

《測驗學刊》
第七十一輯第四期 2024 年 12 月 423~455 頁

整合分析可靠嗎？用結構方程模式處理整合分析中的 檔案櫃效應：以 16 個資優學生的數學自我概念研究 為例

曹傑如¹ 張道宜² 楊智為³

摘要

出版偏差，俗稱檔案櫃效應，是指達統計顯著或符合預期結果的研究才易被接受出版，進而影響整合分析研究結果的可靠性和完整性。出版偏差導致了整合分析研究結果系統性的高估或低估，成為質疑研究結果可靠性的主要因素。因此，發現與校正出版偏差對確保整合分析研究結果的客觀性和準確性至關重要。本研究基於 Cheung (2015) 及李茂能 (2016) 以結構方程模式 (SEM) 軟體處理整合分析的研究成果上，增加 AMOS 軟體進行傳統整合分析的作法，並擴展對出版偏差的偵測與修正。次以 Infantes-Paniagua 等人 (2022) 的 16 個資優學生之數學自我概念整合分析研究成果作為範例，並以 R 軟體的分析結果作為對照，用以驗證本研究的正確性。結果有三：首先，透過 SEM 進行整合分析的方法論，本研究以 AMOS 軟體達成處理傳統整合分析的預期目標；其次，無論是採用固定效果模式還是隨機效果模式，範例中的整合分析結果呈現出版偏差，經本研究藉由 SEM 可偵測效果值高估現象並進行修正；最後，與傳統的整合分析相比，利用 SEM 進行出版偏差的偵測及校正有其便利性，且視覺化結果呈現方式更有助於研究者更直觀地理解複雜的統計資料。

關鍵詞：出版偏差、異質性、資優、數學自我概念

-
1. 曹傑如，國立臺中教育大學特殊教育學系助理教授
 2. 張道宜，國立臺中教育大學博士後研究員
 3. 楊智為，國立臺中教育大學教育資訊與測驗統計研究所助理教授
- 收件日期：2024.03.08；完成修改：2024.05.19；正式接受：2024.05.24
通訊作者：曹傑如；Email：tcj@mail.ntcu.edu.tw
地址：403514 臺中市西區民生路 140 號 國立臺中教育大學特殊教育學系

Is Meta-Analysis Reliable? Addressing the File Drawer Effect in Meta-Analyses Using Structural Equation Modeling: A Case Study of 16 Researches on the Mathematics Self-Concept of Gifted Students

Chieh-Ju Tsao¹ Tao-Yi Chang² Chih-Wei Yang³

Abstract

Publication bias, commonly referred to as the file drawer effect, denotes the preferential publication of studies with statistically significant or anticipated results, compromising the reliability and completeness of meta-analytic research. This leads to systematic overestimation or underestimation of research outcomes, casting doubts on their reliability. Detecting and correcting publication bias is crucial for maintaining the objectivity and accuracy of meta-analytic findings. This study builds upon Cheung (2015) and Li (2016)'s research, utilizing structural equation modeling (SEM) with AMOS software to refine traditional meta-analysis methods and address publication bias. Furthermore, this study re-analyzes 16 studies on gifted students' mathematical self-concept, comparing them to R software results to verify the method's validity. The findings reveal that SEM, particularly with AMOS, effectively detects and corrects overestimated effect sizes, demonstrating its utility in meta-analytic processes. The study confirms the benefits of SEM in providing a more streamlined and intuitive approach to identifying and correcting publication bias, facilitating a better understanding of complex statistical data through visual representation.

Keywords: gifted, heterogeneity, mathematics self-concept, publication bias

1. Chieh-Ju Tsao, Assistant Professor, Department of Special Education, National Taichung University of Education

2. Tao-Yi Chang, Postdoctoral Research Fellow, National Taichung University of Education

3. Chih-Wei Yang, Assistant Professor, Graduate Institute of Educational Information and Measurement, National Taichung University of Education

Received: 2024.03.08; Revised: 2024.05.19; Accepted: 2024.05.24

Corresponding Author: Chieh-Ju Tsao; Email: tcj@mail.ntcu.edu.tw

Address: No. 140, Minsheng Rd., West Dist., Taichung City 403514, Taiwan

Department of Special Education, National Taichung University of Education

壹、緒論

在資訊爆炸的時代，科學研究面臨的挑戰之一是需要從龐大的數據中提取有意義且可靠的結論。隨著研究數量急劇增加，傳統的文獻審閱方法逐漸無法滿足快速總結和評估海量訊息的需求。在這樣的背景之下，整合分析（meta-analysis）作為一種系統性的文獻探討方法，應運而生，其主要目的在於統計整合過去實證研究的結果，以回答特定的研究問題，並據此進行推論。Borenstein 等人（2009）闡述，整合分析能透過標準化結果，高效率的比較不同治療作為或教學方法，整合個別資料彙整出最優治療選擇或教學實踐作法。目前，該方法在醫學（Tanriver-Ayder et al., 2021; Van Aert et al., 2019）和教育（Cook, 2014; Mostert, 2021; Scruggs & Mastropieri, 2021）等領域廣獲應用。

李茂能（2015）指出，整合分析透過明確計算變項之間的效果值（effect size）以及探尋可能調節這些效果變項。Cumming（2012）進一步強調，整合分析透過客觀量化文獻的努力，減少了研究結果的不確定性，並增強了結果的可複製性。這種方法透過整合效果值，不僅改善了虛無假設顯著性考驗（Null Hypothesis Significance Testing [NHST]）的不足，還克服了樣本大小對實驗效果的影響。因此，整合分析是一個強大的總結研究發現工具。

整合分析作為總結研究發現的重要工具，但其結果的可靠性卻常受到異質性（heterogeneity）和出版偏差（publication bias）的影響。儘管異質性可透過諸如隨機效果模型等統計技術進行一定程度的調節和控制，但對於矯正出版偏差的研究相對缺乏，此限制了研究結果的普遍性和可信度，特別是在特殊教育領域，這些問題更加明顯。Gage 等人（2017）指出，在特殊教育領域的整合分析中，有高達 42% 的研究顯示出版偏差，但僅有 33% 的研究試圖檢驗這些偏差，此顯示了一個重要的研究空白。這一現象反映的不僅是研究方法上之挑戰，而且還指向了特殊教育領域常見的小樣本研究困境。小樣本研究為了達到統計顯著性，往往需要較大的效果大小，這不僅增加了研究被出版的可能性，同時也可能導致效果值的高估（Cook, 2014）。

Cheung（2015）認為，運用結構方程模式（Structural Equation Modeling [SEM]）於整合分析中具有特別的優勢。這不僅豐富了 SEM 使用者的統計工具箱，還允許使用者同時進行原本的研究及整合分析。反過來說，SEM 為整

合分析研究者提供了一系列新的研究工具，引入了諸如單變量、多變量和階層整合分析中缺失值處理等新的研究工具。同時，將 SEM 與整合分析整合到一個框架中，對於研究生的統計培訓特別有益。Cheung（2015）亦提及統計軟體 Mplus 與 R（package metaSEM）在 SEM 整合分析中的應用，突出了 SEM 在多變量整合分析中的優勢。然而，目前在使用 SEM 進行出版偏差檢測與糾正方面的研究仍然不足。

AMOS 是另一種常見的 SEM 統計軟體，但這不代表它不能用於整合分析。考慮到 SEM 和整合分析在處理變項間的關係、效果大小及異質性方面有諸多共通之處，使用 AMOS 進行整合分析既可行又有其獨特價值。值得注意的是，雖然 Mplus 與 R 在 SEM 的整合分析應用上已被廣泛討論（Cheung, 2015），由於 AMOS 的圖形界面容易理解，是許多 SEM 研究者主要使用的軟體，但關於其在整合分析中的應用，相關討論卻相對較少。本研究主要探討了 AMOS 在傳統整合分析中的可行性，包括對整合效果值的評估、異質性的分析，以及調節變項的考察。透過這些分析，本研究旨在展示 AMOS 如何幫助研究者更深入地探索和整合現有的研究成果。此外，本研究還進一步探討了 AMOS 在偵測和糾正出版偏差方面的能力。

最後，透過結合 AMOS 的功能與本研究開發的方法，本研究期望能為其他研究者在進行整合分析時提供更大的彈性和深度。本研究預期，這種創新方法將在未來的整合分析研究中扮演愈加重要的角色。

本研究的待答問題如下：

1. 是否可以利用 AMOS 執行傳統的整合分析？
2. 是否可以利用 AMOS 執行出版偏差的偵測與修正？

就傳統整合分析而言，本研究的分析項目涵蓋整合平均效果值，其中包含固定與隨機效果模式，另外還有異質性分析以及調節變項分析。

貳、文獻探討

一、整合分析對於特殊教育的重要意義

在特殊教育領域，整合分析不僅是一項關鍵的研究方法，更是一個不可或缺的资源。以下是整合分析所帶來的幾項核心優勢（Forness, 2001; Scruggs & Mastropieri, 2021）：

1.解決研究不一致性：在特殊教育中，關於教學策略效果的研究報告常呈現不一致的結果。整合分析能夠匯集並分析多項研究數據，提供對教學策略效益和局限性的全面與精確評估。

2.量化策略效能：整合分析使研究者能夠計算策略的效應量，精確呈現策略的平均效果。此一量化數據不僅讓教學策略的影響更為顯而易見，也方便進行不同策略的效果比較。

3.評估多種策略：整合分析不僅用於評估特定教學策略的效果，也是評價和比較特殊教育及相關服務策略效果的重要工具。

4.指引個體化教學：雖然整合分析提供的是一般性指引，但特殊教育專業人士明白每位學生都有獨特的需求。整合分析結果可被用來設計和調整個體化教學策略，幫助教育工作者制定更具針對性和有效性的教學計畫。

