

資訊管理之彙總研究方法

梁定澎

國立中山大學管理學院

洪新原

國立中山大學資訊管理學系

摘要

實證方法是資訊管理研究中常用的一種方法，它能由資料中歸納出結論，或以資料來驗證理論。一般實證研究常用初級方法直接分析所蒐集的資料，例如：個案、調查、與實驗等研究。但有些問題常無法由初級研究中直接獲致結論，或不同研究會有矛盾之結論，此時便需要用進一步的彙總方法加以分析。

彙總研究之主要目的是利用初級研究為樣本，解決各初級研究結果間的矛盾現象。在本文中，我們介紹彙總研究中的敘述研究法與彙總分析法的進行步驟、運用時注意事項，以及它們在資訊管理研究中的應用實例。這類研究的數量雖然不多，但是對於澄清混淆的觀念，以及早日建立有效結論則有相當的幫助。

關鍵字：彙總研究、實證研究、資訊管理研究

ABSTRACT

This paper introduces the procedures and applications of the descriptive methods and meta-analysis. Empirical studies are popular in information systems. It allows theories to be derived or verified from data. Empirical research can use primary methods, such as surveys, case analysis, and experiments, or meta-research that analyzes primary research findings. Meta-research helps clarify inconsistent findings in primary research and is valuable in further our knowledge in information systems.

Keywords: Meta-Research, Empirical Research, Information Systems Research.

一、導論

實證研究在資訊管理之研究中扮演相當重要的角色。過去許多著名的研究都是透過實證研究，蒐集真實資料，以找出具一般性的原則及理論。實證研究中常用的方法是直接自實驗室或田野中蒐集資料加以分析，我們通常稱之為初級研究(primary research)。常用的初級研究方法包括個案分析(case study)、調查(survey)、及實驗(experiment)等。

初級研究方法的優點是能夠掌握住第一手的資料進行分析。理論上而言，其結果的可靠度較高。但是，由於每個研究可能因為研究設計方式或執行過程中管理之不同，有時也使研究結果產生內部或外部效度不足的問題，進而造成針對同一問題的不同研究可能產生互相矛盾的結果，而無法判定何者為正確。造成初級研究產生誤差的原因很多，常見的包括：

- (1)抽樣的誤差：因樣本選擇不當或抽樣方法不當所造成的誤差。例如：研究主管使用資訊系統的態度，卻以沒有擔任主管經驗的學生來作實驗。
- (2)因變數或自變數的界定錯誤：由於因變數或自變數的界定直接影響研究效度，所以界定上的任何錯誤都會造成很大的誤差。例如：在衡量「使用者參與」時，只分析資訊小組中是否有使用者被編入，但卻未研究使用者是否有參與分析設計工作，或心理上的參與程度。
- (3)因變數或自變數的衡量不當：因變數或自變數在實證中需要作正確的衡量，否則可能會導致衡量的誤差。例如：誤把一些有程度上差別的連續變數，利用二分的類別尺度來衡量，或是在衡量使用者對於資訊系統的滿意度時，使用「滿意」與「不滿意」兩種選擇，而未使用李克特尺度(Likert's scale)來衡量。
- (4)因變數或自變數的範圍變異問題：這種誤差產生的原因通常是由於研究的設計不好，使得樣本中因變數或自變數資料的範圍變異，與實際情形有差距，進而降低所得

到結論的效度。例如：在從事影響資訊人員雇用因素的研究時，如果研究的設計不好，則很有可能造成衡量的結果都是從被雇用人員得來的，而未被雇用人員則自然而然被排除於樣本外，如此將會造成研究結果的不客觀。

- (5)因變數或自變數的建構效度不足：所謂建構效度(construct validity)指的是所選定的題目能夠充分衡量出這個變數的程度。產生這種誤差的原因，通常是由於未採用正確的衡量工具或衡量工具還不成熟，所以避免這種誤差產生的方法就是盡量採用發展成熟的衡量工具。例如：欲衡量「使用者滿意」此變數時，可以參考Doll等[8]所發展出來的衡量工具。
- (6)編碼不正確：在從事資料編碼時發生錯誤所造成的誤差。例如：誤將資料中1500萬的營業額之公司，由小型企業誤編為中型企業。
- (7)計算錯誤：計算時的疏忽所造成的誤差，藉由電腦軟體來協助或重複驗算是減少這種誤差的方法。
- (8)結果分析失當：在資料分析時使用了錯誤的分析方法所造成的誤差，這種誤差時常導因於未注意到所使用分析方法背後的假設。例如：未檢定資料的母體分配是否為常態分配，就使用迴歸分析方法來分析資料。
- (9)忽略重要外生變數所產生的變異：所謂外生變數(extraneous variable)指的一些未考慮到，但卻會影響因變數解釋能力的自變數[23, pp. 299]。產生這種誤差的例子，像是探討資訊系統績效評估時，受測樣本中包含有不同文化背景，然而研究模式中卻未考慮「文化差異」這個外生變數。
- (10)研究過程管理不當：實驗或調查過程中控制不夠嚴謹而造成的誤差。例如：調查主管對資訊系統的支持，但回卷者可能包括許多秘書，或是實驗過程未完全一致而影響結果。

上述誤差中任何幾種都可能造成初級研究結果的可靠度下降及彼此之間的矛盾現象。以往資訊管理研究中最著名的例子，就是針對認知型態(cognitive style)在資訊系統設計中所扮演角色的研究時所產生的矛盾。早期在資訊系統設計影響方面的研究許多都考慮到認知型態，並作為一項重要的變數，但是不同研究的結論對其效果的解釋則充滿矛盾的現象，無法正確找出真正結論。這種情形造成Huber[17]在1983年提出兩點關於認知型態研究的看法：(1)現有的認知型態研究成果，無法對於資訊系統設計提出實際的貢獻；(2)未來的類似研究也不會有實際的貢獻。自此以後，有關認知型態之探討大幅減少，幾乎不再有人使用。

