

大型資料庫與後設分析結果之交互驗證性 研究-以新移民與本土子女學習成就差異為例

謝名娟

國家教育研究院測驗及評量研究中心 研究員

中文摘要

在本研究中，以新移民子女與本土子女的學習成就差異之議題為分析主軸，來相互驗證後設分析與大型資料庫研究方法所產生的結果。在後設分析策略中，使用 59 篇文獻進行訊息編碼及後設分析。而大型資料庫則採用 TASA 資料，資料來自 2007 年至 2013 年四年級、六年級、八年級與高中職二年級的實徵數據，研究結果顯示，兩種數據蒐集方法，經統計分析後所得的結果大致符合。本土子女的整體學習表現較新移民子女來得佳，其學習表現差異平均效果量是屬小至中等程度，此外，隨年齡增長，差異有隨之微幅下降的趨勢，兩種數據蒐集方法之優缺點也於文中進行探討。

關鍵字：大型資料庫、新移民子女、學習成就、後設分析、TASA

壹、緒論

實證研究中，蒐集數據為首要之務，尤其是大數據時代的來臨，研究者無不希望能在短時間內，蒐集到大量且隨機的樣本數據來了解研究問題。然而，受限於研究經費與時間，研究者往往只能方便取樣，尋找願意配合的學校來進行小樣本的資料蒐集，若是需要進行教學實驗，也只能挑選一、兩個班級進行，因此，這種研究的結果常會存在效度是否可概化或推論的質疑，換言之，就幾個小型樣本是否能夠推論到其他地區、年齡或形態的學生？即使研究結果顯示對某幾個地區的樣本數據而言，看起來統計結果是顯著的，或是對幾個班級而言，某種教學法是有效的，但對其他的各地的樣本而言，也能得到顯著的統計結果嗎？而對於其他類似的班級來說，同樣的教學法也是有效的嗎？研究結果會不會只是研究者所選取的樣本特殊性使然？諸如此類的議題，為研究者需要克服的議題。

系統性文獻回顧策略與後設分析技術的使用可以適度的回應此議題，而此研究方法主要起源於農業科學和醫學領域，由於此研究方法可針對重要的議題，提出綜合性的建議，因此在教育和心理學領域中，亦成為重要的方法，其譯名是十分多元，有整合分析、彙總研究、統合分析或運用於評量工具中，所使用的效度概化（Murphy, 2003; Schmidt & Hunter, 1977）等，但不論為何者，其概念是意指著為了綜合研究成果而蒐集個別主題中各式相關分析數據，以執行統計分析的過程（Glass, 1976），也就是系統性蒐集過去研究文獻及其相關數據，以進行二次統計分析。

透過整合過去學者的研究成果，以針對某主題提出綜整性研究數據，隨著以證據為導向來形塑教育決策理念近年來逐漸受到國內外學者重視（馬信行，2002；謝進昌，2010；Davies, 1999），教育人員對於證據的需求亦隨之提昇，但科技進步雖使得訊息快速流通與知識不斷累積，亦迫使證據呈現紛雜且品質不一的現象，而與後設分析策略的結合，其優點在於讓研究者能就豐富研究成果，針對不同脈絡或情況，提出概括性及嚴謹性的研究結論或證據，使得決策者能快速掌握決策的方向或是後續的研究者能一窺現況，有效掌握研究整合的相關證據來源。

後設分析雖然是一種有用的統計分析工具，也能針對研究問題進行解答，但是有潛在的限制。首先，由於後設分析將所有可能蒐集到的研究成果進行彙整，研究數據主要是仰賴別的研究者所得到的研究數據，資料的性質是屬於被動性、觀察性而非透過實驗性的歷程所得到的結果，研究者自身沒有去蒐集數據，只能透過文獻中所提供的描述進行編碼紀錄，若是文獻敘述不清，在編碼時則可能會造成錯誤的解釋。第二，在彙整不同的研究數據時，很難將各研究中所產生的各種誤差加以考量，例如同樣是進行學生的閱讀理解研究，A學者使用自編的閱讀測驗，B學者則可能使用標準化流程所編製閱讀測驗，或A學者使用立意抽樣、B學者使用隨機抽樣，這些不同的測驗題數、不同的抽樣歷程，都有可能會有大小不同的誤差，但在計算統合的效果量時，很難將這些不同的誤差來源考量進去。

然而，後設分析為一個較為省錢、省成本的研究方法，研究者只要透過網路與圖書館資源，幾位專家相互檢查編碼的正確性，即可解答研究問題，不需要透過問卷發放的

施測流程，聯絡施測學校，與單位、個人簽署的相關個資同意書，若樣本大還要進行讀卡、數據勘誤等複雜流程。

後設分析過去常用於驗證或是解釋大型測驗所得的結果。例如，Nieuwenstein、Wierenga、Morey、Wicherts、Blom、Wagenmakers與Rijn（2015）同時進行後設分析與大型試驗來檢視潛意識對於受試者做正確決策的影響，透過兩者研究結果的符合程度，來相互驗證後設分析與大型測驗結果的正確性。亦有學者使用後設分析的結果來解釋大型測驗中所得之結果，如Kieffer、Lesaux、Rivera與Francis（2009）從大型測驗中，發現第二外語學習者在數學與科學測驗中常需考試調整的模式來協助，而這些調整模式是否會對測驗效度造成影響，則透過後設分析的文獻歸納來進行解釋。

然而，研究者常有的疑問是後設分析的結果與大規模評量所得的結果是否符合？過去的研究大多是擇一而為之，例如單純以後設分析，或是以大規模評量來進行結果的推論，而較少比較兩者所得到的結果是否符合，然而，透過十篇文章，篇文章100個樣本所集結出來總共1000筆數據，和透過大型試驗，一次蒐集一筆1000個樣本的數據若相同，其實使用較為簡單、省錢的方法，站在前人的肩膀上來進行資料分析或許更符合經濟效益。根據謝進昌（2011）指出，過去的教育政策多以實徵研究為導向，透過證據來決定教育政策的走向，然而，實證數據的來源本質來自不同的過程與背景。例如大型資料庫所強調的是受試者的整體概況，著重可能影響表現之相關因素等，而後設分析著重的是過去經驗的累積，可以是大型資料庫分析或規畫的基礎。

為了驗證後設分析與大型施測兩種不同方法所得之結果符合程度，本研究目的為透過本土與新移民子女的學習成就表現，來檢視大型資料庫所蒐集之數據，所得出之結果與後設分析所得之結果是否相互符合。而在本研究中定義的學習成就表現，則為學生在標準測驗中所獲得之成績，或是學生在校學期成績。大型資料庫的資料來源來自國家教育研究院所建構之「臺灣學生學習成就資料庫」（Taiwan Assessment of Student Achievement, TASA），而後設分析資料則擷取自國內學者所建構之後社分析網站資料庫（網址：<http://meta.naer.edu.tw/>），其資料庫雖過去曾使用於謝名娟與謝進昌（2013）之相關研究，但本研究透過不同的分析步驟，並從中擷取適當的資料，以做為本研究之資料來源。在本文中主要著重探索不同的研究方法分析取徑的異同，探索的議題有二項：