儘管整合分析在特殊教育領域中被廣泛應用和認可，但仍需謹慎評估其結果的可靠性和效度，以確保研究結果的精確性和應用價值。這意味著在評估整合分析效度時，需要特別關注異質性和出版偏差等問題（Mostert, 2021）。

二、結構方程模式（SEM）於整合分析之應用

藉由傳統的整合分析方法可以解決在整合分析上的大多數問題，但 SEM 提供了一個更為全面和靈活的分析框架。Cheung（2015）利用 Mplus 軟件和 R 的 metaSEM 包，以 SEM 的方法來執行整合分析。李茂能（2016）也在其著作中展示了 SEM 在整合分析中的應用，特別是如何使用 AMOS 來計算整體效果值，不過他並未使用 AMOS 進行有關異質性的計算、隨機效果值的計算，以及利用 AMOS 進行出版偏差的偵測與修正。本研究將基於李茂能（2016）的研究成果來進行擴展。在本節中，將詳細介紹傳統整合分析和使用 SEM 進行整合分析的方法。

（一）傳統整合分析

傳統整合分析主要是透過計算各研究的效果值（ d_i ）和其加權量（ w_i 變異量的倒數）來得出整體效果值，如公式(1)所示：

$$\bar{d} = \frac{\sum_{i=1}^k d_i w_i}{\sum_{i=1}^k w_i} \quad (1)$$

測驗學刊，第 71 輯第 4 期

若要在隨機效果模式下計算整體效果值，則需要將加權量進行調整，如公式(2)所示：

$$W_{i(Random)} = \frac{1}{v_i + \tau^2} \quad \tau^2 = \frac{Q - df}{\sum w - \frac{\sum w^2}{\sum w}} \quad (2)$$

在公式(2)中， v_i 代表各研究的變異量， τ^2 是研究間的變異。然而，在一般的異質性檢驗中， τ^2 並不常被用作指標， Q 值和 I^2 被更廣泛地應用於評估研究的異質性。公式(3)中的 Q 用於檢定整合效果值的異質性。 I^2 （如公式(4)所示）則是另一個常用的異質性指標，需通過 Q 值來計算：

$$Q = \sum w_i (d_i - \bar{d})^2 \quad (3)$$

$$I^2 = (Q - df) / Q \quad (4)$$

為了評估整合研究的異質性，一般會檢視 Q 值（也就是卡方值）的 p 值是否大於顯著水準 α 。當使用 I^2 作為異質性的評估標準時，如果 $I^2 < 25\%$ ，則整合研究可視為無顯著異質性； I^2 若在 25% 至 50% 之間表示輕度異質性； I^2 在 50% 至 75% 之間表示中度異質性；而 $I^2 > 75\%$ 則表示高度異質性（Higgins & Thompson, 2002）。本研究以特殊教育整合分析研究為目標。單一個案的研究資料，若所提供資訊無法計算標準化平均差異值（Standardized Mean Difference [SMD]），Parker 等人（2011）指出，尚有 ECL、NAP、IRD、PHI、PEM、Tau-U 與 PND 等方法，來求得單一個案研究的效果值，不過上述方法在整合分析上並不多見。

（二）以結構方程模式（SEM）處理整合分析

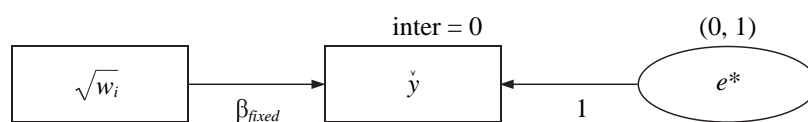
在了解整合分析的原理後，接著來看 SEM 的原理，並討論處理整合分析的作法。SEM 在整合分析中的應用基礎建立在迴歸模式之上，透過此模式能進行效果值的整合和估計。在公式(5)中， y_i 代表該每個研究的效果值， β_{Fixed} 代表母群固定效果值， e_i 代表抽樣誤差（假設為常態分配，平均數為 0，變異數為 σ_i^2 ）：

$$y_i = \beta_{Fixed} + e_i \quad \beta_{Fixed} = \frac{\sum_{i=1}^k w_i y_i}{\sum_{i=1}^k w_i} \quad (5)$$

李茂能（2016）指出，由於整合分析的效果值常無法滿足 SEM 的變異量同質假設，因此需要透過加權轉換（公式(6)）進行調整。加權後的觀察效果值和誤差分別表示為 \tilde{y} 和 e^* 。透過 SEM（如圖 1 所示），經迴歸分析後可用 \tilde{y} 和 $\sqrt{w_i}$ 獲得固定效果 β_{Fixed} 。

$$\frac{1}{w_i^2} y_i = \frac{1}{w_i^2} \beta_{Fixed} + \frac{1}{w_i^2} e \rightarrow \tilde{y} = \sqrt{w_i} \beta_{Fixed} + e^* \quad (6)$$

圖 1 在結構方程模式（SEM）中整合分析之固定模式



在 SEM 中值得注意的是誤差項的變異數處理，由於加權量為 $\sqrt{w_i}$ ，因此誤差 e_i 轉換後成為 e^* 。由於變異數是加權的倒數，理論上需設定為 1（公式(7)）。但在 SEM 模式中，誤差項的變異數設定為 1 與開放估計的兩種設定在分析中有不同的意義。設定誤差變異量為 1 能確保在固定模式下獲得準確的整合效果值標準誤和信賴區間（Cheung, 2015; Raudenbush et al., 1988）。若未設為 1，Cheung（2015）指出，誤差變異量乘以研究數等於 Q 值（公式(8)）。此外如圖 1 所示，在 SEM 上 \tilde{y} 的截距要設定為 0（李茂能，2016）。

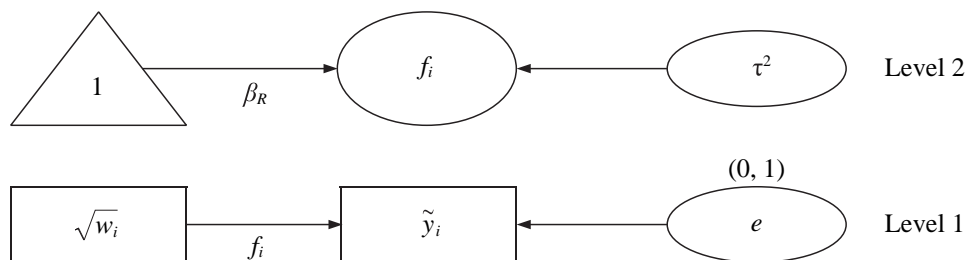
$$\text{Var}(e^*) = \text{Var}(\frac{1}{w_i^2} e) = w \times \text{Var}(e) = 1 \quad (7)$$

$$\begin{aligned} Q &= \sum_{i=1}^k w_i (y_i - \hat{\beta}_{Fixed})^2 \\ &= \sum_{i=1}^k (y_i^* - \sqrt{w_i} \hat{\beta}_{Fixed})^2 = K \times e^* \end{aligned} \quad (8)$$

（三）隨機效果之處理

整合分析需考慮在隨機效果模式下，處理整體平均效果值和計算研究間的變異量。在 SEM 的隨機效果整合分析中，利用兩層的迴歸模式（公式(9)）進行整合估計。隨機模式的 SEM 整合分析如圖 2 所示，由於考慮到每個研究的效果值為獨特分布，每個研究的效果值視為獨特效果值 f_i ，而 β_R 和 τ^2 分別代表母群隨機效果值和研究間的變異量（公式(10)和(11)）。

圖 2 隨機效果模式之徑路圖



整合分析研究在估算 β_R 和 τ^2 上主要採用兩種估算方式（Cheung, 2015），傳統整合分析傾向如公式(2)與(10)所示的 DerSimonian Laird 法（DL），另一種則為公式(11)所示的限制性最大概似法（Restricted Maximum Likelihood [REML]）。兩種方法在估算 β_R 上是一致的，但在 τ^2 的估算上兩者方法會有不一致的結果。在估算上，REML 所估算出的 τ^2 值較 DL 低，且 Tanriver-Ayder 等人（2021）指出在高異質性的研究上，REML 的估算較精準。但在執行上，DL 法有較高的便利性，REML 法的使用則受限於統計軟體是否有提供（李茂能，2015）。

$$\text{Level 1: } \tilde{y}_i = \sqrt{w_i} f_i + \tilde{e}_i$$

$$\text{Level 2: } f_i = \beta_R + \tau^2 \quad (9)$$

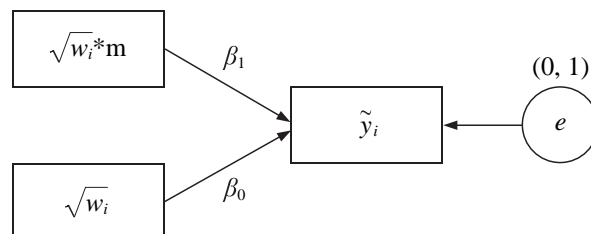
$$\beta_{\text{Random}} = \frac{\sum_{i=1}^k \tilde{w}_i y_i}{\sum_{i=1}^k \tilde{w}_i}, \quad \tilde{w}_i = \frac{1}{\tau^2 + \sigma_i^2} \quad (10)$$

$$-2LL_i(\beta_R, \tau^2) = \log(2\pi) + \log(\tau^2 + v_i) + \frac{(y_i - \beta_R)^2}{\tau^2 + v_i} \quad (11)$$

（四）結構方程模式（SEM）之整合迴歸處理調節變項

整合迴歸與次群體分析不同，它將調節變項和加權量納入預測變項進行分析。就調節變項的處理上，如果異質性不是很高的話，大都利用固定效果模式處理，如公式(12)所示，其中第一層為固定效果值的處理，第二層為調節變項（ M_i ）與其效果值（ β_1 ）的處理。在 SEM 下以整合迴歸處理調節變項，其模式的示意圖如圖 3 所示。

