計時……等）。另一是技術本身的動作，需要透過觀察身體動作，常常是質性的評判。陳鉸澈 (2016) 介紹大專跆拳道品勢動作給分方式：「1. 正確性基本分 4.0 分，動作的正確性及平衡。2. 表現性基本分 6.0 分，含速度與節奏強度與速度、氣三項。」亦即，跆拳道品勢是透過觀察技術動作，也是一種質性評分。

在運動競技的桌球項目上，龍柏安 (2021) 分析 WTT 連兩站奪冠的林昀儒與鄭怡靜擊敗南韓組合的比賽影片後，初步發現其中接發球的制勝武器就是反手擰球技術。類似的觀點在先前研究亦曾被提出，例如：董聲 (2017) 指出張繼科、馬龍與樊振東三位反手擰球精湛的運用，已引起接發球技術革命性改變，被視為主要的接發球技術。因為桌球世界排名第一的中國男子隊，在比賽中接發球時，經常使用反手擰球技術逆轉被動方成為主動方。國內研究陳建彰、游鳳芸與姚漢禱 (2020) 也認為反手擰球不僅能擰出側上旋球及側下旋球，若本身已具備此技術者得於善用，能為自己開啟契機並奪取主導權。隨後一系列的研究，經過文獻分析法探討桌球反手擰球技術，然後進一步，建構

技術動作評量表，完成桌球反手擰球技術動作主觀評量表後，招募 40 位競技選手進行反手擰球技術動作主觀評量，然後利用多層面 Rasch 模式估計資料，最後根據效標參照評量的目的，發展桌球反手擰球技術效標參照測驗。另外，運動技術效標參照評量不多，姚漢禱與林德隆 (2000) 首先利用多層面 Rasch 模式估計，24 位大學女生壘球打擊測驗，編製中學運動績優甄審與甄試之效標參照測驗。此研究以多層面 Rasch 模式將分數轉換為更精確的量尺，因桌球反手擰球技術動作測驗，屬於主觀評分的順序量尺，不具可加性，也是利用 Rasch 模式量化。Wright 與 Masters (1982) 指出 Rasch 測量模式將順序資料轉換成近似等距數線 (連續數線)，透過 logistic 轉換受試者的機率獲得潛在

能力。亦即，順序類別的評分透過 Rasch 模式，轉換為 logit 量尺值，它是單向度線性等距量尺，具有可加性。Linacre (1997) 說明 Facets 程式估計原始測量的量，經過校準後，成為線性量尺結構，推論為等距量尺。

前述分析結果發現：目前國際桌球賽的接發球使用反手擰球技術次數高，也是一項最先進的檯內球進攻技術，能由被動轉成主動的過程，對比賽的得分有直接性的影響，有時也是左右比賽勝負的關鍵。有鑑於此，研究反手擰球技術效標參照測驗有其必要性。所以，本研究的目的是利用多層面 Rasch 模式編製桌球反手擰球技術的效標參照測驗，可提供競技選手技術精熟的判斷，以及大學桌球績優生的甄審或甄試測驗。

貳、方法

一、研究對象

本研究對象選取 40 位，男性 26 位、女性 14 位，年齡 18~31 歲；分為兩組 (事前判斷分類)：曾當選桌球國手 23 位特優選手為「精熟組」及運動績優體育保送生 17 位為「非精熟組」。

二、研究工具

研究工具採用陳建彰、游鳳芸與姚漢禱 (2020) 的桌球橫拍反手擰球介紹與技術分析編製評量表 (經過文獻分析法和運動技術質性分析法兩階段建立) 施測，有 5 位評分者，每位競技選手試做 3 次，項目為評量 6 階段擊球過程 (準備動作、引拍動作、擊球前動作、擊球時動作、擊球後動作、還原動作)。反手擰球技術動

