

## 2

### 觀察與理論評價

科學建立在已知的東西上，此建構過程不僅在於增進我們的具體知識，而且在於擴大我們用新方法觀察自然的能力，增加我們認識自然的能力。

—Dudley Shapere, *Reason and the Search for Knowledge*, p.351

本章旨在表明，實用論對於**理論與觀察**關係的看法，不會威脅「競爭理論可比較優劣」的主張。實用論的「觀察理論」以及「經驗資料不充分決定理論之論題」分別是：觀察已經預設某些東西，不是純粹中立的，它無法充分決定特定理論；理論若遭遇觀察異例，可透過修改其中的任一部份使之與觀察相符。這兩種看法再推遠一點，就走向相對主義或主觀主義的極端。本章將批判此立場並論證，觀察者本身的各種預設，固然會影響觀察陳述的中立性，但由此不必然導出「競爭理論彼此無法比較優劣」。待檢驗理論偶爾可以透過修正其中部分，使之與觀察上的異例相符合，但不是所有理論遭遇觀察異例，皆可透過修正而與競爭理論等價並獲得保留。

以下我將提出四個理論評價特性，並論證(1)傳統理性論把「理論」與「經驗」之間的聯繫，看成僅僅是邏輯關係，觀察述句具有純粹中立的特性。這種看法有其缺失。(2)迪昂、孔恩、韓森等實用論者指出，不存在中立的觀察述句，觀察都是理論負載的 (theory-laden)，不能做為支持特定理論的基礎，這結果挑戰了理論評價的客觀性。「理論負載」的意涵有強弱兩種，「弱意義」的理論負載不會構成評價困難；「強意義」理論負載雖有威脅性，但還不至於導致競爭理論無法比較優劣。(3)回應迪昂與蒯因所提的三個「不充分決定」(under-determinate) 論題：不存在「判決性實驗」否證特定理論；科學知識是整體地受檢驗，某個理論與觀察不一致，無法判定何者應修改；理論永遠可以透過修正其中部份，使之與觀察相符合。

## 2.1 理論評價的特性

不論是理性論或實用論者，幾乎沒有人對理論評價特性做出完整而具體的建議。我們在此勾勒四個理論評價的特性，作為一個方法論「原則」。

**局部整體的** (locally holistic) 理論的評價的單位不是個別的孤立理論，而是一組相關的理論群或理論系列。孔恩稱之為「典範」(paradigms)，拉卡托斯稱「研究綱領」(research programmes)，圖爾敏稱「知性傳統」(intellectual traditions)，勞登稱「研究傳統」(research traditions)，費耶阿本則直接稱之為「理論」(theories)。<sup>1</sup>本文跟隨費耶阿本，以「理論」指稱這類或大或小的理論系列。「局部整體」在本文有兩層含意。一、就科學知識內容而言，理論評價是局部整體的：「整體」意味評價工作，不再是邏輯實證論那種純粹理論結構的分析，「局部」意味科學知識不是整體一塊不可分割，科學變遷也非全然牽一髮而動全身。二、就認知主體（科學家）與知識的關係而言，科學理論的接受不是認知者整個「世界觀」的信念跳躍。科學變遷過程在理論、方法論和目的底各個層次之間，可以是三者逐步轉換，無須僅訴諸心理學上的格式塔轉換 (gestalt switch)。

**動態的** (dynamical) 科學知識隨時間而發展，人們的經驗觀察<sup>2</sup>、科學理論和方法論也隨時間在變動。科學家對於理論的評價，不必限於某個時間點立刻決定，評價工作依賴相關輔助理論、證據、儀器的出現，所需時間並不一定。從科學史和科學哲學家的角度，「評價」永遠是事後的工作。「動態的」意謂理論是在一段時間內，而非在一個時間點上來評價。從科學家的角度，「動態」意指競爭理論必須在發展過程中一起來評價，在相互比較之下，無須等待「判決性實驗」出現才能做出合理的判斷。科學家依賴陸續出現的證據修正其判斷，在此動態意義上，沒有所謂「即時理性」(instant rationality)。

**可錯的** (fallible) 知識都是會出錯的，任何觀察、理論、方法甚至目的等，都不是絕對無疑不可更改的。方法論的規範性不具有先驗的 (a priori) 意涵。沒有「無時間性」(timelessness)、「絕對不變」的方法論規則，科學目的也不是絕對單一的。方法論與科學理論一樣受經驗的檢驗，同屬於經驗的研究，同樣被視

---

<sup>1</sup> Kuhn (1970), Toulmin (1972), Lakatos (1978), Laudan (1977), Feyerabend (1985)。

<sup>2</sup> 「經驗」一詞較為普遍，通常包含「觀察」。本文用法跟隨 H. Brown, *Observation and Objectivity* (1987) 的觀點，將兩者等同視之。

為一種應用的手段和說明工具。我們以對此世界的最佳信念，改進並修正研究方法，使我們方法論更完備；這些方法論反過來得到更有效的應用，因而引導出新的信念，根據這些信念又可使方法論更趨於完備。據此，方法論與科學理論一樣是可錯、變動的。

**比較的**（comparative）自然現象容允多種說明或描述方式，一定會產生很多相互競爭的理論，因此理論優劣的評價是必要的。單一理論當然可以進行評價，但合理的判斷不在於對單一理論的評價，而是對於兩個或多個競爭理論的評價，如此才有「理論選擇」可言。理論優劣的評價，永遠是兩個以上的競爭理論相對比較的情況。所謂觀察驗證（confirm）理論，不是觀察與特定理論兩方對決，而是觀察與諸競爭理論的多方比較。若現有的經驗觀察相對傾向支持某個理論而不是另一個理論，那麼選擇證據比較支持的理論是合理的。