圖 3 固定效果模式下處理調節變項之整合迴歸



$$\text{Level 1: } \tilde{y}_i = \sqrt{w_i} \times d_i = f_i \times \sqrt{w_i} + \tilde{e}_i$$

$$\text{Level 2: } f_i = \beta_0 + \beta_1 \times M_i + u_i \quad (12)$$

若研究經由圖 3 方式處理仍存有高度的異質性，可以結合隨機和固定效果的優勢，使用混合效果模式來進行調節變項的分析（Cheung, 2015），其 SEM 下的混合模式如圖 4 所示，或是將圖 3 中的權重（ w_i ）加上合併後研究間的變異（ τ^2_{pooled} ）。與原來隨機效果模式中 τ^2 的不同之處，在於混合模式合併了調節變項的解釋力， τ^2_{pooled} 的計算可參考公式(13)，其中 QE_{fixed} 為圖 3 開放誤差變異項估計所得出的 Q 值。這種整合迴歸方法提供了一個更靈活和全面的方式來處理調節變項，在整合不同研究和數據時能夠更精確地控制和評估調節變項的影響。公式(13)類似公式(2) τ^2 的算法，其中 m 為調節變項的組別， C_i 則為各組內 $\Sigma W - \frac{\Sigma w^2}{\Sigma w}$ 的計算。若調節變項為連續變項，則 m 值為 1。

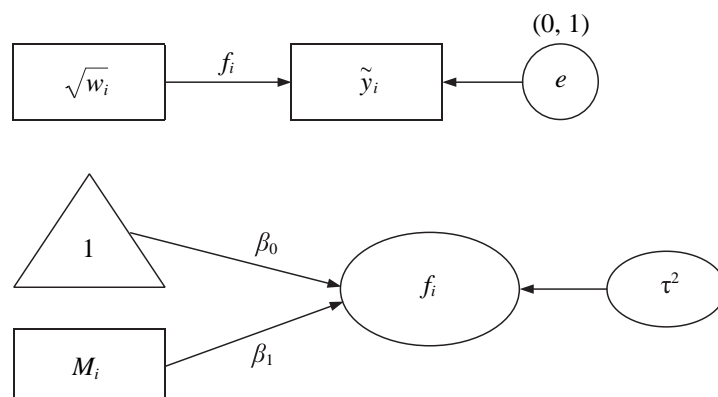
$$\tau^2_{pooled} = \frac{QE_{fixed} - (k - m - 1)}{\Sigma_{i=1}^m C_i} \quad (13)$$

三、出版偏差的處理

出版偏差的處理是整合分析的關鍵步驟之一，主要是因為研究者常無法取得所有相關文獻，從而影響結果的準確性。出版偏差，亦稱為檔案櫃問題，往往源於小樣本研究的不顯著結果未被發表（李茂能，2015）。這種偏差在整合分析中非常常見，特別是當研究結果不顯著或與預期相反時，這些研究可能不會被發表，導致整體研究結果的偏差（李茂能，2015）。在特殊教育領域中，忽略不顯著結果是一個普遍問題，這也常導致研究的效果值被

測驗學刊，第 71 輯第 4 期

圖 4 混合模式下處理調節變項之整合迴歸



高估 (Cook, 2014)。因此，為了有效處理出版偏差，研究者需先檢測其存在，隨後進行修正。

就出版偏差的處理上，傳統整合分析常見的作法有四種，分述如下。

(一) 安全失效估算法 (Fail-Safe Number)

安全失效估算法由 Rosenthal (1979) 所提出，目的是評估需要多少未發表的研究結果，才能使整合分析結果的顯著性受到影響。該方法考慮了整合分析內的所有篇數，FSN 值為 $19S-N$ ，容忍程度值為 $5k+10$ ，若 FSN 值大於容忍程度值，則該研究沒有出版偏差。然而，該方法的一個主要缺點是它假設遺漏的研究平均效果值為 0，且過度依賴 p 值，而無法提供對效果值大小的實質修正 (Orwin, 1983)。

(二) Egger 迴歸對稱性考驗

Egger 迴歸對稱性考驗係根據效果值和其標準差之間的正相關來檢測出版偏差 (Sterne et al., 2006)，如公式(14)所示。當 β_1 的顯著性測試不為 0 時，表示漏斗圖不對稱，存在出版偏差 (Pigott, 2012)。該方法需要至少有六篇研究，且 p 值設定為 .10 較為合適 (李茂能, 2015)。儘管根據公式(14)的 β_0 ，可以得出修正後的效果值，但方法並不精確，Stanley (2008) 以及 Stanley 與 Doucouliagos (2012) 建議使用 PEESE 模式來修正估計的效果值。

$$ES_i = \beta_0 + \beta_1 * SE_i + e_i \quad (14)$$

（三）Begg 與 Mazumdar 等級相關考驗

Begg 與 Mazumdar (1994) 提出利用效果值及其變異量之間的等級相關性來檢測出版偏差。其中，正值表示小樣本研究傾向報告較大的效果值，此暗示可能存在出版偏差。然而，此方法在診斷出版偏差的效能方面不及 Egger's 迴歸對稱性考驗，且不支持進行偏差校正。

（四）刪補法

Duval 與 Tweedie (2000) 提出的刪補法，旨在透過調整漏斗圖的不對稱性來校正出版偏差。該方法首先迭代移除對漏斗圖不對稱影響最大的極端研究，進而創見一個更為對稱的漏斗圖，並計算其效果值，再透過添加鏡像研究填補原有被移除研究的空缺。雖然此方法能有效地糾正出版偏差，但也有其局限性，特別是漏斗圖的不對稱並非全由出版偏差引起時，這在小樣本研究中容易出現 (Coburn & Vevea, 2015)。

Van Aert 等人 (2019) 認為，Egger 迴歸對稱性考驗是最有效的出版偏差檢測工具，而 Stanley 與 Doucouliagos (2013) 則偏好 PEESE 模式，認為它在校正出版偏差時的誤差較小。兩種方法都可以透過 SEM 進行，提供了靈活的應用可能性。相較之下，刪補法則需要依賴特定的整合分析軟體（如 CMA），且無法透過 SEM 來實施，此限制了其使用範圍。因此，研究者在選擇合適的方法時，應考慮研究的具體需求、資料的可用性，以及分析工具的可接受度，以確保出版偏差被有效識別和糾正。

四、結構方程模式（SEM）結構下進行出版偏差的分析

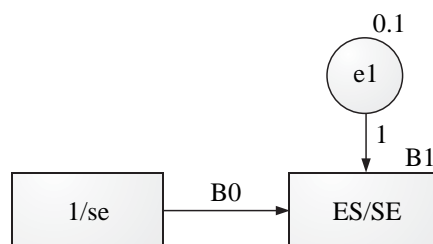
在 SEM 中如果要利用 Egger 迴歸對稱性考驗來檢測出版偏差（公式 (14)），則必須將公式(14)加權〔 $\sqrt{w_i}$ 或是 SE_i （誤差）〕轉換成下列之公式 (15)。在公式中， ES_i 為研究之效果值， β_0 為校正後之效果值， β_1 為出版偏差之估計。如果要考驗是否有出版偏差，則是考驗 β_1 是否顯著不為 0 (Sterne et al., 2006)。公式(15)在 SEM 的模式如圖 5 所示，為 FAT (Funnel asymmetry test) -PET (Precision effect test) 模式，係利用漏斗圖之對稱考驗來檢驗出版

測驗學刊，第 71 輯第 4 期

偏差，並利用淨效果值的考驗來校正出版偏差。事實上，圖 5 與圖 1 相似，差別在於截距上的估計，圖 1 設定為 0，圖 5 則開放估計（ β_1 ）。

$$t_i = \frac{ES_i}{SE_i} = \beta_1 + \beta_0 \times \frac{1}{SE_i} + \tilde{e}_i \quad (15)$$

圖 5 SEM 下之 FAT-PET 模式



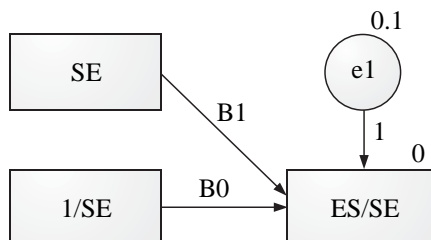
但在使用公式(15)處理出版偏差時，估算出的淨效果值並不穩定，因此 Stanley（2008）建議用 PEESE（Precision-effect estimate with standard error）模式來求得經過出版偏差校正後的真正效果值（true effect），也就是公式(14)上以 SE^2 取代原有的 SE 項（公式(16)）。若要利用 SEM 進行 PEESE 模式，則須將公式(16)加權，也就是將兩端除以標準誤，或是乘上權重的平方根（ $SE = \text{sqrt}W$ ），轉換成公式(17)，該公式在 SEM 上的模式示意圖如圖 6 所示。

$$ES_i = \beta_0 + \beta_1 * SE_i^2 + e_i \quad (16)$$

$$t_i = \frac{ES_i}{SE_i} = \beta_1 \times SE_i + \beta_0 \times \frac{1}{SE_i} + \tilde{e}_i \quad (17)$$

根據 Stanley 與 Doucouliagos（2012）的研究，高異質性的存在會影響出版偏差修正的準確性。若是在面對高異質性的整合結果，在修正出版偏差時，最好是能夠找出調節變項來降低異質性，然後再修正出版偏差，如此才能提高研究的效度（李茂能，2015；Cheung 2015）。因此，在進行整合分析時，為了提高研究的效度，Stanley（2008）、Stanley 與 Doucouliagos（2013），以及李茂能（2015）建議在 PEESE 模式上能夠加入調節變項，然後再藉由出版偏差的修正，得出更精準的整合結果。此又稱為 FAT-PET-PEE-

