

行政院國家科學委員會專題研究計畫成果報告

* 教育學門統合分析之綜合 *

* Synthesis of meta-analyses in education *

計畫類別： 個別計畫

計畫編號：NSC89-2413-H-004-012

執行期間：89年8月1日至90年7月31日

計畫主持人：馬信行

處理方式：可之即提供參考

執行單位：國立政治大學

中華民國90年9月1日

摘要

本文主要以教育學者對教育學之各領域所做統合分析之結果加以綜合，不少統合分析除了總平均效應量之外，還作細部分析，檢定細分變項(moderator variables)是否也達顯著。由於有些細分變項的效應量數目太少(例如少於10個)，故不加以列述。本文研究方法是對重要期刊，尤其是 Review of Educational Research 做地毯式搜尋，並用電腦網路之教育期刊以 meta-analysis 為關鍵自查出教育學門統合分析之研究報告。本文除介紹各統合分析之自變項與依變項之操作型定義之外，還將各研究報告所列之效應量求出總平均效應量、標準差、標準誤及 t 值，看總平均效應量是否顯著的不等於零。如果原來的研究報告已算出 t 值，則從其值。本文結果擬形成「教育科學導論」，以供將來教育學者重複驗證，並對本文所沈澱之教育學知識體加以增修，使教育學真正成為一門科學。

Synthesis of Meta-analyses in Education

Abstract

Hsen-hsing Ma

This study synthesizes the meta-analyses done by educational empirical researchers in various fields of education. A lot of meta-analyses calculated not only average mean of effect sizes, but also average means of moderator variables. But many of moderator variables were not adopted by this study, because of too few number, say less than 10. The method used for synthesizing the results of meta-analyses was using t-test to examine whether the mean effect size was significantly different from zero. The articles of meta-analysis were searched from the important educational journals, such as Review of Educational Research, as well as from computer data bases, such as EBSCOhost. This study has described the operational definitions of each variable and calculated from the available N, mean, and standard deviation, the standard error and t-value of each meta-analysis. If the t-value has been presented originally in article, than it would be adopted directly. The contribution of this study would be the formation of a textbook "introduction to the educational science", which is characterized by the precipitated empirical knowledge body of education.

教育學門統合分析之綜合 (Synthesis of Meta-analyses in Education)

一門科學的知識，最先是研究先鋒將實證研究所發現的知識發表在期刊，經過五年、十年、甚至廿五年的被引用、討論，之後被權威的教科書引用，最後沈澱被融入知識體，而成為知識核心 (Cole, 1983, p.114)。自然科學自十六世紀開始實證性研究，經過三百年的累積已沈澱出可觀的知識體，從普通物理學、化學、生物學等的教科書可見一斑。而教育學的實證研究尚在幼稚階段。

教育科學的實證研究在 1970 年代開始逐漸推展。在 1980 年代則開始有學者以統合分析 (Meta-analysis, 或譯為後設分析) 的方法將教育學門的各領域內之類似的教育實驗結果加以整理，此為一門科學將實證研究結果沈澱為該學門知識體 (knowledge body) 之肇始。由於各個教育實驗所使用的變項單位不一，故 Glass 最先以效應量 (effect size 簡寫為小寫 es) 加以標準化，使不同的研究可放在一起評估，Glass (1976) 也最先將這種方法命名為統合分析。之後，統合分析的方法也逐漸拓展，(參閱 Glass, Mcgaw, & Smith, 1981; Hedges & Olkin, 1985)。Lipsey & Wilson (1993) 將教育方面的統合分析加以歸類研析，諸如電腦輔助教學、個別化教學、視聽輔助教學、合作學習、同儕教學、行為目標、增強、暗示、回饋、精熟學習、開放教育、班級大小、分組教學、能力分班、學生評鑑教師、測試焦慮、數學與科學教育等。Lipsy & Wilson 也對每個研究主題列出平均效應量和篇數，但未列出標準差。本研究擬繼續追蹤，找出平均效應量的標準差，然後算出平均效應量的 95% 信賴區間或 t 值。95% 信賴區間與 t 值一樣可知平均效應量是否達顯著，只要 95% 信賴區間不含零，即表示達 0.05 顯著水準。一般求 ES (EFFECT SIZE, 以大寫表示總平均效應量) 時，都發現效應量並不恆常，隨著各研究而有起伏，以均質性統計考驗 (Q_w) 檢定的結果都易達顯著，但對結果並不會構成太大差異，因為把偏離值 (outlier) 去除，均質性改善後，仍