所謂「理論選擇」，並非意謂從競爭理論中挑選出方法論唯一支持的理論，而排除所有其他可能的競爭者，而是要在現有科學競爭市場上角逐的所有理論中，「選擇」一個得到證據最大支持的理論。我們（科學家）無法知道所有競爭理論是否窮盡了該研究領域的可能狀況，因此，與其說我們在尋找最佳理論，不如說是在挑選**所能找到**的最佳理論。當然，評價競爭理論的過程中，沒有絕對可靠的東西，證據認定和推論都是可錯的，但透過隨時反覆地檢驗與修正，仍可維持理論選擇過程的客觀性。知識社會學家和相對主義者的興趣，大都集中在經驗證據和方法論規則**不充分決定**理論的問題上。他們部分的看法是對的，亦即在特定時間內，若只依證據的支持程度來選擇競爭理論，由於競爭理論可能有相同證據的支持，方法論規則有時候還不足以決定該選擇哪一個理論。

不過，特定時間內的證據無法充分決定一個理論，不表示陸續出現的其他證據無法決定哪一個理論較好，更不意味這些證據**永遠不能**充分決定某個理論。評價理論的過程是動態的，證據無法充分支持某個理論可能是一時的。我們對於哪些證據支持哪些理論，可以透過方法論規則的幫助達成共識。在此，區別傳統上「選擇」（choice）與「偏重」（preference）的意涵與用法是有必要的。<sup>3</sup>傳統方

---

<sup>3</sup> 見 Lakatos (1978), Hempel (1983), Laudan (1984 p.29)。韓沛爾 (1983) 指出方法論規則的「比較」特性，勞登在《科學與價值》強調此特性。拉卡托斯以比較特性論述否證 (falsification) 概念，觀點與孔恩相似，他說：「任何實驗、實驗報告和觀察述句或充分認可的低層否證假設都不能單獨導致否證。在還未有較好的理論出現之前不會有否證。(…) 否證不僅僅是理論與經驗基礎之間的關係，而是相互競爭理論、原先的「經驗基礎」，以及由競爭產生的經驗增長之間的多邊關係。」(1978 p.35)。基切爾則技術性的發展「消去歸納法」(eliminative induction)，

法論者對於「選擇」一詞的用法有較強的意義，通常意謂挑一個絕對最佳理論的判斷。「偏重」則是一種比較的工作，意謂在所有可選擇理論中，相對來判斷哪一個理論較好。因此，雖然我們還無法在競爭理論中（如早期氧氣與燃素理論），作出充分確定的理論選擇時，還是可以挑出比較偏重的理論。在某些不充分決定的情況下，現有證據仍能充分決定一個偏重的理論，亦即，現有證據足以排除某個理論的主張，相對支持另一個理論。孔恩說：「在科學實踐中，真正的驗證問題永遠牽涉兩個理論的比較，以及兩個理論與這世界的比較，而不是單一理論與這世界的比較。」<sup>4</sup>這種驗證上比較的特性顯示，當方法論規則不足以對特定理論做出決定性判斷，依然可展示競爭理論中的某一個是比較受到證據支持，選擇該理論是合理的。這種比較、動態和局部整體的評價特性，傳統理性論不太重視，以下我們就來檢視傳統理性論對於「理論與觀察」關係的看法。

## 2.2 傳統理性論的觀點

傳統理性論的科學哲學研究，大都專注於科學語句的邏輯分析，幾乎不關心理論選擇的問題。理性重建模式也是只描繪科學理論的結構（見 1.1），說明不可觀察的理論實體和術語如何取得經驗意義，並沒有實際說明，經驗述句在何種條件下可成為特定理論的證據或反證；或何種理論與證據的條件關係，才可判定一個理論為可接受的（為真、趨近真等等）。諸理論「理性重建」的金字塔建立在先驗的邏輯基礎上，因此可視為評價理論的普遍適用之架構：只要能在其中找到適合位置的理論，就是值得接受的理論。「理論金字塔」與「經驗泥土」之間的聯繫，只需透過邏輯分析，就能判斷個別經驗對於理論的驗證（confirm）或反證（disconfirm）。寇瞿（Noëta Koertge）指出，金字塔模型至少必須有兩個預設：一是理論術語意義的「毛細作用模型」（Capillary Model），即科學理論的認知意義皆從觀察基礎的泥土滲透上來；二是驗證的「拔靴帶模型」（Bootstrap Model），即支持理論的證據都來自於單稱的觀察述句。<sup>5</sup>邏輯實證論的檢證論

---

同樣以「比較」的概念精鍊傳統的驗證理論（Kitcher 1993 ch.7）。以「比較的」方式「偏重」某個理論的看法，在 Carnap（1968）已略有提及。

<sup>4</sup> Kuhn（1977a p.211）

<sup>5</sup> Koertge（1972）。「毛細作用模型」意指將不可觀察的理論術語翻譯成可觀察語句的機制。例如卡那普（Carnap 1966）所謂「對應規則」（correspondence rules）、布里基曼（P. W. Bridgman 1927）的「操作定義」（operational definitions）或坎貝爾（Norman R. Campbell 1920）所謂的「字典」（dictionary）。「拔靴帶模型」意指全稱理論語句受單稱觀察語句支持或反駁的機制。例如早期邏輯實證論的「可檢證性原則」（principle of verifiability），卡那普（1936-37）的「驗證論」（confirmationism），也包括波柏的「否認論」。

(verificationism) 與波柏的否證論 (falsificationism) 立場，分別以歸納與演繹邏輯作為理論與經驗之間的「拔靴帶」機制，兩者皆依賴某種純粹中立的經驗基礎。依照檢證論觀點，觀察述句被視為無理論預設的中立語言，此中立的「所予」(given) 是所有科學理論的基礎，邏輯實證論者稱此經驗基礎為「記錄語句」(protocol sentence)。<sup>6</sup>當我們觀察到某些事件總是伴隨出現，如果 R 則 S，會試圖將這個伴隨出現的語句普遍化。例如「每當黎明時(Rx)，太陽從東邊升起(Sx)」(假想「 $Rx \rightarrow Sx$ 」的關係)，「現在是黎明(Ra)」那麼「太陽會從東邊升起(Sa)」。檢證的機制依賴歸納法 (induction)，用邏輯式簡單表達：

$$\begin{array}{c} Ra \ \& \ Sa \\ Rb \ \& \ Sb \\ \vdots \\ Rn \ \& \ Sn \\ \hline \therefore (x) (Rx \rightarrow Sx) \end{array}$$