圖 6 結構方程模式 (SEM) 下之 PEESE 模式

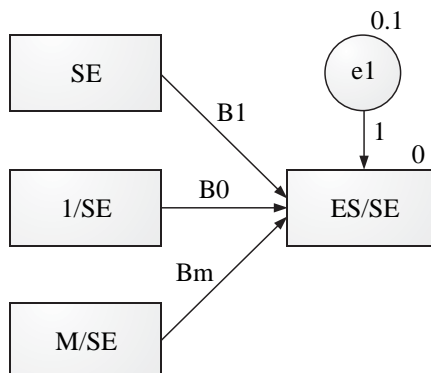


SE-MRA (Funnel Asymmetry Test-Precision Effect Test-Precision Effect Estimate with Standard Error-Meta Regression Analysis) 模式，它結合了 Egger 迴歸對稱性考驗 (FAT) 以考驗漏斗圖的不對稱性，以 PET 來檢查校正後的出版偏差，PEESE 則是利用標準誤來修正出版偏差後的效果值，MRA 則為多元迴歸分析。FAT-PET-PEESE-MRA 模式源自於公式(18)，經由轉換成公式(19)，該公式除了找出修正後的真正效果值外，而且可以了解異質性的來源與分布。如果要將 FAT-PET-PEESE-MRA 模式在 SEM 執行，其模式如圖 7 所示，其中 M_i 為調節變項， β_m 為調節變項的效果值。

$$ES_i = \beta_0 + \beta_1 * SE_i^2 + \beta_m * M_i + e_i \quad (18)$$

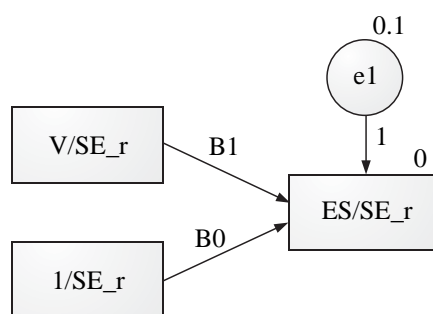
$$t_i = \frac{ES_i}{SE_i} = \beta_1 \times SE_i + \beta_0 \times \frac{1}{SE_i} + \beta_m * \frac{M_i}{SE_i} + \tilde{e}_i \quad (19)$$

圖 7 結構方程模式 (SEM) 下之 FAT-PET-PEESE-MRA 模式



上述所提到的出版偏差模式，不論是 FAT-PET、PEESE，或是 FAT-PET-PEESE-MRA，其應用情境均在固定效果模式下。一般 FAT-PET 模式主要是以 FAT 為主，也就是在固定效果模式下利用 Egger 迴歸對稱性考驗檢測出版偏差（李茂能，2015）。若是要在隨機效果模式下討論 FAT-PET、PEESE 與 FAT-PET-PEESE-MRA 模式，上述公式(15)、(17)與(19)所乘的加權項 $1/SE$ 或 SE 要考慮到 τ^2 。值得注意的是，PEESE 與 FAT-PET-PEESE-MRA 兩個模式在處理隨機效果模式上的方式，由於兩個模式中除了 $1/SE$ 項外，其餘的出版偏差變項（ SE ）與調節變項（ M/SE ）均屬於共變項，因此在研究間的變異 τ^2 值的處理都會採用混合模式，其方法便是利用圖 6、圖 7 模式，將誤差項的變異數開放估計然後乘上研究篇數得到 Q 值，然後再利用公式(13)求得混合模式的 τ^2 值，得出新的變異數與標準誤，於此標註為 SE_r ，所以在混合效果模式下，以 SEM 修正出版偏差的 PEESE 與 FAT-PET-PEESE-MRA 的模式為下列所示，其中圖 8 與圖 9 中的 V 為變異數，也是 SE^2 。

圖 8 結構方程模式（SEM）下之 PEESE 模式（混合模式）

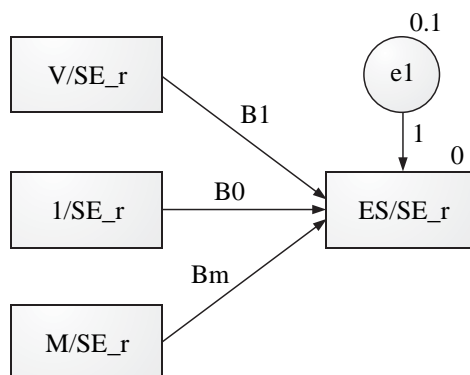


參、研究方法

一、研究設計

本研究將以 AMOS 為工具，利用 SEM 進行整合分析。首要目的在於是否能夠在固定效果模式與隨機效果模式中，進行效果值的整合；再者，針對整合的效果值進行異質性分析，或是針對效果值的異質性進行調節變項分

圖 9 結構方程模式（SEM）下之 FAT-PET-PEESE-MRA 模式（混合模式）



析。最後，同時針對整合結果來偵測出版偏差與修正出版偏差。

本研究以 Infantes-Paniagua 等人（2022）中的資優生與一般生於數學自我概念之整合分析的研究結果作對照，以 AMOS 為研究工具進行整合分析。同時，也利用 R 中的 metafor 所分析出的數據作為本研究結果的參照。本研究選取樣本結果為顯著，且整合結果在漏斗圖的呈現下具有不對稱性，顯示該研究結果存在出版偏差，可以作為校正整合偏差的樣本。

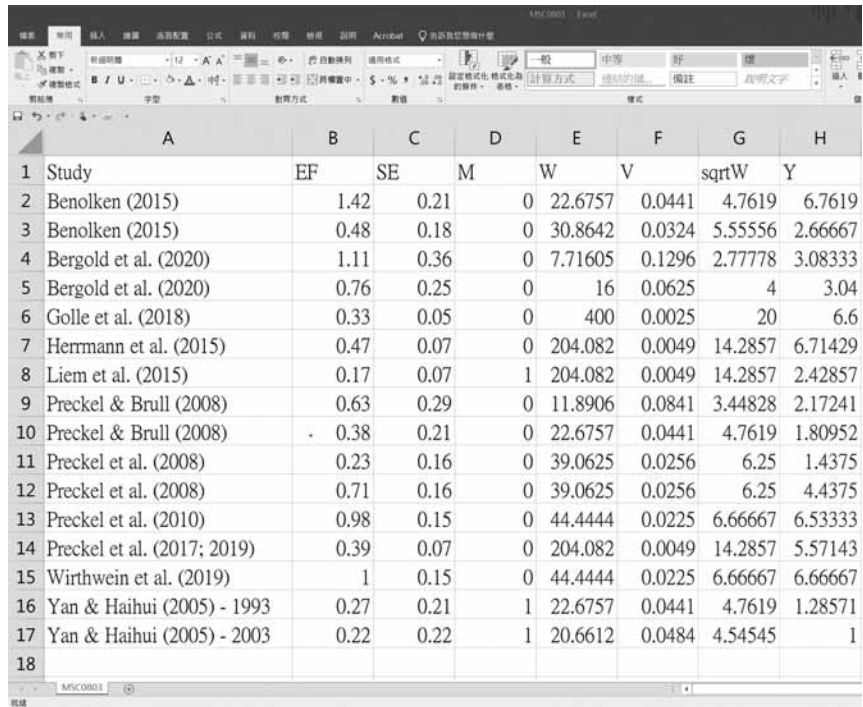
二、研究樣本

本研究進行傳統整合分析與出版偏差的偵測與修正，所採用的資料來源為 Infantes-Paniagua 等人（2022）於數學自我概念的整合分析研究結果。該研究共有 16 篇關於資優生與一般生於數學自我概念的比較，其所提供的效果值（EF）與標準誤（SE）如圖 10 所示。本研究所使用的調節變項（M）為研究地區，在資料上 0 為歐洲地區學生的研究，1 為亞洲地區學生的研究。另外，為了配合 AMOS 進行整合分析所必需，本研究根據上述的標準誤與效果值，增生其他欄位如圖 10 所示。其中，欄位上的 W 為各研究的權重（ $1/SE^2$ ），V 為各研究的變異數（ SE^2 ），sqrtW 為權重的平方根，Y 為各研究的效果值乘上權重的平方根。

本研究主要目的之一是探討 AMOS 在執行整合分析（包括檢測和修正出版偏差）方面的效能，為達此目的，需選擇一個具有顯著出版偏差和至少 .20 整合效果值的樣本，Infantes-Paniagua 等人（2022）的 16 篇關於資優生數學自我概念之研究，即滿足上述條件。本研究以 R 語言將圖 10 中的 16 篇研

測驗學刊，第 71 輯第 4 期

圖 10 Infantes-Paniagua 等人（2022）中 16 個研究與相關新增之欄位



Study	EF	SE	M	W	V	sqrtW	Y
Benolken (2015)	1.42	0.21	0	22.6757	0.0441	4.7619	6.7619
Benolken (2015)	0.48	0.18	0	30.8642	0.0324	5.55556	2.66667
Bergold et al. (2020)	1.11	0.36	0	7.71605	0.1296	2.77778	3.08333
Bergold et al. (2020)	0.76	0.25	0	16	0.0625	4	3.04
Golle et al. (2018)	0.33	0.05	0	400	0.0025	20	6.6
Herrmann et al. (2015)	0.47	0.07	0	204.082	0.0049	14.2857	6.71429
Liem et al. (2015)	0.17	0.07	1	204.082	0.0049	14.2857	2.42857
Preckel & Brull (2008)	0.63	0.29	0	11.8906	0.0841	3.44828	2.17241
Preckel & Brull (2008)	0.38	0.21	0	22.6757	0.0441	4.7619	1.80952
Preckel et al. (2008)	0.23	0.16	0	39.0625	0.0256	6.25	1.4375
Preckel et al. (2008)	0.71	0.16	0	39.0625	0.0256	6.25	4.4375
Preckel et al. (2010)	0.98	0.15	0	44.4444	0.0225	6.66667	6.53333
Preckel et al. (2017; 2019)	0.39	0.07	0	204.082	0.0049	14.2857	5.57143
Wirthwein et al. (2019)	1	0.15	0	44.4444	0.0225	6.66667	6.66667
Yan & Haihui (2005) - 1993	0.27	0.21	1	22.6757	0.0441	4.7619	1.28571
Yan & Haihui (2005) - 2003	0.22	0.22	1	20.6612	0.0484	4.54545	1