- 一、本土與新移民子女差異效果量是否會隨著樣本的學習階段、科目有所不同？
- 二、本土與新移民子女差異效果量是否會隨著母親的國籍差異有所不同？

貳、文獻探討

以下茲就兩大部分來進行文獻探討，第一部分為探討後設分析、大型資料庫、後設分析與大型資料庫之相關性研究，第二部分則著重在新移民子女與本土子女之學習成效差異性之研究，其文獻分述如下：

一、後設分析與大型資料庫之相關研究

（一）後設分析

後設分析（meta analysis）為研究者為整合研究發現而蒐集個別研究主題中各式相

關分析結果，以進行統計分析的過程，自此，此概念就引起許多國內外學者的注意，紛紛將其此方法應用在各種議題之量化研究整合，以企圖尋求概括性的結論或結果。但同時期，系統性回顧（systematic review）策略也受到許多學者的關注，它強調的是一種嚴謹文獻蒐集、評閱過程（Feldman, 1971），直至1980年代初期，才有學者將兩種策略進行融合，其中，系統性回顧策略的過程可以是包含一連串的文獻蒐集、評估的過程，目的就是在降低研究資訊取得、編碼與分析等流程中可能造成的誤差，一般而言，大致包含有主題形成、研究選擇與定位、研究品質的評估、資料收集與編碼、後設分析、結果呈現與詮釋、改善與資料更新等步驟（Cooper & Hedge, 1994; Higgins & Greens, 2008）。主題的形成在系統性文獻回顧流程中扮演的角色，導引著後續的文獻搜集、定位及效果量計算等，透過蒐羅資料庫，可針對相關的關鍵字，在各大資料庫搜尋，其方式可以是利用電子資料庫的檢索、特定期刊的人工搜索或網路（例如：Google）搜尋，以期完整檢閱各式出版或未出版文獻，此外，研究者需透過文獻納入、排除準則，以選擇適合的文獻，作為後續分析標的。

經文獻搜尋、篩選後，研究者對於符合準則的文獻，還必須進行品質的評估，以避免研究設計粗劣的研究，干擾後續的整合結果；研究者再依據各研究特徵、目的、計算效果量的訊息及可能的調節變項，進行編碼的動作，以為執行後設分析與結果呈現。

大致而言，後設分析流程包含：效果量的計算、異質性分析、估計模式選擇與調節變項分析、改善與資料更新步驟，通常是研究者認為後續的整合結果有待更新或擴大研究推論範圍（例如：納入其它國家的研究文獻）所進行的動作。過去後設分析在教育領域中應用廣泛，如 Uysal 與 Sarier（2018）使用後設分析模式來進行學校領導風格對於學生學習成效的影響，Al-Hoorie（2018）則探討動機對於學習成就的影響。

（二）大型資料庫

醫學界定義大型試驗時，有兩種方式。第一種為樣本數大於1000的即可稱為大樣本試驗（villar, Carroli, Belizan, 1995; Cappelleri, et al., 1996; LeLorier, Gregoire, Benhaddad, Lapierre, & Derderian, 1997），第二種為Cappelleri等人（1996）所提出，只要在大型試驗中，其統計檢測力（power）達到80%以上，也可成為大型試驗。然而，使用樣本數來定義是否為大型試驗為較為方便且簡單的方法，但是這樣的方法卻沒考量到實驗設計與執行的嚴謹性。例如大型試驗中可能遇到樣本流失、系統系誤差等，都有可能造成研究結果的偏誤。

而教育的大型資料庫是教育資料的存檔，教育資料內容包含了學生、學校及家長的調查，觀察結果，評量內容及背景狀況。這些資料可用來描述學生或是學校，並說明各種因素之間的關係或結果。透過大型資料庫的建置，相關學者能得到具有全國代表性，即時且完整的資訊，也能透過對資料的分析，提供教育決策者改革或施政的參考。國內外有關學習成就建置之相關資料庫如臺灣教育長期追蹤資料庫（Taiwan Education Panel Survey，簡稱TEPS）、國際數理趨勢研究（Trends in International Mathematics and Science Study，簡稱TIMSS）、國家教育進展評量（National Assessment of Educational Progress，簡稱NAEP）、國際學生評量（the Programme for International Student Assessment，簡稱

PISA)、國際閱讀發展研究(Progress in Reading Literacy Study, 簡稱PIRLS)等,其施測時間間距皆為兩至三年施行一次大規模之學習成就評量,用以檢視教育趨勢與學生成就表現,大型測驗具有嚴謹的標準化施測流程、抽樣、組卷、統計分析等,故可在維持研究水準的前提之下,維持一定的研究水平與數據正確性,為國內外教育相關資料庫之共同特性。

(三) 後設分析與大型資料庫之相關研究

過去大型測驗與後設分析的比較研究可以看出幾種模式,其一為根據大規模實驗的結果,觀察出某些現象,若要找其可能影響之因素,則透過後設分析的文獻來找到可能影響結果之調節變項。其二則是比較兩種方式的所產生的結果,當作交互驗證的工具。大型資料庫的施測流程下,不易找到適合的效標來做關聯效度,但後設分析結果應證的結果,則可當作對於大型施測結果效度的證據之一。

第一種模式如Kieffer、Lesaux、Rivera與Francis(2009)的研究。其研究先從大型資料庫的結果中,發現第二外語學習者由於英語能力較差,因此在參與測驗時,考試單位在考試時會提供輔助的方式,例如延長考試時間,提供英語辭典等模式來幫助第二外語學習者來進行測試,然而,透過考試輔助的模式,這批學生的成績是否能提升,大型測驗的結果並無法得知。研究者則想探究這個議題,因此蒐集了11篇文獻,共23,999位受試者來進行後設分析。其結果發現大多數的考試調整對於測驗成績是沒有影響的,只有給考生使用英語辭典,才能稍微拉近第二外語學習者和本國人之間的學習差異。Hopfenbeck、Lenkeit、Masri、Cantrell、Ryan和Baird(2017)與Kieffer等人(2009)的研究透過後設分析來探究大型測驗中所發現的問題。如Hopfenbeck則探討PISA的大型測驗中學者利用其資料所做出的研究結果,並所探討從數據中看到的各種教育政策問題。