「 $Ra \ \& \ Sa, Rb \ \& \ Sb, \dots$ 」為觀察述句，「 $(x) (Rx \rightarrow Sx)$ 」為一個普遍假設。我們觀察到 R 與 S(「 $Rx \rightarrow Sx$ 」)總是伴隨出現，會企圖歸納出一個普遍或全稱(universal)述句「 $(x) (Rx \rightarrow Sx)$ 」當作一個假設。不過，此結論的普遍述句，理論上蘊含無限多組觀察述句「 $R \ \& \ S$ 」，科學實踐上無法以有限的觀察充分檢證「 $(x) (Rx \rightarrow Sx)$ 」，亦即，現實上無法窮盡這世界所有可能的  $R \ \& \ S$  (n 為無限)。因此，我們只能以歸納所得檢證至某種程度，檢證的機制永遠不能絕對確定，亦即邏輯上不允許我們得出普遍述句的結論。<sup>7</sup>

<sup>6</sup> 關於記錄語句的特性，邏輯實證論內部雖存有不同意見，但大致都將其視為建構知識的基礎。見 Carnap (1932, 1934) ; Neurath (1932 ch.7)

<sup>7</sup> 檢證 (verification)、驗證 (confirmation)、檢驗 (test) 等語詞翻譯，皆參考 Robert Audi 英文版主編/王思迅主編，《劍橋哲學辭典》(台北：貓頭鷹出版社，2002)。在此大略交代檢證與驗證的區別。檢證論的原則會將所有科學普遍法則排除於有意義命題之外，如同排除形上學命題一樣。早期檢證論的原則只能針對單稱命題或述句作檢證，若有人拿了一張石蕊試紙聲稱「這張石蕊試紙是紅色」，那麼這述句直接可用觀察來檢證。而所有科學的普遍法則都是全稱命題或述句的形式，例如「所有結晶鹽粒皆可溶解於水」，當我們觀察一匙鹽溶解於水，這只是前述全稱述句中的一個個例(instance)，而全稱述句所指涉的個例數目卻是無限多的，再多的觀察個例也無法窮盡檢證一個全稱述句。換言之，一個普遍的假設邏輯上蘊含無數個經驗觀察結果，實際的經驗觀察並無法窮盡所有這些結果。因此科學法則無法符合檢證原則，也無法依其要求化約成為經驗述句。為了避免科學普遍法則被檢證論認定為無意的命題，卡那普發展驗證理論取代檢證原則，成為一種較為寬鬆廣義的檢證論立場。在〈可檢驗性與意義〉(“Testability and Meaning”)一文中，卡那普主張：「在某種意義下，一個語句的意義等同於我們決定它為真或假的方法；而

波柏反對歸納推論，他認為只有演繹法（deduction）能提供科學客觀性與確定性。「確定性」不是指可確定一個科學理論這類全稱述句為真，而是指可確定它為假。換言之，我們雖不能以歸納推論確定一個理論為真，但卻可用演繹推論確定其為假。波柏以演繹法為基礎提出否證論觀點，檢驗基礎就是大家一致同意的「基本述句」（basic statement）。邏輯形式為「否定後件因而否定前件式」（*Modus Tollens*），在此用語句邏輯式較能清楚瞭解：

$$(H \rightarrow O) \ \& \ \sim O \ \rightarrow \ \sim H$$

H 為一假設的全稱述句，O 為單稱述句。原本 H 應該蘊含 O，但若得到一個觀察  $\sim O$ ，就足以反駁 H。這種確定性由邏輯有效性賦予，一個理論與多次的實驗結果相符，不表示該理論就更加可靠，因只要出現一個反例，理論就該拋棄。<sup>8</sup>

在此，不論是「記錄語句」還是「基本述句」，都被當作是客觀中立的基礎，理性重建的金字塔有此基礎才能穩固。卡那普將之視為中立純淨的語言，波柏則把它當成是科學家之間的約定（convention）。雖然兩人都承認此經驗基礎並非絕對而不可矯正，但都認同觀察的單稱述句比理論的全稱述句更純粹且中立，經驗意義可由此基礎「滲透上」理論金字塔。他們幾乎不考慮觀察述句是否預設科學家心理與社會因素，以及由此產生的評價理論的問題。實用論者就在此關節指出一個問題：「觀察」既不純粹也不中立，由此導出的檢證或否證都無法充分支持或反駁一個理論。簡言之，「理論金字塔」與「經驗泥土」之間的橋樑並非一條客觀清晰的「邏輯毛細管」，除了邏輯以外還有其他成分，例如科學家的心理和社會的主觀因素。波柏看出這個缺陷，他將科學客觀性的基礎寄託在一致同意的

---

一個語句之具有意義僅當如此的決定為可能。（……）如果檢證（verification）就是指真（truth）的限定與最後的建立，那我們將發現沒有任何一個（綜合）語句是可以永恆檢證的。我們只能不斷驗證（confirm）一個語句而已。」（1936 §1, p.420）「普遍語句永不可能被檢證，例如物理與生物學法則。（……）我們雖不能檢證該法則，但卻可藉由檢驗（test）此法則的個例以檢驗其本身。（……）如果在此檢驗實驗的繼續序列中沒有反例出現，而只有正面例證數目增加，則我們對該法則的信心將可逐步地建立。因此，我們寧願稱此檢驗過程為逐漸增進該法則的驗證，而不稱為檢證。」（§3, p.425）如此，科學法則這類全稱述句在卡那普的驗證理論中就具有意義了。全稱述句與單稱述句之間的可檢證性差異，因而沒有本質上的不同，只有程度上的差別。科學法則這類述句，需要更多的正面檢驗例證以增加我們對它的信念。亦就是，每通過一次實驗觀察的檢驗，就更增加該法則或理論的驗證程度（degree of confirmation），該理論因而更可靠。

<sup>8</sup> 我在此刪略「檢驗條件」的表示。若跟隨前例以述詞邏輯表達為： $(x) (Rx \rightarrow Sx)$ ,  $Ra \ \& \ \sim Sa$  /  $\therefore \sim(x) (Rx \rightarrow Sx)$ 。只要得出一觀察事件（ $Ra \ \& \ \sim Sa$ ），就足以反駁全稱述句  $(x) (Rx \rightarrow Sx)$ 。

