

國立臺灣師範大學教育心理與輔導學系
教育心理學報，民 96，38 卷，3 期，271 — 290 頁

類比與證偽思考技能之學習

吳庭瑜

明新科技大學
工程管理研究所

黃麗分

靜宜大學
觀光事業學系

洪瑞雲

交通大學
工業工程與管理學系

黃永昌

本研究在探討以類比的方式學習證偽的法則測試邏輯。實驗中 227 個大專學生被隨機分派至三種教導情境（證偽解題實例、解題實例 + 邏輯結構圖示、及控制組）。教導情境以四個練習題來操弄。法則測試作業的測試題含 16 個有具體經驗內容的法則規定敘述及 2 個抽象法則。具體法則測試題中，一半的法則規定語意結構與練習題相同，另一半則不同。此外，具體法則測試題中，半數的法則中至少有一變項名稱與練習題相同（表面相似）；另一半法則則否（無表面相似度）。結果發現，當法則規定的測試題與練習題間為相同語意結構時，證偽之解題實例確可以顯著提高證偽該法則的測試行為，且當無表面相似性的干擾時，不論是否附有問題之邏輯結構圖示，證偽之先前實例均皆可以提高不同語意結構測試題的證偽次數。由此推論，邏輯結構之掌握與排除表面相似性是以類比進行科學思考的要件。

關鍵詞：法則規定的測試、類比學習、證偽思考技能之獲得

類比是人們有意、無意的套用過去的經驗來解決目前遭遇到的新問題的一種歷程。結構相似的問題解法相似，若人可以正確的掌握當前問題的結構訊息，並以此結構訊息為線索去提取記憶中與此問題結構相似的先前經驗，就有機會套用先前經驗中的解題實例來解當前的問題，甚至創新或有新的發現。類比似乎是人類學習與思考的一種機制。學習可以遷移至新的情境，所依賴的即是學習者會將在某一特殊情境中學到的經驗推廣到與此學習相似的情境中，如，在認識「黃」檸檬後，看到與其外型相似但顏色是「綠」色的果子時，也會認為它是檸檬，同時還會認為它是同一種水果，味道是酸的、含豐富的維他命 C...等。一個老師教導學生「求學就如逆水行舟」時，目的即是希望學生能由「逆水行舟，不進則退」的知識或經驗來理解「求學也是不進則退」。又如 Bohr 在教學時將原子類比成微型的太陽系，又用「撞球」的類比來說明量子力學的不確定性，如，「撞球般的荷電物體依機械與電磁理論的法則互動」（引自李怡嚴，民 89）。類比因此是讓人可以聞一知十，或舉一反三的一種推理機制。

科學知識之所以可靠是來自一套嚴謹的科學思考方法，此思考方法的獲得是科學教育的核心工作。本研究的主要目的即在探討科學方法的核心推理特徵之一，法則測試作業中所需的證偽思考技能，是否可以透過類比的歷程來獲得，以及由此類比學習獲得的證偽思考技能是否可以遷移至抽象的問題中。

一、科學思考與證真、證偽的思考技能

歸納推理是人根據部份或有限的經驗去推論出其中的一般性法則，因此可以超越個人有限的經驗而產生新的知識。舉例來說，若我見過的金屬，如金、銀、鐵等在加熱之後體積會膨脹，則我可以推論所有的金屬遇熱會膨脹。此知識可以應用在預測我不曾加熱試驗過的金屬上。亦即，我由有限的經驗中歸納出一個新的、可以廣泛應用的一般性法則。Francis Bacon (1960) 因此認為歸納法是一種科學的方法，透過它我們才能超越經驗去創造出新的知識來。

但問題是，由有限的經驗歸納出來的法則並非必然無誤。基本上，「若我碰到的蚊子都會叮人，則所有的蚊子都會叮人」的推論在邏輯上是不周延的。此由歸納法所得的法則因此只是個假設，必須經過嚴格的測試、檢定才能被視為知識。Popper (1959) 即主張，由歸納形成的法則要以證偽 (falsification) 的方式加以檢定。亦即，若假設 $P \rightarrow Q$ (如：金屬遇熱會膨脹) 的法則為真時，此法則的意義為若 P 出現， Q 必然隨之出現，因此 P 出現為 Q 出現之充分條件；而 Q 出現為 P 出現之必要條件，亦即，若 Q 不出現， P 即不可能出現 ($\sim Q \rightarrow \sim P$)。以實驗來驗證此法則時，我們可由操弄 P (如：將銅加熱)，以觀察 Q 的預測 (體積膨脹) 是否出現。但當實驗顯示銅遇熱會膨脹時，由於銅只是金屬的一類，我們無法由此支持性證據來推論 $P \rightarrow Q$ 一定為真。但若實驗資料顯示銅遇熱不會膨脹 (已知「 $P \rightarrow \sim Q$ 」)，由於邏輯上 $P \rightarrow Q$ 意涵著 $\sim Q \rightarrow \sim P$ ，此觀察到的證據 ($\sim Q \cap P$) 為法則的負向案例，根據此負向案例便足以質疑「 $P \rightarrow Q$ 」法則的有效性 (Johnson-Laird, Girotto, & Legrenzi, 2004)。科學思考的核心特徵是質疑，藉由證偽邏輯的應用，科學家才能將不適當的法則淘汰或修正。科學家的工作之一即是不斷的嘗試去檢驗現有理論、法則是否為偽的活動。證偽思考因此是科學思考的重要技能之一。

但以證偽的方式去檢驗自己的想法是否容易呢？答案是否定的。自 60 年代起 Wason (1960, 1966) 就發現一般大學生對十分簡單的法則 (如： $P \rightarrow Q$) 進行驗證時，使用證偽策略 ($\sim Q \rightarrow \sim P$) 的人偏低；相反的，卻呈現出所謂的證真偏好 (confirmation bias)。以他所設計的翻卡片作業 (selection task) 為例，他將「如果卡片的一面是母音 (P) 另一面則為偶數 (Q)」這個法則中， P 和 Q 的四可能性 (P 、 $\sim P$ 、 Q 、 $\sim Q$) 以四張卡片呈現 (如：E K 4 7)，並詢問大學生若要驗證此法則是否為真要檢查那幾張卡片。研究發現，大多數人選擇同時測試 P 和 Q ，其次是只測試 P ，正確的選擇 P 及 $\sim Q$ 兩張卡片者通常少於 10%。由實驗中獲得的語文資料顯示，受試者測試時主要的目的在證明其假設是對的，因此稱為證真偏好。亦即，人在測試一個想法時，證真的傾向遠大於證偽的傾向。Oaksford 和 Chater (1994) 的後設分析則指出典型的卡片選擇是 $P (89\%) > Q (62\%) > \sim Q (25\%) > \sim P (16\%)$ ，選擇測試 $\sim Q$ 的人嚴重偏低，顯示一般人很不容易由證偽的角度去思考一個法則。

以證真的方式來檢視一個法則除了邏輯上的缺失外，亦可能使錯誤的、不周延的法則被錯誤的保留下來，以致延遲了真相被發現的時機，影響學術的發展，甚至造成一些錯誤的決策。人為何會有證真偏好呢？Johnson-Laird 與 Wason (1970) 認為受試者在翻卡片作業上的行為顯示他們對問題當中的邏輯結構無法完全掌握。Klayman 與 Ha (1987) 的分析則指出證真測試策略顯示的是一種正向測試 (positive test, 指測試條件句中的 P 或 Q)。人偏向使用正向測試方式的理由，除了不擅於負向測試方式 (測試 $\sim P$ 或 $\sim Q$) 外，是因為正向測試在許多情形下也有機會推翻錯誤的法則。例如，在 Wason 的 2-4-6 作業中，受試者的工作是要找出產生 2、4、6 三個案例的法則。若受試者的假設為「以 2 往上遞增」而實際的法則是「偶數以 2 的等距往上遞增」時，即使他使用正向測試，以符合法則的正向案 (例如：1、3、5) 測試，他就有機會可得到否定的回饋而被迫推翻原先的法則。亦即，當受試者的法則涵蓋面較真實法則大、或重疊、或無關時，使用正向測試即有機會證偽一個法則。證真的法則測試因此並不完全是錯誤的。但是，若受試者所形成的法則是包含在真實法則之內時，如法