究之數據呈現在一個漏斗圖（如圖 11 所示）中。漏斗圖的非對稱性顯示出版偏差存在，此證明小樣本研究的效果值往往較大。漏斗圖也顯示整合效果值為 .42，超過 .20 的門檻。

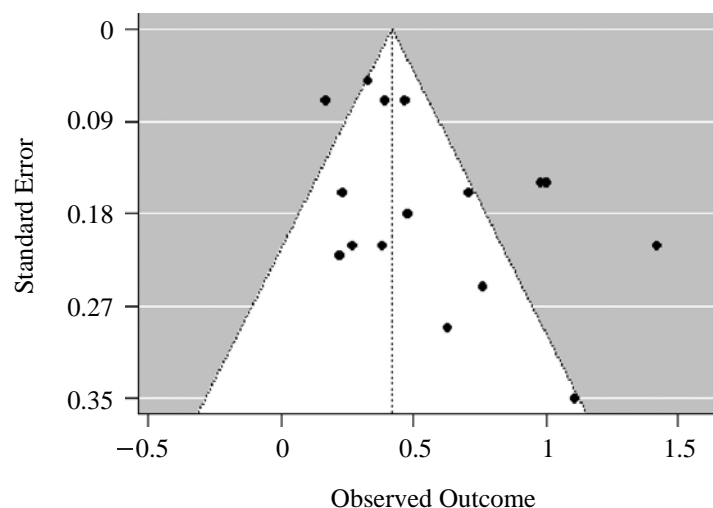
三、研究工具

本研究選擇使用 AMOS 20 軟體進行 SEM 分析，其主要目的是評估利用 AMOS 進行整合分析的可行性和準確性。在本研究中，研究者使用 AMOS 20 進行以下操作，分別為：(1)效果值的整合；(2)異質性分析；(3)調節變項分析；(4)出版偏差的偵測與修正。同時，本研究也使用 R 語言的 metafor 套件進行整合分析，將其結果作為效標來檢驗 AMOS 的分析結果。

肆、研究結果

本研究於下以 AMOS 為工具利用 SEM 進行整合分析，根據研究問題將

圖 11 以漏斗圖呈現出版偏差



逐一展開討論與呈現結果。

一、以 AMOS 進行傳統整合分析

就整合平均效果值的計算，利用 AMOS 之結果如圖 12 所示，其中 \sqrt{W} 為 $\sqrt{\text{加權}}$ ， Y 為效果值乘 $\sqrt{\text{加權}}$ ，固定效果模式下的整合效果值為 .42。本研究將 AMOS 所得的結果與 R 語言執行傳統整合分析的結果進行比較（如圖 13 所示），其結果是一致的。

就異質性而言，根據公式(8)， Q 值為開放誤差所得值乘上整合研究篇數，由於 AMOS 不是專門設計進行整合分析，因此在 Q 值的計算必須利用手動計算得出，其值為 $5.03 \times 16 = 80.48$ ，這個結果與 Infantes-Paniagua 等人（2022）所報告的 Q 值 79.81 相差不大。

本研究進一步以 R 的結果（如圖 13 所示）來驗證本研究的結果，發現 R 所報告的固定效果值與 Q 值均與本研究利用 AMOS 得出的結果一致。就效果值的標準誤而言，在圖 12 的固定效果模式下必須要將誤差的變異數設定為 1，根據這個模式可以得出整合效果值的標準誤為 .028，與 R 所得出的結果幾乎一致。就信賴區間估計而言，AMOS 無法直接提供，必須利用整合效果值 ± 1.96 乘上標準誤得出上下區間，分別為 .365 與 .476，與 R 所報告的幾乎一致。就固定效果模式的 I^2 而言，和 Q 值一樣無法直接算出，必須利用公式

測驗學刊，第 71 輯第 4 期

圖 12 利用 AMOS 的固定效果模式

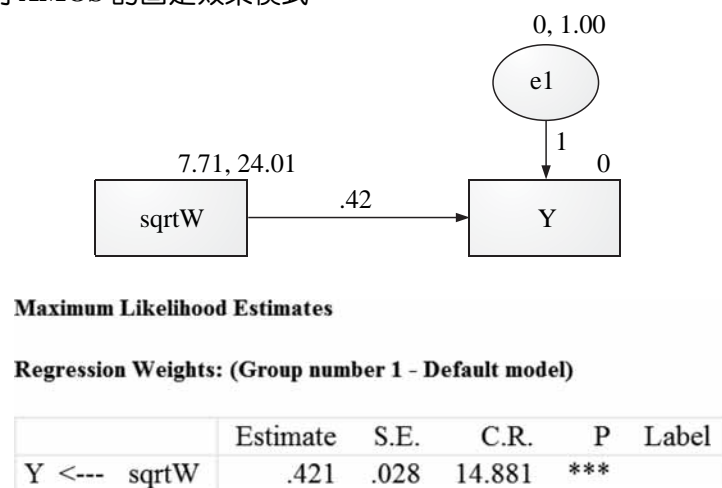


圖 13 利用 R 的固定效果模式

```
I^2 (total heterogeneity / total variability): 81.36%
H^2 (total variability / sampling variability): 5.37
```

```
Test for Heterogeneity:
Q(df = 15) = 80.4798, p-val < .0001
```

Model Results:

```
estimate      se      zval      pval      ci.lb      ci.ub
0.4207    0.0274   15.3694   <.0001   0.3671   0.4744   ***
```

(4)手動計算得知，其 df 為 15， $(Q - df) / Q = 81.36\%$ ，該值與 R 的結果一致。

就固定效果模式中加入調節變項而言，在 AMOS 中進行整合分析，其模式如圖 14 所示，其中調節變項為 M ， sqrtWM 為 sqrtW 乘上 M 。本研究以 Infantes-Paniagua 等人（2022）文中的地區，作為數學自我概念的調節變項，而調節變項中的地區，歐洲代碼為 0，亞洲為 1。本研究進一步利用 AMOS 在此固定模式下分析，可發現歐洲資優學生與一般生在數學自我概念上的差距為 .47（ $SE = .03$ ），而亞洲為 .18（ $SE = .069$ ），上述的結果與 R（如圖 15 所示）所得的結果一致。

圖 14 AMOS 之調節變項分析（固定效果模式）

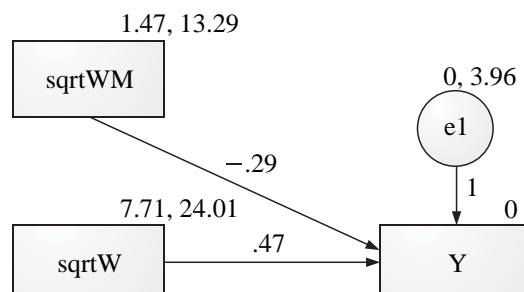


圖 15 R 之調節變項分析（固定效果模式）

I² (residual heterogeneity / unaccounted variability): 77.90%
H² (unaccounted variability / sampling variability): 4.53
R² (amount of heterogeneity accounted for): 15.65%

Test for Residual Heterogeneity:
QE(df = 14) = 63.3623, p-val < .0001

Test of Moderators (coefficient 2):
QM(df = 1) = 17.1175, p-val < .0001

Model Results:

	estimate	se	zval	pval	ci.lb	ci.ub	
intrcpt	0.4748	0.0303	15.6531	<.0001	0.4153	0.5342	***
M	-0.2914	0.0704	-4.1373	<.0001	-0.4295	-0.1534	***

就調節變項對於Q值的解釋程度而言，本研究根據圖12的模式並開放該圖誤差變異數的估計（5.03），減掉圖14中的誤差中之變異數（3.96）之間的差距，乘上研究篇數（16），可得出調節變項所解釋的Q值為17.12，該值在整合分析中均以 Q_{between} 或是 Q_B 呈現，且未能解釋的Q值則為63.36，而是以 Q_{within} 或是 Q_E 呈現，上述不論是 Q_E 或是 Q_B ，均與R的結果一致。就上述無調節變項在固定模式Q值一般以 Q_{total} 呈現。

就隨機效果模式下進行整合分析而言，在AMOS中並沒有提供計算隨機效果模式 τ^2 值的程式（不論是DL或是REML法），因此無法直接得出。間接地可以藉由圖10中的Excel表格進行 ΣW 與 ΣW^2 加總，再利用圖12中的 Q_{total} 值，套入公式(2)來得出該值。經初步計算隨機效果模式的 τ^2 值為.0588，求出來的值與R一致。在AMOS下進行隨機效果模式之分析，其模式

測驗學刊，第 71 輯第 4 期

如圖 16 所示。其中， sqrtW_r 是利用各研究之變異數加上 τ^2 ，將資料欄位上進行倒數處理為權重，並再開根號所求得。而 Y_r 則是藉由效果值乘上 sqrtW_r 。AMOS 執行後隨機效果模式之效果值為 .56，與 Infantes-Paniagua 等人（2022）中的報告一致。同時，本研究以 R 進行隨機效果模式的效果值估計，可得到相同的對照結果（如圖 17 所示）。