「基本述句」。基本述句的接受是一種決定（decision），無法再給理由證成，如果還有人不同意公認的基本述句，只能將之忽略不計。<sup>9</sup>波柏認為，只有這類大家都同意的底層經驗才是客觀的，「客觀性」意謂「互為主觀地可檢驗」（inter-subjectively tested）。波柏說：「科學述句的客觀性在於它們能夠被互為主觀地檢驗這一事實」。<sup>10</sup>理論金字塔建立在此客觀的經驗基礎上，人們才能斷言整個科學事業是理性、進步的。不過，實用論者如蒯因、孔恩等人批判理性重建模式，質疑所謂「客觀的」觀察，直接威脅了科學的理性基礎。

蒯因很早就指出理性重建模式的缺陷。他認為理論與證據之間（信念/感覺與料）的鴻溝無法以傳統知識論的作法彌補，主張從心理上的因果關係來看兩者。心理學的經驗研究顯示，每個人的感覺受器（sensory receptors）即使受相同的外在刺激，不一定每個人都產生相同的反應。<sup>11</sup>孔恩、費耶阿本則認為，觀察必須透過某些理論或背景知識來詮釋，這種「受理論污染」的經驗不能成為科學事業客觀性的基礎。就算理性論的分析沒錯，理論與觀察純粹是邏輯推論的關係，但若無法確定觀察的中立性，也不能充分判定哪一個理論才是可接受的。現在問題在於，若觀察是中立的，那麼所謂「經驗證據」即可充分決定，「理論」與「經驗」兩者之間就無疑只是邏輯推論的關係。但是，「經驗」或「觀察」並非是清楚的概念，「證據」便無法充分確定。我們都是以實驗觀察來決定「證據」，但觀察必須預設某些背景知識，否則無法辨別呈現的經驗內容為何，更別說決定何者為「證據」。一個世紀之前，法國物理學暨科學史家迪昂（Pierre Duhem）舉了一個影響深遠的例子，揭露觀察的不確定與模糊性：

走進這個實驗室，接近這張擺滿各種儀器的桌子：一個電池、裹著絲綢的銅線、裝滿水銀的管子、一些線圈、一根裝上鏡子的小鐵棒。觀察者將鐵棒裝上橡皮圈，投入小洞中，這鐵棒將會上下震盪，並藉由其上鏡片，把一道光線送上賽璐珞直尺（celluloid ruler）上，而觀察者則注視其上那道光線的運動。無疑地，在此有一個實驗；透過光點的振動，這

<sup>9</sup> 波柏認為基本述句不等於是「真」，必須經過多次的檢驗才能接受。這種述句的特色是，存在著某種數學或邏輯的技術，任何人學會它都能斷定該陳述是可以接受的。但這不表示經驗基礎的決定，我們能夠給出判定規則。如果有人對已廣為接受的陳述提出質疑，波柏也只能建議此人，指出陳述錯誤之處或再去檢驗看看。例如，若有人在一個實驗情境下，無法同意公認的基本述句「這張石蕊試紙是紅色的」，波柏只能建議他再檢驗；若他還有懷疑，只能忽略他的意見。詳見 Popper（1968 pp.98-99）

<sup>10</sup> Popper（1968 p.44）

<sup>11</sup> 見“Epistemology Naturalized”, in Quine（1969）

位物理學家精細地觀察這根鐵棒的擺盪。如果問他正在做啥。他難道會回答說：「我正在研究這根帶有鏡片的鐵條的震盪運動」嗎？當然不會，他會告訴你，他正在測量線圈的電阻。如果你感到詫異並問他，這些名詞是什麼意思，而這跟他和你所同時知覺到的現象又有什麼關聯，他將會告訴你這說來話長，並建議你先去修完電學再說。(1991 p.145)

如果沒有電學知識，旁觀者如何知道此實驗者正在做電阻的實驗？若沒有預設相關的知識訓練，不同的觀察者對於相同的現象會有不同的詮釋。顯然，每個人的觀察必然預設程度不一的背景知識。

### 2.3 觀察的理論負載與不充分決定

迪昂指出：「物理學中的實驗是對現象的精確觀察，同時伴隨對這些現象的詮釋；這種詮釋藉助觀察者接受的理論。」<sup>12</sup>迪昂、韓森、蒯因、孔恩、費耶阿本、赫西 (Marry Hesse) 對此雖各有不同的表述，但都抱持類似的看法，即理論與觀察之間的驗證或否證，不如傳統理性論所言那樣簡單，甚至造成理性論的困難。他們的論點雖不盡相同，但可粗略地通稱為「整體論」(holism) 觀點。他們認為，經驗觀察的描述都是理論負載 (theory-laden, theory-loaded) 或理論依賴 (theory-dependent)，也就是說，科學知識所依據的經驗基礎，都是我們使用某些理論所做的詮釋。若任何觀察都負載了理論，那麼我們就無法以此觀察來支持或反駁待檢驗的理論，因為在這種理論與觀察衝突的場合，我們不知道是觀察所負載的理論有問題，還是待檢驗的理論有問題。他們認為，知識無法從絕對清晰的基礎開始，而是各種觀點的集合體；構成知識的內容是整體的，理論與經驗述句難以清楚劃分。科學家的研究方法、實驗儀器及所受的教育訓練，與個人成見、社會和歷史背景緊密結合，他的信念無法與公認的背景知識截然分開，當他陳述觀察現象或實驗結果時，已經將所有背景知識摻入。蒯因比喻，我們接受的知識就像一張信念網 (web of belief)，赫西則形容，我們接受的理論如同一個網路系統 (network)。<sup>13</sup>所謂觀察述句，事實上是由觀察對象和科學家接受的理論共同決定，沒有「純粹的」觀察述句。韓森稱這種觀察述句必然伴隨普遍理論為「理論負載的」<sup>14</sup>。費耶阿本則是更激進地主張：「觀察述句不僅僅是理論負載

---

<sup>12</sup> Duhem (1991 p.147)

<sup>13</sup> Quine (1951), Quine & Ullian (1978), Hesse (1980 pp.83-87)