而在第二種模式中,則比較兩種方法所得之結果。如Nieuwenstein、Wierenga、Morey、Wicherts、Blom、Wagenmakers與van Rijn(2015)使用大型試驗與後設分析對應的結果來交互驗證潛意識對於人類做決策的影響,在大型的實驗中有399位受試者,其結果顯示潛意識對於人類的決策並沒有顯著的影響,但是對於不同的性別產生的效果不同。在後設分析部分,則採計了61篇文獻,後設的結果顯示潛意識對於人類決定有顯著性的影響,但透過funnel plot卻顯示所蒐集的文獻中出現了出版偏誤的現象。其中在後設分析的文獻中,作者發現研究者偏好發表有顯著效果的文章,然而,有些文章的樣本數很小,另外,判斷潛意識是否有影響的指標有許多種,但作者卻大多選擇具有顯著的指標來發表,或是重複進行多次的實驗,直到有顯著性的結果才發表出來。因此作者對於後設的結果抱持著存疑的態度,並建議未來的研究者應謹慎檢視其研究方法與樣本大小。

這種比較性研究在醫學領域亦受到重視。可能是由於醫學領域實驗成本較高,因此若採後設分析,而不用實際去花經費、花人力來做大型實驗,對於研究者而言是較方便的做法。然而,彙整一堆文獻的小樣本的結果和做一次大規模樣本的實驗結果是否相同,亦為醫學界重視的要點之一。就過去研究指出,後設分析和大規模實驗的結果是否一致要視議題來看,例如在檢視哪種藥物對於治療心肌梗塞有幫助,後設分析的結果對

於使用早期 β 受體阻滯劑 (β blockers) 和血栓溶解劑 (thrombolytic therapy) 具有療效的結論與大規模的隨機試驗的結果符合 (Antman, Lau, Kupelnick, Mosteller, & Chalmers, 1992; Yusuf, Collins, Peto, Furberg, Stampfer, & Goldhaber, 1985; Yusuf, Peto, Lewis, Collins, & Sleight, 1985)。但是對於硝酸鹽和硫酸鎂的結果則有爭議。後設分析的結果指出這兩種藥物對於治療心肌梗塞有幫助 (Domanski & Friedman, 1994)，但是兩項大型隨機試驗 (ISIS-4與GISSI-3)，蒐集了數萬筆樣本，所得出的結果卻指出使用這兩種藥物和用安慰劑的效果是差不多的。

雖然過去這些研究主要著重在醫學領域，但是在教育領域應亦可能會產生類似的結果，綜整來看後設分析與大規模實驗可能造成結果差異的原因有以下幾點：

1、研究的樣本數大小和研究年份不同

後設分析小規模累積的人數，和大型施測的總樣本數仍有所差距，例如在心肌梗塞的研究中，所累積的總樣本數為2000人左右，然而在大型施測的樣本人數有高達近八萬人，總樣本數的差異可能造成結果的不同。另外，不同的研究年分，也可能造成結果的落差時，例如Domanski等人 (1994) 的大型施測時間是1990年後，而他的後設分析文獻蒐集時間卻在1980年代，後設分析所納入的文獻時間和大型施測的時間點不同，而醫學在1980到1990年間的快速進步，有可能為研究效果造成差異的原因。

2、出版偏誤的影響

研究者在發表文章時，大多都優先發表結果顯著的文章，也就是所使用的藥物是對病情有所幫助的才會拿出去投稿或發表，若結果不顯著，發表 (或被接受) 的情形較少，謝進昌與陳敏瑜 (2011) 也針對這種出版偏誤的現象來進行討論，顯示出版偏誤出現於國內的後設分析研究，並造成結果的誤差。而這也造成文獻選擇上的侷限性，因為在後設分析的選擇上希望能廣納各種進行中的實驗，但是若研究者選擇不發表，則會造成所謂的抽屜效應。

3、文獻品質或是抽樣誤差的影響

後設分析若是納入一些品質不佳、或實驗流程較不嚴謹的實驗都可能造成結果的偏誤。相對而言，大型試驗沒有進行嚴格的品質管控，針對不符合納入原則樣本予以刪除，都可能造成誤差，產生結果的差異性。

比較後設分析和大型資料庫的兩種資料蒐集方式，可綜整歸納為表1。就經費成本來說，後設分析為一種省錢省力的方式，但是其誤差來源較難控制，研究者必須追根究柢，探索文獻中作者蒐集樣本的方式與相關評量工具適切性，其誤差來源不易掌控，另外，許多研究者偏好發表有顯著差異結果的文章，也容易造成出版偏誤的現象。在研究主題的針對性而言，受限於資料蒐集的完整度，有些文獻可能有提供相關訊息但是卻沒有可分析的數據，也會造成研究者的困難。

表1

後設分析與大型資料庫之比較

	後設分析	大型資料庫
經費與人力成本	較低，可由研究者一兩位助理完成	較高，需使用較為龐大的研究經費與人力。
誤差來源	出版偏誤、部分文獻樣本過小會有抽樣誤差，誤差來源較不易控制。	抽樣誤差與測驗誤差，較能透過嚴謹的流程來控制相關誤差
研究主題針對性	不見得所有的主題都有足夠的文獻可以進行後設分析，可能遇到文獻不足無法分析。	可針對研究主題進行相關問卷設計。
執行困難性	較低，大多透過文獻蒐集的方式，只要有網際網路與圖書資源即可完成。	較高，要和受試者聯絡、簽定個資同意書、保密協定，還需監試人員協助發送與回收問卷。
資料類型	次級資料	原始資料

二、新移民子女與本土子女學習表現之相關影響因素

過去的研究指出，家庭的相關背景會影響到學生從幼稚園到大學的學習成就（許誌庭、姜添輝，2015），文化再製的學者Bourdieu（1990，2000）表示學生不同的教育成就表現，為社會結構所衍生的結果，而非學生自身的心智能力或是喜好所產生。家長在家中所建構的社會空間，其家庭成員間彼此分享生活型態，會建造出團體性的認知結構，做為與外界交流的依據，而學生的理解模式則是受這個社會空間的影響，進而影響到其學業表現。高社經地位的學生具有比較多的經濟來源與文化掌控能力，而這種文化資本一代傳承一代，在某種程度上產生具有優勢的地位，而其發展出的社會空間較為抽象、有多元的文化刺激，因此高社經的孩童會有較好的學業表現（林湘慧，2003；陳珊華，2004）。低社經的家庭所建構的社會空間以具體、實務性為主，相關文化與教育資源也較少，因此在學習上處於較為弱勢的地位。

東南亞或大陸地區的新娘嫁至台灣，形成現在新移民，他們多為家庭社經地位與文化資本較低的群體（張芳全、王平坤，2012；林美惠、王奕貞、莊財福，2010）。許誌庭、姜添輝（2015）的研究指出，新移民的家庭學習問題的主因為社會結構的因素，而非個人的心智程度較差，如果新移民的家庭能夠發揮功能，學生的學習表現也未必落後，而教師若能夠協助擔任轉化社會結構的責任，則有助於新移民子女學習。