維持原來的顯著水準。另外 Lipsey & Wilson 並未詳細列出變項的操作型定義，故本研究擬找出原始的研究，以確定其操作型定義，此在沈澱知識時甚為重要。

方法

本研究除追蹤找出(Lipsy & Wilson)1993年所綜合之原始統合分析之標準差與變項之操作型定義之外，在現有1993年以後的教育期刊做地毯式搜索：

1. Review of Educational Research(此為大宗)
2. American Educational Research Journal
3. Journal of Research in Science Teaching
4. International Journal of Educational Research
5. Journal of Educational Research

另外，以ERIC及EBSCO網路上以"Meta-analysis"為關鍵字找出其他有關教育的統合分析論文。

本研究效應量呈現的主要方式是：

1. 單個研究的效應量以小寫"ef"表示

$$es = \frac{M_E - M_C}{SD_C} \dots\dots\dots (1)$$

(1)式中 M_E 為實驗組的依變項平均值， M_C 為控制組的平均值， SD_C 為控制組的標準差。

$$\text{或者，} es = \frac{M_E - M_C}{SD_{pooled}} \dots\dots\dots (2)$$

(2)式中 SD_{pooled} 為實驗組與控制組的合併標準差

公式(1)是由Glass(1976)所提，公式(2)是由Schmidt & Hunter(1977)所提。這兩種方法都可用，也都有人用。Glass之所以要用控制組的標準差當分母，是因為控制組未受實驗處理的影響。而Schmidt & Hunter用兩組的標準差合

併計算為分母，是因為他們認為在等組時這樣做，會使誤差減少（Hough & Hall, 1994）。

2. 多個研究效應量的的平均效應量是以大寫 SE 表示

$$ES = \frac{\sum_{i=1}^k es_k}{k} \dots\dots\dots (3)$$

(3) 式中 ES 為 k 個效應量之平均效應量， es_k 為第 k 個研究的單個效應量。 SE 的 95% 信賴區間為：

$$SE \text{ 的 } 95\% \text{ 信賴區間} = ES \pm (t_{.95, k-1}) \left(\frac{SD_{es}}{\sqrt{N}} \right) \dots\dots\dots (4)$$

(4) 式中自由度 = $k-1$ ， $t_{.95, k-1}$ 為在 $k-1$ 自由度時 t 值在 $\alpha = .05$ 時之臨界值（也就是查表值）。 SD_{es} 為 k 個效應量之標準差。雖然有些學者認以統合分析時，所有相關研究已一網打盡，故相當於母群體了，因此不需再作推論統計，如 t -考驗，之分析。筆者覺得雖然所收集的研究極可能完全納入，可是那只是時間上及空間上的點，隔年仍有新的統合分析的研究出現，且美國以外的地區所發表的統合分析未必會被收集到，故統合分析的研究母群體應是無限的，故仍宜作推論統計。故本研究仍可將未作 t -考驗的統合分析求出 t -值。

如果統合分析的對象不是實驗研究而是相關係數的研究，則以 $\bar{r} \pm (t_{.95, k-1}) \frac{SD_r}{\sqrt{k}}$ 表示平均相關係數之 95% 信賴區間。其中 \bar{r} 為 k 個研究的平均相關係數， $t_{.95, k-1}$ 為平均相關係數在 $\alpha = .05$ ， $df = k-1$ 時的 t 臨界值。 SD_r 為 k 個相關係數的標準差。如果要將 r 轉換為 es ，則用公式 (5)

$$es = 2 \sqrt{\frac{n}{n-1}} \cdot \frac{r}{\sqrt{1-r^2}} \dots\dots\dots (5)$$

(5) 式中， n 為單個研究的樣本數， r 為單個研究的相關係數。(5) 式中在轉換時有考慮到樣本數，故可適用大樣本及小樣本的研究之轉換。當樣本數接近無限

大時， $n - \frac{1}{n}$ 接近於 1，而成大樣本之轉換公式。

在方法論上，使用相關係數矩陣轉換為效應量，容易引起觀察值是否獨立的問題。在一個相關矩陣裡，無法從一個相關係數預測下個相關係數，這支持相關係數間是獨立的假設。但因相關矩陣裡的相關係數皆來自同一群體，故支持相關係數間是不獨立的之假設。