另外一個例子是關於電子會議系統(electronic meeting systems, EMS)使用績效的研究，EMS相關研究在1990年代相當豐富，實證研究有數十篇，然而卻也面臨相同的矛盾問題。例如：有些研究結果[9,11]認為EMS的使用會顯著地提高會議生產力；然而也有研究[12]指出其效果並不顯著。

為了解決上述初級研究所可能產生的結果矛盾問題，從1970年代開始，研究者逐步發展並使用許多次級研究方法(secondary research methods)，稱為彙總研究法(meta-research)。彙總研究的精神很簡單：「既然個別的初級研究無法找出令人信服的結論，那麼把大量相關的初級研究放到一起，再進一步分析，似乎應該可以找到較可信賴的結果」。因此，它的做法就是使用敘述性方法或統計分析方法，來整合針對相同問題的許多研究結果，找出最可信賴的結論。因此，它在本質上有別於個案、調查、及實驗等這些常見的初級研究方法，初級研究是針對一個問題實地去蒐集原始資料來進行研究，而彙總研究則是利用這些初級研究所得到的結果來從事進一步的研究。

雖然彙總研究在教育、心理、及其他管理領域中的應用已經相當多，而且扮演的角色也越來越重要[19, pp.40]，然而在資訊管理研究領域中的運用卻還在起步階段，本文之目的即在介紹這種研究方法，並且探討它在資訊管理研究上的應用情形及可能面臨的問

題。在本文第二節中將概略說明彙總研究的演進及分類；第三節介紹彙總研究的進行步驟；第四節及第五節分別詳述敘述性研究與彙總分析之方法及應用實例；最後再討論這種方法的優缺點及面臨的問題。

二、彙總研究之演進與分類

知識累積的科學方法有兩個步驟，第一步是經由整合各別研究成果來導引出事實；第二步再將這些事實以有條理且有用的方式加以組織，進而產生出理論。彙總研究即著重在上述第一個步驟的工作。它自從1976年Glass[13]在學術期刊上正式發表此方法至今，隨著應用的方法越來越多(至今已有五種進行方法)，發展也越來越臻成熟。根據Guzzo等[15]的調查結果顯示，其在學術研究上的使用率以每年2倍的數量在增加，而且應用領域也從原先的心理、教育、與行為科學等領域，擴展到各個適用的領域。

由於彙總研究是根據初級研究結果再綜合整理，因此關鍵乃在於如何將初級資料加以整理，是直接把各結果列表比較，或是進一步做統計分析。Hunter and Schmidt[19]依分析方法之不同列出五種方法，分別是：

- (1)敘述法(narrative methods)：最直接的方法就是利用定性或敘述性的方法將各初級研究的重點加以摘要，再彙總比較初級研究的結果，並找出共通的結論。
- (2)投票法(voting methods)：就研究結論來看，可以將每個研究的結果歸入顯著正向、顯著負向、或不顯著等三類之一，然後以相同結論的研究個數最多的結果來作為總結論。
- (3)p值累計法(cumulation of p-values methods)：著重在研究結果的顯著性，因而將每個研究的顯著水準累計，而產生一個全體的顯著水準p值，並據此來歸納結論。
- (4)統計修正投票法(statistically correct vote-counting methods)：將前述投票法的結果再利用統計方法加以修正，一般又可分為產生統計顯著水準和提供平均效果規模等兩

種方法。

- (5)彙總分析法(meta-analysis methods)：利用統計方法，累計各研究的效果規模來找出結論，並從事相關的統計檢定。

上述五種方法各有利弊，其中以敘述法最為簡單，而彙總分析法最為嚴謹。因此，在實務上比較常被使用的是定性的敘述研究法與計量的彙總分析法。目前可蒐集到資訊管理的相關研究中主要也是利用這兩種方法，因此在本文中我們主要著重在這兩種方法之介紹。

三、彙總研究之進行步驟

前面談到彙總研究可以利用定性的敘述法或計量的彙總分析來進行，敘述研究法(narrative research)利用定性方式作歸納比較；而彙總分析(meta-analysis) [13]則利用統計分析，以達到整合初級研究之目的。不論是使用定性或計量的方式，其研究的進行都有幾個主要步驟，包括問題形成、資料蒐集、資料編碼、與分析解釋等四個階段。以下我們就每個階段的進行分別說明。

3.1 問題形成階段

由於彙總研究是針對初級研究的結果，進一步加以探討，因此適合彙總研究的問題，應該已經累積有相當的初級研究結果，而且各結果之間又存在有矛盾的現象，當然研究的問題也應該是大家所關心的主要課題，才有研究的價值。因為結果已有矛盾，而且造成矛盾的原因眾多，所以研究開始之初，需要參考相關的理論，針對問題界定研究範圍，提出合理的架構與假說，並且根據理論將研究變數之間的關係標明出來。其中需要明確地界定研究中的自變數與因變數，然後再將這些自變數與因變數的操作性定義，明確地界定出來，才能正確地設定合適的研究範圍。

3.2 資料蒐集階段

在研究範圍確定之後，就可以依據所界定之範圍來蒐集所需要的初級研究文獻。由

於初級研究之資料蒐集對於研究結果正確與否有相當的影響，因此資料之蒐集應該務求周詳。在蒐集相關文獻時，可以運用下列技巧[6,22]：

- (1)參考相關研究所引用的文獻，也就是透過幾篇探討此問題的研究中所列出的參考文獻，來找出更多的其他相關研究。
- (2)先找出一篇重要的文獻，再經由查詢文獻光碟資料庫，據此找出引用它的其他相關研究。
- (3)利用關鍵字(keywords)搜尋文獻光碟資料庫，找出所有相關的研究。
- (4)將目標鎖定在一些從事此方面研究的特定人員，蒐集他們的研究成果。
- (5)找出與研究問題相關的重要學術期刊，然後搜尋刊登在上面的文章。

這個階段最重要的工作就是必須確定得到的樣本要具有代表性。因此，有兩點需要注意的地方：第一是應當明確訂定文獻納入與排除於研究樣本中的判定準則，以避免有選擇性納入所造成的偏差。第二是由於結果不顯著的研究，比較不容易被公開發表，以致於所蒐集到的樣本常常是結果顯著的研究，所以為求樣本具有代表性，在蒐集文獻時應透過適當管道(如博士論文資料庫等)，取得一些未發表或較易被忽略的研究。