作給分依達成動作要點：給予 0 至 6 分 (如附錄一)。

三、多層面 Rasch 測量模式分析資料

使用 FACETS 測驗軟體估計桌球反手擰球技術效標參照測驗，模式設計含競技選手、評分者、試做和試題等四個層面，使用多層面 Rasch 測量模式。線性模式如下：

$$\log(P_{nijl}/P_{nijl-1}) = B_n - R_i - T_j - S_l - F_k$$

符號說明為： P_{nijl} 是競技選手 B_n 在第 l 題、試做第 j 次、評分者 i 給評分 k 的機率。 P_{nijl-1} 則是競技選手 B_n 在第 l 題、試做第 j 次、評分者 i 給評分 $k-1$ 機率。 B_n 是競技選手 n 的能力。 R_i 是評分者 i 的評分嚴謹度。 T_j 是試做第 j 次的難度。 S_l 是試題第 l 題的難度。 F_k 是給分 F 在從類別 $k-1$ 到類別 k 的階段難度。

四、效標參照測驗的信度計算

效標參照測驗信度採 Looney (1987) 的不一致的閾值 (threshold loss agreement)：為兩次測驗分配至精熟或不精熟組的一致性。姚漢禱 (1995) 進一步說明三項信度指標如下。

(一) 一致性比率 (Proportion of Agreement)：兩次測驗判斷結果的一致性比率。

(二) 柯恆 (Cohen) kappa 係數

(Cohen kappa, 簡稱 κ 值)：指校正一致性比率受機會影響的機率。

(三) 修正 kappa 係數 (modified kappa, 簡稱 κ_q)：由於精熟比率大、機率也將變大，結果使 kappa 係數變小；機率應該隨類別數目 (q) 而定，亦即邊際自由 (Free Marginals，意指當某一類別數量較大時，用於調整信度的失真)。意即修正 kappa 係數是用 $1/q$ 做為機率 (效標參照測驗分為精熟和非精熟兩組，故 $q = 2$)；修正 kappa 係數將大於 kappa 係數，其上限為 1，類別增加時，修正 kappa 係數也隨著增加。

五、效標參照測驗的效度計算

姚漢禱 (1995)：「建議效標參照測驗的效度採用正確決斷 (Decision Accuracy)，它是指效標參照測驗分類的正確性。」為 Safrit (1989) 提出的 ϕ 相關 (ϕ vc) 為效度係數。

六、分界分數的計算

在體育運動研究上，何榮桂與姚漢禱 (1995) 分析：「設定分界分數理論雖多，但專家們指出三種實徵方法較適合運動技能領域，其中評價最高的是 Berk (1976) 的效標組法 (Criterion Groups)。」後來謝進昌 (2006) 改進 Berk (1976) 的方法提出「對照組 M-SD 法」，計算分界分數公式為

$$CS = M_1 - [S_1 (M_1 - M_2) / (S_1 + S_2)]$$

符號說明為：CS 是分界分數， M_1 是精熟組平均數、 M_2 是未精熟組平均數， S_1 是精熟組標準差、 S_2 是未精熟組標準差。

參、結果

一、估計桌球反手擰球技術效標參照測驗

本研究利用 FACETS 的 4 層面模式估計測驗，重要結果摘要如下：

(一) 測量資料分析摘要

競技選手、評分者、試做、試題等 4 個層面估計結果 (如表 1)。

1. 估計潛能 (單位 logit) 的平均數與標準差：競技選手 $5.65 \pm .78$ 、評分者 $.00 \pm .47$ 、試做 $.00 \pm .15$ 、試題 $.00 \pm 5.11$ 。模式設定第一層面浮動，其他三層面固定，利於估計收斂；故競技選手潛能平均高達 5.65，反應實際差距，而其他三層面潛能平均為 .00。試題難度標準差高達 5.11，顯示具有各種高低難度的試題，能滿足競技選手潛能適配。

2. 母群的 Rasch 模式個別信度：競技選手 .91、評分者 .97、試做 .00、試題 1.00。顯示競技選手信度優良、試題接近完美信度；評分者和試做應該接近 .00 才是完美信度，因為他們一致才是良好，故評分者信度不良、試做接近完美信度。

3. 同質性考驗結果機率是：競技選手 .00、評分者 .00、試做 .92、試題 .00。達 $p < .05$ 顯著水準時，裁決拒絕虛無假設；表示競技選手、評分者、試題等 3 個層面的潛能皆非同質，而試做層面是同質。意即競技選手和試題層面的估計潛能有高低差異，就 AERA, APA 與 NCME (2014) 認為高分組較低分組在高級課程有較好的表現。這是具有效度的證據之