除了社經背景的影響之外，過去關於影響新移民子女的學習成就表現，可綜整有幾項因素，其中常見的影響因素包括移民的國家來源、學習階段、與學習領域。以下分別針對這些影響因素來進行說明。

第一從新移民的來源來看，台灣的文化中，對於歐美的移民子女較為友善，因此以社會適應的角度來看，歐美的新移民會比大陸或東南亞子女來的好。但吳毓瑩、蔡振州（2015）則對這樣的說法存疑，其研究認為母親的移民背景並非影響學業成就的主要因素，而文化資本（如父親的教育背景）才是。若控制了文化資本的變因，本土與新移民子女的學習表現並無差異。陳建州（2010）使用「台灣教育長期追蹤資料庫」（TEPS）進行分析，其結果亦顯示母親的外籍身分，對於學生的學習成就表現不具影響力。

第二，就學習領域來看，過去關於本土與新移民子女學習表現領域的探討以國語文、數學為主，根據謝佩蓉、曾建銘、王如哲、郭工賓（2015）的研究指出，國小四、

六年級的新移民子女在國語文的學習表現顯著低於非新移民子女。其中四年級的差距高於六年級。其原因可能與知覺班級常規、同儕關係、母親教養風格、或自我概念有關。亦有研究發現，大陸子女的國語文的能力、閱讀能力和本土子女相近，而東南亞子女，國語文成績則較本土子女落後（陶宏麟、銀慶貞、與洪嘉瑜，2015；莫藜藜、賴珮玲，2004；陳佩足、陳小云，2003）。

語言上的障礙是否會影響數學學習表現，則在文獻上有所爭議（Suleiman, 1997; Gelman & Butterworth, 2005）。姚如芬（2009）針對雲林、嘉義、台南地區為研究地區，探討92位新移民子女在數學的學習表現，其成果發現新移民子女在術語計算、統計及代數得分和本土子女差不多，但是在量與實測、圖形與空間與總分則顯著低於本土子女。

趙善如、鍾鳳嬌、江玉娟（2007）以高雄縣的外籍與大陸配偶子女為研究對象，探究影響其學習表現的影響因素，其結果發現在國語文方面，影響的因素包括理解與表達、注意與記憶之學習行為、家中的經濟狀況、性別、母親的國語文程度等，而影響數學的因素則為理解與表達、注意與記憶之學習行為、家中的經濟狀況、社會適應等。

第三為學習階段，過去有關新移民子女的研究之學習階段，大多集中於幼兒及國民小學階段，其次，則涉及國民中學階段，但在過去研究中，卻發現不同的學習階段，本土語新移民子女在數學與國語文的差異有大有小（巫翠芬，2008；黃琮智，2009）。陶宏麟等（2015）的研究更指出這樣的差距會隨者年紀增加而變大，尤其是英語文最為嚴重，其中在國二英語文的差異為小六的近兩倍，視小四的三倍以上。而在國語與數學的差距則沒有達顯著水準，類似的結論也由吳毓瑩、蔡振州（2015）的三年追蹤調查看出，不管數學或國語文，本土子女的成績排名在第一年的排名約高於新移民子女三個名次，不過到了第三年之後，和新移民子女的差距已經不顯著。

受限於研究人力與經費，研究者大多集中在某些固定的區域進行新移民樣本的立意取樣，無法以全國樣本為根基，進行大規模、全國性的樣本取樣，此外，學生所調查的學習成就，主要使用學習總成績，或是研究者本身自己出的題目來當作學習成就表現的依據，然而每個學校都有不同的評分系統，而研究者所出的題目也多非標準化測驗。

本研究除了進行現有文獻的後設分析，透過後設分析策略，有系統的整合新移民子女與本土子女於各項學習表現差異之相關文獻之外，亦使用TASA的大型資料庫進行結果的比較。TASA資料庫除了樣本為全國抽樣之外，用來評估學生學習成就的試題也經過一定的標準化流程，數據採TASA 2007年至2013年的實徵數據，學生樣本年齡涵括小學至高中，並以國、數兩個科目為主體，用以檢視是否大型資料庫所分析出的結果，和過去研究的成果是否符合。

參、研究方法

以下茲就大型資料庫流程與後設資料庫的研究對象、研究實施程序與資料分析方法進行說明：

一、大型資料庫分析

本研究主要針對TASA資料庫進行次級資料分析（Secondary Data Analysis）。首先整

理資料庫中與異國婚姻相關的資料，再定義出哪些資料可以代表新移民子女的特徵，藉以分析新移民子女學習成就表現。

TASA資料中，除了有TASA學生在各科的學習表現之量尺分數之外，亦針對每一位學生進行背景問卷調查，其調查範圍相當廣泛，包括社經背景、親子互動、同儕相處、師生互動、學習態度、學習習慣等各層面，而這些層面都可能影響到新移民子女的學業成就表現，在本研究中，將著重異國婚姻對於學習成就的影響。也就是探討是否學生的母親來自於東南亞國家，學生的學習成就會比較低落。

(一) 研究樣本

在TASA資料庫中，主要使用的研究樣本為2007-2013年的新移民與本土子女的相關資料。而在TASA的學生問卷中，主要辨識是否為新移民子女的資料則透過TASA的學生問卷題目，其題項為：

您的親生媽媽出生於？

- ①台灣地區（含台灣本島及離島）②中國大陸③香港、澳門④越南
- ⑤印尼⑥泰國⑦菲律賓⑧馬來西亞⑨東埔寨⑩緬甸
- ⑪新加坡⑫其他⑬我不知道

若學生選①，則判定此學生為本土子女，若為②~⑩則為新移民子女。本研究將新加坡與其他（如歐美國家）移除，主要因臺灣社會中，對於新加坡與歐美國家的外籍配偶，較其他東南亞或大陸配偶來得友善及禮遇，致使其延伸的學習差異之面向之差異來得不同。依此，本研究首先將先依此題項判定學生新移民之身分，而後利用學生量尺分數，比較新移民子女和本土子女在數學、國語文上的差別。

(二) 研究實施程序

受限於TASA釋出資料的可獲得性，本研究主要使用2007與2009至2013年的資料當作主要分析來源，各年段的資料透過定錨題進行等化，以確保量尺分數的可比性，TASA的學科內容包括課綱中大多数的能力指標，在每個學習領域都有上百題的試題，理論上應該將所有試題給所有學生來作答，但是受限於考試時間，採不同的題本設計，並使用定錨不等組設計（non-equivalent groups with anchor test design, NEAT）的模式來進行等化，不同題本均包含共同區塊，共同區塊的題目為定錨試題。其中2008年TASA並未進行調查。2007年的施測對象為小四、小六、國二、高中職二等四個年段，而2009年的施測對象為小四、小六，2010年則為國二，2011年為高中職，2012年為小四、小六，2013年為國二。每個年段則取國語文、數學兩個主要考科進行分析。由於TASA原先的測驗目的乃了解全台灣學生的學習成就表現，並沒有針對新移民子女進行加重抽樣，在某些年段，尤其是高年段的學生，TASA資料庫中，新移民人數較少，因此在分析時，研究者將資料進行不同資料蒐集年段的合併。只要樣本年級相同，則進行合併（如將2007與2009年小四的資料合併），以檢視新移民子女與本土子女在每個年段的差異性。