3.3 資料編碼階段

文獻蒐集完成後，我們必須根據文獻的內容，依照每篇文獻與研究問題的切題性，以及所提供資料的完整性來作篩選。篩選之後，再由合適的文獻中，將敘述性研究所需要的研究變數值，以及彙總分析所需要的統計量加以編碼。

這個階段的工作，有兩點需要注意的事項：第一是在從事篩選工作時，需要仔細審核每篇文獻對研究變數的操作性定義與衡量方式，確保不同研究之結果是在相同基礎上比較，以免造成把橘子和蘋果加在一起比較

的問題。第二是在從事編碼工作時，往往會受到編碼者主觀的判斷所影響，所以最好由兩位以上的人員同時進行，然後再比較討論結果，以提高研究的信度。

3.4 分析解釋階段

資料編碼完成後，接下來進行分析與解釋的工作。此階段由於敘述研究法是採用定性分析，而彙總分析法則採用計量分析，所以分析過程不太相同。敘述研究法是將編碼後所得的研究變數值，經過分類、排序、或加總等運算，然後應用簡單的推理法則來比較各研究之間的差異與綜合結果。

至於彙總分析則是將編碼後所得之統計量，經過公式計算決定來找出結論及可靠度。整個過程中的分析計算，信度與效度的考量要注意到四個構面：

- (1)變數間關係假設的強度：通常使用平均效果規模(average effect size)來衡量。平均效果規模表示研究結果拒絕虛無假設的整體強度，數值越大代表自變數與因變數之間的關係越強。
- (2)變數間關係假設的顯著性：利用統計顯著檢定(statistical test)來檢定自變數與因變數之間的整體關係是否夠顯著。
- (3)研究樣本間是否相似而可以比較：利用同質性測試(homogeneity test)來檢定研究樣本中各研究的同質性是否足夠，這個測試結果可以作為上述資料篩選工作優劣的一個判斷指標。同質性越高，彙總分析所得結論的可靠度也越高。
- (4)彙總結論的穩定性：好的彙總結論應該很穩定而不易被額外一兩個初級研究所推翻。穩定性的衡量常用失敗安全係數(fail-safe N)，它反應此彙總分析結論有多少個額外不顯著的研究(也就是效果規模為零的研究)加入，才能使得原有的平均效果規模降低到使原來的彙總結論不成立。所以這個係數可以反應上述資料蒐集工作中樣本代表性的問題。失敗安全係數通常越大越好。

經由上述分析計算，就可以找出彙總結論並據以作出解釋。四種統計量的詳細說明見附錄。

四、敘述研究法之應用

在資訊管理研究中常用的方法為敘述研究法及彙總分析法。敘述研究法的應用可追溯至1990年，本節中我們以四個研究實例加以說明。前面三個是針對EMS研究的矛盾現象，而最後一個則是期望經由整合DSS績效方面的研究，以達到建立DSS理論之目的。內容上，我們主要以Dennis等[7]來逐步說明，各研究的進行步驟與結果之摘要可參見表一。

在1990年時，Dennis等[7]發現EMS實驗室(laboratory experiments)與田野研究(field studies)結果之間存在矛盾現象。例如：EMS的使用是否會提升會議的效能、效率、與使用者滿意程度，在實驗室的研究中並沒有一致的結論；但是田野研究的結果卻顯示EMS的使用對於提升會議的效能、效率、與使用者滿意程度，有顯著的助益。所以他們的研究主題是：「在實驗室與田野研究中，組織環境(4個變數)、群組特性(7個變數)、工作特性(8個變數)、與EMS因素(5個變數)等變數，有哪些會影響EMS的效能、效率、與成員滿意度」。因此，在資料蒐集階段，研究者依據經驗(非系統性方法)蒐集了25個EMS實驗室研究與10個EMS田野研究作比較研究。

在資料編碼方面，他們將每個研究的組織環境、群組特性、工作特性、EMS環境、以及最後產出共24個變數值擷取出來。以Bui and Sivasankaran[5]之研究為例，它們是以學生為受測者，沒有獎勵措施，這些學生間有共同的目標，而且採行的工作中沒有相互關聯的問題。所以其組織環境中的組織文化、動機有無、是否有共同目標、問題是否相互關聯等四個變數之變數值被編碼為(學生、無、是、否)；而Vogel等[29]的研究則是以私人企業中的員工為受測者，有獎勵措施，部份員工間具有共同的目標，而且採行的工作中存在有相互關聯的問題。故其編碼為(私人企業、有、是/否、是)。

上述編碼結果以定性比較分析後，發現

並非它們之間對EMS研究的結果有衝突，而是由於這些實驗室研究與田野研究在7個自變數的控制上不同所造成。例如，實驗室研究都採用人數較少的群組，而且群組成員都比較沒有實際工作經驗；但是田野研究的群組規模則較大，而且群組成員都有實際工作經驗等這些差異造成其結論的不同。

整體而言，這個研究提供了對於EMS研究結果之間矛盾情形相當合理的解釋，然而其限制在於沒有採用比較有系統性的資料蒐集方法，以至於容易產生選擇性納入樣本的問題，進而可能影響研究結果的客觀性。

隨後同樣於1990年時，Pinsonneault and Kraemer[27]則認為不同研究可能因為所使用的實驗工具不同而有不同結果，因此利用實驗工具作自變數來解釋EMS實驗研究結果的矛盾現象。他們把EMS分為輔助決策導向，主要是支援決策過程的GDSS(group decision support systems)；以及輔助資訊導向，主要是支援溝通過程的GCSS(group communication support systems)。經由敘述分析法比較12個GDSS與19個GCSS的實驗結果，將這些研究的情境變數(包括個人因素、情勢因素、群體結構、科技支援、與工作特性等變數)、群組過程變數、與結果變數等分別展列出來。結果發現所使用實驗工具的不同，可以用來解釋EMS實驗研究結果之間的衝突，也就是說GDSS與GCSS對於群組過程變數與結果變數的影響不同。