Tracz, Elmore, & Pohlmann (1992) 用模擬的方式來探討此問題。在變相間的母群體相關係數 ρ 是非零時， r 之中數常大於 r 之平均數。因 r 是偏態統計。當 ρ 是正時，其分配是負傾斜，但是當 N 增大時，則 r 之中數與 r 之平均數趨近於 ρ 。例如當 $N=20$ 時，在 $\rho=0.7$ 之情況下， r 的平均數為 0.688， r 的中數為 0.708。當 N 增到 50 時， r 的平均數為 0.697， r 的中數為 0.703。當 N 再增到 100 時， r 的平均數為 0.696， r 的中數為 0.700，越接近 0.7。Tracz, et al. (1992) 用五個自變項模擬，當 $N=20$ 時，在 $\rho=0.7$ 的情況下，0.90, 0.95, 0.99 之信賴區間各為 0.898, 0.950, 0.989。在 $N=50$ 時，則 0.90, 0.95, 0.99 之信賴區間各為 0.900, 0.948, 及 0.987。當 N 增到 100 時，則 0.90, 0.95 與 0.99 之信賴區間各為 0.899, 0.951, 及 0.99，可說很接近。故使用相關係數做統合分析時，違反資料獨立性似乎不會構成嚴重的錯誤推論機率。

有不少統合分析的論文在分析總平均效應量後，再作細部分析，例如細分真實驗設計與準實驗設計、研究是出版的或未出版的、新近出版的或早期出版的、研究是學位論文或是期刊論文、受試者是那個教育階段、受試者的智力是正常的還是啟智班的、實驗期間長短、實驗處理類型、依變項類型、實驗情境是在實驗室或在自然情境等等。但由於在細部分析裡，有些項目的效應量數目太少，筆者覺得效果會不穩定，故除非有必要，否則不再作細部分析之詮釋。只待未來研究數量增多到相當多的數目時才考慮。

有些作統合分析的研究者在判斷平均效應量時，依 Cohen (1977, p.24-27) 的建議將平均效應量正負 0.2 視為小的平均效應量，將正負 0.8 視為大的平均效應量。筆者覺得平均效應量的大小選得依效應量的數目及標準差而定，故在本文中，平均效應量的大小以其 t -值而定。

結 果

本研究對各子題的排序，依 Fraser, Walberg, Welch, & Hattie (1987)之排序為骨幹，重新整理，也有新增部分再加上。Fraser, et al. (1987) 有一些子題之效應量如取自 Dissertation Abstract International，本研究不採用。因光從 Abstract 無法知道其變項的操作型定義，故僅取其採用期刊論文部分。如下所列：

1. 教育政策
 - 1.1. 對有發展遲緩危機之幼兒早期矯治訓練之效果
 - 1.2. 美國中小學教育法案第一標題(對經濟處境不利學生之補救教育方案)之實施效果
 - 1.3. 回歸主流(特殊教育學生編在正常班與特殊班效果之比較)
 - 1.4. 有中輟危機 (high-risk) 的大學生之輔導效果
 - 1.5. 雙語教育之效果
2. 影響學生學業成就的學校因素
 - 2.1. 影響學習的學校因素
 - 2.2. 開放教育
 - 2.3. 性別在中小學校長領導風格上的差異
 - 2.4. 能力分班
 - 2.4.1. 校內能力分(組)班
 - 2.4.2. 班內分組教學
 - 2.4.3. 不分年級的小學
 - 2.4.4. 資賦優異教育
 - 2.5. 資優生加速教育之效果
3. 家庭因素對學業成績的影響
 - 3.1. 社經地位對學業成績的影響
 - 3.2. 在家看電視時間對學業成就的影響
 - 3.3. 以學校為主的家庭教育
 - 3.4. 針對殘障的、不利的及瀕臨危機的孩子之父母參與
4. 教學法
 - 4.1. 動機對學業成就的影響
 - 4.2. 各論
 - 4.2.1. 電腦輔助教學之效果
 - 4.2.1.1. 電腦輔助教學對學業成就的影響
 - 4.2.1.2. 電腦輔助教學對認知能力的影響
 - 4.2.1.3. 超媒體教學對學業成績的影響