一，顯示測量的競技選手和試題層面具有鑑別度。而評分者有評分寬嚴差異的不一致情形，對照前面高信度，證明測量分析發現評分者有問題，但受試者在三次試做的表現，估計結果則為一致的試做難度。

4. 常態性考驗結果機率是：競技選手 .54、評分者 .27、試做 .69、試題 .29 皆未達 $p < .05$ 顯著水準，裁決接受虛無假設；表示競技選手、評分者、試做、試題等 4 個層面樣本的潛能皆符合常態分配，統計母群理論為常態分配，樣本符合母群的常態分配，因此推論 4 個層面樣本皆具有代表性。

5. 適合度考驗的平均數與標準差。(1) 訊息加權適合度：競技選手 $.96 \pm .52$ 、評分者 $.78 \pm .41$ 、試做 $1.01 \pm .02$ 、試題 $1.08 \pm .29$ 。(2) 偏離反應適合度：競技選手 $.93 \pm .77$ 、評分者 $.93 \pm .48$ 、試做 $.93 \pm .06$ 、試題 $.93 \pm .22$ 。這兩個指標的期望值為 1.00、資料適合模式的範圍是 .05~1.5。整體而言，本研究 4 個層面皆適合模式；符合 Wright 與 Masters (1982) 假使受試者表現吻合適合度統計，則此測量有效。意即此測驗具有效度。

6. 評分者一致性：Facets 估計評分者間兩配對的一致性，即評分者間信度，得到 74.5%，評分者一致比率七成多，屬於普通的評分者信度。

表 1 全體競技選手、評分者、試做和試題等 4 層面測量分析摘要表

層面	估計潛能(logit) 平均±標準差	Rasch 模式 個別信度	同質性考驗 (<i>p</i> 值)	常態分配考驗 (<i>p</i> 值)	訊息加權 適合度	偏離反應 適合度
選手	5.65 ± .78	.91	.00	.54	.96 ± .52	.93 ± .77
評分者	.00 ± .47	.97	.00	.27	.78 ± .41	.93 ± .48
試做	.00 ± .15	.00	.92	.69	1.01 ± .02	.93 ± .06
試題	.00 ± 5.11	1.00	.00	.29	1.08 ± .29	.93 ± .22
評分者一致性= 74.5%			總數 7200		一致性 5361	

(二) 精熟組與非精熟組競技選手測量結果

競技選手兩組的測量分析摘要(如表 2)。

1. 人數「精熟組」23 位,「非精熟組」17 位。因為參與的受試者是志願的,無法使兩組人數完全相等,幸好差異不大,比較時 *t* 考驗仍然可行。

2. 估計潛能的平均數與標準差:「精熟組」6.10±.48、「非精熟組」5.04±.73。*t* 考驗比較差異結果:Levene's 變異數同質性考驗 *F* 值 4.62、*p* 值 .04,判定兩組變異數不同質,選擇調整自由度為 25.39,則 *t* 值 5.11、*p* 值 .00,裁決拒絕虛無假

設,表示「精熟組」和「非精熟組」估計潛能不相等,明顯的「精熟組」高於「非精熟組」。

3. 母群的 Rasch 模式個別信度:「精熟組」.74,「非精熟組」.91。顯示「精熟組」信度普通,「非精熟組」信度優良。

4. 同質性考驗結果機率是:「精熟組」.00,「非精熟組」.00。顯示兩組的潛能皆是不同質,即具有鑑別度,也證明具有區別效度。

5. 常態性考驗結果機率是:「精熟組」.66,「非精熟組」.47。顯示兩組的樣本符合常態分配,進而推論具有代表性,即兩組的樣本可以代表母群。

表 2 精熟組與非精熟組競技選手測量分析摘要表

競技選手	人數	估計潛能 (logit) 平均±標準差	Rasch 模式 個別信度	同質性考驗 (<i>p</i> 值)	常態分配考 驗 (<i>p</i> 值)
精熟組	23	6.10 ± .48	.74	.00	.66
非精熟組	17	5.04 ± .73	.91	.00	.47

二、編製桌球反手擰球技術效標參照測驗

本研究採效標組法編製桌球反手擰球技術效標參照測驗，並進一步加上臨界組法 (The Borderline Group Method) 針對 13 位評量結果有疑慮選手，進一步再請一位評者確認是否精熟，然後再進行分析，得到結果如下：