最近，洪新原等[1]發現EMS相關實驗中所探討的實驗變數，除了EMS使用與否被前述學者探討過之外，其他的匿名、協助者角色、群組大小、領導者、與接近程度等變數，研究結果也諸多矛盾，但則尚未被探討；而且1993-1995年間在知名學術期刊上發表的EMS實驗研究篇數也非常多，而這些研究並未被納入以前EMS彙總研究的樣本中。所以他們應用敘述研究法將35篇EMS實驗研究所探討的情境變數、群組過程變數、與結果變數分別提列出來，並且比較其差異與形成原因。結果顯示匿名的效果多不顯著；協助者的效果多不顯著；領導者的效果多不顯著；而且大群組會議的決策時間較長、生產力較高、決策品質較好。

另外一個實例是Eierman等[10]應用敘述分析法來整合DSS的研究成果，以建立DSS理論。他們利用敘述研究法從15篇探討DSS績效的相關研究中，找出DSS研究的8個構面(包括環境、工作、施行策略、DSS能力、DSS結構、使用者、使用者行為、與績效等)，然後再探討這些構面之間的關係。研究結果顯示獲得顯著支持的關係包括(1)施行策略對於使用者行為與績效的影響；(2)環境因素對於績效的影響；與(3)DSS結構對於績效的影響。同時作者也發現8個構面排列組合所得到的56種關係中，有39種關係尚未被探討。最後他們指出有許多關係的相關研究間會產生不一致的結論，並認為可能的解釋是由於這些研究未考慮重要變數間的交互作用。

五、彙總分析法之應用

除了應用敘述研究法來從事資訊管理相關研究外，也有學者將彙總分析運用在資訊管理之相關研究上。在本節中介紹六個應用實例：第一個例子是彙總比較表格與圖形的展示方式對於決策支援績效的影響；隨後的三個例子是彙總GSS(或稱EMS)實驗室研究的結果；接著是彙總DSS實施成功的研究結果；最後則是彙總系統有用性(usability)的主觀與客觀衡量方式之研究結果。關於彙總分析法的進行過程說明，我們主要以McLeod[25]為例，這些研究的進行步驟與結果歸納於表二中。

在1990年時，Hwang and Wu[20]首先採用彙總分析法來解釋長期以來，表格與圖形等不同展示方式對於決策績效影響之研究結果間互相矛盾的現象。他們根據研究的切題性與實驗設計，選出12篇相關的論文。然後，由這些論文中的17個研究，擷取出計算平均規模效果與變異數所需要的相關統計量。最後他們發現在整體不分群的情況下，表格與圖形不同的展示方式對於決策績效的影響並不顯著。然而，將這些研究以其所選定的工作複雜度與工作型態分群後，則在採行工作是不太困難也不太簡單時，圖形的展示明顯地對於決策績效有較佳的影響。但是在採行工作是很困難或很簡單的情形下，展示方式是表格或是圖形，對於決策績效的影響則不顯著。

表一：四個資訊管理相關敘述性研究之進行步驟與結果

	問題形成	資料蒐集	資料編碼	分析解釋
Dennis, Nunamaker, and Vogel [7]	EMS效果研究在實驗室與田野研究中會有不同結果，可能導因於組織環境(4個變數)、群組特性(7個變數)、工作特性(8個變數)、與EMS因素(5個變數)等變數的不同，會影響EMS的效能、效率、與成員滿意度。	依據經驗蒐集25篇EMS實驗室研究與10篇EMS田野研究。	將每個研究的24(=4+7+8+5)個情境變數值與3個結果變數值抓取出來，並進行實驗室與田野研究之間的差異比較。	由於實驗室與田野研究在群組建立、群組過去與未來、群組大小、工作特性、資訊提供、EMS特性、與激勵因素等情境變數的不同，造成所得到的結論不同。
Pinsonneault and Kraemer [27]	EMS實驗結果之不同，可能由於GDSS與GCSS在特性與支援目的上之不同。	依據經驗蒐集12篇GDSS研究與19篇GCSS研究。	將每個研究中的5個環境變數值、6個自變數值、3個群組過程變數值、與5個結果變數值分別編碼，然後將GDSS與GCSS研究分別歸納結果，再綜合比較。	GDSS與GCSS研究在某些過程與結果變數上有相同的結論(例如：都會增加參與度以及決策品質等)，然而在其他變數方面則有不同的效果(例如：GDSS會減少決策時間，但GCSS會增加決策時間等)。
Eierman, Niederman, and Adams [10]	以DSS研究的8構面來彙總DSS績效研究之結果，消除矛盾發現，導出異同結論。	從1975年至1990年MIS Quarterly, Communications of the ACM, Management Science等六種學術期刊的200多篇文章中，選擇出合適的15篇研究作為樣本。	將這15個研究中有關8個構面之間關係(例如：DSS能力對於DSS績效、DSS能力對於使用者行為等關係)的研究結果分類整理出來，並且加以比較討論。	八個構面排列組合所得的56種關係中，只有17種關係被研究過。其中呈現顯著支持的關係包括：(1)施行策略對於使用者行為與績效的影響；(2)環境因素對於績效的影響；與(3)DSS結構對於績效的影響。
洪新原、黃錦祥、許孟祥、梁定澎 [1]	為了瞭解EMS實驗室研究成品的全貌，廣泛地分析既有初級研究中匿名、EMS使用、協助者、群組大小、領導、與接近程度等效果對於會議過程(包括參與平等性、3個變數)與結果(包括生產力等3個變數)的影響。	從1985年至1995年MIS Quarterly, Communications of the ACM, Management Science等六種學術期刊的35篇EMS實驗室研究作為樣本。	由兩位研究人員依照每篇研究的問題加以分類(例如：是屬於討論匿名與否對參與平等性的影響、或是屬於討論EMS使用對於生產力的影響等)，並且將問題的結論記錄下來，然後歸納比較。	匿名、領導、與接近程度的效果大多不顯著；大群組相較於小群組而言，有較高的生產力與決策品質。同時在研究中也整理出值得後續研究參考的研究方向。