- 4.2.2.個別化教學
- 4.2.3.精熟學習
- 4.2.4.編序教學法
- 4.2.5.合作式的學習
- 4.2.6.同儕個別教導
- 4.2.7.科學教育教學法
 - 4.2.7.1.探究式科學課程
 - 4.2.7.2.活動為主的小學科學教育課程
 - 4.2.7.3.對控制變項能力的訓練(學生實驗精神的培養)之效果
 - 4.2.7.4.大學數、理、工、科技教育小組教學之效果
- 4.2.8.模擬教學之效果
- 4.2.9.前導組織體
- 4.2.10.概念構圖對學習的效應量
- 4.2.11.教科書中附加圖片之效果
- 4.2.12.回饋對測驗式個別化學習之效果
- 4.2.13.家庭作業對學習之效果
- 4.2.14.野外探險教育之效果
- 4.3.教學方法之應用
 - 4.3.1.作文教學之效果
 - 4.3.2.教導學生使用電腦文書處理對寫作的影響
 - 4.3.3.字彙教學法
 - 4.3.4.對學習障礙學生之教學策略之效果
- 4.4.教師發問技巧訓練之效果
- 4.5.教師期望之效果
- 4.6.學生評饋之效果
- 5. 特殊訓練方案的效果
 - 5.1.減肥訓練的效果
 - 5.2.電腦程式撰寫訓練對認知能力的效果
 - 5.3.對測驗之補習效果
 - 5.3.1.對學業成就測驗之補習效果
 - 5.3.2.對性向測驗之補習效果
 - 5.3.3.應試技巧訓練對成就測驗成績之效果
 - 5.3.4.減輕測試焦慮訓練之效果
 - 5.4.1.對學習技巧訓練之效果
 - 5.4.2.訓練學生自己擬出問題以促進閱讀理解之效果
 - 5.4.3.閱讀之前給予引導的效果
 - 5.5.犯罪理論之統合分析
 - 5.5.1.遺傳與犯罪關係之統合分析
 - 5.5.2.學校裡學生問題行為之防治

- 5.5.3. 犯罪矯治之效果
- 5.5.4. 對性犯罪矯治之效果
- 5.5.5. 毒品濫用之矯治效果
- 5.5.6. 暴力犯之矯治效果
- 5.6. 使用行為改變技術以減少公立學校學生搗亂行為之效果
- 5.7. 採用社會技巧訓練以減少情緒困擾與行為困擾之效果
- 5.8. 觀看電視暴力節目對反社會行為的影響
- 5.9. 學校性教育的效果
- 5.10. 道德教育措施對道德判斷發展的效果
- 5.11. 功能交換理論的道德觀
- 6. 性別差異
 - 6.1. 性別差異對科學學業成就的影響
 - 6.2. 性別差異對認知能力的影響
- 7. 研究方法論之問題
 - 7.1. 酬謝金對問卷回收率之效果
 - 7.2. 在教育實驗的霍桑效應
 - 7.3. 假實驗處理之效果

以下是依照教育領域內之各子題所做統合分析的初步綜合：

1. 教育政策

1.1 對有發展遲緩危機之幼兒早期矯治訓練(early intervention program)的效果

對於有發展遲緩危機之幼兒愈早矯治則愈能及早發展其學習能力。1986年美國通過殘障教育法修正案(Education of Handicapped Act Amendments)，1990年通過失能者教育法(Individuals with Disabilities Education Act)為失能嬰兒與幼兒提供早期矯治計劃。事實上所服務的兒童是出生到六歲，需要被早期矯治的幼兒是那些有發展遲緩危機的幼兒(例如出生時體重太輕、出生時有呼吸困難者、腦溢血、有傳染病、營養不良、及受到虐待或忽略等)。所謂發展遲緩是指經過測量後測量值低於平均值 1.5 或 2 個標準差者(各州標準不一)。達此標準者其父母便可申請接受早期矯治計劃，加入這種計劃的失能兒童便可接受特殊教育或/及醫療服務(Spiker, et al., 2000)。