TASA抽樣設計主要採用二階分層叢集抽樣設計，第一階段抽樣單位為學校，並使用了兩個分層變數：行政區域和學校規模。而後採用系統機率比例抽樣（systematic

probability proportional to size, PPS) 抽樣方法來抽取樣本學校。第二階段抽樣單位為受測樣本學校內的班級，依據學校抽樣抽出樣本學校，再從樣本學校裡面採簡單隨機抽樣原則抽取受測班級，而抽樣班級數依據學校規模大小而有所不同。抽樣出受測班級後，被抽到受測班級的學生都需要接受測驗，惟身心障礙類別、分校學生、在家教育三大類別不施測。其詳細抽樣方式請參閱國家教育研究院「臺灣學生學習成就評量資料庫」建置計畫的相關網站，網址為<https://www.naer.edu.tw/files/11-1000-1408.php?Lang=zh-tw>。

(三) 資料分析方法

本研究進行兩樣本比較時，除提供描述統計值，像是學科測驗的平均、標準差、樣本數與顯著性檢定指標等之外，亦根據這些數值，進行效果量檢視，此效果量可為兩個組之間所觀察的變項平均值差異（本研究採 g ），或是兩個變項之間的關連大小等等，在計算效果量時，最常代表兩群組差異的效果量公式即為Hedges的 g 值（Hedges & Olkin, 1985）。其計算值算式可如下所示：

$$g = \frac{\bar{X}_1 - \bar{X}_2}{s_p}$$

其中， \bar{X}_1 為本土子女的平均數、 \bar{X}_2 為新移民子女平均數，而 s_p 為考量兩組所加總的標準差（pooled standard deviation），可用下列公式計算：

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 - 1) + (n_2 - 1)}}$$

其中， n_1 為本土子女的樣本數、 n_2 為新移民子女的樣本數、 s_1^2 為本土子女的變異數、 s_2^2 為新移民子女的變異數。

比較本土子女與新移民子女的差異時，當 g 值為正數時，代表本土子女的表現比新移民子女表現為佳，相反的，若是 g 值為負數時，代表本土子女的表現比新移民子女的表現來的差。效果量的大小以Cohen（1992）的方式來進行效果量大小判斷，其中0.8為大效果量、0.5為中效果量、0.2則代表小效果量，而此指標和統計上的各項指標如相關係數 r ，迴歸的 R^2 等，均能作相互的轉換。

二、後設分析資料來源

(一) 研究樣本

本研究有關本土與新移民子女學習表現之後設分析資料來源，主要採用謝名娟、謝進昌（2013）關於本土與新移民子女學習表現差異與所蒐集之後設分析資料。

(二) 研究實施程序

文獻資料乃是以「新移民、新住民、新臺灣、外籍配偶、外籍新娘、東南亞籍、大陸配偶」等關鍵字於國家圖書館的臺灣博碩士論文系統、及臺灣期刊論文索引系統、及華藝線上圖書館進行檢索，覆經檢核與刪除重覆文獻，總共從系統中蒐集出481篇文獻，其中依據準則排除412篇之文獻（含7篇作者不公開全文），剩餘69篇文獻中，尚有10篇文獻來自於同一份研究的調查，最後以59篇文獻作為分析資料，其中包括50篇博碩士論文與9篇期刊論文。累積的總樣本人數為54,852人，而本土子女為46,792人、新移民子女為8,060人，就出版年代而言，文獻涵蓋從2004年至2011，並以2008年、2009年出版之文

獻較多，共有23篇，其次則為2011年及2007年的9篇。以就獲得之文獻研究進行編碼與後設分析。

其中，其訂立之納入與排除準則如下：

1. 該文獻研究必須同時納入新移民及本土子女樣本，並進行學習相關變項之比較。
2. 排除研究是採用大型資料庫進行上述議題探討者。
3. 排除以分析質性資料（如訪談、觀察）為研究分析焦點。
4. 排除未提供充足資料以進行效果量計算之研究。
5. 納入期刊、及博碩士論文類型文獻研究。

文獻研究訊息編碼方面，所含之資料變項包含有國籍、學習科目、學習階段、及相關作為計算效果量之訊息。其中，有關國籍大致分為兩類，分別是母親來自中國大陸（含港澳）或東南亞地區與臺灣人民結婚所生子女。此外，有關學習領域是大致區分為國語文、數學；學習階段則區分為幼兒、國小低年級、中年級、高年級、國中、及高中職等。編碼一致性Kappa係數為0.829，編碼不一致處，皆會再次重新檢視與校正，以力求百分之百正確性。

（三）資料分析方法

根據所得之後設分析資料，研究者進行效果量的計算，為對前述大型資料庫議題分析結果進行比較，同樣採用Hedges的g值(Hedges & Olkin, 1985)。研究者經漏斗圖(funnel plot)來檢視，其文獻分布並沒有出現不對稱現象，且透過線性迴歸考驗來檢測圖形的不對稱性，亦未達顯著水準，代表其研究應不存在出版偏誤的情形。而在同質性檢驗部分，其Cochran (1954)的 Q_r 達顯著水準，代表文獻具有異質性。在加權及參數部分，以Hedges與Olkin (1985)所認為最佳的加權方式進行估計，各調節變項的估計採用固定效果模式。

三、研究變項

本研究主要探討三個變項對本土與新移民學生學習表現的影響，其中學習表現為學生在標準化評量的成績，或是學生在學校的學習表現，或是學者採用各種學習量表（如中文理解測驗）所之成績。包括學習階段在本研究探討的包括幼稚園階段，國小低年級（小一、小二）、國小中年級（小三、小四）、國小高年級（小五、小六）、國中階段、與高中階段。學習科目則包括國語文與數學。新移民學生的母親國籍分為母親來自中國大陸或母親來自東南亞兩類。

肆、研究結果與討論

有關結果與討論，以下茲就不同議題進行說明如下：

一、從學習階段、科目來看不同本土與新移民學生間學習落差

研究分析結果如表2所示，從分析的資料中，可看出低年級的新移民樣本數較多，隨著年紀增長，所收到的樣本數遞減。例如在TASA的資料中，四年級的所蒐集到的新移民學生樣本人數約有兩千多人，其次為六年級、八年級，而以高中職二年級最少，只