表二：六個資訊管理相關彙總分析之進行步驟與結果

問 題 形 成	資 料 蒐 集	資 料 編 碼	分 析 解 釋
探討圖形或表格等資料展示形式對於決策績效的影響，以及工作型態與複雜度對於資料展示效果是否有影響。	經由對切題性與實驗設計的詳細挑選，選出12篇論文中的17個實驗室研究，作為研究樣本。	由每個研究中取得相關統計量，然後計算平均規模效果與變異數，以作為分析解釋的依據。	當這17個研究不分群時，表格與圖形的差異並不顯著；但是若依從事的工作複雜度與型態來分群，則在中度複雜的工作時，圖形顯著優於表格，但對於高度與低度複雜的工作，其間差異也不顯著。
探討GSS的使用對於生產力、參與平等性、決策時間、決策品質、與共識達成、與成員滿意度的全體影響效果。	從1980年至1990年相關期刊與研討會論文中蒐集了42篇研究，然後再檢查研究內容的適當性與切題性，最後以12篇為研究樣本。	由每個研究中取得相關統計量，然後計算平均規模效果與檢定統計顯著性，以作為分析解釋的依據。	GSS的使用可以增加決策品質、決策時間、參與平等性、與會議生產力，但是會減低共識程度與成員滿意度。
以大樣本來重複McLeod (1992)的研究。	由文獻庫、相關期刊與論文、諮詢相關研究人員、與博士論文等四個來源，蒐集到1970年至1992年共31篇相關研究。	二個研究人員分別各研究中取得相關統計量，計算出平均效果規模與抽樣誤差，並且檢定統計顯著性與其同質性，以作為分析解釋的依據。	GSS的使用對於決策品質、構想產生、以及參與平等性有正面效果，而對於決策時間、共識達成、與成員滿意則有負面效果。
探討EMS實驗室研究成果的全貌，廣泛地分析匿名、EMS使用、協助者、群組大小、領導、與接近程度等效果對於會議過程(包括參與平等性等3個變數)與結果(包括生產力等3個變數)的影響。	從1985年至1995年MIS Quarterly, Communications of the ACM, Management Science等六種學術期刊的35篇EMS實驗室研究作為樣本。	三個研究人員分別各研究中取得相關統計量，計算出平均效果規模與抽樣誤差，並且檢定統計顯著性與其同質性，最後計算各個主題的失敗安全係數，以作為分析解釋的依據。	(1)匿名可以提高會議的生產力，但是會降低與會者的滿意程度；(2)EMS的使用可以讓與會者的參與程度更加平等，會議生產力也會提高，但是需要較長的開會時間；(3)會議中協助者的角色越積極，則越不容易達成共識；(4)小組組會議中的與會者會有較高的滿意程度；(5)面對面會議相較於遠距會議，其與會者有較高的滿意程度。
探討DSS使用者認知型態(2個變數)、個人屬性(8個變數)、背景變數(3個變數)、與使用者情境變數(3個變數)等會影響DSS實施效果與使用者感覺的程度。	參考相關文獻，並且尋找電子文獻資料庫，取得1973年至1988年間共33篇相關研究。	三個研究人員分別各研究中取得相關統計量，並計算出平均效果規模、抽樣誤差、信賴區間、與失敗安全係數，來分析結果。	使用者的參與、訓練、以及經驗等使用者情境因素對於DSS實施成功的影響，相較於心理因素來得重要，而且高出三十個百分點。
探討系統有用性的兩種衡量工具，主觀地詢問使用者偏好與客觀地測量系統績效，兩者之間的關係。	由相關期刊與論文集中的1937篇相關研究中，再找出同時利用主觀與客觀衡量方式的57篇研究。	由於此類研究多沒有報導彙總分析所需要的統計量，所以他們發展出一套彙總比較程序，先將主觀與客觀衡量資料編碼，然後經過標準化，最後再比較分析。	主觀的詢問使用者偏好與客觀的測量系統績效，所得到的結果之間存在正相關的關係。

在1992年時，McLeod[25]採用彙總分析法來解釋關於群組支援系統(group support systems, GSS)之矛盾研究結果。他們的研究主題是要探討「GSS的使用對於生產力、參與平等性、決策時間、決策品質、共識達成、與成員滿意度等六個構面的全體影響效果」。針對此一主題，他們從1980年至1990年相關期刊與研討會論文中蒐集了42篇研究論文，然後再對研究內容的適當性與切題性從事篩選，最後以12個GSS實驗室研究為研究樣本，其中包括有7個探討工作生產力、5個探討參與平等性、3個探討決策時間、9個探討決策品質、4個探討共識程度、以及6個探討成員滿意度等6個主題的GSS研究。

在資料編碼階段，他們由每個研究中取得相關統計量，然後計算平均規模效果與檢定統計顯著性，作為分析解釋的依據。以探討GSS對於共識程度影響的4個研究之彙總分析為例：

- (1)Gallupe等[11]的實驗群組數為24、效果規模為-0.40、統計顯著性為-1.96。
- (2)George等[12]的實驗群組數為30、效果規模為-0.64、統計顯著性為-3.50。
- (3)Lewis[24]的實驗群組數為30、效果規模為-0.09、統計顯著性為-0.47。
- (4)Steeb and Johnson[28]的實驗群組數為82、效果規模為-0.18、統計顯著性為-1.65。

所以經由公式(詳見附錄)可以計算出，GSS的使用對於共識程度影響的平均效果規模為-0.35，而總合統計顯著性為-3.79($p=0.00007$)，表示GSS的使用會顯著地負面影響共識程度的達成。

最後，在分析解釋階段，經由上述的步驟來從事6個主題的彙總分析，結果顯示GSS的使用可以增加決策品質、參與平等性、與會議生產力，但是會減低共識程度與成員滿意度，同時也會拉長決策時間。此研究的結果讓我們對於GSS使用的整體效果有進一步的瞭解，不過其限制在於納入彙總分析的研究樣本過小，而且有一些重要的彙總分析統計量(如同質性檢定和失敗安全係數等)並沒有在研究中探討。