則為「偶數，等差 2，上升」，而真實法則是「任何上升的數字」，以符合法則的正向案例（如：2, 4, 6）則只會得到正向的回饋，無助於真實法則的發現。在這種情形下，只有使用證偽的負向測試方能產生與法則不符的負向案例（如：2, 6, 10）而得以排除多餘的屬性。「偶數，等差 2，上升」此過窄的法則與真實法則「任何上升的數字」間的差異對非以追求「真相」為職志的人而言也許差不多，但對追求真知的科學工作而言，則是偏離真相。

Gorman (1995) 由發明電話的個案分析結果提出，證真偏好是一種支持研究者在證據不充分前，繼續前進的機制。人在從事原創性的研究、發明工作時，在觀念萌芽的最初階段若無證真式的偏向，則科學家很可能在少數幾次測試卻得不到支持性回饋之後就放棄原先的構想。Gorman 指出，使用證偽的思考方式的較好時機是在有了大量的支持性資料之後，研究者在接受其想法之正確性之前，再嘗試以負向案例來驗證其法則之正確性；如此，則有機會擴展其原先法則不足之處，而得到更好的替代方案。證偽式思考對科學發現的重要性由此可見。

然而，法則測試時的策略運用不一定全然在個人的意識的掌控之內。一般人往往不清楚自己思考運作的方式，也不易掌握、或控制自己的思考於理想的狀態下。例如，法則中的給定條件往往會限制了人們的認知過程，而將問題中的語意內容（如： $P \rightarrow Q$ 中的 P 及 Q ）當作最直接的認知線索，據以激發、搜尋記憶中的相關知識，題目中未提到的知識經驗（如： $P \rightarrow Q$ 中的 $\sim P$ 及 $\sim Q$ ）通常很難成為當事人考慮的對象 (Evans, 1989; Johnson-Laird & Byrne, 2002)。因此證真偏好（測試 P 與 Q ）也反映出一個人在資訊處理上選擇性登錄 (selective encoding) 及資訊處理不適當的現象。

此外，在重複處理同類事物的過程中，人會發展出一些實用的推理架構 (pragmatic reasoning schema) 以迴避複雜的邏輯分析。所謂實用的推理架構是指人從實際生活情境中建立起來的一套處理某些類似問題的通用知識架構。若問題與當事者先前經驗過的問題情境相似，則他就可以用實用推理架構推知答案，例如在驗證「如果喝酒，則年齡必大於 20 歲」的法則是否為真時，一個人若知道法律上「喝酒」與年齡間的關係是一種「允許」的知識架構 (permission schema)：「若要做某事 (P ，喝酒) 必須先具有某條件 (Q ，20 歲)」，以及反過來：「若某條件不存在 ($\sim Q$ ，小於 20 歲)，則不能做某事 ($\sim P$ ，不准喝酒)」，則可以比較輕易的提出「如果抽查年齡小於 20 歲的人而發現他在喝酒 ($\sim Q \cap P$) 就是違法的證據」來證偽此法則。反之，若問題內容是以抽象的型式表達（如：若 A 則 B ），或不是當事人所熟悉、或瞭解的「權利」或「義務」關係法則，則人因缺乏此通用的知識架構的輔助，證偽就很難了 (Cheng & Holyoak, 1985)。如研究發現，以英國受試者測試「若信封已封口，則反面一定貼有 5 元郵票」時，雖然此法則的內容與英國郵政規定不同，但英國郵政確有信封封口與否郵資不同的相關規定，英國的受試者因此有「信封是否可以允許封口和付的郵資高低有關」的「允許法則」的推理架構，因此證偽率很高；但因美國的郵政無此相關規定，美國受試者測試此法則時即出現較高的證真偏向 (Johnson-Laird, Legrenzi, & Legrenzi, 1972)。因此，證真或證偽的法則測試策略的使用往往會因問題的內容、呈現方式、受試者相關知識的多寡等因素而異。

二、提高證偽思考技能的研究

由於證真偏好，人往往不易質疑自己的思考內容，因此如何增進一般人使用證偽式思考便是一個重要的課題。研究指出，以教導的方式增進大學生使用證偽式思考的嘗試大多無效，例如，Mynatt, Doherty 與 Tweney (1977) 用指導語要求受試者使用證真策略、證偽策略、或控制組，去進行一類似 2-4-6 作業的電腦遊戲，並找出其中的法則，結果發現證偽指導語並無法扭轉受試者的證真偏好。

比較成功的教導例子來自 Tweney, Doherty, Warner 與 Pliske (1980) 的研究，在他們的實驗四中他們告訴受試者實驗者心中有二個法則，其中一個法則產生 DAX 的數字，另一個法則產生 MED 的數字，而 2-4-6 則是 DAX 的一個案例。因此受試者用新案例測試時，得到的回饋是此為 DAX 的案

例、或此為 MED 的案例。結果發現，在此方式之下雖然證偽偏好仍在，卻顯著的提高了受試者找出正確法則的機率，顯示一般人偏好測試與法則相符的案例可能是因為他們並不知道一個法則必然存有「對立假設」；為測試為負向的案例設計一個法則名稱（如：MED），可以幫助受試者意識到對立假設的存在。由此推論，透過作業的設計讓受試者知道每組案例後都有一個以上的法則，也許是提高證偽行為的一種方式。

Kareev 與 Halberstadt (1993) 的研究則是讓受試者評估別人在導出法則時所產生的案例。由於受試者已經知道真實法則是什麼，他們便能判斷那些能推翻不適當法則的案例（即負向案例）才是有用的測試，因此給它們最高的評價。Kareev 與 Avrahami (1995) 也發現，當讓受試者扮演教師的角色教人進行法則導出的作業時，由於受試者知道真實法則是什麼，若發現他的學生所測試的法則是包含於真實法則內時，受試者也會用證偽的方式來教導學生，以推翻其正在測試的法則。這些結果顯示，當人知道目前的法則與真實法則不同時，或知道他所觀察到的案例可以被兩個以上的法則加以解釋時，就會知道證偽測試的必要性。

然而，在實際的法則發現作業中，人是無法知道真正的法則的。上述現象只顯示，面對一組資料時，要同時設想有兩個以上的法則才可能促使人去使用證偽的測試方式。但是，Adsit 與 London (1997) 的研究也指出，在受試者不知道真正法則是什麼的 2-4-6 作業中，先給受試者一段時間去產生可能的法則，或是提供不同的法則給受試者參考，對證偽行為都未產生作用。唯一的例外是，當受試者已知供其參考的法則中含有真正的法則時，才可以顯著地提昇他們使用證偽策略的比率，並發現較多正確的法則。由此推論，人對證偽的邏輯雖並非完全無知，但只有當作業環境中含有強有力的線索時，如，已知正確的法則、已知參考法則中含有真正的法則、或以詳細的實例說明必須分別測試法則的充分與必要條件 (Ahn & Graham, 1999) 等等，才能激發其使用。

三、類比與證偽思考技能之學習

人在法則測試作業上不易證偽，且證偽思考技能很難教導，但卻比較容易透過作業環境的安排而激發其使用，這些現象顯示法則測試策略或證偽思考技能是屬於一種程序性的知識 (procedural knowledge)。相對於語意性知識 (semantic knowledge)，程序性知識不易完全靠一般的語文教學即可習得，且一般人對此種程序性知識的學習歷程通常無法十分清楚、亦很難以語文的方式去清楚的表達它。那麼，究竟證偽思考技能要以何種方式方可獲得呢？

研究指出，一般複雜的程序性認知技能，單以語文的教導方式來學習通常不足，但可以觀察或「做中學」的實作的方式，讓受試者從反覆處理同類的案例中導出其運作的模式。法則測試作業牽涉到複雜的正向測試與負向測試的程序性知識，既然透過邏輯的教學去學習的成效不理想 (Cheng, Holyoak, Nisbett, & Oliver, 1986)，那麼以觀察實例的實作方式來學習是否會有較好的效果呢？

類比推理是在實作過程中運用過去的經驗以達成觸類旁通的一種學習機制。這種在實作過程中將先前經驗類推以解決現在問題的能力似乎是人天生的一種能力，亦是人類學習的一種重要方式 (江淑卿，民 90；黃幸美，民 90；Goswami, 1992)。本研究的目的即在探討以類比的方式教導證偽的法則測試技能的效果及其限制。

