

行政院國家科學委員會專題研究計畫成果報告

國科會專題研究計畫成果報告撰寫格式說明

Preparation of NSC Project Reports

計畫編號：NSC 89-2413-H-002-040

執行期限：89年8月01日至90年7月31日

主持人：連韻文 執行機構及單位名稱：台大心理系

中文摘要

近年來我們發現影響假設檢驗推理作業成敗的關鍵在於是否能夠產生新角度的假設。在本研究中，我們進一步澄清與過去文獻與我們的發現不一致之處，藉以探討上述發現的普遍性。兩個實驗分別探討規則難易度與不同形態的作業的影響。

Abstract

We found that, rather than searching for disconfirmatory information, the ability to view the data from a different viewpoint was more critical for better performance in a hypothesis-testing task (246 task). The current research discuss whether the above finding would be hold for problems with various levels of difficulty and those rule-discovery tasks with different format, in which too many alternative hypotheses waited to be ruled out (e.g., concept learning tasks composed of visual features.)

緣由與目的

發現事物背後的關係或規則(因果的或類別的)是預測或理解事物的依據，因此我們可以說規則發現能力是人類一個基本且重要的認知能力。規則發現又可分為規

則產生(或稱之為假設產生, hypothesis generation)與規則測試(或假設檢驗, hypothesis testing)兩個部份。過去心理學在這方面研究最常被提起的發現是受試者常傾向於尋找支持自己目前假設的例子來測試假設是否為真。Popper (1959)曾指出得到否證才是一個科學理論或假設得以修正的要素，因此人們這種尋求正例的傾向常被心理學家視為是一種肯證偏誤 (confirmation bias)，也用來解釋受試者在某些規則發現的作業的失敗 (e.g., Wason, 1960)。許多研究者因而試圖增加推理者得到否證的機會 (e.g., Tweney, 1980; Gorman, 1984, 1987; Klayman & Ha 1989)，但並沒有一致的效果。我們近年的研究則指出這是因為以往研究所重視的否證訊息並不是假設檢驗推理的成功唯一要件，得到否證後如何修改原先的假設可能更為關鍵，因此我們進一步分析受試者所產生的假設。若受試者常僅對原假設 H 做出擴大或縮小範圍的修正，則我們稱新假設 H' 為互包性另有假設 (新假設包含或被包含於舊假設)。若是受試者跳脫原假設的範疇，從新的角度或面向審視所有的證據，則我們稱新假設為「交疊性另有假設」 (overlapped alternative hypothesis 新舊假設沒有包含的關係)。推理者在得到否證之後常只是部分修改原假設，在某些情況下這種傾向可能使推理者無法脫離既有窠臼而增加推理失敗的機率。在測試策

略方面，除了過去研究者著墨許多的正例與反例法 (positive- and negative-testing strategy) 之外，我們也發現有些推理者會同時測試兩個假設 (雙假設測試)。根據所產生的兩個假設之間的關係，雙假設測試又可分為交疊性測試法 (同時測試兩個有部份疊合關係的假設) 與互包性測試法 (同時測試兩個有包含關係的假設)。(詳見連韻文，1998；林緯倫與連韻文，2001)

簡言之，我們區分出交疊性另有假設與互包性另有假設，並發現工作記憶的負擔會影響前者的產生，但不會影響後者或總假設個數，顯示此區分有其意義，更重要的是推理成功者穩定地產生較多交疊性另有假設。本研究計畫則進一步探討上述發現的普遍性。有幾個因素可能會影響我們先前發現的普遍性，分別是規則的難易度以及不同形態的作業，以下分述之。

我們過去的研究結果雖然顯示否認或互包性另有假設的多寡並不是成功歸納出規則者與失敗者主要的分野，但這並不是意涵否認對於規則的發現全然無足輕重，或是互包性另有假設就是沒有用的另有假設。不論是肯證或否認都是在假設檢驗推理過程中必要的訊息。如果沒有肯證，推理者無法有任何依據形成可能的假設；否認則可以指出原假設的不足，提供修正假設的契機，也可淘汰掉不正確的假設；而當假設已經與目標範疇十分逼近時 (亦即在推理的後期或所找的規則較簡單時)，推理者僅需在原假設的範圍上稍作修改 (即產生新的互包性的假設) 就可找出正確的規則，因此產生交疊性另有假設是否是成敗的關鍵可能和問題的難易度有關。在過去探討假設檢驗的研究中，多採用 Wason(1960)所設計的「246 作業」。在他原始作業中的規則 (由小排到大的數