圖 16 利用 AMOS 進行隨機效果模式分析

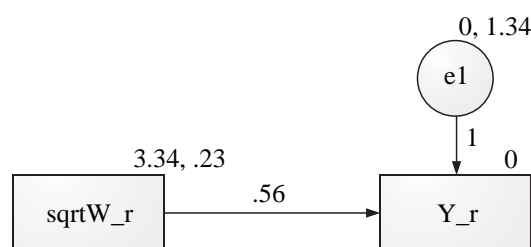


圖 17 利用 R 進行隨機效果模式分析

Random-Effects Model (k = 16; tau^2 estimator: DL)

logLik	deviance	AIC	BIC	AICc
-6.2944	43.1621	16.5889	18.1341	17.5120

tau^2 (estimated amount of total heterogeneity): 0.0588 (SE = 0.0375)

tau (square root of estimated tau^2 value): 0.2425

I^2 (total heterogeneity / total variability): 81.36%

H^2 (total variability / sampling variability): 5.37

Test for Heterogeneity:

Q(df = 15) = 80.4798, p-val < .0001

Model Results:

estimate	se	zval	pval	ci.lb	ci.ub	
0.5561	0.0742	7.4948	<.0001	0.4107	0.7015	***

就混合效果模式而言，即利用固定效果模式加上調節變項的 Q （63.36）來求得合併的 τ^2 ，由於涉及調節變項，必須利用公式（13）的 DL 法。經 Excel 計算後，可以求得混合模式下的 τ^2 為 .0531。根據混合模式下的 τ^2 值，本研究將權重的均方根欄位標註為 sqrtW_r1 ，且效果值乘上權重均方根值欄位為 Y_r1 ，調節變項乘上權重均方根值標註為 MsqrtW_r1 。在 AMOS

下混合模式的結構，如圖 18 所示。在混合模式下的結果顯示，歐洲資賦優異學童與一般生在數學自我概念上的差距為 .635 (SE = .082)，而亞洲地區學生的差距為 .211 (.635 - .424, SE = .187)。本研究的結果與 Infantes-Panigua 等人 (2022) 接近 (歐洲為 .64，亞洲為 .18)，但是與 R 分析所得的結果是一致的 (如圖 19 所示)。

圖 18 AMOS 下進行整合分析之調節變項分析 (混合模式)

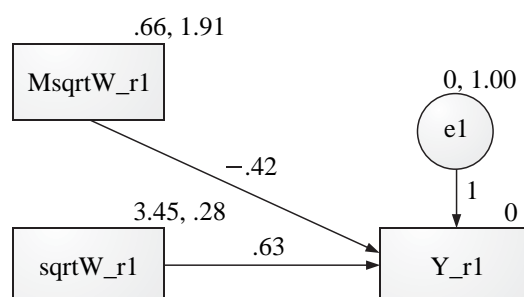


圖 19 R 下進行整合分析之調節變項分析之結果 (混合模式)

Mixed-Effects Model (k = 16; tau^2 estimator: DL)

logLik	deviance	AIC	BIC	AICc
-3.7687	38.1106	13.5374	15.8552	15.5374

tau^2 (estimated amount of residual heterogeneity): 0.0531 (SE = 0.0362)
tau (square root of estimated tau^2 value): 0.2304
I^2 (residual heterogeneity / unaccounted variability): 77.90%
H^2 (unaccounted variability / sampling variability): 4.53
R^2 (amount of heterogeneity accounted for): 9.75%

Test for Residual Heterogeneity:
QE(df = 14) = 63.3623, p-val < .0001

Test of Moderators (coefficient 2):
QM(df = 1) = 5.4313, p-val = 0.0198

Model Results:

	estimate	se	zval	pval	ci.lb	ci.ub	
intrcpt	0.6345	0.0795	7.9771	<.0001	0.4786	0.7904	***
M	-0.4238	0.1819	-2.3305	0.0198	-0.7803	-0.0674	*

二、以 AMOS 進行出版偏差的偵測與修正

(一) 出版偏差的偵測

就傳統整合分析在出版偏差的偵測上，只用 Egger 迴歸對稱性考驗才能在 SEM 下進行。在 AMOS 中以 SEM 進行 Egger 迴歸對稱性考驗，其模式如圖 5 所示。就出版偏差的偵測而言，如果是利用 SEM 進行 Egger 迴歸對稱性考驗，其主要觀察的項目在於依變項（Y）上的截距是否為 0。出版偏差的偵測都是在固定效果模式下進行，以 Infantes-Paniagua 等人（2022）的研究為例，其出版偏差偵測的結果如圖 20 所示，就 FAT 部分的截距顯著不為 0（截距 = 2.24， $p = .014$ ），表示該整合研究存在出版偏差。且 PET 的結果為 .21， $p = .032$ ，顯著不為 0，該訊息透露出效果值需要被修正。本研究進一步以 R 進行出版偏差的偵測，發現不論是在 FAT 或是 PET 上，兩者的結果一致，如圖 21 所示。

圖 20 以 AMOS 進行出版偏差的偵測

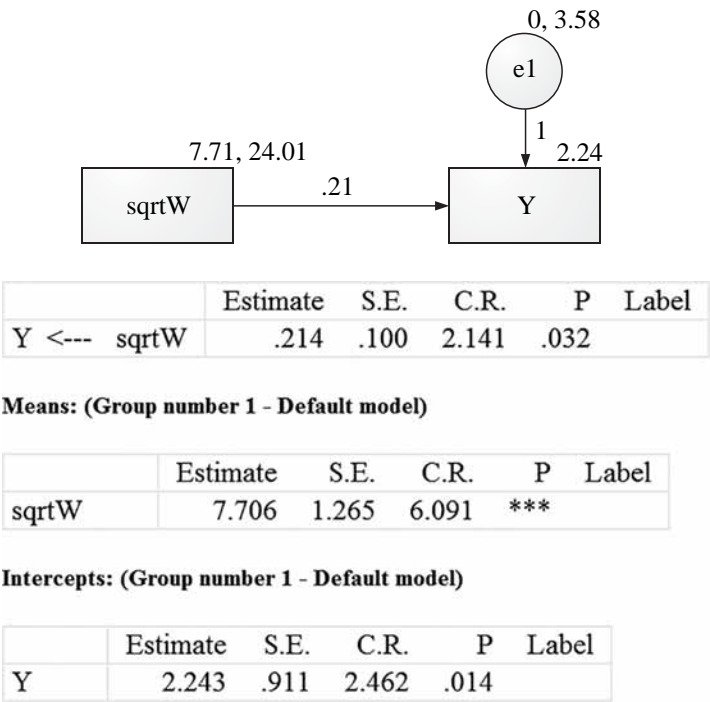


圖 21 以 R 進行出版偏差的偵測

```

> egger_test <- metabias(res)
> print(egger_test)
Linear regression test of funnel plot asymmetry

Test result: t = 2.38, df = 14, p-value = 0.0321

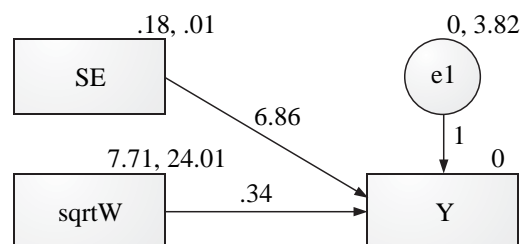
Sample estimates:
      bias se.bias intercept se.intercept
2.2426  0.9427    0.2135      0.1032

```

(二) 出版偏差的修正

就出版偏差的修正而言，本研究以 AMOS 進行 PEESE 模式的效果值修正，在固定效果模式下的結構如圖 22 所示。以 Infantes-Paniagua 等人（2022）的研究進行出版偏差的修正，其結果如圖 23 所示，在固定效果模式下，經修正出版偏差後的整合效果值為 .34。這個結果與 R 中使用 metafor package 所得的結果一致。本研究也利用 R 進行填補法來修正出版偏差，藉由在右下角填補五個研究後，在固定效果模式下所修正後的結果為 .35，如圖 24 所示。

圖 22 以 AMOS 利用 PEESE 進行出版偏差的修正（固定效果模式）



就修正出版偏差而言，若模式如圖 22，PEESE 在固定效果模式下的結果存在高異質性，可以考慮加入整合研究間的變異（ τ^2 ），例如：圖 8 以 PEESE 混合模式下進行出版偏差的修正。該混合模式利用上述在固定效果模式修正出版偏差的 Q 值（ $3.82 \times 16 = 61.12$ ），然後再利用公式(13)的 DL 法來計算混合效果模式下的 τ^2 值。以 Infantes-Paniagua 等人（2022）的研究為例，

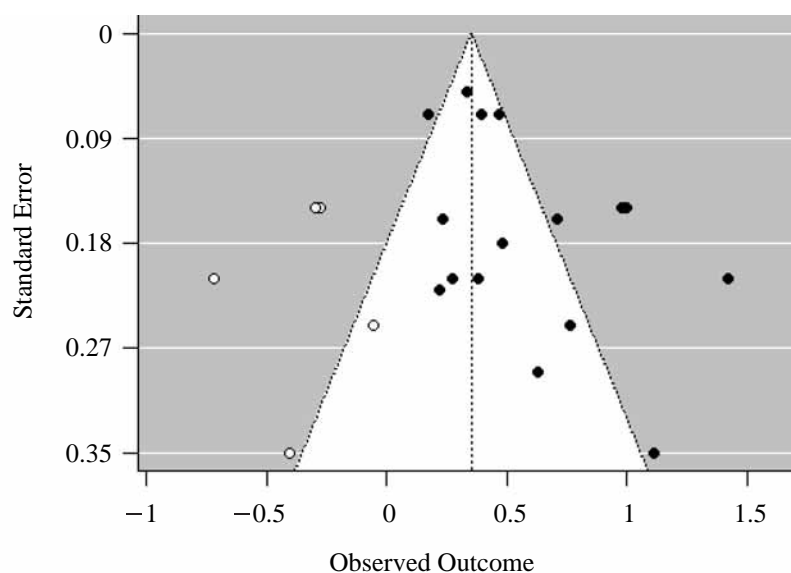
測驗學刊，第 71 輯第 4 期

圖 23 以 R 利用 REESE 進行出版偏差的修正（固定效果模式）

Model Results:

	estimate	se	zval	pval	ci.lb	ci.ub	
intrcpt	0.3385	0.0332	10.2091	<.0001	0.2735	0.4035	***
I(V)	6.8594	1.5601	4.3967	<.0001	3.8016	9.9172	***