有近200多人的樣本。其主因可能為外配婚姻啟自九零年代初期，而大部份的子女在資料蒐集的年段都還在就讀國小階段以下，因此國高中的樣本較少。同樣的在後設分析上的文獻也有類似的情形，大多數的新移民研究鎖定在國小以下的學童，國中以上為樣本的研究較為少見。

以後設分析來看國語的學習落差而言，過去的文獻以幼兒為主，幼兒的平均效果量達0.48，代表新移民子女和本土子女的學習表現在幼兒階段差距達最大，接續國小低年級到中年級差距有下降的趨勢，但是到國小高年級差距又上升。其中，國小中年級的樣本數雖然不少，但是篇數僅有一篇，因此在結果推論上應謹慎。而在大型資料庫的結果顯示，國小中年級與國中的國語文新移民和本土子女的效果量為0.21與0.18，為小效果量（Cohen, 1988），但到高中階段，新移民子女和本土子女在國語文的學習表現上幾乎是沒有差異。然而，在資料蒐集的完整性部分，後設分析在幼兒部分的研究較多，但新移民蒐集的人數以國小低年級較多。而幼兒和低年級的數據剛好是大型資料庫所缺乏的。從大型資料庫和後設分析的結果來比較，可看出結果是具有落差的，例如在國小高年級，後設分析的平均效果量達0.31，但是TASA的效果量卻顯示為0.09。

表2

國語文表現差異之後設分析與大型資料庫比較

學習階段	後設分析				大型資料庫 (TASA)		
	篇數	本土人數	新移民人數	差異平均效果量	本土人數	新移民人數	差異平均效果量
幼兒	11	257	576	0.48	---	---	---
國小低年級	6	571	708	0.2	---	---	---
國小中年級	1	565	362	0.04	21,175	2,694	0.21
國小高年級	3	1,001	219	0.31	24,963	1,602	0.09
國中	---	---	---	---	23,090	1,280	0.18
高中職	---	---	---	---	28,246	262	-0.02

以數學來看，表3可看出不管在後設分析與大型資料庫的結果都顯示新移民子女的表現較差於本土子女，而兩群體表現差異程度大致是位於Cohen(1992)所認為的中(0.5)至小(0.2)的效果量間。後設分析的國中樣本雖有三百多人，但均出自同一個研究，經檢視其研究再選擇樣本時，有控制其社經背景因素，並非隨機抽樣選取，因此可能造成其效果量差異只有0.01。

表3

數學表現之後設分析與大型資料庫比較

學習階段	後設分析				大型資料庫 (TASA)		
	篇數	本土人數	新移民人數	差異平均效果量	本土人數	新移民人數	差異平均效果量
幼兒	---	---	---	---	---	---	---
國小低年級	2	716	122	0.32	---	---	---
國小中年級	9	5,988	1,607	0.54	21,008	2,895	0.29
國小高年級	2	803	353	0.25	24,274	1,579	0.26
國中	1	328	317	0.01	22,456	887	0.24
高中職	---	---	---	---	27,256	238	0.27

二、從母親國籍來看本體與新移民子女學習落差

針對不同國籍學生於不同學習領域進行深入探究時，其結果如表4所示，可發現母親來自於中國大陸之子女，其國語文的學習表現幾乎與本土子女沒有任何差異，在大型資料庫的分析中，甚至可看出大陸子女的表現比本土子女還要好。以數學來看，後設分析的結果中，本土子女表現比大陸子女好，屬小效果量，而國語的後設分析結果可看出，大陸子女和本土子女幾乎沒有差異，但是大型資料庫卻顯示大陸子女國語表現比本土來的好。就母親來自於東南亞地區子女而言，不論是國語文或者是數學，其表現差異均不如本土子女，其平均效果量均為中等程度附近，而後設分析的效果量比大型資料庫還來的大。

表4
不同新移民學生的母親國籍於不同學習領域之比較

背景變項		後設分析			大型資料庫 (TASA)	
		篇數	新移民人數	差異平均效果量	新移民人數	差異平均效果量
母親來自中國大陸 (含港澳)	國語文	6	195	0.06	2,042	-0.19
	數學	10	221	0.27	1,992	0.00
母親來東南亞	國語文	22	799	0.43	3,796	0.19
	數學	15	787	0.60	3,607	0.32

由表5可看出在大型資料庫中，母親來自大陸的表現和本土子女差不多，但是母親來自東南亞地區，則在各階段表現均較本土子女來的差，部分後設分析所使用的文獻，並沒有針對母親是大陸籍與東南亞籍進行區分，因此這些文獻不納入分析，而從後設分析的結果可看出，其效果量大於大型資料庫的效果量，而從大型資料庫可大致看出，母親來自東南亞地區之子女，與本土子女的學習表現差異隨著年級增加而減少。但後設分析則看不出明顯的趨勢。

表5
不同新移民學生的母親國籍於不同學習階段之比較

背景變項		後設分析			大型資料庫 (TASA)	
		篇數	新移民人數	差異平均效果量	新移民人數	差異平均效果量
母親來自中 國大陸 (含 港澳)	國小低年級	3	43	0.04	---	---
	國小中年級	7	81	0.23	1,747	-0.03
	國小高年級	4	37	0.14	1,297	-0.05
	國中	---	---	---	788	0.02
	高中職	---	---	---	202	-0.07
母親來自東 南亞地區	幼兒	13	230	0.51	---	---
	國小低年級	20	266	0.31	---	---
	國小中年級	18	564	0.66	3,842	0.38
	國小高年級	19	94	0.38	1,884	0.33
	國中	---	---	---	1,379	0.31
	高中職	---	---	---	298	0.24

伍、結論與建議

一、結論

本研究主要探討透過大型資料庫與後設分析兩種研究方式，來檢視以下兩個研究問題（1）本土與新移民子女差異效果量是否會隨著樣本的學習階段、科目有所不同？（2）本土與新移民子女差異效果量是否會隨著母親的國籍差異有所不同？

雖然使用不同的方法，但大致可以得到以下兩個結論。第一，本土子女的在國語文與數學的學習表現較東南亞子女來得佳，而數學上的差異大於國語文的差異。本土子女與新移民子女之間學習表現的差異，隨著年紀遞增而差異縮小。大型資料庫與後設分析的結果有些符合，但仍存在差異。以國語文來看，在大型資料庫的結果顯示，國小中年級與國中的新移民與本土子女具有小效果量的差異，但是到了高中階段的差異幾乎消彌。而在後設分析的結果可以看出在幼兒階段的差異最大，到了低中年級的差距下降，但是到了高年級卻又上升，由於缺乏國中以上的資料，因此看不出高年級以後的趨勢。數學部分則均顯示新移民子女表現比本土子女來的差，而隨著年紀上升，後設分析的結果從國小中年級到國中階段兩者的差異變小，但大型資料庫的結果看不出明顯的趨勢。第二，本土子女的學習表現和大陸子女表現上差異不大，有些科目和年段，大陸子女表現甚至優於本土子女，然而，東南亞子女表現則明顯低於本土子女。而這部分的結果在後設分析與大型資料庫的資料都能支持。