隨後到了1993年時，Benbasat and Lim[4]嘗試以更大的樣本來解釋GSS的有效性，他們蒐集了1970年至1992年間31篇GSS相關研究，將每個研究的工作特性、群組特性、環境因素、科技因素、與績效、滿意度、結構產出等變數編碼，經由彙總分析後發現GSS的使用對於決策品質、構想產生、以及參與平等性有正面效果，而對於決策時間、共識達成、與成員滿意則有負面效果。他們的結論與上述McLeod[25]的研究結果一致。

在1996年，Hung等[18]基於仍有許多課題未經彙總探討，因而把[1]中的結果以35篇的EMS實驗研究作進一步彙總分析。結果顯示：(1)匿名可以提高會議的生產力，但是會降低與會者的滿意程度；(2)EMS的使用可以讓與會者的參與程度更加平等，會議生產力也會提高，但是需要較長的開會時間；(3)會議中協助者的角色越積極，則越不容易達成共識；(4)小群組會議中的與會者會有較高的滿意程度；(5)面對面會議相較於遠距會議，其與會者有較高的滿意程度。

另外一個實例是Alavi and Joachimsthaler[2]在1992年應用彙總分析來探討影響DSS實施成功之因素。經由蒐集33篇相關研究，並且從中找出144個研究發現來作彙總分析，研究結果顯示：使用者的參與、訓練、以及經驗等使用者情境因素對於DSS實施成功的影響比使用者認知型態的影響來得重要，而且高出有三十個百分點。「使用者的認知型態對於DSS績效的整體影響效果很小」，此點結論也呼應了Huber[17]對於認知型態這個研究變數的觀點。

最後的一個應用實例是Nielsen and Levy[26]在1994年，以彙總分析來探討利用主觀與客觀方式衡量系統有用性之差別。以往探討系統有用性的衡量方式有兩種：第一種是測量受測者的主觀偏好(preference)；第二種是客觀地衡量使用者的績效(performance)。為了瞭解這兩種衡量方式結果的差異，他們利用57篇同時採用這兩種衡量方式的研究來從事彙總分析，結果發現主觀的偏好與客觀的績效之間存在著正向的關係。

根據Hwang[21]的研究指出，目前彙總分

析在資訊管理研究的應用上，存在著相當的潛力，其原因包括：(1)彙總分析的研究樣本數，開始時並不需要很多，而是透過相關基礎研究成果的累積，慢慢地擴大樣本數，並且重複地驗證研究發現。(2)由於彙總分析中效果規模(effect size)計算方式的多樣化，所以彙總分析的研究對象並不需要侷限在兩群體(例如：研究樣本中包括提供工具支援的一群，以及沒有工具支援的另一群)的基礎研究，也可以包含單一群體的另類基礎研究。(3)彙總分析提供了研究中介變數(mediator variable)，以達到進一步解釋研究結果的功能。不過目前這類應用相當缺乏，GSS的研究是一個相當可行的應用方向。

六、分析與討論

前述敘述研究法與彙總分析法在資訊管理研究中都有被應用到不同課題，而且各有優劣(見表三之比較)。敘述研究法是利用定性方式的歸納比較來進行彙總研究，它也是最早被使用的一種彙總研究方法。其優點在於容易使用，不需要複雜的訓練，而且當可供分析的初級研究總數不多時，不需要複雜的分析運算，可以很快地整理出結論。然而正因如此，有許多學者認為應用此方法所得到的結果過於主觀[3]，而且進行方法上較不嚴謹[22]，以致於研究的結果是否可以重複得到不易確定。另外，當初級研究的數量很多時，其間差異往往不容易直覺地比較出來[3,22]。最後的一個限制是可能太過依賴顯著水準，而忽略了不同研究的不同重要性[22]，也就是說，只注重結果顯著或不顯著，而不重視各個初級研究不同的顯著程度。

至於彙總分析則是集合個別研究的結果來從事統計分析，以達到整合研究發現之目的。這種方法的優點包括：(1)各研究之間的平均處理效果可以被估算出來；(2)平均處理效果的一致性可以被測試；(3)造成研究之間差異的解釋變數效果可以被測試；以及(4)解釋變數對於處理效果變異的顯著性可以被測試，以決定是否所有處理效果變異都被充分解釋。由於它具有充分的數理背景，所以彙總分析法的結果較客觀，而且研究的可重複性也較高[16, p.9]。

然而，彙總分析在研究進行上也存在著許多的陷阱，值得研究人員在應用時注意，這些限制包括[14]：(1)衡量工具、變數定義、與受測者差異太大的研究，不應該放在一起比較或下結論；(2)好的實驗設計與不好的實驗設計之研究結果，在彙總分析中無法被區別開來；(3)公開發表的文章多半是結果顯著的，而結果不顯著的研究則通常不容易被發表，因而會造成抽樣誤差；(4)由同一研究所得到的多重發現，由於這些發現間彼此並不獨立，所以往往會造成彙總分析結果的誤差，也就是會高估其信度。

表三 敘述研究法與彙總分析法之比較

	敘述研究法	彙總分析法
被提出的使用歷史	較早	較晚
需要複雜計算的程度	較低	較高
所依靠的決策型態	直覺	分析
背後的數理基礎	很弱	很強
對文獻權重的衡量	相等	不等
結果的客觀性	主觀	客觀
研究的可重複性	很低	很高
適用的時機	相關文獻數量不多或 其間關係不很複雜	相關文獻數量很多或 其間關係非常複雜

當然，要改善這些問題需要在研究過程中遵循一定的原則。Wolf[30]舉出以下13點原則，提供研究人員進行彙總分析時參考：

- (1)明確定義與說明特定研究樣本，應該被列入或排除的判斷準則。
- (2)搜尋未出版的研究，將其納入研究樣本，以降低「型一誤差」。
- (3)進行資料編碼時，儘可能從各研究中多抓取一些變數，以確保分析時能考慮到更多的層面。
- (4)除了著重主效果之外，也應注意主效果間的交互作用。
- (5)將結果的分布情形繪圖，以進一步檢驗異常值。
- (6)檢視資料編碼時的信度，一種改善的方法是由兩個以上的研究人員同時來進行資料編碼，並且針對不一致的部份，作特別處理。
- (7)結合統計顯著性的測試與規模效果的估計