圖 24 以 R 利用填補法進行出版偏差的修正



Model Results:

estimate	se	zval	pval	ci.lb	ci.ub	
0.3507	0.0261	13.4439	<.0001	0.2996	0.4018	***

利用 Excel 計算出混合效果模式下的 τ^2 值為 .045。將此 τ^2 值加入每個研究的變異數，求得新的加權平方根，並利用該項乘上 $SE^2(V)$ 與效果值兩欄位，以 AMOS 進行混合效果下的出版偏差修正，其結果如圖 25 所示，修正後的效果值為 .41。同時本研究也利用 R 進行 PEESE 在混合模式下的出版偏差修正，如圖 26 所示，兩者所得出的結果均一致。

在修正出版偏差上考慮到調節變項，若在固定效果模式上以 AMOS 進行

圖 25 利用 AMOS 以 PEESE 法修正出版偏差（混合模式）

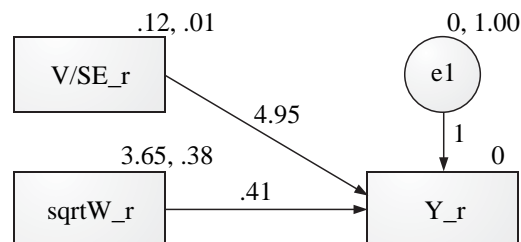


圖 26 利用 R 以 PEESE 法修正出版偏差（混合模式）

```

tau^2 (estimated amount of residual heterogeneity):    0.0450 (SE = 0.0303)
tau (square root of estimated tau^2 value):           0.2120
I^2 (residual heterogeneity / unaccounted variability): 77.10%
H^2 (unaccounted variability / sampling variability):  4.37
R^2 (amount of heterogeneity accounted for):           23.57%
  
```

```

Test for Residual Heterogeneity:
QE(df = 14) = 61.1485, p-val < .0001
  
```

```

Test of Moderators (coefficient 2):
QM(df = 1) = 3.7990, p-val = 0.0513
  
```

Model Results:

	estimate	se	zval	pval	ci.lb	ci.ub	
intrcpt	0.4103	0.0981	4.1845	<.0001	0.2181	0.6025	***
I(V)	4.9522	2.5408	1.9491	0.0513	-0.0276	9.9320	.

出版偏差的修正，必須利用模式 FAT-PET-PEESE-MRA（如圖 9 所示）。本研究以 Infantes-Paniagua 等人（2022）的研究為例，利用效果值、加權平方根、標準誤與調節變項乘上加權平方根等變項，以 AMOS 進行 FAT-PET-PEESE-MRA 模式的出版偏差修正，其結果如圖 27 所示。可以發現在效果值的部分上為 .39，也就是在歐洲地區資優生與一般生在數學自我概念上的差距。然而，調節變項為 $-.29$ ，也就是亞洲地區資優生與一般生在自我概念的差距大約為 .10。標準誤（SE）項的斜率 6.88（SE = 2.09）為顯著不為 0，表示 Infantes-Paniagua 等人（2022）在數學自我概念上的整合研究確實存在出版偏差，而該斜率的顯著性表示本研究由該模式下所得到的效果值是經由出版偏差修正後的結果。本研究進一步利用 R 進行 FAT-PET-PEESE-MRA 模式在固定效果模式下的分析（如圖 28 所示），發現 AMOS 與 R 兩者的結果一致。

測驗學刊，第 71 輯第 4 期

圖 27 AMOS 下的 FAT-PET-PEESE-MRA 模式（固定效果模式）

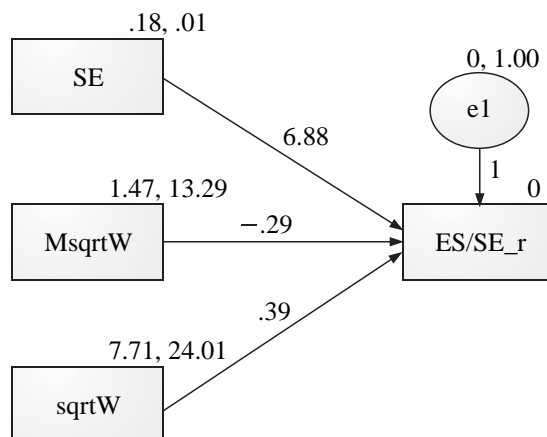


圖 28 以 R 進行 FAT-PET-PEESE-MRA 模式分析（固定效果模式）

Model Results:

	estimate	se	zval	pval	ci.lb	ci.ub	
intrcpt	0.3925	0.0356	11.0207	<.0001	0.3227	0.4622	***
I(V)	6.8832	1.5601	4.4119	<.0001	3.8254	9.9409	***
M	-0.2926	0.0704	-4.1535	<.0001	-0.4306	-0.1545	***

若在 FAT-PET-PEESE_MRA 模式中考慮到研究間的變異，在 SEM 上則需要利用圖 9，也就是 FAT-PET-PEESE-MRA 之混合效果模式下（如圖 9 所示）進行出版偏差的校正。本研究利用圖 27 固定效果模式開放誤差項效果值的估計來求出 Q 值，再以公式(13)的 DL 法來計算 τ^2 值，得出該值為 .0357。經由 AMOS 下進行分析，其效果值為 .49，也就是說若以歐洲地區為對照組，該地區資優生在數學自我概念上與一般生的差距為 .49。調節變項為 -.41，表示在亞洲地區這個差距就降下來了，該地區資優生在數學自我概念上與一般生的差距為 .08。在出版偏差上以 AMOS 進行 FAT-PET-PEESE-MRA 模式分析，發現出版偏差值為 4.9，但 p 值顯示為 .105。

就出版偏差的顯著性而言，Van Aert 等人（2019）認為在出版偏差的檢驗上，由於篇數少的整合分析研究容易導致統計考驗力不足，所以應該要將 p 值向上調整。本研究以 R 進行 FAT-PET-PEESE-MRA 在隨機效果模式下的分析，研究發現兩者在效果值、調節偏差與出版偏差上有一致的結果，但是在

標準誤的估計上兩者有不一致之現象。儘管差異不是很大，但是這個結果造成兩者在出版偏差顯著性上的考驗有不一致之結果。研究者使用 ESS 增益集（李茂能，2015），分析該筆資料，ESS 與 R 一致，但 AMOS 與 metaSEM（Cheung, 2015）的結果比較接近。儘管在 AMOS 以 FAT-PET-PEESE-MRA 進行出版偏差的校正時，出版偏差有出現不顯著的現象，但在這個模式校正之下，研究者發現效果值與圖 18 在隨機模式估計調節變項的模式相比，效果值經出版偏差的校正由 .63 降至 .49。

圖 29 AMOS 下的 FAT-PET-PEESE-MRA 模式（隨機效果模式）

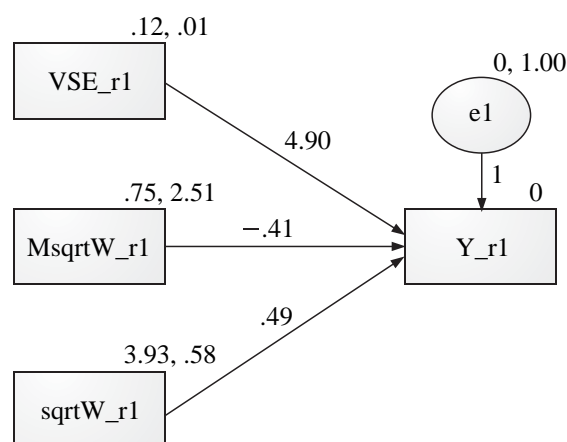


圖 30 以 R 進行 FAT-PET-PEESE-MRA 模式分析（隨機效果模式）

```

tau^2 (estimated amount of residual heterogeneity):    0.0357 (SE = 0.0265)
tau (square root of estimated tau^2 value):          0.1889
I^2 (residual heterogeneity / unaccounted variability): 70.39%
H^2 (unaccounted variability / sampling variability):  3.38
R^2 (amount of heterogeneity accounted for):          39.31%
  
```

```

Test for Residual Heterogeneity:
QE(df = 13) = 43.8970, p-val < .0001
  
```

```

Test of Moderators (coefficients 2:3):
QM(df = 2) = 10.9432, p-val = 0.0042
  
```

Model Results:

	estimate	se	zval	pval	ci.lb	ci.ub	
intrcpt	0.4883	0.0954	5.1177	<.0001	0.3013	0.6753	***
I(V)	4.8997	2.4040	2.0382	0.0415	0.1881	9.6114	*
M	-0.4054	0.1589	-2.5517	0.0107	-0.7168	-0.0940	*

測驗學刊，第 71 輯第 4 期

本研究基於 Infantes-Paniagua 等人（2022）對資優學生數學自我概念的 16 項研究，以 AMOS 進行了整合分析與出版偏差分析並以 R 進行驗證，將結果整理如表 1 所示。綜合分析顯示，未經修正的整合結果存在被高估的情況，不論採用固定效果模式或隨機效果模式。同時，表 1 也顯示，該研究在校正出版偏差後，異質性不論在一般分析模式或是調節變項模式上均有下降的趨勢。但是，該研究上仍存有顯著的異質性，若其他研究者要以該結果進行推論時，建議採取隨機效果模式的結果。