類比是當前問題中的線索激發長期記憶中的相關訊息以提取相關的解題程序的歷程。類比推理成功的要件是前後問題間的結構對應關係，結構相同的問題即使表面特徵不同，仍可使用相同的解題方法 (如，Gentner, 1983；Reeves & Weisberg, 1994；Ross, 1987, 1989；Bowdle & Gentner, 2005)。如，麻雀和雞的外型雖不盡相同，但生理結構卻有共同點，因此當遇到生病的麻雀時，我們可以利用已知的治療雞的疾病的方式幫麻雀治療。又如，看到「已知等腰三角形 ABC 的三邊長，計算其面積」，若知道等腰三角形可分解成兩個相等的直角三角形時，便可套用畢氏定律求出三角形 ABC 的高，即可以

計算其面積了。然而，一般人在面對此類問題時，往往覺得困難，困難的原因除了缺乏相關知識外，更常見的理由是不知道測試題 (target problem) 跟記憶中的先前問題 (source problem) 是同構但不同形 (isomorphic) 的問題。因此，成功的類比往往要藉助測試題與先前問題間的表面相似性來作為回憶的線索，例如，看到一個等腰直角三角形時，一般學生很容易的就可以認出它是直角三角形，就可以套用直角三角形的知識來解題。但缺乏表面相似的線索人就不易產生類比，例如：計算連續數列（如 1 至 100）的和的算數問題可直接套用梯形面積的計算公式【(上底+下底) X 高 ÷ 2】而迅速的算出解答，但是「1 至 100 的連續連續數列的和」和「梯形面積」外觀不同因此就不易產生類比。然而，問題間表面相似並不能保障其內部結構相似。例如，機械錶和石英錶雖然錶面看來相同，但其實內部的機械構造不同，此時，套用機械錶的知識來修理故障的石英錶是不會成功的。

事物的表面訊息似乎是人最初級的記憶提取線索。研究即指出，人往往會以目前的問題的表面特徵做為線索去提取記憶中的先前例題。表面特徵相似的問題其結構不一定相似，若提取到表面相似但結構不同的先前例題，套用此結構不同的先前例題的解題程序，便成了類比錯誤的主要來源 (王亞同, 民 91; 陳茹玲、蘇宜芬, 民 94; Ross, 1987; Reeves & Weisberg, 1994; Gentner, Rattermann, & Forbus, 1993)。研究也指出，只有當受試者發現表面線索所激發的先前例題不恰當時，問題的結構訊息才會被注意到 (Gentner, 1989)。例如，Wharton 等人 (1994) 以偶發學習的方式來探討類比中的提取線索的問題。他們先讓受試者閱讀一串句子後，另給受試者一系列的「主詞 - 動詞 - 受詞」的簡單句子做為提取線索，要受試者儘可能的去回憶出先前讀過的句子。那些做為回憶線索的句子與先前讀過的句子間的結構 (句型) 或表面 (主題) 關係分成：一致、不一致、或無關三種。研究結果指出，當做為回憶線索的句子與先前讀過的句子間沒有不一致的線索時，影響回憶的主要因素是表面相似性。但當提取線索與先前讀過的句子間有不一致的地方時，結構相似性才可能成為提取線索。

在類比歷程中受試者往往會因表面相似性而造成類比錯誤。這現象與 Wason 的選卡片作業中受試者只偏向選擇測試 P 與 Q 等法則中提到的選項而未能選法則中未提到的 $\sim P$ 或 $\sim Q$ 的情形十分類似。由於邏輯推理是一種形式思考，理論上不應受法則中變項的表面特徵所影響。因此，證真偏好中選擇 P、Q 多過於 $\sim P$ 、 $\sim Q$ 這個現象反映的極可能是受試者以表面特徵 (而非邏輯結構) 來提取解答的現象。另一方面，在熟悉的作業中受試者也可能可以由表面特徵直接提取可證偽的負向案例 (如：未成年的小孩不能喝酒)，或作業的設計可以讓受試者掌握到目前問題中的結構時 (如：雙類別回饋)，證偽的行為便十分容易。這些現象顯示，當受試者在從事這類的法則測試作業時，可能有時是以經驗類比的方式進行的，而非全靠抽象的邏輯分析。我們因此預測，經驗類比 (如：提供證偽的先前例題) 可以是教導證偽思考技能、減輕證真偏好的一種方式。

本研究中將以在練習題中提供證偽的解題實例來作為可供類比的先前例題。實驗中類比對證偽思考技能的學習的效果將與沒有任何教導的控制組比較。此外為加強類比學習的效果並有一類比情境會在練習題中加入該問題中之法則的邏輯結構圖以幫助受試者掌握題目當中的邏輯結構，以降低類比錯誤。然而，若證偽的思考技能可透過類比的歷程習得，我們懷疑其亦將表現出類比歷程的典型特徵。亦即，表面相似性將是影響受試者是否可由解題範例中習得證偽思考技能的主要因素。若缺乏表面相似的提取線索時，受試者即使遇到與練習題結構相同的問題也不易有證偽行為。但練習題中若附有問題的邏輯結構圖示時，我們預測類比正確率將會相對地提高。此外，由於類比習得的知識係由直接套用先前的例子以處理新問題的實作，過程中並不易形成抽象的法則 (Berry & Dienes, 1993)，我們也預測，當問題是以抽象法則的方式呈現時，由先前例題習得的證偽行為將無法遷移。值得注意的是，由於邏輯形式不同，問題的難易度也隨之不同 (如： $P \rightarrow Q$ vs. $\sim P \rightarrow Q$ 或 $P \text{ and } Q \rightarrow R$ vs. $P \text{ or } Q \rightarrow R$)，因此，研究中使用的所有問題均是 $P \rightarrow Q$ 的邏輯結構。練習題與測試題的結構相似度僅以問題的語意結構來區分。練習題及與練習題語意結構相同的法則測試題其 $P \rightarrow Q$ 的語意表達是屬於「行

動→條件」的許可法則；與練習題語意結構不同的法則測試題其 $P \rightarrow Q$ 的語意表達是屬於「條件→行動」的義務法則 (Cheng & Holyoak, 1985)。

方 法

一、受試者

本研究的受試者為 227 位主修管理的大學生，根據他們可以參加實驗的時間形成 1 至 15 人間的小組。實驗時每一個小組被隨機指派至 (1) 有先前例題的類比學習情境；(2) 有先前例題+邏輯圖示的類比學習情境；或 (3) 無先前例題的控制組，三者之一。由於每個小組的人數不等，去掉有遺漏之資料後，三組人數分別為：類比學習，71 人；類比+結構圖，76 人；控制組，80 人。

二、法則測試作業及其表現之衡量

法則測試作業係以 Wason 之選卡片作業為設計依據，共 16 個含具體內容之法則測試作業及 2 個抽象之法則測試作業。

(一) 具體內容的題目。每個法則測試作業中描述的是一個故事，故事引含一個「 $P \rightarrow Q$ 」的法則規定。法則規定中的 P 、 Q 項均是一般人熟悉的事件，但 P 與 Q 間的關係法則則是虛構的，如「擔任採購職務者需每年公布財產」。受試者的工作是去檢定故事中規範 P 、 Q 間關係的法則規定是否有被確實執行。檢定的方式是由分別代表法則中的 P 項、 $\sim P$ 項、 Q 項、及 $\sim Q$ 項的四個案例 (如，採購人員、品管人員、今年有公布財產、今年未公布財產) 中圈選受試者要測試的案例，圈選後並以九點量表對該案例測試的結果可以判斷法則規定確實有被執行的程度給予一個主觀的信心評估，此信心評估的目的一方面是希望藉以防止受試者四個選項都不加思索的全都圈選；一方面也強迫受試者在做多重選擇時，將他們所做的選擇測試項目間的資訊量大小加以區分。由於法則中的規定不一定與常理相容，因此受試者在每一題之前都會被要求以 9 點量表來對法則的合理程度進行評量。此作業的目的是在每一個法則測試作業前提供一個混淆作業 (filler task)，讓有些看來不合常理的法則可以受到受試者同等的考量。