<sup>14</sup> Hanson (1965) p.19ff。韓森在同一本書中使用“theory-laden”和“theory-loaded”。



的，而且是完全理論的。」<sup>15</sup>人們觀察到的「東西」不管如何簡單，都只是一個鬆散的整體，必須藉助理論概念將其分割、詮釋，經過分類以後，再由人們將它陳述出來，因此所陳述的觀察句早已受到理論的污染。觀察的理論負載觀點似乎威脅科學的客觀性基礎，納格爾憂心地指出：

每個觀察述句的含意因此由研究者所接受的某些理論來決定，所以一個理論的適當性就不能經由理論中立（theory-neutral）的觀察述句來判定。如果這些批判是正確的，它們顯然導致一個影響深遠的「知識相對主義」（relativism of knowledge），一種關於實現保證自然知識可能性的懷疑主義。這個立場遠比曼海姆（Karl Mannheim）與其他知識社會學的相對主義更為徹底。<sup>16</sup>

若我們堅持研究者的觀察皆依賴某些理論，那麼就會導出一個結論：依賴不同理論的研究者對同一現象會有不同的詮釋。亦即，相同的現象由不同的人來觀察，可能產生不一致的觀察述句，如此便不會有普遍認可的證據。因此，這種觀察理論再走遠一點，等於斷言競爭理論中不存在可作為裁判的中立證據，理論之間彼此是不可共量的（incommensurable）。<sup>17</sup>

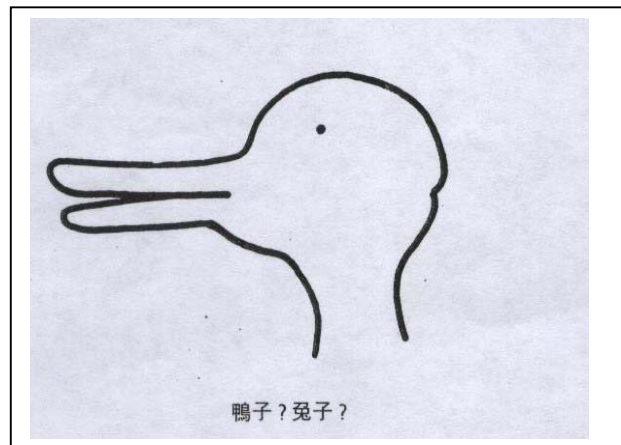


圖-2

以認知心理學的實驗「鴨兔（duck-rabbit）圖」為例（圖-2），有人將之看成鴨子，有人將之看成兔子。以此實驗類推至科學知識的建構上，可以宣稱：觀察

<sup>15</sup> Feyerabend（1988 p.229）加底線字部份是費耶阿本所強調。

<sup>16</sup> Nagel（1979 p.31）

<sup>17</sup> 「不可共量」在此章意指：競爭理論之間沒有客觀證據或判準可比較優劣。

者依照各自的信仰、背景知識、世界觀來詮釋觀察到的現象，以此觀察為基礎建構出不相容的競爭理論，沒有中立的觀察能提供我們在競爭理論中決定一個較好的理論，造成競爭理論彼此不可共量。如同孔恩論述：「不同典範的支持者在不同的世界中執行它們的行業」，「兩群在不同世界中執業的科學家從同一點注視同一方向時，他們看到不同的東西。」<sup>18</sup>拉瓦謝（A. L. Lavoisier）當作是氧氣的東西，普里斯特里（J. Priestley）卻看成是除去燃素的空氣（dephlogisticated air）。<sup>19</sup>韓森則認為，克普勒（Johannes Kepler）和第谷（Tycho Brache）一起站在山丘上看日出，「克普勒將太陽當作是靜止的，地球在運動。但第谷幾乎是跟隨托勒密和亞里斯多德：地球是固定的，而其他天體繞著地球運動。」<sup>20</sup>今天，雖然我們共享關於太陽的知識，而且都具備相同的背景知識。「然而第谷和克普勒看相同的事物，卻看到不同的東西。這些不同的東西可說是依賴於他們的知識、經驗以及理論。」<sup>21</sup>換言之，在第谷、克普勒以及拉瓦謝、普里斯特里等人各自的時代裡，每個人理論預設不同，產生不同的觀察述句，而這些「觀察」皆可支持各自的理論，相互競爭的理論彼此不能比較優劣。

在此必須強調，部分（後來的）理性論者也承認觀察負載了「某些東西」，不是所有理性論者都徹底堅持觀察句的「純粹中立」性，他們僅反對由此可以推導出「理論是不可共量」的主張。面對實用論者的觀察理論，我們要先問：所有觀察述句皆負載「理論」，此「理論」是什麼意思？首先，主張「觀察完全是理論負載的」是有問題的。海金（I. Hacking）指出，「當我們說及任何東西時，我們都有各種期待、偏見、意見、工作假說和習慣。」<sup>21</sup>一個外科醫師描述某人身體上的創傷，的確隱含某種情境，例如想像此人是在打架或作戰中受傷，但這種假想絕對稱不上是具有「理論」預設，頂多只是一種經驗上的猜測或期待。部分哲學家使用理論負載的「理論」一詞過於模糊，所謂「所有觀察都是完全理論的」難以清楚界定所指為何。<sup>22</sup>費耶阿本經常將「理論」泛指觀察者的背景知識、信

<sup>18</sup> Kuhn (1970 p.150, p.118)

<sup>19</sup> Hanson (1958 p.5)

<sup>20</sup> Hanson (1958 p.18)

<sup>21</sup> Hacking (1983 p.176)

<sup>22</sup> Shapere (1982; 1984 ch.16) ; Hacking (1983)。加里森（Peter Galison 1987）反對理論負載觀點的無限擴張，他考察現代高能物理的實驗活動研究，認為「觀察負載理論」言過其實。他批評：「格式塔心裡學的模糊暗示，或許已成為反抗獨斷實證論的有效戰略，但實驗學家真正關心不是全體世界觀的改變。在實驗室中，科學家企圖找尋局部方法以排除或至少定量背景基礎，並瞭解信息是否失去，然後改正系統錯誤。格式塔轉換的語言太過簡化地描述二十世紀早期實驗物理學的每天實踐工作。（……）特別是當哲學家宣稱實驗數據是理論負載的，他們有一種將所有觀念和成一團的傾向。」（1987 p.245）一個廓清這類理論預設的觀點，就是去分析「理論」名