表 1 進行整合分析與出版偏差修正的結果：以 Infantes-Paniagua 等人（2022）之資優學生數學自我概念研究為例

整合分析之 SEM 模式	未校正出版偏差		校正出版偏差後	
	固定效果	隨機效果	固定效果	隨機效果
效果值／異質性：一般模式	.42 Q = 80.4798 ($p < .001$)	.56	.34 Q = 61.1485 ($p < .001$)	.41
效果值（調節變項效果）／異質性：調節變項分析	.47 (−.29) Q = 63.3623 ($p < .001$)	.63 (−.42)	.39 (−.29) Q = 43.8970 ($p < .001$)	.49 (−.41)

註：括弧內為調節變項之效果估計。

伍、討論與結論

本研究旨在以 AMOS 進行傳統整合分析以及出版偏差的偵測與修正，並以 Infantes-Paniagua 等人（2022）整合分析研究成果作為對照以展示成果。本研究發現，利用 AMOS 的確可以進行效果值的整合，異質性與調節變項等分析，而 AMOS 所呈現的結果，也與利用 R 進行整合分析驗算的結果一致。由於 AMOS 並非專門設計用來進行整合分析的工具，因此在某些項目上的結果並不是直接產出。在整合效果值上，AMOS 無法直接提供整合平均效果值的信賴區間，需要藉由計算器或是另用 EXCEL，以 AMOS 估計出的標準誤求得。

其次，在隨機效果值的計算上，由於 AMOS 無法提供 REML 法計算 τ^2 值，所以必須利用權重的欄位，以 Excel 求得進行 ΣW 與 ΣW^2 ，並利用 Q 值與自由度，以 DL 法來計算 τ^2 值，再利用改變的權重來求出在隨機效果模式

下的整合平均效果值。就異質性而言，AMOS 並無法直接產出異質性中的 Q 值， I^2 值與 τ^2 值，但是由結果顯示，可以利用計算器或是 Excel 經簡單的計算得出。以 Q 值為例，利用模式中依變項誤差的變異數開放估計，得出該值後，再乘上研究篇數即可求得。

就出版偏差的偵測而言，由於沒有牽扯到隨機效果，因此可以藉由 AMOS 直接產出結果。在校正出版偏差上，若是在固定效果模式上進行出版偏差的修正，不論是 PEESE 或是 FAT-PET-PEESE-MRA 兩類上，可以直接求得修正值。但此兩類在混合模式上，則無法直接求得，必須要先求得 Q 值，再利用 Q 值與公式(13)的 DL 法，求得混合模式的 τ^2 值進行效果值的修正。

本研究所用的樣本來源為 Infantes-Paniagua 等人（2022）於數學自我概念的整合分析研究結果，該研究共有 16 篇關於資優生與一般生於數學自我概念的比較。就樣本數的限制性，儘管本研究僅有 16 篇研究，但是根據李茂能（2016）ESS 計算器，以效果值 .40，Cronbach Alpha 設定為 .05，異質性設定為高異質性的條件下，若要達到統計考驗力 .80 以上，該整合分析必須要達到 9 篇以上，才有辦法偵測出出版偏差與其進行出版偏差的修正。此外，根據 Egger 等人（1997）的研究，進行出版偏差之整合分析之蒐集篇數不能少於 6 篇。根據上述兩個條件，本研究之研究樣本在進行出版偏差的偵測與修正時，不會因為統計考驗力不足或是篇數過少，而導致產出的結果不具備效度。

另外，從表 1 可知，本研究發現 Infantes-Paniagua 等人（2022）未經修正的整合結果存在被高估的情況，不論是採用固定效果模式或隨機效果模式。此發現強調了在進行整合分析時，對出版偏差檢視的重要性。尤其在特殊教育領域，在研究樣本與篇數明顯偏低於其他研究領域的情況下，出版偏差對整合結果的準確性影響尤為顯著。因此，在整合分析的處理上更應該注重出版偏差的處理。本研究之範例在經過調節變項分析後，異質性仍然顯著。使用固定效果時容易產生偏差，因此在推論上建議使用隨機效果（李茂能，2016）。

不管是傳統或是 SEM，在進行整合分析的偵測與修正出版偏差時，均是利用漏斗圖的不對稱性。傳統整合分析大都是利用填補法（trim and fill）來進行，該種方法的缺點在於以直觀方式來判定出版偏差，同樣的也是以直觀方式填補到圖形達到對稱。在本研究中使用 SEM 時，則採用 Egger 迴歸對稱性考驗偵測出版偏差，以及利用 PEESE 來校正出版偏差。兩者均利用迴歸方

測驗學刊，第 71 輯第 4 期

式來判定漏斗圖的不對稱性以及校正其不對稱性。就精準度而言，後者採用統計的方式當然較為精準，但該方法的精準度並不適合於研究篇數較少的整合分析，尤其是小於 6 篇的研究。

此外，若整合研究具高異質性，不論是填補法或是 SEM 進行偵測與校正出版偏差，其精準度也會受到影響。因此，若能藉由調節變項分析，則可降低異質性來提高其精準度。就執行調節變項分析上，SEM 以 PEESE 為基礎，以 AMOS 為例，很容易可以在介面上進行 FAT-PET-PEESE-MRA 法。

綜合以上，本研究使用 AMOS 軟體以 SEM 進行整合分析，能夠達到預期的成效。由於特殊教育領域學門特性，蒐集到的整合分析研究篇數不多，且未達到顯著差異的研究常常難以出版，而易產生出版偏差。建議未來能在此類學門的整合分析研究上，能善用出版偏差的偵測與校正來完善整合結果的效度。此外，相較於傳統整合分析方法，以 SEM 進行出版偏差的偵測與修正不但有其便利性，更可以利用視覺化清楚地呈現結果，可以讓一般研究者降低解讀複雜統計結果報表的困境。

參考文獻

中文部分

- 李茂能（2015）。**傳統整合分析理論與實務：ESS & EXECL**。五南。
- [Li, M.-N. (2015). *Traditional integrated analysis theory and practice: ESS & EXECL*. Wu-Nan Book.]
- 李茂能（2016）。**當代整合分析理論與實務：ESS、Meta-SEM、Mvmeta & WinBUGS**。五南。
- [Li, M.-N. (2016). *Contemporary integrated analysis theory and practice: ESS, Meta-SEM, Mvmeta & WinBUGS*. Wu-Nan Book.]

英文部分

- Begg, C. B., & Mazumdar, M. (1994). *Operating characteristics of a rank correlation test for publication bias*. *Biometrics*, 50(4), 1088-1101.
- Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hodges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221-235). Russell Sage Foundation.
- Cheung, M. W.-L. (2015). *Meta-analysis: A structural equation modeling approach*. John Wiley & Sons.
- Coburn, K. M., & Vevea, J. L. (2015). *Publication bias as a function of study characteristics*. *Psychological Methods*, 20(3), 310-330. <https://doi.org/10.1037/met0000046>
- Cook, B. G. (2014). *A call for examining replication and bias in special education research*. *Remedial and Special Education*, 35, 233-246.
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455-463.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629-634. <https://doi.org/10.1136/bmj.316.7129.469>

- Forness, S. R. (2001). Special education and related services: What have we learned from meta-analysis? *Exceptionality*, 9(4), 185-197.
- Gage, N. A., Cook, B. G., & Reichow, B. (2017). *Publication bias in special education meta-analyses*. *Exceptional Children*, 83(4), 428-445.
- Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558.
- Infantes-Paniagua, Á., Fernández-Bustos, J. G., Ruiz, A. P., & Contreras-Jordán, O. R. (2022). Differences in self-concept between gifted and non-gifted students: A meta-analysis from 2005 to 2020. *Anales de Psicología*, 38(2), 278-294. <https://doi.org/10.6018/analesps.461971>
- Mostert, M. P. (2021). Characteristics of meta-analyses reported in mental retardation, learning disabilities, and emotional and behavioral disorders. In K. A. Kavale (Ed.), *Meta-analysis research in special education* (pp. 199-225). Routledge.
- Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, 8(2), 157-159.
- Parker, R. I., Vannest, K. J., & Davis, J. L. (2011). Effect size in single-case research: A review of nine nonoverlap techniques. *Behavior Modification*, 35(4), 303-322.
- Pigott, T. D. (2012). Missing data in meta-analysis: Strategies and approaches. In *Advances in meta-analysis* (pp. 79-107). Springer.
- Raudenbush, S. W., Becker, B. J., & Kalaian, H. (1988). Modeling multivariate effect sizes. *Psychological Bulletin*, 103(1), 111-120.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638-641.
- Scruggs, T. E., & Mastropieri, M. A. (2021). How to summarize single-participant research: Ideas and applications. In K. A. Kavale (Ed.), *Meta-analysis research in special education* (pp. 227-244). Routledge.
- Stanley, T. D. (2008). Meta-regression methods for publication selection bias: Simulations and Heckman regression. *Bulletin of Economic and Meta-Analysis*, 70, 103-127.
- Stanley, T. D., & Doucouliagos, H. (2012). *Meta-regression analysis in economics and business*. Routledge.
- Stanley, T. D., & Doucouliagos, H. (2013). *Better than random: Weighted least squares meta-regression analysis*. Economics Series.
- Sterne, J. A., Becker, B. J., & Egger, M. (2006). The funnel plot. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment and adjustments* (pp. 75-98). John Wiley & Sons.

曹傑如 張道宜 楊智為

整合分析的可靠性：結構方程模式處理檔案櫃效應

- Tanriver-Ayder, E., Faes, C., van de Castele, T., McCann, S. K., & Macleod, M. R. (2021). Comparison of commonly used methods in random effects meta-analysis: Application to preclinical data in drug discovery research. *BMJ Open Science*, 5(1), e100074. <https://doi.org/10.1136/bmj-2020-100074>
- Van Aert, R. C., Wicherts, J. M., & Van Assen, M. A. (2019). Publication bias examined in meta-analyses from psychology and medicine: A meta-meta-analysis. *PloS one*, 14(4), e0215052.

測驗學刊，第 71 輯第 4 期