Ottenbacher & Petersen(1985)收集了 38 篇實證研究作統合分析。這些研究是檢定早期治療對器官受損的幼兒之效果，共得 118 個效應量。依變項有五類：(a) 整體發展的評估；(b) 肢體動作/反射機能之測量；(c) 語言之測量；(d) 社會人際交往及自助表現之測量；(e) 認知能力的測量。矯治地點有以家庭為主，也有以醫院為主。這些接受早期矯治服務的幼兒大都是在三歲前就開始接受矯治，並都至少有一項器官受損者。統合分析的結果，總平均效應量為 0.97，標準差為 0.90，因有 118 個效應量，故推算出標準誤為 0.08285， $t_{(117)}=11.7$ ， $p<.01$ ，達非常顯著，表示接受早期矯治計劃的器官受損幼兒在身心發展的測量值方面比未接受矯治計劃的器官受損幼兒有顯著的改善。Ottenbach & Petersen(1985)再根據依變項作細部分析，發現五類依變項都達顯著：

1. 整體發展量表方面，平均效應量為 0.86，標準誤為 0.14， $t_{(25)}=6.14$ ， $p<.01$ ，達非常顯著。
2. 肢體動作與反射：平均效應量為 0.90，標準誤為 0.17， $t_{(29)}=5.29$ ， $p<.01$ ，

達非常顯著。

3. 語言：平均效應量為 1.29，標準誤為 0.20， $t_{(22)}=6.45$ ， $p<.01$ ，達非常顯著。
4. 社會/人際及自助：平均效應量為 0.86，標準誤為 0.23， $t_{(18)}=3.74$ ， $p<.01$ ，達非常顯著。
5. 認知：平均效應量為 0.98，標準誤為 0.22， $t_{(19)}=4.45$ ， $p<.01$ ，達非常顯著。

顯出接受早期矯治計劃的器官受損幼兒在整體發展、肢體動作/反射、語言、社會/人際及自助、與認知上都會顯著的優於未接受矯治計劃的器官受損幼兒。

Casto & Mastropieri(1986)也對檢定早期矯治計劃的研究作統合分析。他們收集 74 篇實證研究，共得 215 個效應量。這些實證研究論文是在 1937 到 1984 年間刊出，但大部分是在 1970 以後刊出的。受試者有 44%是心智發展遲滯(mentally retarded)，29%是多重障礙，10%是器官受損，8%是語言能力受損，4%是有情緒困擾(emotionally disturbed)，有 3%是普通的發展遲緩(general developmental delays)，有 2%是聽力受損。所用依變項有 IQ、社會能力、語言能力、肢體動作技巧、學業成就等。統合分析的結果，總平均效應量為 0.68，標準誤為 0.05， $t_{(214)}=13.6$ ， $p<.01$ ，達非常顯著，表示對於有發展遲緩危機的幼兒，若施以早期矯治，則在心智發展上會有顯著的優於未接受早期教育的發展遲緩幼兒。Casto & Mastropieri(1986)作細部分析發現早期矯治的期間愈長，矯治時間超過 500 小時者，平均效應量為 0.86，標準誤為 0.12， $t_{(38)}=7.17$ ， $p<.01$ 。矯治期間在 50-100 小時者，平均效應量為 0.62，標準誤為 0.12， $t_{(20)}=5.17$ ， $p<.01$ ，矯治時間在 50 小時以下者，平均效應量為 0.56，標準誤為 0.16， $t_{(21)}=3.5$ ， $p<.01$ ，雖然都達非常顯著，但矯治期間 500 小時以上者，自由度與 t 值皆最高，故效應量也最大。在依變項的細部分析方面，雖然 Casto & Mastropieri(1986)有列出平均效應量及樣本數，可是

沒列出標準差或標準誤，故尚無法判斷其顯著水準。

1.2 美國中小學教育法案第一標題（對經濟處境不利學生之補救教育方案）之實施效果

在美國，中小學教育法案之標題 1 (Title 1 of the Elementary and Secondary Education Act) 是在 1965 年實施。該法案每一標題 (title) 代表一個領域，例如標題七代表雙語教學等。標題一在提供經費補助給地方教育機構去擴展與改善低收入戶子弟之教育。特別是要滿足教育不足兒童 (educationally deprived children) 在教育機會與學業成就上的需求。每年聯邦的補助款總共大約 80 億美元 (Borman & S'Agostino, 1996; Shaul, 2000)。