(二) 抽象題。目的在檢驗受試者是否可由具象題的類比經驗中形成抽象的證偽邏輯。兩題抽象的題目敘述不含任何故事， $P \rightarrow Q$ 法則中變項的內容均屬抽象符號，其中一題為「卡片的一面為 a, b, c, d 的英文字母時，另一面則是一個個位數的阿拉伯數字」；另一題是「甲 \rightarrow 乙」。

法則測試的正確反應。由於絕大多數受試者皆會選擇測試 P ，因此我們以受試者同時選擇測試 P 及 $\sim Q$ 案例的正確反應次數來衡量他們證偽的思考技能。實驗中，受試者被告知在 P 、 $\sim P$ 、 Q 、 $\sim Q$ 的四測試案例中，至少要選一個，但可複選。由於受試者原始的反應中同時測試 P 及 $\sim Q$ 的正確反應數極端的少，因此，我們的法則測試正確率是採取較寬鬆的標準，亦即，測試 P 及測試 $\sim Q$ 的正確反應的判斷方式分別是：(1) 受試者選 P 的信心不可低於或等於選 $\sim P$ 或 Q 的信心。若 P 、 $\sim P$ 、 Q 三種方案中， P 的信心最高，則得 1 分，否則即為 0 分；(2) 同樣的，在 $\sim Q$ 、 Q 、 $\sim P$ 三者中受試者選 $\sim Q$ 的信心不可低於或等於其選 Q 或 $\sim P$ 的信心。若選 $\sim Q$ 的信心在 $\sim Q$ 、 $\sim P$ 及 Q 三者中最高，則得 1 分，否則 0 分。據此評分原則，在每一問題中同時選擇測試 P 及 $\sim Q$ 案例才得 1 分。因此，2 (語意結構相似性) \times 2 (表面相似性) 四類具體法則測試題中，每一類具體測試題的正確得分全距為 0-4。練習題及抽象題的評分方式與具體測試題同。

三、自變項之操弄

實驗中的自變項有三個，一為證偽思考技能的教導方式；二為法則測試問題與練習題間之表面相似度；三則為法則測試題與練習題間之語意結構相似度。

(一) **證偽思考技能之教導方式**。以 4 個有具體內容的法則測試作業為練習題來操弄三種不同的證偽思考技能教導方式：(1) 類比學習，指練習題中有證偽之解題實例。此組之受試者在每一練習題之後會看到一個完整的證偽的解題實例（附錄一）。解題實例是以一個人自言自語的在解此問題的方式呈現的。其中，第一段是解題者對該問題中法則（ $P \rightarrow Q$ ）的合理性的評論。接下來是解題者的解題說明，主要在說明法則測試過程中證偽的必要性，解題者並依此原則同時選擇 P 及 $\sim Q$ 的測試方案，但在信心評估上則是 $\sim Q$ 高於 P 。(2) 類比 + 邏輯結構圖示。在此組中的受試者除了看到的 4 題解題實例與類比組完全相同外，在每一練習題中並附有此題中法則的結構圖示（附錄一），受試者可由圖示中推論出 P 、 $\sim P$ 、 Q 與 $\sim Q$ 之間的關係。在解題實例中亦有此圖及參考此圖之提示。(3) 控制組。此組受試者直接以自己的方式回答四個練習題，沒有任何的教導活動。實驗操弄是以小組個別進行，每一組內的受試者所屬的實驗情境相同，接受的實驗材料與指導語也完全相同。為了確定在類比學習兩組的受試者會詳細閱讀每一練習題的解題說明，實驗者在指導語的說明時並要求受試者務必逐題仔細閱讀解題說明及每個練習題後的解題實例。

(二) **表面相似度**。指具體測試題中之法則的 P 或 Q 項的內容是否至少有一項與練習題的 P 或 Q 的內容相同。以例題一為例，此題中的 P 項為開車， Q 項為年齡。測試題的第三題中的法則為 P 項為參選， Q 項為年齡；如此，此題與練習題間便具有某種程度的表面相似度。依此原則，12 題有具體內容之測試題中有一半與練習題間具有表面相似度，另一半則否（參見附錄二）。

(三) **語意結構相似度**。邏輯表達的是變項（arguments）間的一種抽象的、形式的關係（formal relation）。其形式關係的效度與變項的語言敘述方式或內容應無關。 $P \rightarrow Q$ 的法則中 P 與 Q 可以是任何事物， $P \rightarrow Q$ 所表達的意義與 $Q \rightarrow P$ 完全不同。 $P \rightarrow Q$ 顯示的是 P 為 Q 的充分條件， Q 為 P 的必要條件（亦即， $\sim Q \rightarrow \sim P$ ），因此 P 與 Q 的位置是不可顛倒的。邏輯上，與 $P \rightarrow Q$ 同構而不同形（isomorphic）的法則為 $\sim Q \rightarrow \sim P$ 。由於不同的邏輯形式代表不同的結構（如，一為肯定命題，一為否定命題），且難度不同，因此 4 題練習題及 16 題具體法則測試題均是 $P \rightarrow Q$ 法則，法則的結構相似與否僅以不同的語意結構來操弄。實驗中，4 題練習題中所描述的故事均屬許可法則（permission rule），其語意結構是：「若要做**行動 A**，則必須有**B 的前置條件**。」（ $\sim B$ 條件 $\rightarrow \sim A$ 行動）亦即 $P \rightarrow Q$ 法則中， P 的內容是行動（action）， Q 的內容是情境條件（state condition），意義為凡是從事 P 行動者，必具有 Q 的條件才可以。在 16 題具體測試題中有 8 題是屬於許可法則，與練習題屬相同語意結構。另 8 題中的故事則是屬於義務法則（obligation rule）。義務法則也是 $P \rightarrow Q$ 的一例，但是其中的 P 、 Q 項的語意特徵剛好與許可法則相反，其語意結構為：「若 **A 條件** 存在，則 **B 行動** 必然發生。」（ A 條件 $\rightarrow B$ 行動），亦即， P 的內容為一前置條件，此條件滿足時，則 Q 之行動隨即產生。若以 A 與 B 分別代表行動及情境條件的名稱時，許可法則的語意結構為「行動 $A \rightarrow$ 條件 B 」，義務法則為「條件 $B \rightarrow$ 行動 A 」。雖然二者的抽象邏輯法則相同，但語意結構不同。

整體而言，16 個法則測試題（附錄二）中有 4 題與練習題為表面及語意結構皆相似，4 題與練習題為語意結構相似但表面不相似，另 4 題與練習題間具有表面相似度，但語意結構不同，最後 4 題的表面與語意結構均與練習題不同。為了衡量由 4 個練習題與 16 題具象題的答題經驗之後受試者是否可掌握到 $P \rightarrow Q$ 的邏輯法則，在 16 題具象題後的 2 抽象題，其問題形式不含故事內容，因此與練習題無表面或語意結構相似性的問題。

四、程序

受試者是按照他可以參加實驗的時間組成 1-15 人的小組在實驗室中進行法則測試作業。實驗時，每個受試者會拿到一份含有指導語、4 題練習題、16 題具體法則測試題、及 2 題抽象法則測試題的題本，以及另外一份作答用的答案紙。類比所需的解題實例及邏輯結構圖示的操弄均是透過實驗材料的安排進行。4 題練習題的呈現順序固定，但 16 題具體法則測試題及 2 題抽象法則測試題則分別有兩個不同順序的版本。練習題與所有的法則測試題中所提供的四個選項 (P、Q、~P、~Q) 的順序則是完全隨機安排。

在實驗開始時，實驗者先說明法則測試作業的性質及作答方式。接下來，請控制組以 4 個練習題進行作答練習；兩組類比組則說明每一練習後實驗者會提供過去成功解題者的解題實例給她們參考，請他們在每題之後詳細閱讀解題實例。四題練習題都做完之後，所有的受試者皆自行回答 16 個具體法則測試題及 2 題抽象法則測試題。整個活動並無時間限制，所需時間約 30 ~ 60 分鐘。