列) 是一個推理者較不容易一開始就想到的 (亦即是一個較困難的題目)，但後來的研究者在採用此作業時常加入其他題目，有可能因為他們並未控制題目的難易度，而無法發現交疊性另有假設的重要 (例如 Klayman & Ha, 1989)，這是實驗一所要探討的。

此外，在上述作業中，推理者所要測試與歸納出的是一個和數字概念有關的規則，而且推理者必須自己產生要測試的例子，在這種作業中並不容易想到很多不同角度的假設。但有些規則發現的作業並不需要受試者產生測試例子，而是主動給予例子或讓受試者在所有可能例子中挑選所要測試的，例如許多對聯列表 (contingency table) 學習或概念學習 (concept formation) 的作業都屬於這一類。在日常生活中也有這樣的情形，例如小孩在學習狗的概念時，他並無法自己產生一個例子去測試，而是接受環境所給予的例子。

Bruner et al. (1956) 曾採用的一個作業可作為代表。此作業呈現 81 卡片，每張卡片包含四個面向的視覺特徵 (圖案形狀、顏色、數目，以及邊框的數目)，這些卡當中有某些卡符合實驗者心中的概念，受試者的任務是找出此概念的規則。由於規則都是由具體知覺刺激所組成，推理者很容易形成不同面向的假設 (例如顏色方面或形狀方面)。不同於 246 作業，推理者在這類知覺作業中所面對到的問題反而是可能的假設太多，無法輕易的得知何者才是正確的規則 (排除多餘假設在日常生活中也常碰到)。我們過去的發現在這種情形下是否仍然成立？或者在這種情況下測試策略會比產生不同面向的新假設更為關鍵？實驗二將探討這個問題。

實驗一 難易度的影響

方法

採「2 4 6 作業」。除了上述數字題外，又增加姓名題（起始正例為「黃小英、蕭小紅、藍小美」）與撲克牌題（起始正例為「梅花 6、磚塊 3、黑桃 6」）。在前測中，我們分別給予 37 位台大的受試者上述三個起使正例，要求他們分別寫下所有可能包含此一正例的規則，以每一個規則被寫出的比例來決定其難易度。簡單題的正確規則約 50% 的受試者會提及，困難題的規則少於 20% 的受試者提到。實驗設計和各個狀況組中的正確規則見下表一。規則下方的數字則為各組受試者的人數，共 78 名。

受試者在開始時都先被告知他們的任務就是要去猜出某預設規則（例如：「由小排到大的數列」）。實驗者會先給予受試者一個符合此規則的正例（例如：「2、4、6」），並告知此預設規則所衍生的例子都由三個項目所組成。受試者先根據此正例形成第一個假設，然後進行 12 次的測試，在每一次測試中，受試者先提出一個自認為對猜出正確規則有幫助的例子，然後寫下他測試此例子的理由，接著實驗者會根據所測試的例

表二、各組成功率表（%，刮號中為成功人數 / 各組人數）

難易度 題目類型	數字題	姓名題	撲克牌題
易	61.54 (8/13)	71.43 (10/14)	50.00 (7/14)
難	23.08 (3/13)	33.33 (4/12)	33.33 (4/12)

子是否符合預設的規則而給予「是」或「否」的回饋。接著受試者寫下他目前認為最有可能的規則。完成作業的時間沒有限制，但不管猜出的規則對否，受試者要做完 12 個測試後才結束作業。

結果與討論

成功率

各組的成功率如下表三所示，除了撲克牌簡單題與困難題的成功率沒有顯著差異外，數字題與姓名題難易題目的成功率皆有顯著差異 ($p < 0.05$)。

假設種類、假設總數與測試策略

我們根據過去研究中的標準區分受試者的假設種類與測試策略（詳見林緯倫和連韻文，2001），並進行三因子變異數分析（題目類型 3× 難易度 2× 成敗 2）。結果題目類型的主效果在交疊性測

表一、難易度實驗各組起始正例與正確規則

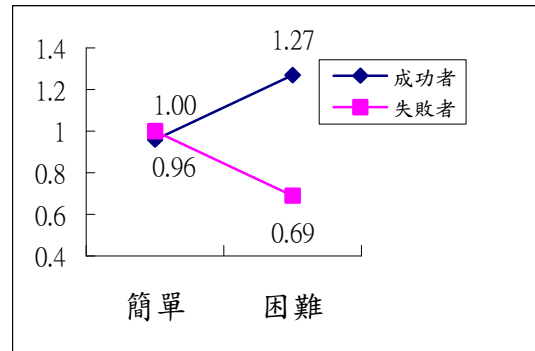
起始正例	數字題	姓名題	撲克牌題
	「2、4、6」	黃小英、蕭小紅、藍小美	梅花 6、磚塊 3、黑桃 6
易	正整數等差數列 N=13	三個字的姓名 N=14	第 1、3 張牌數字相等 N=14
難	由小到大排列的數字 N=13	任一字相同的姓名 N=12	中央數字最小的撲克牌 N=12