1994 年以前補助的對象主要是經濟上與教育上處境不利的學生 (disadvantaged students)，1994 年以後補助的對象逐漸擴大到所有學生 (Shaul, 2000) 接受補助的學校一般是為處境不利學生提供額外的補救教育 (supplementary education)。例如利用放學後時間、週末、或暑假提供教學時間，有些用個別指導 (tutoring)，有些用小組教學、或全班教學。補救教育的重點在核心課程，尤其是數學與閱讀。有些學校除了補救教學之外，還提供學生點心及交通。有些還舉辦親職教育，教導父母如何在家協助其子弟的課業。有些則提供電腦及軟體，進行個別化教學，以讓學生依其自己的學習速度決定進度。接受補助的學校可以全校性實施補救教學或只針對處境不利學生作補救教學。

原則上，在 1994 年以前，一個學校貧窮率超過 75% 者才可選擇全校性實施補救教學的方式。1994 年標題 1 的再授權 (reauthorization)，鼓勵學校採取全校性實施補救教育方式，規定貧窮率超過 50% 即可實施全校性實施方式。所謂貧窮率是一個學校的貧窮學生 (其父母所得在全國貧窮標準線以下者) 所佔的百分率。全校性實施的優點是成本較便宜，缺點可能使處境不利的學生反而沒受到應有的重視。如只針對處境不利學生作補救教學，則成本較高，尤其是採個別教學或小組教導的方式，師資成本要提高 (Shaul, 2000)。

1974 年，美國國會採取主動，提議要發展系統性的評核工具，以對標題 1 方

案實施的效果作全國性的評核。尤其要求用客觀的標準來評量，以供作州與全國性的比較。Borman & D'Agostino (1996) 收集 17 篇實證研究作統合分析。這些實證研究是在 1966-1993 年間刊出。共有 657 個效應量，統合分析的結果，加權的總平均效應量為 0.11，標準誤為 0.004， $t_{(656)}=27.5$ ， $p<.01$ ，達非常顯著。這些效應量大都是實驗組的後測優於前測之效果，以後測平均值減去前測平均值，之後除以前測之標準差，而非實驗組後測優於控制組後測之效果。因 Borman & D'Agostino(1996) 並未檢定控制組的前後測是否達顯著，故效應量應有所保留，不過 t 值達 27.5，扣除前測驗效果，真正效應量應仍可達顯著。這些效應量大都是以閱讀或／與數學為依變項。Borman & D'Agostino 的細部分析雖然顯出低年級學生的效應量大於高年級（從 1 年級到 12 年級），在寒假前所測的效應量大於在暑假前所測者，但由於未列出標準差，故尚不能斷定其顯著機率。在依變項上，應該還可用出席率及輟學率，但在 Borman & D'Agostino 的統合分析中未見列出。

1.3 特殊教育學生編在正常班與特殊班效果之比較

在特殊教育早先是集中在特殊班上課。之後回歸主流 (mainstreaming) 的運動，主張特殊教育的學生應與正常班級的學生一起上課，以便與正常班的同儕有較多的互動機會。1960 年代回歸主流的運動蔚為風潮，1970 年代特殊教育班的成長便有顯著的下降。Carlberg & Kavale(1980)收集 50 篇實驗研究，共有 322 個效應量，比較特殊教育學生編在正常班與特殊班的效果，統合分析的結果，得總平均效應量為-0.12、標準差為 0.65、標準誤為 0.036223、 $t_{(321)}=3.31$ ， $p<.01$ 。表示將特殊教育學生編在特殊班的話，效果顯著的比編在正常班的差。他們又根據依變項作細部分析，發現如以學業成就為依變項，則平均效應量為-0.15、標準差為 0.59、標準誤為 0.052354、 $t_{(128)}=2.865$ ， $p<.01$ 達非常顯著，表示特殊教學生如被編在特殊班，則其學業成就 (成績) 會顯著的比被編在正常班者差。如以社會適應分數為依變項，則平均效應量為-0.11、標準差為 0.74、標準誤為 0.05832、 $t_{(160)}=1.886$ ， $p>.05$ ，不達顯著。顯出特殊教育學生如被編在正常班，則在社會適應上尚無顯著的優於被編在特殊班。