五、實驗設計及資料分析

本研究是個 3 (教導情境) × 2 (表面相似度) × 2 (語意結構相似度) 的三因子實驗設計。其中教導情境為組間變項，表面及語意結構相似度為組內變項。受試者正確測試 P 及 ~Q 的次數分別以 SPSS 的一般線性模式中的重複量數變異數分析進行統計檢定。當研究變項的主效果或交互作用顯著時，接下去是以 LSD 法來進行平均數間的多重比較，統計檢定的顯著水準 (α) 為 0.05。

結 果

一、實驗處理的檢驗

為了檢定類比學習的有效性，受試者在 4 題法則測試練習題中選擇同時測試 P 及 ~Q 的平均正確反應量 (表 1) 以學習情境為自變項的單因子變異數分析結果，學習情境的主效果顯著， $F(2, 224) = 10.23$, $MSE = 1.80$, $p < .001$ ，以 LSD 進行多重比較之後發現，練習題後附有解題實例，或解題實例 + 邏輯結構圖示的兩組選擇同時測試 P 及 ~Q 的反應次數分別為 2.06 及 1.95，均顯著高過無解題範例組的 1.16，但例題中有無邏輯結構圖示所造成的差異並不顯著。

表 1 練習題及抽象題信心修正後測試 P & ~Q 的平均次數

測 試	控 制 組		解 題 實 例		實 例 + 結 構 圖	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
練習題	1.16	1.20	2.06	1.47	1.95	1.36
抽象題	0.29	0.51	0.35	0.56	0.36	0.58

二、具體法則測試題同時測試 P 及 ~Q 的反應次數

受試者在 16 題具體法則測試題中選擇同時測試 P 及 ~Q 的反應量列於表 2。以教導情境為組間

變項，表面相似度及語意結構相似度為組內變項， $3 \times 2 \times 2$ 重複量數變異數分析的結果，教導情境主效果顯著， $F(2, 224) = 16.86, MSE = 4.13, p < .001$ 。表面相似度， $F(1, 224) = 8.27, MSE = 0.55, p < .01$ ；及語意結構相似度， $F(1, 224) = 65.77, MSE = 0.49, p < .001$ ，等主效果也顯著。

表 2 具體題信心修正後測試 P & ~Q 的平均次數

結構	表面	控制組		解題		實例+		結構圖	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
同	同	0.96	1.12	1.96	1.50	2.01	1.63		
	異	0.76	1.05	1.70	1.40	1.70	1.34		
異	同	0.59	0.81	1.07	0.85	0.95	0.89		
	異	0.81	0.92	1.63	1.41	1.78	1.36		

就教導情境的主效果而言，以 LSD 法進行事後的平均數比較發現，兩個類比學習情境的受試者同時測試 P 及 ~Q 的次數（類比組， $M = 1.59, SE = .12$ ；類比 + 圖示， $M = 1.61, SE = .12$ ）均顯著高過控制組（ $M = 0.78, SE = .11$ ），但此二組間的差異則不顯著，顯示具體測試題整體的類比效果與練習題相似，在先前例題中提供法則的結構圖示對證偽思考技能的學習似乎沒有預期的加分效果。就表面相似度的主效果而言，與先前例題表面相似的測試題的正確反應數（ $M = 1.26, SE = .07$ ）反而低於與先前例題表面不相似的測試題（ $M = 1.40, SE = .08$ ），此結果與我們的預期及大部分的研究不一致。結構相似度的主效果則與預期相同，與先前例題語意結構相同的允許法則測試題的正確反應數（ $M = 1.52, SE = .08$ ）高於與先前例題語意結構不同的義務法則測試題（ $M = 1.14, SE = .06$ ）。然而由於變項間存有交互作用，顯示類比學習的效果會受到先前例題與測試題間的表面相似度與語意結構相似度的限制，下面進一步分析它們間的確切關係。

類比與語意結構相似度間的交互作用顯著， $F(2, 224) = 5.52, MSE = 0.49, p < .01$ ，LSD 事後多重比較發現，語意結構相似度的效果只出現在類比加圖示組中；在此組中，與練習題語意結構相似的測試題的證偽反應顯著高於與練習題不同語意結構的測試題（相似， $M = 1.86, SE = .14$ ；不相似， $M = 1.36, SE = .10$ ）。控制組（相似， $M = 0.86, SE = .14$ ；不相似， $M = 0.70, SE = .10$ ）與沒有圖示的類比組皆無語意結構相似度的效果（相似， $M = 1.83, SE = .15$ ；不相似， $M = 1.3, SE = .11$ ）。由此推論，語意結構相似的先前實例若要成為類比時的參考經驗還需要當事人能掌握問題中的邏輯結構。

表面相似度與語意結構相似度間的顯著交互作用， $F(1, 224) = 49.89, MSE = 0.72, p < .001$ ，LSD 事後多重比較發現，交互作用是來自受試者在語意結構與練習題不同，但題目中卻具有表面相似度的義務法則測試題中的正確率偏低的緣故（ $M = 0.87, SE = .06$ ），其他三組測試題型間的差異不明顯（ M 's, 1.39 -1.64; SE 's, .08 -.10）；顯示當記憶中找不到語意結構相同的先前實例時，表面相似度便成了人尋找記憶中之參考實例的主要線索；由於本研究中的問題當語意結構不同時，若使用相似的表面線索會造成法則中 P、Q 項的位置錯置，因而產生了類比錯誤。

然而，類比、表面相似度與語意結構相似度三者間的交互作用也顯著， $F(2, 224) = 3.53, MSE =$

0.72, $p < .03$ ，由圖 1 可清楚的看出，在控制組中表面與語意結構相似度對同時測試 P 及 $\sim Q$ 的次數完全沒有影響（語意結構相同 + 表面相同， $M = 0.96$, $SE = .16$ ；語意結構相同 + 表面不同， $M = 0.76$, $SE = .14$ ；語意結構不同 + 表面相同， $M = 0.59$, $SE = .10$ ；語意結構不同 + 表面不同， $M = 0.81$, $SE = .14$ ）。反之，事後多重比較顯示，在兩類比情境中當測試題與練習題語意結構相同時，不論題中是否有相同的表面線索，受試者同時測試 P 及 $\sim Q$ 的次數（解題實例，表面相同， $M = 1.96$, $SE = .17$ ；表面不相同， $M = 1.70$, $SE = .15$ 。解題實例 + 結構圖示，表面相同， $M = 2.01$, $SE = .16$ ；表面不相同， $M = 1.70$, $SE = .15$ ）均顯著高過四組控制組，顯示先前實例對證偽行為的類比學習效果可遷移至與解題實例屬相同語意結構的新問題中。但當法則測試題與練習題的語意結構不同時，不論先前實例中有無邏輯結構圖示，只要題目中有與先前實例相似的表面提取線索（僅有解題實例， $M = 1.07$, $SE = .11$ ；解題實例 + 結構圖示， $M = 0.95$, $SE = .10$ ），受試者證偽測試的次數都與控制組相似，沒有產生類比學習。只有當法則測試題與與先前實例不具表面相似度時受試者同時測試 P 及 $\sim Q$ 的次數（僅有解題實例， $M = 1.63$, $SE = .15$ ；解題實例 + 結構圖示， $M = 1.78$, $SE = .14$ ）才會顯著的高於四組控制組，出現類比的效果。這些現象顯示，以類比的方式運用證偽思考技能的條件為問題與先前實例間需具語意結構相似，表面相似並非必要；更甚者，當先前證偽實例的學習要遷移至語意結構不同的法則測試問題時，法則中的表面相似性若被當成記憶提取線索還可能因條件項與行動項的錯置而造成類比的錯誤。此外，上述分析結果也顯示，先前解題實例中是否附有法則的邏輯圖示對類比學習的效果沒有造成差異，此與我們的假設不符。可能的原因之一是練習題中所附的法則的邏輯圖並不容易看懂；另一個可能是受試者可能低估了問題的難易度，因而沒有認真的去參考練習題中的邏輯圖。