(一) 計算分界分數

效標組法的分界分數是判斷非精熟組和精熟組次數分配兩曲線的交叉點，如表 2 估計的潛能，「精熟組」平均數 6.10、標準差 .48，「非精熟組」平均數 5.04、標準差 .73。謝進昌 (2006)「對照組 M-SD 法」是根據前述兩組的平均數和標準差，產生兩條常態分配曲線，再估計兩條曲線的交叉點，此點就是判定精熟組和非精熟組的標準，又叫做分界分數 (Cut-Off Score)，計算分界分數如下：

$$\text{分界分數} = 6.10 - [0.48 \times (6.10 - 5.04) / (0.48 + 0.73)] = 5.69。$$

故多層面 Rasch 模式估計能力的平均數和標準差，推算得到分界分數為：估計潛能 5.69 (分界分數 5.69 介於原始分數 422 分和 421 分之間)。好的分界分數在於正確分辨是否精熟？也就是測驗分類的正確性，由效度 (ϕ_{vc}) 來看，效度係數愈高的分界分數愈適當，所以，效標參照測驗的正確決斷依賴分界分數決定。

(二) 效標參照測驗的信度和效度

實際測量值分類和事前判斷分類結果歸類如表 3，然後計算信度和效度於下：

1. 一致性比率：

$$PA = (20 + 15) / 40 = .88。$$

2. 柯恆 kappa 係數：

$$\Sigma P_i.P_i = (22 \times 23 + 18 \times 17) / 40^2 = .51$$

$$\kappa = (0.875 - 0.5075) / (1 - 0.5075) = .75$$

3. 修正 kappa 係數：

$$\kappa_q = (0.875 - 0.5) / (1 - 0.5) = .75$$

4. 效標參照測驗效度：

$$\phi = (20 \times 15 - 2 \times 3) / (23 \times 17 \times 22 \times 18)^{0.5} = .75$$

表 3 編製桌球反手擰球技術效標參照測驗雙向度類別表

		實際測量值分類		
		精熟組	非精熟組	合計
事前判斷分類	精熟組	20	3	23
	非精熟組	2	15	17
	合計	22	18	40

總結計算結果，桌球反手擰球技術效標參照測驗的信度：一致性比率.88，柯恆 kappa 係數 .75，修正 kappa 係數 .75；「柯恆 kappa 係數」和「修正 kappa 係數」兩者皆是修正「一致性機率」的信度指標，當我們表示效標參照測驗信度指標時，一致性比率適合說明效標參照測驗分類結果的信度，若受機率影響時，應該再加上柯恆 kappa 係數和修正 kappa 係數輔助說明。依據 Cohen (1960) 解釋柯恆 kappa 係數的一致性強度

(Strength of Agreement)， κ 值在 .61~.80 時，一致性強度是「良好 (Good)」。效標參照測驗效度 .75，計算為 2×2 列聯表的 ϕ 相關，它是「實際測量值分類」和「事前判斷分類」，兩者的相關係數，做為效標參照的效度 (Safrit, 1989. p129)。此效度係數，表示效標參照測驗的正確性分類 達 74.72%，也就是說，本測驗正確判定：精熟組或非精熟組，有近七成五的機率。

肆、討論

運動技能測驗有兩種計分：第一種是可用度量衡計量，結果直接可以計算。第二種是透過觀察判斷，給予計數或評分，其原始得分量尺未達等距量尺，不宜直接運算。曾盟堡 (2002) 研究：「國際標準舞 (簡稱：國標舞) 的評分，利用多層面 Rasch 模式分析。」國標舞的評分是採用強迫排名法，八位裁判分別給予六位選手 1 至 6 名，名次不得重複；名次是順序量尺，直接加總計算，造成惡意裁判偏差給分，有機可乘，唯有透過多層面 Rasch 模式量化，研究找出「評分者概念錯誤的隨機誤差，以及評分者偏見造成的系統性誤差，兩種偏差類型。」本研究的桌球反手擰球技術效標參照測驗，就是由評分者針對技術動作的好壞，進行質性評定；為提升