Carlberg & Kavale(1980)將特殊教育學生細分為 (a) 可教的智能遲鈍 (educable mentally retarded) 學生，其 IQ 在 50-75 之間，(b) 學習遲緩 (slow-learning) 學生，其 IQ 在 75-90 之間，(c) 學習困難 (learning-disabled) 學生、行為失常 (behaviorally disordered) 學生及情緒困擾 (emotionally disturbed) 學生。分別求出分班的效應量。結果發現：

1. 對於 IQ 在 50-75 之間可教的智能遲鈍生，平均效應量為-0.14、標準差為 0.61，筆者算出標準誤為 0.038657、 $t_{(248)}=3.62$ ， $p<.01$ 達非常顯著，表示 IQ 在 50-75 的可教的智能遲鈍生如被編在特殊班上課，其成績會顯著的較被編在正常班，與正常生一起上課者差。
2. 對於 IQ 在 75-90 之間的學習遲緩生，平均效應量為-0.34、標準差為 0.85，筆者算出標準誤為 0.1378882、 $t_{(97)}=2.466$ ， $p<.05$ 達顯著，表示 IQ 在 75-90 之

間之學習遲緩生如被編在特殊班，其成績會比被編在正常班，與正常生一起上課者差。

3. 對於學習困難、行為失常、及情緒困擾的學生，平均效應量為 0.29、標準差為 0.54，筆者算出標準誤為 0.091277、 $t_{(34)}=3.18$ ， $p<.01$ ，表示患有學習困難、行為失常、或情緒困擾的學生如被編在特殊班上課，其成績會顯著的比被編在正常班，與正常生一起上課者為優。

1.4. 有中輟危機 (high-risk) 之大學生之輔導效果

Kulik, Kulik, & Shwalb(1983)收集 60 篇在 1930-1979 年間刊出之有關有中輟危機之大學生之輔導效果之實證研究作統合分析。有些大學生在其學術基礎與技巧方面有所不足，例如測驗分數低，在大學的成績低。如果不給予補救教學，則可能面臨離開學校。在 1930 年代與 1940 年代，在美國大學中有為入學測驗成績、高中在學成績、及大學第一年成績較差者提供無學分的補救教學，這種補救教學一般是提供閱讀與學習技巧之課程。在 1950 年代與 1960 年代開始對有中輟危機的學生注重情意與認知的發展。提供個別諮商或團體諮商，諮商方式也由指導性諮商轉為非指導性諮商。1960 年代民權運動，注重社會上、經濟上、及教育上弱勢團體之教育機會均等問題，教育機會開放與擴張之後對弱勢團體之學生提供全學期的、有學分的基礎課程，像數學、英文、閱讀、及學習技巧等。由大學教師或諮商人員組成團隊教導。Kulik, et al.(1983)統合分析的結果，在學生不中輟，繼續唸到畢業的平均加權效應量為 0.19、標準差為 0.23、標準誤為 0.04、 $t_{(28)}=4.75$ ， $p<.01$ ，達非常顯著，表示可能中輟的大學生若給予輔導補救教育，則其繼續唸到畢業的百分率會顯著的高於沒接受輔導補救教育的中輟危機大學生。在對學期成績的影響方面，總平均效應量為 0.27、標準差為 0.32、標準誤為 0.04、 $t_{(56)}=6.75$ ， $p<.01$ ，達非常顯著，表示有中輟危機的大學生若接受輔導補救教育，則其學期成績會高於未接受輔導補救教育的中輟危機大學生。在輔導補救教育中，大部份是教導學術技術 (academic skills)，平均效應量為 0.29、標準誤為 0.06、 $t_{(31)}=4.83$ ， $p<.01$ ，達非常顯著，表示對有中輟危機的大學生，若施與學術技術的教導，則其學期成績會顯著的優於未接受學術技術教導的中輟危機生。另外有輔導諮商及綜合性的支持服務也達顯著，由於效應量數目尚少，故不予討論。