試與交疊性另有假設個數上達到邊際顯著的差異($p=.06$)。LSD 事後比較顯示，在姓名題中使用交疊性測試法顯著地多於數字題與撲克牌題；而在交疊性另有假設方面，撲克牌題的受試者顯著地多於數字題。此結果顯示不同的題目類型可能對交疊性另有假設是否容易產生有所影響。

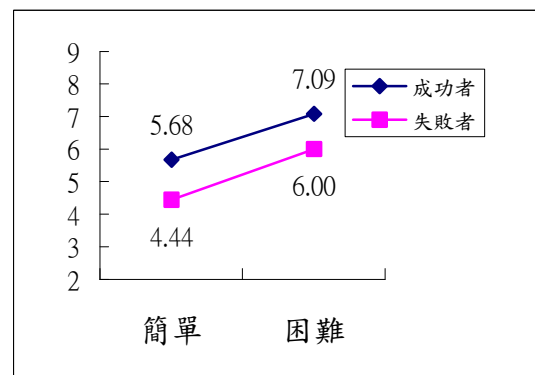
在難易度方面，受試者在困難題目上產生較多互包性另有假設($p<.05$)、有較多假設總數($p<.05$)，也較常採用互包性測試策略($p<.05$)與雙假設測試策略($p<.05$)。此結果可能是由於在簡單的題目中，大部分受試者很容易便能找出正確的規則，並不需要繼續再尋找其他的規則，但在困難的題目中，受試者較常發現自己原有假設有誤，必須繼續想假設，因此得到如上的結果。但有趣的是，困難題目的受試者雖然想更多的假設，也多半想出互包性的假設，這或許也是困難題成功率不高的原因之一，誠如我們之前的分析，在困難的題目中產生互包性的假設對成功與否並沒有幫助。

但難易度與成敗在假設種類或測試策略上並未有我們所預期的交互作用出現。雖然如此，我們仍進一步分析簡單與困難的題目中成功者與失敗者是否有所差異。在假設總數與不同假設種類方面，無論在簡單或困難的題目中，成功者與失敗者在假設總數、交疊性另有假設個數與互包性另有假設個數上均不達統計上的顯著差異，但我們發現其傾向並不違反我們原先的預期，即在困難題上，成功和失敗者在交疊性另有假設個數的差異大於簡單題，但在互包性假設與假設總數方面，成功與失敗者的差異不因題目難度而有改變(如圖一所示)。上述

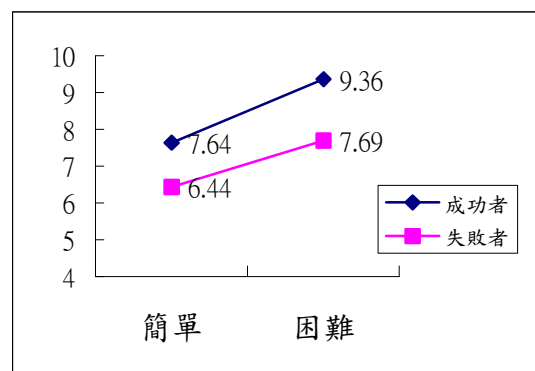
傾向並未達統計上的差異有可能在不同的題目中，受試者想到交疊性另有假設的容易度不同，致使交疊性另有假設個數的變異量增加，致使在統計上便較不易達到顯著的效果。



交疊性另有假設



互包性另有假設



假設總數

圖一、在不同難易程度題目中，成功與失敗者之假設總數與不同假設種類個數

實驗二 不同作業的影響

方法

本實驗所採用之規則發現作業改自 Bruner(1956)。採個別施測，共 24 名台大學生參與實驗。作業一開始呈現 192 張卡片，每張卡片包含五個向度的視覺特徵，分別為顏色(紅色或綠色)、中央圖形形狀(六邊形或五邊形或菱形或三角形)、中央圖形個數(一或二或三)、邊線(有或無)、角落星形(八角星或五角星或十字星或無星形)；卡片依序排列成矩陣。受試者被告知這些卡片分符合或不符合一個預定的規則，此規則與卡片上的圖案特徵有關，可能牽涉到一個或多個圖案特徵的組合，受試者的任務是找出預定的規則。