，統計顯著性主要提供整體效果是否顯著的資訊，而規模效果則是顯示整體效果的大小。

- (8)同時計算經過調整與未經調整的統計量，以衡量樣本大小對於結果的影響；
- (9)進行有母數方法時，也應考量無母數方法的適用性。
- (10)有疑問時隨時參考彙總分析文獻，求得正確答案。
- (11)可以同時利用定性與計量的分析方法。
- (12)研究中必須詳細說明研究之限制，以及根據研究結果對後續研究提供建議；
- (13)切記一個觀念：「資料分析是幫助思考，而非取代思考」。

七、結論

在本文中，我們分別介紹了彙總研究方法的演進分類與進行程序，同時也列舉出一些資訊管理研究相關的應用實例，最後並且比較定性與計量兩種常用彙總研究方法的優缺點。希望經由這樣的介紹能夠讓國內資訊管理相關研究人員，對於這種研究方法能夠有更深一層的認識。最後，我們提供下面七點建議給研究人員在實際應用時作為參考：

首先，進行研究之初要判斷該課題運用彙總研究是否適當。運用彙總研究的時機在針對一個問題的初級研究已經相當多，而且其研究結果之間有矛盾的情形發生，因此並不是適用於所有問題。

其次，編碼工作是決定彙總研究是否有效的關鍵。為求提高編碼的信度，資料的編碼最好由兩個以上的人員分開來作，最後再比較討論編碼的結果。編碼過程更應以科學方法嚴謹的進行。

第三，資料分析有些涉及複雜計算，為求提高計算規模效果的信度，以及降低人力成本的花費，利用現有的彙總分析套裝軟體(例如：Hunter and Schmidt[19]提供的BASIC程式，或是Johnson[22]的DSTAT電腦軟體等)來取代人工計算，是一個很好的方法。

第四，為瞭解研究樣本間是否可以比較，應當避免把不同性質的研究，放在同一個

彙總研究中來比較討論，彙總分析中的同質性測試是一個可供參考的指標。

第五，由於實際上很多結果不顯著的研究，都不容易公開發表，所以在彙總研究中，要注意彙總結論被推翻的可能性，計算運用失敗安全係數是一個很好的驗證指標。

第六，在從事彙總分析時，可以多加考慮除了研究重點的自變數與因變數之外，是否存在第三者的中介變數(mediating effects)會對於研究結果產生影響。這種分析可以經由中介變數值的不同來對研究加以分群，然後比較其間的差異，如此可以得到更深入的結果。例如：除了將所有25個研究一起彙總分析之外，我們還可以依照「受測者性別」這個中介變數，來將研究分為13個受測者為男性與12個受測者為女性的兩群研究，然後分別對於這兩群來進行分析，如此可以更加瞭解造成結論不同的原因。

最後，誠如Guzzo等[15]所言，一個進行嚴謹的敘述性研究，其解釋能力與研究成果並不比利用複雜統計技術進行的其他方法來得差。所以，定性的敘述研究法應與計量的彙總分析法相輔相成，在不同的研究場合選擇適當的方法來應用。

參考文獻

- 1.洪新原、黃錦祥、許孟祥、梁定澎，「電子會議系統相關實驗之比較研究」，第七屆國際資訊管理學術研討會論文集，民國85年，856-867頁
- 2.Alavi, M., and Joachimsthaler, E.A., "Revisiting DSS Research: A Meta-Analysis of the Literature and Suggestions for Researchers," MIS Quarterly, Vol.16, No.1, March 1992, pp.95-116.
- 3.Aldag, R., and Stearns, T.M., "Issues in Research Methodology," Journal of Management, Vol.14, No.2, 1988, pp.253-276.
- 4.Benbasat, I., and Lim, L.H., "The Effects of Group, Task, Context, and Technology Variables on the Usefulness of Group Support Systems," Small Group Research, Vol.24, No.4, November 1993, pp.430-462.