量尺品質，採用多層面 Rasch 測量模式，將順序資料轉換成等距尺度 (Wright & Masters, 1982)，原技術動作的評分（順序尺度）經過校準 (calibrations)，成為線性量尺推論為等距量尺 (Linacre, 1997)。即經多層面 Rasch 測量模式估計，得到各項估計值 (logit)，估計值它是線性且具可加性 (Linear and additivity)，故多層面 Rasch 模式所得的估計值，用於計算測驗品質的分界分數、效度、信度等，在統計和測驗的運算上獲得支持。

對照組 M-SD 法是利用統計數推估分界分數，屬於母數統計 (謝進昌, 2006)。因此，樣本分配應該符合常態分配；結果的第一節、第二小節精熟組與非精熟組競技選手測量結果：常態性考驗顯示兩組樣本皆符合

常態分配，除了樣本具有代表性，最重要的是適合對照組 M-SD 法的假定，能夠推估分界分數。研究對象事前依成就分為兩組：曾當選桌球國手的特優選手為「精熟組」及運動績優體育保送生為「非精熟組」。t 考驗兩組的估計潛能，得到 t 值 5.1、p 值 .00，裁決拒絕虛無假設，明顯的「精熟組」估計潛能 6.10 高於「非精熟組」5.04。此乃以事前主觀的效標組法為主、然後再輔以臨界組法分組，兩道主觀分類；透過客觀測量和「對照組 M-SD 法」算出分界分數，此部分是客觀的量化分析，判斷是否精熟。也就是說，以客觀量化的分類，來檢討主觀的分類法，結果顯示研究事前的分類是合理的。

本研究桌球反手擰球技術效標參照測驗，是一種技術動作質性好壞的評判，較大部分依賴評分者主觀的評定；對照姚漢禱與林德隆 (2000) 的壘球打擊測驗，結果為球的落點遠近，有可見的距離給分，屬於成績表現 (performance) 對照落點計分，主觀成份的涉入很小。測驗品質比較，結果都是本研究略低，效度係數 .89 vs .75，信度的一致性比率 .96 vs .88、柯恆 kappa 係數 .89 vs .75 和修正 kappa 係數 .92 vs .75。此現象是合理的，顯示客觀的成績表現測驗的測驗品質，會優於主觀的技術動作的質性評判；在 Xiomara, Fernando 與 Camilo (2017) 運動功能量表 (motor function measure, MFM) 的研究，也報

告了 32 個項目的柯恆 kappa 係數從 .51-.94，評價為中等至優秀 kappa 統計數據。歸納效標參照測驗的效度和信度，柯恆 kappa 係數和修正 kappa 係數：客觀的成績表現測驗在 .90 以上、主觀的技術動作測驗在 .70 以上，都屬於良好、成功的測驗 (Cohen, 1960; Cohen, 1968)。良好的運動技術效標參照測驗，更重要的是要能夠在實際的競賽中使用，且有良好的效果；換言之，技能類別的測驗，並不計較測驗的分數，理論和認知所佔的份額也不高。考試通過獲得駕照，不一定或不敢開車，所以有路考，或學習駕駛證可以暫時充當駕駛執照使用，在一段實務經驗後再頒發正式執照。延伸許多技能為主的工作都有實習制度。回顧前面「記者龍柏安報導，然後觀察比賽影片」，要說明接發球時，反手擰球技術的使用和得分情況，只有直接觀看比賽影片；更嚴謹的是比賽分析研究，故引用敖可 (2017)「以張繼科 2014-2015 年的 10 場比賽影像分析中發現，張繼科在接發球時使用反手擰球得分的比例達 81.05%。」這非常高的數據，呈現一項新技術出現，具有驚人的威力。這也是本研究選擇競技選手，編製桌球反手擰球技術效標參照測驗，希望提升桌球的競技水準。

桌球反手擰球技術是一項新發展的技術，未來隨著時間的推移，選手的技術能力，將不斷的精進；而且，桌球是一種開放性運動，特性是複雜

多變，必然會演進更多親族技術，類似病毒會不斷變種，故本研究的效標參照測驗，可能短時間內就會衰退，需要及時修訂測驗，這是研究面臨的主要問題和限制。不斷產生的問題和限制，正是研究者最有興趣挑戰的課題；關於未來的研究建議，可以擴大收集比賽中經常使用球反手擰球技術的選手，結合教練持續利用「桌球反手擰球技術效標參照測驗」來檢測選手，觀察選手是否精熟反手擰球技術？診斷選手技術動作的問題，以及技術演化精進，克服不同來球的要領。如此將有助於反手擰球技術的推廣和進步，不但能促進桌球水準的提升，另一方面，也提供修訂資料和機會。