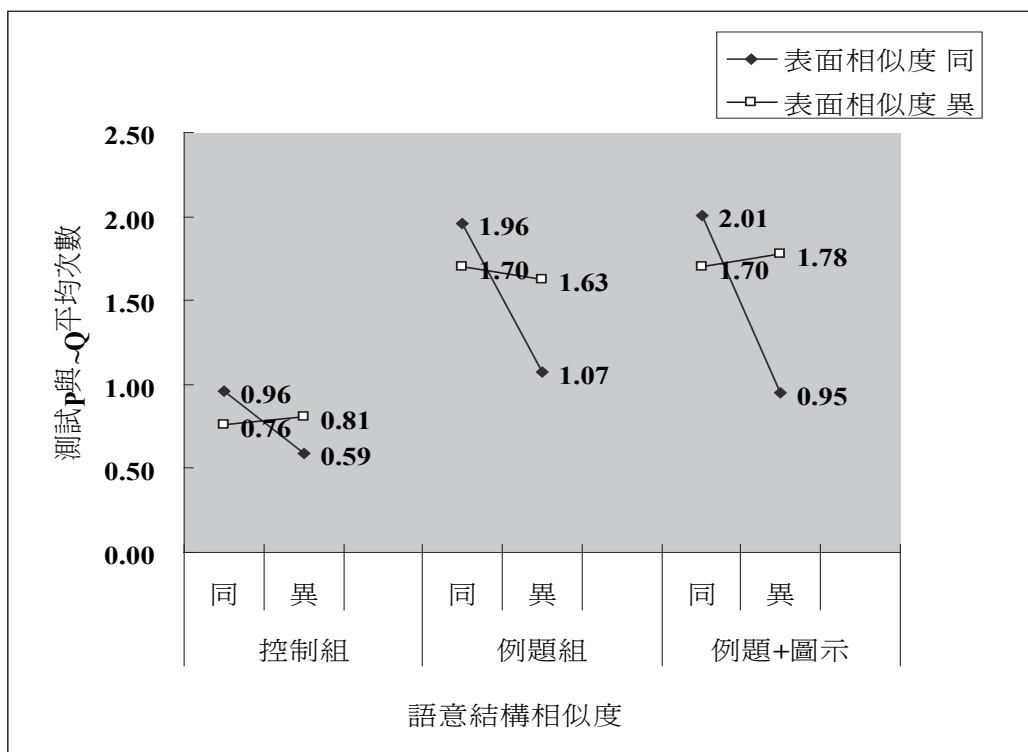


圖 1 表面相似度、語意結構相似度、與類比學習情境三者間之交互作用

三、抽象題法則測試 P 及 $\sim Q$ 的次數

受試者在兩題抽象法則測試題中同時測試 P 及 $\sim Q$ 的次數 (表 1) 以單因子 (教導情境) 變異數分析結果, 教導情境效果不顯著, $F(2, 224) = 0.38, MSE = 0.30$, 顯示不論事先前例題是否有附結構圖示, 類比學習對形成抽象的證偽邏輯沒有幫助。

結論與討論

本研究的目的是探討科學思考的推理特徵之一, 證偽的思考技能, 是否可以透過類比的歷程來獲得, 以及由類比獲得的證偽思考技能是否可以遷移至抽象的問題中。類比是一種套用先前的經驗來解新問題的歷程。本研究中, 在類比情境下的受試者會在每個練習題後看到一個使用證偽策略的解題實例, 實驗中有一組受試者的練習題及其解題實例中還附有法則的邏輯結構圖示。類比學習的效果是以受試者在與練習題有、無表面或語意結構相似度的法則測試題上的證偽反應來推論。我們預測若證偽的思考技能可以透過類比的歷程來獲得的話, 由受試者在測試題上的證偽表現 (同時測試 P 與 $\sim Q$) 將可以觀察到解題實例的效果。由於類比推論的成功與否受表面相似度的影響極大, 我們也預測在缺乏表面相似度時, 解題實例產生成功的類比的機會將較小; 此外, 由類比歷程獲得的程序性知識技能可能不易形成可以用語文表達的抽象法則, 因此遷移到抽象題目中的可能性很小。

實驗結果顯示, 在有具體內容的法則測試題中, 除了類比的主效果顯著, 整體而言, 不論問題中是否有法則邏輯結構圖的輔助, 證偽的解題實例都可促進證偽思考行為。然而, 類比、測試題與練習題間的表面及語意結構相似度三者的交互作用也顯著, 顯示由解題範例類比習得的認知技能若遷移至新的問題情境時, 若缺乏語意結構相似度的支援, 其效果將受限。類比、表面相似度、語意結構相似度三因子交互作用分析顯示, 類比學習的效果只能遷移至與先前實例有相同語意結構的測試題中, 且不受表面相似度影響。若測試題與先前例題間沒有語意結構相似度時, 類比學習的效果只能遷移至與先前例題表面線索不同的情境中, 而在與先前例題具有相似表面線索的情境中, 則可能因提取線索的干擾, 沒有出現類比的學習遷移效果。由此我們推論, 類比在證偽的思考技能的學習上是有效的, 但類比學習遷移至新的問題時, 類比成功的限制條件是在記憶中的先前實例與目前問題中的法則要是屬於相同的語意結構。若問題的語意結構與先前實例不同時, 則只有當問題與先前實例沒有混淆的表面線索時, 過去的解題經驗才可能被正確引用。

這些現象顯示, 以類比方式學習證偽邏輯的有效性在於學習者是否能將問題中的邏輯結構正確表達並儲存在記憶中。表面相似性雖是搜尋相關經驗的線索, 但若問題本身即為一推理活動時, 人還得懂得要分析問題中法則的語意結構, 以尋找相同語意結構的問題以作為解題的依據。但若記憶中找不到相似語意結構的解題實例且沒有相似的表面線索的誤導時, 人才會進一步分析問題的邏輯結構; 此時, 記憶中的解題實例是否含有其正確的邏輯結構表達法就成了類比成功與否的依據了。此研究發現顯示在科學發現過程中, 若過去的經驗要順利發揮類比的作用, 研究問題的解答必須落在當事人已有的理論架構下才行; 若研究問題的解答是落在當事人原有理論架構之外時, 先前經驗若要發揮類比的作用的話則有賴研究者需刻意排除表面線索的誤導, 及正確掌握原有理論隱藏在語意結構下的基本邏輯架構, 如此方能對新問題的邏輯分析有所助益。

本研究中我們所預期的表面相似度對類比學習的效果並沒有得到支持, 反而出現了語意結構相似的效果, 這和過去的研究不同。我們推論這是因本研究中的問題結構指的是介於表面屬性與邏輯結構間的語意結構, 而非真正的邏輯結構。研究中不同語意結構的具體問題其邏輯結構其實是相同的。其中, 許可法則以自然語言描述是「若要從事行動 A 則先要獲得條件 B 」(行動 $A \rightarrow$ 條件 B); 而義

務法則則是「若 B 條件存在，則 A 行為必然出現」(條件 B → 行動 A)，二者的邏輯形式皆是「P → Q」，不同的只是 P 與 Q 的內容 (argument) 是個條件或行動。此研究中語意結構相似性的效果顯示，語意結構可以支援推理，沒有慣用的語意結構的指引時，人的思考能力可能會受限。然本研究中也發現，當先前解題實例與測試題的語意結構不同時，若問題中沒有不恰當的表面線索時，即可能出現類比的學習遷移，此現象顯示即使是大學生，其推理行為有部分仍架構在經驗或記憶所及的語意概念世界中；而非依賴可完全抽離情境的抽象或形式法則。

實驗中發現類比對證偽邏輯的學習效果並無法由具體題遷移到抽象法則題，顯示由類比的方式獲得的認知技能不易形成抽象的法則表達形式。我們推論科學思考法則的學習與掌握一方面有賴提供學習者在其生活經驗中實際觀察與應用 (練習) 證偽的推理，以累積其記憶中可供進行類比的解題實例；一方面則需對其經驗過的解題實例進行反省及邏輯分析以掌握其中的基本結構，如此方可在未來遇到新問題時可正確的找到類比的實例。