1.5. 雙語教育之效果

美國移民問題較大，移民之子弟之母語非英語時，在上課時會因語言的障礙而感到學習有困難，如何幫助這些移民子弟克服語言的障礙，是中小學教育法案標題VII的目標。中小學教育法案標題7(Tital VII)是由聯邦政府補助地方教育機構從事中小學的雙語教育。目標設定在「英語非其母語」的學生提供均等的教育機會。希望透過加強這些學生的英語能力，使其在學業表現上不致因語言的關係而受負面的影響。

美國的雙語教育是使用兩種語言當作教學媒介，而非像加拿大所實施的，對以英語為母語學生教法語直到3-4年級後，學生依其志願以兩種語言繼續學習。在美國，雙語教育的對象一般是低社經地位的非以英語為母語之學生。一旦英語的能力能適應學校要求時，便可離開，去接受專以英語授課的課程。故這種方案也稱為過渡性的雙語教育方案，目標是在學習第二語言(英語)的同時也能增進其學業表現(Willig,1985)。

在美國辦理雙語教育的方式有很多種，Valverde & Armendariz(1999)將之歸為五類：

1. 調出式(pull-out program)：英語非其母語的學生從其原來班級被調出來去接受英語補教教學，不同母語的非英語母語學生集中在一班上課，全部以英語教學。
2. 結構性的浸入式(structured immersion)：在各科目上課中教師用英語上課，教師具有學生母語的能力(例如在一個多數學生來自拉丁美洲，西班牙語是其母語的班級，教師也會西班牙語)，允許學生在課堂上以西班牙語溝通。
3. 過渡性的雙語教育(transitional bilingual education)：有些課的教師以學生的母語(如西班牙語)教學，甚至使用學生母語的教材，希望使學生盡快過渡到全使用英語上課。
4. 維持／發展式(maintenance/developmental)：在使學生發展英語能力的同時，也維持甚至加強學生的母語，目的要使學生精通兩種語言(母語與英語)，例如一班

拉丁美洲裔的學生，要學英語，也要學西班牙語。一般上西班牙語至少要上到小學六年級甚至到中學。一般原以為雙語教育會阻礙學生的英語學習，其實不然。用學生已了解的語言去學習、去閱讀，會比較容易。在知識內容的學習上也會有正面的遷移。

5. 雙元的語言(two-way/dual language)：在一個雙語教育的班級上，學生的組成刻意的安排一半是英語母語學生，另一半是非英語母語學生，目的在使這兩類學生精熟兩種語言。這種方式一般要至少持續 4-6 年。這種方式注重語言的學習，也注重學科內容與社會發展。上課時，兩種語言分開使用，有些課用英語上，有些課用學生母語上。例如在美國 North Philadelphia 州的一所雙語中學 Julia de Burgos 所採用的方式就是雙元語言方式(Freeman,2000)。在該校 30%的課是用西班牙文，在 70%是用英文上課(理想上原本希望各半)。所有英語能力不足的學生都被編到所謂的溝通小組(communication team)，在 1996-97 年有兩班七年級及兩班八年級。1997-98 有三班六年級及一班八年級。教師與學生大都也是波多黎哥人(Puerto Rican)，都會說西班牙語，學生在英語與西班牙語的程度參差不齊，也有以英語為母語的波多黎哥裔學生，也有非裔美籍學生。一班的學生有一半或一半以上是英語能力不足的學生。上課時，社會與溝通技巧是用英語上課，數學與科學則用西班牙語上。為使學生了解學科內容，常發現寫在黑板上的單字，常是兩種語言一起出現。在用英文上課的科目，教師也常用兩種語言以解釋生字或複雜的觀念。只要是在設法讓學生對學科內容有較好的理解，當英語非其母語的學生逐漸精熟英語時，西班牙語的出現逐漸的減少。但所有學生目標還在學好兩種語言。

Willig(1985)共收集 23 篇在 1971-1980 刊出之實驗研究作統合分析，這些研究是在檢定標題 7 雙語教育之效果。Willig(1985)將各篇研究根據主要依變項列出平均效應量，筆者再將這些平均依變項之平均效應量投入 t 考驗計算，共有 80 個效應量，得出總平均效應量為 0.34，標準誤為 0.0755657， $t_{(79)}=4.53$ ， $p<.0001$ ，顯出標題 7 的雙語教學達非常顯著的效果，表示如對非英語為母語的學生施以雙