念、習慣與成見，甚至是一些不成熟的猜想等。例如在《反對方法》中，他主張科學研究應該如同人類學的研究方式，科學家的「觀察」，牽涉文化、社會因素，包含科學家個人所承受的教育、宗教、神話、心智狀態甚至文學、藝術薰陶等等。<sup>23</sup>在此，我稱此為「弱」意義的理論負擔，而與以下將談論的「強」意義區分開來。在「弱意義」上，每個觀察述句負擔了所有可能影響觀察者的因素；而「強意義」是指，觀察負擔了特定的理論體系，該理論體系自成一個研究傳統。在弱意義上的理論負擔，並不會威脅理論選擇的客觀性，它甚至與科學實際研究不相干。不可否認，科學家是在特定社會、文化和歷史下做研究，會受到各種影響。但科學家身處不同背景脈絡，不表示他們無法跨越特定脈絡取得觀察上的共識。夏佩爾指出「觀察受污染」的說法過於誇大，關鍵在於：「認為科學的觀察是有背景信息『負擔的』，這事實不蘊含該觀察被『裝填了』贊成任意的（arbitrary）或相對的，或任何實用意義上『不確定』（uncertain）觀點。」<sup>24</sup>類似費耶阿本弱意義的理論負擔觀點，忽略現實狀況：實踐的科學家經過專業訓練，各社群科學家們的觀察背後的預設，還不至於南轅北轍；而且有些觀察甚至無須預設理論。<sup>25</sup>科學家的觀察除了文字和口頭上的陳述、報告，還有實際操作，這些觀察不見得必須得預設儀器和設備的知識，才能有「觀察」上的共識。例如，望遠鏡剛發明之初，在沒有共同理論預設情況下，也能達成觀察上的共識（見 5.4）。而要求學生至野外觀察一隻猩猩的活動，或觀察某個行星的運動，這類觀察也不一定需要專業知識引導，才能獲得具有共識的陳述。

「弱意義」的理論負擔攻擊理性論的火力較弱，因為觀察負擔了影響科學家本身的所有因素，但這些因素不必然會左右科學家的研究工作，進而影響理論評價的客觀性。但「強意義」的理論負擔，是理性論需要面對的挑戰。強意義的觀點認為，不同觀察者依據「融貫的理論系列」（孔恩的「典範」、拉卡托斯的「綱領」）詮釋他所觀察到的現象，而各自依據的理論彼此又不相容，造成競爭理論沒有客觀中立的「觀察」來評價優劣。以韓森的例子來說，第谷與克普勒在相同境況下，觀察同一現象（拂曉日出），第谷抱持亞理斯多德和托勒密體系，克普勒則相信哥白尼體系；因而第谷看到運動中的太陽，克普勒眼中的太陽卻是靜止的，兩人不同的「觀察」，都能支持各自的理論。

---

稱之下所進行的活動。

<sup>23</sup> Feyerabend (1988 ch.16)

<sup>24</sup> Shapere (1982 p.516)

<sup>25</sup> 例如海金 (I. Hacking) 指出冰島晶石 (Icelandspar) 的雙重折射和偏振光 (polarized light) 現象例子，在十七世紀發現此現象時，不能說負擔了「理論」(1983 ch.9)。

強意義的理論負擔也非全然困擾理論評價的客觀性，我將「強意義」區分成兩種情況。理論選擇在一種情況下，科學社群並不會產生太大的歧見：某個理論預測了「新現象」，而其他競爭理論做不到。亦即，先產生理論預測，而後該理論引導實驗，觀察到新事實。一個理論蘊含某些未知的觀察結果，而後觀察上實現所預測的現象，科學家對這種結果的歧見不大。這種情況可能發生在舊理論無法說明某些現象，而急於以其他理論替代，或新理論戲劇性地預測出前所未見的「事實」。傳統理性論大多認同這種理論引導實驗觀察的情況，例如波柏說：

理論家將某些特定問題交給實驗家，而實驗家便僅僅針對這些問題，以實驗嘗試找出答案。他試圖排除其他問題。(……) 但以為實驗家「是為了要減輕理論家的工作」，或為了賦予理論家一個歸納普遍化的基礎，這是錯誤的想法。相反地，理論家必須早就將理論完成，或至少完成其中最重要的部分：他必須已經盡可能精確地形式化他的問題。因此，理論家才是真正指引實驗的人。但實驗家並非完全投注於精確觀察；他的工作大部分是屬於理論的，在實驗室裡，從最初的設計到完工，理論一直主宰著實驗工作。<sup>26</sup>

波柏這種「理論優先」的觀點強調，理論主宰實驗，理論必須預測某些現象將發生，檢驗的步驟才能進行。而我們對那些預測現象的觀察，就負載了該理論。因此，若所預測的現象真的發生，那麼也就能決定理論的接受與否。例如，愛因斯坦的相對論預測：萬有引力場會迫使光線彎曲，而且可在日蝕的情況下觀察到此現象。這個理論指引愛丁頓 (Arthur Eddington) 觀察到此現象，戲劇性地決定了愛因斯坦理論，並宣告牛頓理論時代的結束。<sup>27</sup> 其他例子如牛頓理論預測的海王

---

<sup>26</sup> Popper (1968 p.107)。巴斯德 (L. Pasteur) 在一份演講稿中表達類似立場：「沒有理論，實踐 (practice) 就不過是習慣上的例行公事。唯有理論能夠產生發明的精神並使之發展。」「在觀察的場合中，機會只利於素有準備的心靈。」 (Vallery-Radot 1923 p.76, p.79)