5. Bui, T., and Sivasankaran, T.R., "Relation between GDSS Use and Group Task Complexity, Proceedings of the Twenty-Third Hawaii International Conference on System Sciences, 1990, pp.69-78.
6. Cooper, H.M., The Integrative Research Review: A Systematic Approach, 1984, Sage Publications.
7. Dennis, A.R., Nunamaker, J.F. Jr., and Vogel, D.R., "A Comparison of Laboratory and Field Research in the Study of Electronic Meeting Systems," Journal of Management Information Systems, Vol.7, No.3, Winter 1990-91, pp.107-135.
8. Doll, W.J., Xia, W., and Torkzadeh, G., "Confirmatory Factor Analysis of the End-User Computing Satisfaction Instrument," MIS Quarterly, Vol.18, No.4, December 1994, pp.453-461.
9. Easton, A.C., Vogel, D.R., and Nunamaker J. F. Jr., "Interactive Versus Stand-Alone Group Decision Support Systems for Stakeholder Identification and Assumption Surfacing in Small Group," Decision Support Systems, Vol.8, No.2, 1992, pp.159-168.
10. Eierman, M.A., Niederman, F., and Adams, C., "DSS Theory: A Model of Constructs and Relationships," Decision Support Systems, Vol.14, No.1, 1995, pp.1-26.
11. Gallupe R.B., DeSanctis G., and Dickson, G.W., "Computer-Based Support for Group Problem-Finding: An Experimental Investigation," MIS Quarterly, Vol.12, No.2, June 1988, pp.277-296.
12. George, J. F., Easton, G. K., Nunamaker, J. F. Jr., and Northcraft, G. B., "A Study of Collaborative Group Work With and Without Computer-Based Support," Information System Research, Vol.1, No.4, December 1990, pp.394-415.
13. Glass, G., "Primary, Secondary, and Meta-Analysis of Research," Educational Researcher, Vol.5, No.10, 1976, pp.3-8.
14. Glass, G., McGaw, B., and Smith, M.L., Meta-Analysis in Social Research, 1981, Sage Publications.
15. Guzzo, R.A., Jackson, S.E., and Katzell, R.A., "Meta-analysis analysis," in Cummings, L.L., and Staw, B.M. (eds.) Research in Organizational Behavior (Vol.9), 1986, Greenwich Press.
16. Hedges, L.V., and Olkin, I., Statistical Methods for Meta-Analysis, 1985, Academic Press.
17. Huber, G.P., "Cognitive Style as a Basis for MIS and DSS Designs: Much Ado about Nothing?," Management Science, Vol.29, No.5, May 1983, pp.567-579.
18. Hung, S.Y., Huang, J.S., Hsu, M.H., and Liang, T.P., "Revisiting the Effect of Electronic Meeting Systems: A Meta-Analysis of Literature," Proceedings of The Third Pacific Asia Conference on Information Systems, Australia, April 1997.
19. Hunter, J.E., and Schmidt, F.L., Methods of Meta-Analysis: Correcting Error and Bias in Research Findings, 1990, Sage Publications.
20. Hwang, M.I., and Wu, J.P., "The Effectiveness of Computer Graphics for Decision Support: A Meta-Analytical Integration of Research Findings," The DATA BASE for Advances in Information Systems, Vol.21, No.2&3, Fall 1990, pp.11-20.
21. Hwang, M.I., "The Use of Meta-Analysis in MIS Research: Promises and Problems," The DATA BASE for Advances in Information Systems, Vol.27, No.3, Summer 1996, pp.35-48.
22. Johnson, B.T., DSTAT: Software for the Meta-Analytic Review of Research Literatures, 1989, Lawrence Erlbaum Associates.
23. Kerlinger, F.N., Foundation of Behavioral Research, 1986, Holt Rinehart Winston.
24. Lewis, L.F., "A Decision Support System for Face-to-Face Groups," Journal of Information Science Principles and Practice, Vol.13, No.4, 1987, pp.211-219.
25. McLeod, P.L., "An Assessment of the Experimental Literature on Electronic Support of Group Work: Results of a Meta-Analysis," Human-Computer Interaction, Vol.7, No.3, 1992, pp.257-280.
26. Nielsen, J., and Levy, J., "Measuring Usability: Preference vs. Performance," Communi-

cations of the ACM, Vol.37, No.4, April 1994, pp.66-75.

27. Pinsonneault, A, and Kraemer, K.L., "The Effects of Electronic Meetings on Group Processes and Outcomes: An Assessment of the Empirical Research," European Journal of Operational Research, Vol.46, No.2, 1990, pp.143-161.
28. Steeb, R., and Johnson, S.C., "A Computer-Based Interactive System for Group Decision Making," IEEE Transactions on Systems, Man, and Cybernetics, Vol.11, No.8, 1981, pp.544-552.
29. Vogel, D.R., Martz, W.B., Nunamaker, J.F. Jr., Grohowski, R.B., and McGoff, C., "Electronic Meeting System Experience at IBM," Journal of Management Information Systems, Vol.6, No.3, Winter 1990, pp.25-43.
30. Wolf, F.M., Meta-Analysis: Quantitative Methods for Research Synthesis, 1986, Sage Publications.

附錄：彙總分析中常用統計量之介紹

在進行彙總分析時，我們經常會利用幾個統計量，來作為分析判斷的依據。以下我們將介紹這些常用到的統計量[30]：

(1) 平均效果規模(average effect size)

表示虛無假設（亦即自變數與因變數之間是無關的假設）是錯誤的程度，如果一個研究的結論是接受虛無假設，則表示其效果規模為0。根據Cohen[30, pp.27]建議的判斷法則，0.2算是弱效果，0.5算是中等效果，而0.8則可算是強效果。其計算公式如下：

$$d_u = \frac{\sum_i w_i d_i}{\sum_i w_i}$$

其中： d_i 是各研究的規模效果，

$$d_i = \frac{\bar{X}_1 - \bar{X}_2}{SD}$$

w_i 是各研究的權重， $w_i = \frac{2N}{8 + d_i^2}$ 。

(2) 統計顯著測試(statistical test)

對於一個假設，提供一個測試全體效果顯著性的指標，針對研究結果中報導不同的統計量（包括： p 值、 t 值、與 Z 值），Fisher、Winer、與Stouffer等人提出不同的計算方式，其公式分別如下：

$$\text{Fisher's: } \chi^2 = -2 \sum_i \log_e(P_i)$$

$$\text{Winer's: } Z_c = \frac{\sum_i (t_i)}{\sqrt{\sum_i (df_i / df_i - 2)}}$$

$$\text{Stouffer's: } Z_c = \frac{\sum_i (Z_i)}{\sqrt{N}}$$

其中： df_i 是各研究的自由度； N 是研究總數。

(3) 同質性測試(test of homogeneity)

它的功用在於檢定一些個別的研究結果，是否適合放在同一個彙總研究來從事分析。它所分析的對象又可分為上述的兩種統計量，其計算公式分別列出如下：

檢驗效果規模的同質性：

$$\chi^2 = \sum_i (w_i (d_i - d_u)^2)$$

檢驗統計性測試的同質性：

$$\chi^2 = \sum_i (Z_i - \bar{Z})^2$$

其中： $\bar{Z}$ 是各研究的平均 Z 值，

$$\bar{Z} = \frac{\sum_i Z_i}{N}$$

(4) 失敗安全係數(fail-safe N)

由於可公開發表的研究，其結論多半是比較顯著的，而許多研究結論不顯著的，往往比較不容易發表，所以此係數的功用即在於反應此彙總分析結果被推翻的可能性。亦即此係數在反應還有多少個效果不顯著($d_i = 0$)的研究加入，才能使得平均規模效果(d_u)低到某個預定的極小值(d_c)。

$$N_{fs} = \frac{K(d_u - d_c)}{d_c}$$

其中：K為彙總分析中的研究個數。