歸納上述的討論，共計五個重點：

- 一、採用多層面 Rasch 測量模式，提升量尺品質，能合適的編製測驗品質。
- 二、考驗樣本為常態分配，除具有代表性，且滿足對照組 M-SD 法的假定，以推算分界分數。
- 三、 t 考驗結果驗證研究的事前分組合理。
- 四、效標參照測驗的效度和信度，客觀性測驗在 .90 以上、主觀性測驗在 .70 以上，為良好的測驗。
- 五、技能類的測驗，以實務上的應用表現為主。

從這五項討論的重點來看，本研究結論為桌球反手擰球技術效標參照測驗具有良好的測驗品質，能鑑別競技選手反手擰球技術的精熟。

引用文獻

- 何榮桂、姚漢禱 (1995)。運動技能效標參照測驗分界分數設定之新方法—以控制組法及技術抽離法為例。**測驗年刊**，**42**，187-196。
- 林清山、程炳林 (1996)。國中生自我調整學習因素與學習表現之關係暨自我調整的閱讀理解教學策略效果之研究。**教育心理學報**，**28**，15-58。
doi:10.6251/BEP.19960901.2
- 吳璧純、詹志禹 (2018)。從能力本位到素養導向教育的演進、發展及反思。**教育研究與發展期刊**，**14**(2)，35-64。
- 姚漢禱 (1995)。效標參照測驗和體育教學。**台大體育**，**26**，75-82。
- 姚漢禱、林德隆 (2000)。利用多層面 Rasch 測量模式編製中等學校運動績優甄審甄試測驗—以壘球打擊測驗為例一。**中華民國大專院校八十九年度體育學術研討會報告書**，41-45，國立雲林科技大學。
- 敖 可 (2017)。張繼科反手側擰技術在比賽中的應用分析。**體育時空**，**2**，122。
- 陳鉸澈 (2016)。大專跆拳道品勢動作難易度之主觀認知分析。**跆拳道學刊**，**3**，11-23。
- 陳建彰、游鳳芸、姚漢禱 (2020)。桌球橫拍反手擰球介紹與技術分析。**大專體育**，**152**，1-9。
doi:10.6162/SRR.202003_(152).0001
- 張宏亮 (1995)。效標參照測驗及其在體育上的應用。**中華體育季刊**，**8**(4)，12 -22。
doi:10.6223/qcpe.0804.199503.1702
- 曾盟堡 (2002)。是誰評判不公。**測驗統計年刊**，**10**，121-133。
doi: 10.6773/JRMS.200212.0121
- 董 聲 (2017)。反手擰拉技術發展回顧及其對乒乓球運動的重要意義。**山東體育科技**，**39**(3)，50-53。
- 龍柏安 (2021，03 月 13 日)。林昀儒 鄭怡靜擊敗南韓組合 WTT 連兩站奪冠。中央社電子報，取自 <https://www.cna.com.tw/news/firstnews/202103135008.aspx>
- 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**，157-193。
- 魏豐閔、施登堯 (2014)。體育成績考查標準的選擇與設定。**臺灣教育評論月刊**，**3**(9)，133-144。

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *The standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Berk, R. A. (1976). Determination of optimal cutting scores in criterion-referenced measurement. *The Journal of Experimental Education*, 45, 4-9. doi: 10.1080/00220973.1976.11011567
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46. doi: 10.1177/001316446002000104
- Cohen, J. (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70(4), 213-220. doi: 10.1037/h0026256
- Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, 18, 519-521. doi: 10.1037/h0049294
- Looney, M. A. (1987). Threshold loss agreement indices for criterion-referenced measures: A review of applications and interpretations. *Research Quarterly Exercise and Sport*, 58, 360-368. doi:10.1080/02701367.1987.10608113
- Linacre, J. M. (1997). *A user's guide to Facets Rasch measurement computer program*. Chicago, IL: MESA Press.
- Safrit, M. A. (1989). Criterion-Referenced Measurement: Validity. In M.J Safrit & T.M. Wood (Ed.), *Measurement Concepts in Physical Education and Exercise Science* (119-135). Champaign, IL: Human Kinetics Publishers.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement*. Chicago, IL: MESA Press.
- Xiomara, R-C., Fernando, O-C., & Camilo, M-P. (2017). Reliability of home-based, motor function measure in hereditary neuromuscular diseases. *Journal of International Medical Research*, 45(1), 261-271. doi:10.1177/0300060516674608