<sup>27</sup> 依照牛頓力學，萬有引力對光線路徑沒有影響。但 1911 年愛因斯坦發表一篇〈萬有引力對光線傳播的影響〉的論文指出，光線在萬有引力場會被迫彎曲。不過光線曲折程度由於本身的巨大速度而顯得非常微小，沒有什麼實際可行的天體實驗可以觀測出來。愛因斯坦猜測，這個效應經由在固定恆星射到地球的光線經過太陽表面時也許觀察得到。太陽這個重力場必定會產生一種偏折使光線朝內彎曲。由於耀眼的光芒，所有的星體靠近太陽都無法看見。愛因斯坦在他的論文中預言：在日全蝕時，靠近太陽的恆星就可以看見，因此用實驗結果來驗證這個理論的結果是可能的。1917 年，天文學家指出 1919 年三月將會有日全蝕發生，這給了愛因斯坦的理論驗證提供一個機會。當時科學界都想知道愛因斯坦的預言是否會實現，倫敦皇家協會組織了兩支探險隊到可以觀察到日蝕的地方。最後由愛丁頓爵士領導的探險隊在普林西群島 (Principe Island)，成功地觀測到愛因斯坦所預言的現象。

星 (Neptune)<sup>28</sup>，以及哥白尼理論預測恆星視差等等。這些例子表明，「理論」先產生並作出「觀察」預言，可指引科學家觀察，並不會導致科學家無法經由觀察做客觀評價，反而更有助於評價工作。**必須強調，這類（波柏所指）「理論主宰實驗」當然也是「觀察的理論負擔」的情況之一，實驗家根據理論預言所做的觀察，必然多少負擔了該理論。**此時實驗家不是一個任意的觀察者或像一個靠運氣的監視器，他是背負著該理論的預期去觀察，例如愛丁頓爵士的觀察行動當然背負愛因斯坦理論的預期。

不過，「強意義」的另一種情況卻經常是理論選擇的爭議來源：當觀察尚未決定性地支持某個理論時，競爭理論都可對**這些已知現象**作出解釋，亦即，現有的「經驗證據」可同時支持相互競爭的理論。在此情況下，某個理論所詮釋的現象，永遠可以透過另一個競爭理論做出不相容的詮釋。就如同第谷與克普勒依照不同背景信念，詮釋出不同的天文景象。所以「第谷和克普勒看相同的事物，卻看到不同的東西」；「拉瓦謝當作是氧氣的東西，普利斯特里卻看成是除去燃素的空氣」。以下要論證的是，這種觀察上的差異不會造成相互競爭理論無法比較優劣的結果。

首先，不同觀察者觀察同一個現象，觀察者不必然依據不同理論來詮釋，也不必然會得到不可共量的結果。競爭理論的支持者對於其「證據」的看法，可透過相互檢視與批判來達成某種共識，例如，看成鴨子的人向看成兔子的人說明，為何他看到的是鴨子，然後看誰的理論比較好。而且，進一步的共識也非全然以經驗資料為判準；理論系統的一致性，解決問題的能力等等，也是決定的因素。例如，第谷和克普勒分別支持「地中說」與「日中說」，以「保全現象」判準來說，兩者的爭議或許不小，但就當時兩體系的理論和諧性而言，「日中說」可能比「地中說」有優勢（詳細討論見 5.4）。其次，我們可以說觀察者將鴨兔的線條圖案「看成」（seeing as）鴨子或兔子，但不能認定有任何觀察者「看出」（seeing that）那是鴨子或兔子。假設，觀察者在陰暗的動物養殖場，他將牆壁上的某種動物的頭部**影子**「看成」是鴨子的頭（或其他種動物的頭），我們不會說此觀察者「看出」那就是鴨子，只有當他在充分的知覺條件下，識別此對象與其他動物不同的特徵時，例如充足光線下看見羽毛和扁長的嘴，才能確定說他看出了鴨子。<sup>29</sup>第谷與克普勒能夠將太陽分別「看成」運動和靜止的，但是，第谷無法「看

<sup>28</sup> 海王星發現過程可參考 Norton Grosser, *The Discovery of Neptune* (New York: Dover, 1979)。或參林正弘（1991 ch.3）的概述。

<sup>29</sup> Greenwood（1990 p.561）

出」太陽正在運動，因為太陽的確沒有在運動；而克普勒也無法「看出」太陽是靜止的，因為他做不到。只有在理論與儀器科技發展到一定程度之後，科學家們才能做得到。簡言之，韓森等實用論者將「看成」當作是「看出」的個案，例如「把一個物體**看成**一個電壓計，就意謂**看出**它是一個電壓計」。<sup>30</sup>但我們主張「看成」不是「看出」的個案，「看成」是「無法看出」(not being able to see that)所經常發生的情況。<sup>31</sup>因為在觀察技術與儀器尚貧乏的階段，我們不能斷言「看出」了什麼，不同的看法正在競爭，都只是嘗試性待檢驗的觀點。<sup>32</sup>

當然，這種說法可能遭致「輝格」(Whig)<sup>33</sup>的批評。因為科學家當下無法知道，理論和測量儀器是否發展到足以判定某個觀察是否適當。而且，在同一時空，若競爭理論沒有其他認可的(corroborated)輔助理論，可裁判哪個「觀察」

<sup>30</sup> 伯朗歸納出實用論的此特徵，見 H. Brown (1977 p.179 n.11)。陳瑞麟〈科學現象的觀察與建構〉(2005)將韓森的觀察理論作了精細分析：韓森提及三個概念「看成」(see as)、看到(see)、「看出」(see that)，三者有蘊含的邏輯關係。「『看成』是『看到』的必要條件，『看到』則是『看成』的充分條件」；「『看出』蘊涵『看到』，『看到』也必定某一定種類的『看出』，但無法包含所有的『看出』」。我們在此僅區分「看成」與「看出」。根據伯朗分析，「看成」是指，我們處理識別對象的知覺及其知覺意義。「看出」通常是指，我們識別出對象而且也知道了事實。例如，迪昂實驗室的例子，物理學家從實驗室中「看出」電阻的實驗，而外行人則「看成」是一連串行為和許多器具的綜合。又如，心智正常人看出牆上的鐘知道現在幾點，心智尚未成熟的小孩或不正常的人看見牆上的鐘，並不瞭解其意義，可能將鐘「看成」一個不知名的東西。