每位受試者做兩個問題(即找尋兩個不同的規則)，簡單題的規則為單一向度(例如：綠色)、困難題的規則為兩向度的聯集(例如：菱形或十字星)，題目的順序採隨機分派。每個問題開始時給予受試者一張已知符合規則的卡片，受試者先根據這張卡片形成第一個假設，接著受試者可依次選取多張卡片來做測試，實驗者每次都會根據這張卡片是否符合規則給予「是」或「否」的回饋。受試者每選取一張卡片都要寫下測試的理由，實驗者給予回饋之後再寫下受試者認為目前最有可能的規則，並評估此規則是否正確的信心程度(0-100%，0 為沒有信心，100 為非常有信心)。每一題在三十分鐘的測試時間內，受試者可做多次測試，直到受試者覺得非常有信心的時候便可公布規則，受試者事先被告知只有一次公布規則的機會。

結果與討論

由於受試者在簡單題中全數在時間內找

到正確規則，在此先不予討論。在困難題中(規則為聯言形式，A 或 B)，則有 37.5% 的受試者在時間內找到規則。如所預期，受試者在一開始就可同時形成多個假設(以「和」、「或」組合不同面向特徵)，但我們在計算假設個數時是以受試者真正拿來測試的假設為準。此外，由於受試者假設的修正多為增加或減少特徵數或是改變同面向內特徵的值，一般皆視為只是原假設範圍的修正，只有當新假設包含不同於原有假設的面向，且被當作獨立的假設來測試以與舊假設區辨時，才視為新角度的假設。

結果顯示成功者所得到的反例數(「否」回饋的次數)顯著多於失敗者($p=0.06$)。交疊性另有假設數(新角度的假設)與交疊性測試法也皆顯著多於失敗者(兩者皆 $p<0.03$)。互包性假設數、假設總數、其餘的測試法、以及否證數等皆無差異。

上述結果和我們過去的發現雷同，顯示在不同性質的假設檢驗作業時，能產生與測試新角度假設仍是推理成功者與失敗者主要的差異。仔細觀察成功者的策略，可以簡單分為兩個關鍵點：1) 找到一個反例(若規則為 A 或 B，則反例為 $\sim A$ 且 $\sim B$)，這是不同於以往的發現。2) 若確定一個變項為規則後，在測試其他變項時會將其排除(亦即去除混淆變項。若受試有此測試，我們會將其視為交疊性測試，因為兩個假設沒有互包的關係)。相反的，失敗者則常是沒有去除混淆變項的動作或是找到一個可能假設後，沒有進一步測試此假設的反例(若得到否證，則知道有第二個規則)。

有趣的是受試者的測試順序多像 Bruner (1956) 所描述的有系統地一次改變一個變項值，來測試該變項是否是正例

的必要條件。Bruner 及其他研究者多認為這是一個對且有效的測試策略，但若規則是聯言形式的，採用此策略的受試者會錯誤地推論出所有變項都是無關的或是只找到一個可能的假設，除非他能有意或無意地得到反例，將其視為比較基礎，再逐次改變各個變項。事實上，在因果推理或實驗法中，這就是在排除其他所有可能原因後，比較一個變項出現與不出現時效果的差異。以此角度觀之，有系統的一次變一個變項並非規範性的策略。它在規則是單一變項或變項的交集時，不失為一簡便的策略，但在規則包含兩個變項的聯集時（亦即有多重互相獨立的原因時）便容易導入死胡同。

參考文獻

- [1] 林緯倫和連韻文(2001), 科學教育學刊, 第九卷第三期, 1-24。
- [2] 連韻文(1998): 科學資優生哪裡資優? --以假設檢驗能力為例。資優教育二十五週年研討會論文專輯, 135-147。
- [3] Bruner, J. S., Goodnow, J. J., & Austin, G. (1956, 1986). A study of Thinking. New York: Wiley.
- [4] Gorman, M. E., & Gorman M. E. (1984). A comparison of disconfirmatory, confirmatory and control strategies on Wason's 2-4-6 task. *The Quarterly Journal of Experimental Psychology*, 36A, 629-648.
- [5] Gorman, M. E., Stafford, A., & Gorman M. E. (1987). Disconfirmation and dual hypotheses on a more difficult version of Wason's 2,4,6 task. *Quarterly Journal of Experimental Psychology*, 39A, 1-28.
- [6] Klayman, J., & Ha. Y.W. (1989). Hypothesis testing in rule discovery: strategy, structure, and content. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 596-604.
- [7] Popper, K. R. (1959). *The logic of scientific discovery*. New York: Basic Books.
- [8] Tweney, R. D., Doherty, M. E., Wornor, W. J., Pliske, D. B., Mynatt, C. R., Gross, K. A., & Arkkelin, D.L. (1980). Strategies of rule discovery in an inference task. *Quarterly Journal of Experimental Psychology*, 32, 109-123.
- [9] Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12, 129-140.