其中後設分析在比較國小中年級本土、新移民子女國語文表現時，文獻並不只有一篇，但是由於作者在分析時，並沒有針對低年級、中年級、或高年級的樣本單獨相關數據，或是有提供，但是並沒有針對數學或國語文，而是做綜整性的學業分析，因此在採納這些文獻進行後設分析時，會受到原作者樣本分類與數據提供的狀況而產生限制。但其優點是執行起來較為容易，基本上只要有圖書資源即可進行研究。反過來看在大型資料庫的資料蒐集過程中，所要耗費的成本較高，執行起來也不容易，尤其在現在個資法的保護下，研究者還必須透過研究倫理的審查才能進行資料蒐集，但是在誤差來源部分或是研究主題上，較能由研究者自行掌握。

二、建議

茲針對本研究分析結果，提出以下幾點建議：

第一，分析結果是發現本土與新移民子女學習表現具有隨著年齡增長，呈現微幅下降趨勢。但嚴格來說，此結論仍是由數個橫斷面觀察研究所整合而得，因此，研究者可嘗試以本研究分析結果所提供立論基礎的功能，在未來大規模新移民子女調查或追蹤研究中，以此方向再次檢測前述假設及可能原因探索，目前國內各大型調查研究多是以一般學生為抽樣目標，較少是以新移民子女為焦點，甚至是長期追蹤，因此，若是相關單位（如移民署）未來想針對新移民家庭子女學習表現與因應對策進行調查或研究，可考慮將此面向納入探究。

其次，根據過去的研究指出，在進行大型實驗與後設分析比較性研究時，若是兩種

方法所蒐集的文獻年份不同時，可能會因為政策變動（如推動新的新移民政策），而造成結果的不同。例如在2003年，行政院針對新移民訂定「外籍與大陸配偶照顧輔導措施」，2015年成立行政院層級的新移民事務協調單位，都可能會造成研究成果的影響。因此在本研究中，以大型資料庫與後設分析蒐集年分盡可能相近為原則，以減少相關變因。然而，後設分析的精進以及重要性在於新資料的加入積累，未來可以考量持續補充2013年以後之新的文獻資料跟現有大型資料庫的內容做比較，以進行更深入的分析。

而就研究結果來看，母親來自東南亞子女與本土子女的學習表現，相較於來自中國大陸落差較大。其可能原因為語言不通、文化不利等因素，因此若在有限的經費與資源下，可以針對年紀較小、母親來自東南亞國家的學童給予較多的資源，例如相關的語文課程、補救教材等，而在學校相關政策推動上，可再提供新移民子女的相關文化活動、甚至可以開發相關的新移民文化教材，讓他們能展現自己母國的文化，而學校老師的態度也應再加調整，勿因為母親的新移民身分就與學童學習成就低落做連結，而是能在學校教育中積極的提供相關輔助與教導。

最後，彙整本文的結果，可推論出後設分析與大規模調查結果之差異，可能受到不同誤差因素的影響，例如在其他條件（如測驗題目）均相同的情況下，若後設分析所蒐集到之各研究樣本匯集起來，和大規模調查的樣本特性差異性越大，則可能會造成後設分析和大規模調查的結果差異。大規模調查在進行抽樣時，為全台隨機抽樣的樣本，而後設分析所蒐集到的文獻來源，為學者以便利取樣所得之樣本，其樣本的代表性侷限在某些大縣市或偏遠地區，而沒有中小型都市的樣本，如此一來後設分析的分析結果則可能會受此影響。另外，即使兩邊的樣本來源均具有全台代表性，但是其結果之差異程度仍須檢視各文獻之施測題目是否與大規模之題目具有相似性，越相似則結果差異越小。

這兩種方法可以相輔相成，學者們可以先進行後設分析，探討過去影響的成因可能有哪些因素，並以此為據來進行大規模抽樣的問題設計。而大規模調查之結果，其發現之問題，已可透過後設分析的文獻來進行解析。然而，後設分析與大規模調查結果是否能交互驗證，其影響之因素很多，除非能使用實驗設計來控制這些因素，方能明確解析後設分析與大規模調查結果之差異程度。

誌謝

感謝兩位審查者之意見與科技部專案計畫（MOST 104-2511-S-656-004）補助經費，謹此致謝。

參考文獻

一、中文部份

- 王振世、蔡清中（2008）。臺灣外籍配偶子女學習適應與學業成就之間的關係：東南亞、大陸與本國籍配偶子女之比較。**教育政策論壇**，**11**（2），75-105。
- 吳毓瑩、蔡振州（2014）。東南亞籍新移民女性之子女的學業成就真的比較差嗎？與本地對照組比較之三年追蹤探究。**教育研究集刊**，**60**（1），77-113。

- 巫翠芬 (2008)。國小學童之學習態度與學業成績：彰化縣新台灣之子與非台灣之子的比較 (未出版碩士論文)。立德大學，臺南市。
- 林美惠、王奕貞、莊財福 (2010)。新移民女性參與子女學校教育：以台灣雲林縣個案為例。臺北：秀威資訊。
- 林湘慧 (2003)。家庭社會空間與兒童習性關聯性之研究 (未出版碩士論文)。國立臺南師範學院，臺南市。
- 許誌庭、姜添輝 (2015)。提升人力資本以強化國際競爭力的新移民子女教育政策之探討。嘉大教育研究學刊，34，25-59。
- 姚如芬 (2009)。探究低年級「新台灣之子」的數學學習－以雲嘉南地區的學童為例。教育與社會研究，18，71-99。
- 馬信行 (2002)。教育科學導論。臺北：元照。
- 莫藜藜、賴珮玲 (2004)。臺灣社會「少子化」與外籍配偶子女的問題初探。社區發展季刊，105，55-65。
- 陳珊華 (2004)。小學生文化資本之累積與作用 (未出版碩士論文)。國立台灣師範大學，臺北市。
- 張芳全、王平坤 (2012)。新移民與非新移民子女的文化資本、閱讀動機與閱讀行為之研究。臺中教育大學學報，26 (1)，55-89。
- 黃琮智 (2009)。新移民女性子女中文閱讀能力及其與家庭閱讀環境之相關研究 (未出版碩士論文)。國立臺南大學，臺南市。
- 陶宏麟、銀慶貞、洪嘉瑜 (2015)。臺灣新移民與本國籍子女隨年級的學習成果差異。人文及社會科學集刊，27 (2)，289-322。
- 陳建州 (2010)。影響跨國婚姻子女學習成就之因素：父母「外籍身分」的效果。臺東大學教育學報，21 (2)，61-89。
- 陳佩足、陳小云 (2003)。外籍新娘子女的語言發展問題。國小特殊教育，35，68-75。
- 趙善如、鍾鳳嬌、江玉娟 (2007)。影響外籍暨大陸配偶學童子女學業成績關鍵因素：以個人人口特徵、學習行為能力、家庭環境探討之。教育心理學報，39 (1)，129-147。
- 謝名娟、謝進昌 (2013)。本土與新移民子女學習表現差異之後設分析研究。教育與心理研究，36 (3)，119-149。
- 謝進昌 (2010)。國內教育學門系統性文獻評閱策略及後設分析發展現況與建議。國立臺南大學教育研究學報，44 (2)，130-153。
- 謝進昌、陳敏瑜 (2011)。國內教育、心理後設分析研究出版偏誤檢定之實徵分析。測驗學刊，58 (2)，391-422。
- 謝進昌 (2011)。國內後設分析應用研究發展之概況。教育研究月刊，212，75-82。
- 謝佩蓉、曾建銘、王如哲、郭工賓 (2015)。新移民子女國小國語文學習成就大型評量調查研究，教育科學研究期刊，60 (4)，93-127。