作者簡介

第一作者：陳建彰

服務單位：國立體育大學體育研究所、明志科技大學體育室

第二作者：姚漢禱

服務單位：國立體育大學競技與教練科學研究所

第三作者：游鳳芸（通訊作者）

服務單位：國立陽明交通大學體育室

通訊地址：300 新竹市大學路 1001 號

連絡電話：0932074359

E-mail：fenny660125@gmail.com

A construction of Criterion-Referenced test for backhand twist skills in table tennis

Chien-Chang Chen

Graduate Institute of Physical Education, National Taiwan Sport University
Office of Physical Education, Ming Chi University of Technology

Han-Dau Yau

Graduate Institute of Athletics and Coaching Science, National Taiwan Sport University

Feng-Yun Yu*

Office of Physical Education, National Yang Ming Chiao Tung University

Abstract

Purpose: The objective for this study was to construct of criterion-referenced test for backhand twist skills in table tennis using many-facet Rasch measurement model. **Methods:** 40 athletes were selected as the study subjects and was separated to two groups: 23 of them in Masters and 17 in Non-masters. The backhand twist technique criterion-referenced tests were applied to all of the subjects. The athletes, raters, trials and items were evaluated using many-facet Rasch measurement model. **Results:** The test data were fitted for Rasch model; The ability of cut-off score for criterion of validity reference test was 5.685 logit that corresponds to the original score between 422 and 421, and the validity coefficient was 0.75. The reliabilities of the agreement proportion was 0.88, the Cohen kappa was 0.75 and the modified kappa was 0.75. **Conclusion:** The criterion-referenced test for backhand twist skills in table tennis did generate good quality test results, and can be used to distinguish the master of backhand twist skills in athletes.

Keywords: Proportion of Agreement, cut-off score, validity coefficient, the backhand twist skills in table tennis, many-facet Rasch model

一、桌球反手擰球技術動作評量內容

(一) 準備動作 (Preparation period)：

1. 站立—腳的位置。
2. 站立—軀幹位置。
3. 重心位置—重心要低。
4. 重心位置—身體中心線。
5. 球拍位置—球檯面上方。
6. 眼睛—注視對方來球。

(二) 引拍動作 (Back Swing period)：

7. 步法—腳的移動。
8. 重心—重心移至右腳。
9. 手肘位置—手肘抬高。
10. 手腕—依來球調整角度。

(三) 擊球前動作 (Impact Front period)：

11. 手腕—球拍端部朝身體。
12. 手腕—旋內轉向來球左側。
13. 手肘位置—高手肘迎球。
14. 手肘—手肘與肩部的的位置。
15. 重心運用—放在右腳上。
16. 重心運用—放在腹部上。

(四) 擊球時動作 (On Impact period)：

17. 手肘—高手肘為軸心。
18. 重心—重心放在右腳。
19. 手腕—旋內向右前上方。
20. 球拍位置—擊球瞬間。
21. 球拍位置—擊球過程。

(五) 擊球後動作 (Impact Back period)：

22. 揮拍—手臂的揮動。
23. 揮拍—手腕的位置。
24. 墊步—右腳的移動。

(六) 還原動作 (Follow Way period)：

25. 收拍—手肘的位置。
26. 收步—腳的迅速還原。
27. 站位—腳的位置。
28. 站位—軀幹的位置。
29. 重心位置—重心要低。
30. 球拍位置—球檯面上方。

二、計分：各擊球階段動作要點，每達成一個動作，給予一分，未達成計零分。

三、試做：每人三次，錄影以供評分。

四、評分：五位評分者依計分方式，進行錄影觀察評分。