整體而言，本研究的發現支持類比學習可促進證偽思考技能的獲得，但效果距理想仍有一段距離。此發現有三個理論與實務上的涵意：1、科學思考、推理等認知技能可以透過提供解題實例的類比學習方式獲得。由於單以語文的方式教導推理的邏輯對證偽思考技能的養成效果有限，在科學教育、教材設計中提供科學思考的實例及親身觀察到老師或周圍的人使用證偽的思考便是個重要的思考技能的教導模式，例如講解科學家當初在科學發現的過程中究竟經歷過那些推理歷程、設計實驗、或參與科學家的實驗等，老師並要藉機分析以協助學生瞭解其中的推理邏輯。2、一般在從事思考、推理的活動時，其解答的產生可能是仰賴如類比之類的歷程去套用過去的解題程序；而非經由邏輯法則的推演以導出解答。證真偏好即可能就是此非邏輯法則思考歷程的產物。人靠記憶來「思考」，而非刻意地使用法則去思考，這和一般人的直觀經驗可能不同。研究指出，許多的解題或決策錯誤都可能源自人直接由記憶中套用解答而來 (Barrett, Tugade, & Engle, 2004 ; Kintsch, 1998 ; Sloman, 1996)。直接由記憶中套用解答的認知過程類似直覺，較快速、輕鬆；使用邏輯法則來分析、統整經驗的認知過程則相對上緩慢且辛苦。但慎思明辨不僅是科學家所需具備的條件，也是一般人在處理問題或決策時所不可或缺的要件。如何釐清以記憶為基礎的思考和以法則為基礎的思考間的差異便是個可進一步研究的理論議題。就教育、應用方面而言，如何才能激發人使用以邏輯法則為基礎的思考推理也是個值得探討的議題。3、類比學習雖是獲得證偽的思考技能的方法，但類比有賴問題間的結構相似性，缺乏分析結構的能力或過度依賴表面相似線索或語意結構時，類比的效果不僅侷限在經驗範疇內，且可能帶來錯誤。人如何詮釋一個法則的意義已逐漸受到重視 (如：Stenning & van Lambalgen, 2004)；如何將由如類比等學習的方式獲得的知識技能轉換到語文、法則的知識表達層次，也將是研究的重要方向。

法則測試是科學思考的核心工作。它是種複雜的認知技能，其中包含了意識及無意識的記憶激發、搜尋、推理與決策歷程。在一個資訊化的社會中，環境中無所不在的是各種宣傳、廣告、文宣，對資訊加以質疑、否證已是人人必備的素養。如何有效的提供證偽思考的實例與訓練，不僅是從事科學研究及科學教育者所該關心，也是一般教育所該關心的議題。

參 考 文 獻

- 王亞同 (民 91)：例示作用促進類比問題解決的研究。教育與心理研究，25 卷，179-193 頁。
- 江淑卿 (民 90)：兒童類比推理能力的學習潛能評估研究。教育心理學報，33 卷，1 期，47-64 頁。
- 李怡嚴 (民 89)：隱喻—心智的得力工具。國立清華大學哲學研究所主辦系列專題演講宣讀之論文 (新竹)。

- 黃幸美 (民 90)：兒童解決數學及自然科學問題的問答討論與類比推理思考之研究。教育心理學報，32 卷，2 期，121-141 頁。
- 陳茹玲、蘇宜芬 (民 94)：類比故事的表面特徵與結構特徵在不同年齡層對故事回想表現的影響。教育心理學報，37 卷，2 期，123-146 頁。
- Adsit, D., & London, M. (1997). Effects of hypothesis generation on hypothesis testing in rule discovery tasks. *Journal of General Psychology*, 124, 19-34.
- Ahn, W. K., & Graham, L. M. (1999). The impact of necessity and sufficiency in the Wason's four-card selection task. *Psychological Science*, 10, 237-242.
- Bacon, F. (1960). *The new organon and related writings*. New York: Liberal Arts Press.
- Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, 130, 553-573.
- Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hillsdale, NJ: Lawrence Erlbaum.
- Bowdle, B. F., & Gentner, D. (2005). The career of metaphor. *Psychological Review*, 112, 193-216.
- Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17, 391-416.
- Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology*, 18, 293-328.
- Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. Hove, East Sussex: Lawrence Erlbaum.
- Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Gentner, D. (1989). The mechanisms of analogical reasoning. In S. Vosinadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). Cambridge, England: Cambridge university Press.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, 25, 524-575.
- Gorman, M. E. (1995). Confirmation, disconfirmation, and invention: The case of Alexander Graham Bell and the telephone. *Thinking and Reasoning*, 1, 31-53.
- Goswami, U. (1992). *Analogical reasoning in children*. Hillsdale, NJ: Lawrence Erlbaum.
- Johnson-Laird, P. N., & Wason, P. C. (1970). Insight into a logical relation. *Quarterly Journal of Experimental Psychology*, 22, 49-61.
- Johnson-Laird, P. N., Legrenzi, P., & Legrenzi, M. S. (1972). Reasoning and sense of reality. *British Journal of Psychology*, 63, 395-400.
- Johnson-Laird, P. N., & Byrne, R. (2002). Conditionals: A theory of meaning, pragmatics and inference. *Psychological Review*, 109, 646-678.
- Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (2004). Reasoning from inconsistency to consistency. *Psychological Review*, 111, 640-661.
- Kareev, Y., & Halberstadt, N. (1993). Evaluating negative tests and refutations in a rule discovery task. *Quarterly Journal of Experimental Psychology*, 46, 715-727.
- Kareev, Y., & Avrahami, J. (1995). Teaching by example: The case of number series. *British Journal of Psychology*, 86, 41-54.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing.

- Psychological Review*, 94, 211-228.
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, 29, 85-95.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101, 608-631.
- Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchison.
- Reeves, L. M., & Weisberg, R. W. (1994). The role of content and abstract information in analogical transfer. *Psychological Bulletin*, 115, 381-400.
- Ross, B. H. (1987). This is like that: The use of earlier problems and separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629-629.
- Ross, B. H. (1989). Distinguishing types of superficial similarity: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 456-458.
- Slovan, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3-22.
- Stenning, K., & van Lambalgen, M. (2004). A little logic goes a long way: Basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science*, 28, 481-529.
- Tweney, R. D., Doherty, M. E., Warner, W. J., & Pliske, D. B. (1980). Strategies of rule discovery in an inference task. *Quarterly Journal of Experimental Psychology*, 32, 109-123.
- Wason, P. C. (1960). On the failure to eliminate hypothesis in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129-140.
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135-151). Harmondsworth: Penguin.
- Wharton, C. M., Holyoak, K. J., Downing, P. E., Lange, T. E., Wickens, T. D., & Melz, E. R. (1994). Below the surface: Analogical similarity and retrieval competition in reminding. *Cognitive Psychology*, 26, 64-101.

收 稿 日 期：2006 年 03 月 03 日

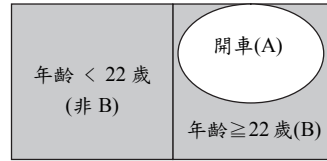
一稿修訂日期：2006 年 08 月 14 日

二稿修訂日期：2006 年 09 月 13 日

接受刊登日期：2007 年 01 月 13 日

附錄一 假設檢定作業及解題示範實例¹

根據資料顯示，22 歲以下的人開車肇事的比例最高，為了減低年輕人開車肇事的比例，W 市最近通過一項交通法，規定年滿 22 歲的人才可以開車。



白色部分：開車(A)
灰色部分：不開車(非 A)

請問你認為這樣的規定合理嗎？

不 合 理					普 通				很 合 理
1	2	3	4	5	6	7	8	9	

目前已知 W 市甲、乙、丙、丁四人的資料為：



如果你想知道該項規定是否有被認真執行，你會對那個人或那些人進行蒐證活動？(可複選)
並請針對你選擇的每一項活動，評估它可以讓你看出執法與否的信心程度。

全 無 信 心					中 度 信 心				高 度 信 心
1	2	3	4	5	6	7	8	9	

- _____ 1. 抽問甲 (開車的人)，看他是幾歲？
- _____ 2. 抽問乙 (沒開車的人)，看他是幾歲？
- _____ 3. 抽問丙 (25 歲)，看他有沒有開車？
- _____ 4. 抽問丁 (18 歲)，看他有沒有開車？

¹ 有解題實例但無邏輯結構圖示組之練習題與此相似，只是練習題中未附邏輯結構圖示。證偽邏輯組及控制組的練習題中則均無邏輯結構圖示，也無解題實例。

解 題 實 例

首先，我覺得這個規定並不完全合理¹。雖然可以降低年輕人開車出事的情形，但我覺得這也會讓許多需要開車上班或上學的人十分不方便，二十二歲以下的人也不能找開車方面的工作。