二、西文部份

- Al-Hoorie, A. (2018). The L2 motivational self system: A meta-analysis. *Studies in Second Language Learning and Teaching*, 8(4), 721-754.

- Antman, E. M., Lau, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. *Journal of the American Medical Association*, 268, 240-248.
- Borzak, S. & Ridker, P. M. (1995). Discordance between meta-analyses and large-scale randomized, controlled trials. *Annual International Medicine*, 123, 873-877.
- Bourdieu, P. (1990). *In Other Words: Essays towards a Reflexive Sociology*. Stanford, California: Stanford University Press.
- Bourdieu, P. (2000). *Pascalian Meditations* (Richard, N, Trans.). Stanford, CA: Stanford University Press
- Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, 10, 101-129.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cappelleri, J. C., Ioannidis, J. P., Schmid, C. H., de Ferranti, S. D., Aubert, M., Chalmers, T. C., & Lau, J. (1996). Large trials vs meta-analysis of smaller trials: how do their results compare? *Journal of the American Medical Association*, 276(16), 1332-1338.
- Cooper, H., & Hedges, L. V. (1994). Research synthesis as a scientific enterprise. In H. Cooper, & L. V. Hedge (Eds.) *The handbook of research synthesis* (pp. 3-14). New York, NY: Russell Sage Foundation.
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108-121.
- Domanski, M. & Friedman, L. (1994). Relative role of meta-analysis and randomized controlled trials in the assessment of medical therapies. *American Journal of Cardiology*, 74(4), 395-396.
- Feldman, K. A. (1971). Using the work of others: Some observations on reviewing and integrating. *Sociology of Education*, 4, 86-102.
- Glass, G. V. (1976). Primary, secondary and meta-analysis of research. *Educational researcher*, 5, 3-8.
- Gelman, R., & Butterworth, B. (2005). Number and language: How are they related? *TRENDS in Cognitive Sciences*, 9(1), 6-10.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.
- Hopfenbeck, T., Lenkeit, J., Masri Y., Cantrell, K., Ryan, J., & Baird, J A. (2017). Lessons Learned from PISA: A Systematic Review of Peer-Reviewed Articles on the Programme for International Student Assessment. *Scandinavian Journal of Educational Research*, 1-21.

- Higgins, J. P. T., & Greens, S. (2008). *Cochrane handbook for systematic reviews of interventions*. Chichester, England: John Wiley & Sons, Ltd.
- Kieffer, M. J., Lesaux, N. K., Rivera, M., & Francis, D. J. (2009). Accommodations for English language learners taking large-scale assessments: A meta-analysis on effectiveness and validity. *Review of Educational Research, 79*(3), 1168-1201.
- LeLorier, J., Gregoire, G., Benhaddad, A., Lapierre, J., & Derderian, F. (1997). Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *New England Journal Medicine, 337*, 536-542.
- Murphy, K. R. (2003). *Validity generalization: A critical review*. Mahwah, NJ: Lawrence Erlbaum.
- Nieuwenstein, M., Wierenga, T., Morey, R., Wicherts, J., Blom, T., Wagenmakers, E. and van Rijn, H. (2015). On Making the Right Choice: A Meta-Analysis and Large-Scale Replication Attempt of the Unconscious Thought Advantage. *Judgment and Decision Making, 10*, 1-17.
- Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology, 62*, 529-540.
- Suleiman, M. (1997). *A balanced formula for math and science education in diverse setting*. Paper presented at the National Association for Multicultural Education Conference, Albuquerque, NM.
- Uysal, S. & Sarier, Y. (2018). Meta-analysis of school leadership effects on student achievement in USA and Turkey. *Cypriot Journal of Educational Science, 13*(4), 590-603.
- Villar, J., Carroli, G. & Belizan, J. M. (1995). Predictive ability of meta analyses of randomised controlled trials. *Lancet, 345*, 772-776.
- Yusuf, S., Peto, R., Lewis, J., Collins, R., & Sleight, P. (1985). Beta blockade during and after myocardial infarction: An overview of the randomized trials. *Progress in Cardiovascular Disease, 27*(5), 335-371.
- Ysusuf, S., Collins, R., Peto, R., Furberg, C., Stampfer, M. J., Goldhaber, S. Z. and Hennekens, C. H. (1985). Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: Overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials. *European Heart Journal, 6*, 556-585.

投稿日期：2019年01月02日

修正日期：2019年03月25日

接受日期：2019年03月29日

A Cross Validation Study of Differentiated Performance between Native and Immigrant Children with Results from Meta-Analysis and Large-Scale Assessment-Taking the Difference in Learning Achievement between New Immigrants and Taiwanese Students as an Example

Ming-Chuan Hsieh

Research Fellow, Research Center for Testing and Assessment, National Academy for Educational Research

ABSTRACT

The purpose of this study was to assess the difference in learning achievement between the local children and the new-immigrant children in Taiwan. A large-scale database analysis and meta-analysis were adopted to analyze. TASA data which taken from four, six, eight, and eleven graders from 2007 to 2013 were adopted as the large-scale database. Fifty-nine research papers were used for coding as a meta-analysis. This result was in good consistency with both analyses. The Taiwan local children's learning performance is better than new-immigrant children's. However, there was some discrepancy between performances. In addition, with the students' age increase, the difference in the performance became smaller. The advantages and disadvantages of both study methods were discussed as well.

Keywords: large scale assessment, new immigrant students, academic achievement, meta analysis, TASA