『要開車 (A)，則要二十二歲以上 (B)』，這樣的規定相當於

『 $A \rightarrow B$ 』的法則，A 和 B 的關係可由上面的圖形看出²。A 是包含在 B 裡頭，而 A 與非 B 則是互斥。換句話說，1) 開車的人 (A) 必定是 22 歲以上 (B)；但 2) 22 歲以上的人 (B) 則不一定開車 (A)；且，3) 22 歲以下者 (非 B) 一定不可開車 (非 A)。

由圖可以判斷，開車的人年齡應該全在二十二歲以上。所以，

- 1) 若我問開車的人 (A)，發現他們的年齡全在二十二歲以上 (B)，這也無法保證就是認真在執法。因為我並無法窮盡所有的開車的人，所以這樣並不夠。
- 2) 若我問開車的人 (A)，只要發現有人的年齡在二十二歲以下 (非 B)，那我就可以說這邊的警察執法並沒有完全。換另一個角度來看，
- 3) 若我問年齡全在二十二歲以下的人 (非 B) 有沒有開車的經驗，只要有一個人說他曾經開過車 (A)，那就足以證明這裡的警察執法不力了。

所以，要判斷警察執法認真與否，我會選擇：

- 1。抽驗開車的人 (A)，問他們幾歲，以及，
- 4。抽驗二十二歲以下的人 (非 B)，問他們開車嗎？

我有信心由 1 正確地判斷出警察是否有認真執法的程度約是… 8。

我有信心由 4 正確地判斷出警察是否有認真執法的程度約是… 9。

¹ 在 4 個解題實例中對法則做出合理或不合理的評論的情形各兩題。

² 除了這句話外，僅有解題實例組的解題示範內容完全相同。

附錄二 實驗材料

練習題

(許可法則：) 行動 A → 條件 B

1. 根據資料顯示，年紀在二十二歲以下的人開車肇事的比例最高，為了減低年輕人開車肇事的比例，A市最近通過一項交通法，規定年滿二十二歲的人才可以開車。【開車→年齡 > 22】
2. B市的教育局為了提升該市受過高等教育者的人數，規定凡該市高中學生三年學業成績在班上前十五名者，即可免試取得該市轄區內的大學之入學許可。【免試取得入學許可→學業成績在班上前十五名】
3. 為了避免投票時出現大量的幽靈人口，某市規定要在該市取得投票權的人，必須在該市居住滿二年以上。【投票→在該市居住滿二年以上】
4. 某大學的女子籃球隊是國際明星球隊，該校規定要加入該籃球隊的女生身高至少要190公分。【加入籃球隊→身高至少要 190 公分】

具體測試題

語意結構相似 (許可法則： 行動 A → 條件 B)，表面相似

為了鼓勵並保障人民從政的機會，憲法明文規定，凡是年滿 20 歲的國民就可以參加民意代表選舉。【參選→> 20 歲】

國內某職業棒球隊，為了取得較佳的戰績，規定球隊的球員若要取得出賽資格，則其平均的打擊率至少要在 2.0 以上。【取得球隊出賽資格→打擊率 > 2.0】

美國一些名校的校史室內設有名人榜，以表揚對該校有特殊貢獻的人士或傑出校友。某校的校規規定，要列入該校名人榜的最低標準是畢業時成績是該系畢業生第一名。【列入名人榜→該系畢業成績第一名】

某大學規定，大學生若要取得校務會議的代表資格 (可對校務有投票/表決權)，則需要在校內選舉時要能取得該校學生 1/20 以上的選票。【學生可在校務會議投票表決→得票率 > 1/20 以上的學生】

語意結構相似 (許可法則： 行動 A → 條件 B)，表面不相似

在台灣，公營機構的職員是不能被任意解聘的，因此是大家爭著要捧的鐵飯碗，但按規定若要進入公營機構當職員，則一紙高普考證書是少不了的。【擔任公營機構職員→高普考證書】

爲了讓低收入的家庭亦可享有良好的休閒活動，政府規定只要年收入低於 50 萬元的家庭，每年可申請免利息的旅遊貸款 5 萬元。【免利息的旅遊貸款→家庭年收入低於 50 萬元】

某電影公司，爲了要維持旗下演藝人員之號召力，要求若要取得演出合同，任何影星在簽約期間對外得宣稱其爲單身。【簽約→單身】

爲了讓職業婦女的幼兒得到妥善的照顧，某國政府設有免費的托兒所，凡是具有該國公民身份的人，其幼兒就可免費就讀這種公立的托兒所。【免費托兒→該國公民】

語意結構不相似（義務法則： $\boxed{\text{行動 B} \rightarrow \text{條件 A}}$ ），表面相似

南美洲有一個國家的足球隊居世界之冠，該國的兵役法中設有足球役一類，役男身高在 190 公分以上的自動分入此類。【 $\text{身高} > 190 \text{ 公分} \rightarrow \text{足球(隊)役}$ 】

爲了加強組織人力的汰舊換新，目前世界各國大多有強迫退休的規定，但退休年齡不盡相同。某國規定，凡是年滿 60 歲的人就得退休。【60 歲→退休】

爲了提昇國家的學術研究水準，教育部規定只要碩士第一年的成績在班上前 50%，就可以取得攻讀博士學位之入學許可。【 $\text{成績前} 50\% \rightarrow \text{直攻入學許可}$ 】

美國的法院採陪審團制度，某州規定在當州居住一年以上的成人，每年得義務擔任陪審員一次。【 $\text{居住} > 1 \text{ 年} \rightarrow \text{每年得義務擔任陪審員一次}$ 】

語意結構不相似（義務法則： $\boxed{\text{行動 B} \rightarrow \text{條件 A}}$ ），表面不相似

爲了優生學的理由，某國家的法律規定身心有遺傳性問題的人要強迫結紮，智障者爲其中一類。【智障→結紮】

某國家爲了管制流浪動物，規定寵物的主人，一定要爲他的每一隻寵物蒙面後才可讓寵物出門。【寵物→蒙面才可出門】

某國將電影分成五級，三級以上（3、4、5 級）的電影含有暴力或色情，兒童不宜，該國並規定兒童若要看第 3 級的電影，一定要有父母作伴。【3 級電影→兒童要有父母伴陪】

某大企業，爲了有效預防其職員利用職權而有貪污、收紅包的行爲，特別制定一個辦法，要求任職該公司的採購人員每年要公佈其財產一次。【採購人員→一年公佈財產一次】

抽象題

如果某人告訴你一個法則的邏輯表示法是 $P \rightarrow Q$ （若 P 則 Q ），那麼，若你要測試這個法則的正確性時，你要如何進行？【 $P \rightarrow Q$ 】

你的朋友告訴你，在他手上的一疊卡片中，凡是正面出現英文字母 A、B、C、D 中的一個，反面則是一個個位數的阿拉伯數字，若你要試試看他說的是不是真的，那麼下面的幾張卡片中你要翻哪幾張才能確定？【A、B、C、或 D → 一個個位數的阿拉伯數字】

Bulletin of Educational Psychology, 2007, 38 (3), 271-290
National Taiwan Normal University, Taipei, Taiwan, R.O.C.

Analogical Process in the Acquisition of Falsification Reasoning Skill

Ting-Yu Wu

Graduate Institute of
Engineering
Management

Ming-Hsin University of
Science and Technology

Li-Fen Huang

Department of Tourism
Providence University

Ruey-Yun Horng Yung-Chang Huang

Department of Industrial
Engineering and Management
National Chiao Tung University

The purpose of this study was to investigate the effect of analogical learning on the acquisition of falsification reasoning skill. Two hundred and twenty-seven college students were randomly assigned to three analogical learning situations (worked-out examples, worked-out examples + logical structure illustration, control) to work on four practice items with or without falsification worked-out examples. Their performance on 16 test items of deontic rule testing showed that worked-out examples could enhance the use of falsification reasoning only when test items shared the same semantic structure with the worked-out examples. When test items differed from the worked-out examples in the semantic structure, the success of analogical learning occurred only when there was no surface similarity between them. These findings suggest that in scientific reasoning a scientist's challenge lies in extracting a problem's underlying logic of rules and meanwhile excluding the misleading effect of surface similarity.

KEY WORDS: acquisition of falsification reasoning skill, analogical learning, deontic rule testing.