<sup>31</sup> Greenwood (1990 p.563) 釐清此差異。

<sup>32</sup> 寇第 (Kordig 1971 ch.1) 以另一種進路駁斥韓森的觀點。韓森說：「看到一物體 x 是看出它以我們知道 x 的確如此行動的方式而行動」(1958 p.22) 寇第認為，韓森傾向將「看到」意指「相信」(believing that)；亦即當人們看到什麼，就意謂人們相信什麼。不過，我們關於知識主張(knowledge-claim)必定蘊含信念，但「看到」不一定需要信念。韓森混淆「看到」「相信」兩者語詞用法。寇第分析，當我們在使用「相信」時，大部分基於特定內涵(intensional)(對象所包含的各種屬性)，在使用「看到」時則否。例如，我相信「它是一盞檯燈」，不一定要相信「它是一堆原子的集合，而且是我阿姨最愛的收藏」。「看到」的語詞用法就不同。如果我**看到**檯燈，意思就是我**看到**一堆原子的集合；如果這盞檯燈是我阿姨最愛的收藏，那麼我就**看到**我阿姨最愛的收藏。因此，人們能夠**看到**一堆原子的集合，而無須**相信**它是一堆原子的集合；人們可以**看到**我阿姨最愛的收藏，而無須**相信**它是我阿姨最愛的收藏。(Kordig 1971 p.11) 換言之，第谷和克普勒一起觀日出，並非看到不同的東西，而是看到相同的東西，只是表述方式不同。另一類似的分析進路是將「看」區分為「無認知看」(non-epistemic seeing)與「認知看」(epistemic seeing)。參見 F. Dretske, *Seeing and Knowing* (Chicago: The University of Chicago Press, 1969) 以及 H. Brown (1987 ch.4)

<sup>33</sup> 「輝格」完整地應稱為「歷史的輝格詮釋」(Whig interpretation of history)，意謂科學史家意欲將過去的歷史事件視為「進步」與「保守」的對立，而「進步」的因素最終都贏得最後勝利，才會產生現代世界。此術語最早由 Herbert Butterfield 在其 1931 出版的 *The Whig Interpretation of History* 所確立。他認為這種詮釋態度錯估現在與過去人們的行為的因果關係，因為以現在觀點看過去人類行為，不能據以認定現在人類相似行為放在過去來看，在意欲上或事實上總導致相同結果。

是正確，那就表示理論評價的不可共量性確實存在。我對此意見的回應是，一時之間存在著觀察上的不可共量性，不表示所有競爭理論**永遠是**不可共量的。理論在動態的發展過程，本來就沒有判斷上「立即理性」，這是模糊寬容的允許空間（見 1.5）。重點在於，觀察即使負載不同理論，我們仍可能設想出某種檢驗方法，來比較這些「負載理論」的優劣。我借用赫西舉的「思想實驗」(thought experiment) 例子來說明這點。古希臘哲學家阿納克西米尼斯 (Anaximenes) 認為，地球是**圓盤狀**的，地球上高處物體落下，意謂單一方向往圓盤面地球掉落；亞里斯多德 (Aristotle) 則認為，地球是**圓球狀**，物體由高處落下，意謂可多方向地「往地球中心掉落」。顯然，**這兩人對於「物體落下」的觀察，或「落下」(fall) 術語，有著不同的理論負載**。從絕對空間以固定點看地球，阿納克西米尼斯認為物體「落下」只能單一方向（從上到下）運動，亞里斯多德認為可多向運動（從球面任一點到中心點）。假定兩人要來評價誰的觀點（理論）才正確，於是他們乘船到地球的下部，希臘的正下方，準備做實驗來檢驗。至於，是否確定在希臘正下方，可由星象觀測來決定。於是，他們在船的桅桿頂端準備丟石頭，「石頭落下」這現象仍是可評價兩人理論的**中立觀察基礎**。當石頭由桅桿頂端落下，亞里斯多德說：「我說得對，石頭落下」，而阿納克西米尼斯則說：「我錯了，石頭居然往上升」（因依照阿氏理論，此時石頭應該**落向**空中）。<sup>34</sup>這個例子是要表明，我們對「觀察」雖無立即的共識，但不同的「觀察」詮釋，也不必然導致競爭理論不可共量。一方面，負載理論的觀察也可能充分決定某理論接受與否，這種情況必須以個案來看，不可一竿子打翻一條船，以為觀察負載了理論，就完全不能作為中立的裁判。<sup>35</sup>另一方面，人們**一時之間**或許會對相同現象有不同詮釋，但隨著時間的演進，相關輔助理論與和檢測技術的發展，競爭理論間的不可共量或無法比較優劣的情況將會消失。如同萊爾 (Gilbert Ryle) 所述，「觀察」就像「感知」、「聽見」是一種動態、延續的「成就動詞」(achievement verbs)<sup>36</sup>，我們觀察的能力是隨著科學的發展而進步。同樣地，我們在競爭理論底觀察上的辨別能力，也是隨著各種理論的發展而進步。夏佩爾 (D. Shapere) 指明：「科學學習如何觀察自然，而且它的觀察能力隨著知識的增長而增長。」「在獲取知識的過程中，我們不僅瞭解自然，而且經由學習是什麼構成訊息以及如何獲得訊息，來學習如何瞭解自然。」<sup>37</sup>也就是說，人們的「觀察」隨著背景知識的進步而進步，當觀察到的現象逐漸成為某個理論的支持證據，或成為排除某個理論的反證時，就可

<sup>34</sup> Hesse (1980 pp.97-98) 此例子稍異於赫西的。赫西以神話中的波斯魔毯，作為兩人去希臘正下方的交通工具，此假設似乎比較不抵觸阿氏的觀點。

<sup>35</sup> H. Brown (1993)

<sup>36</sup> Ryle (1949 p.238)

<sup>37</sup> Shapere (1984 pp.347-348)。底線部分為夏佩爾所強調。

以避免理論間的不可共量。克普勒與第谷觀察日出的例子，或許可認為兩人觀察到不同「現象」，可用以支持不同理論。但此**單一孤立的**觀察事件並不會導致「太陽中心說」與「地球中心說」兩者在當時無法比較優劣，人們還可以透過其他方面的既有資料對兩者進行評價。而且觀測技術也在進展，當望遠鏡發明後，許多觀察上的證據都指向支持太陽中心說而排除地球中心說（見 5.4）。堅持「觀察皆有理論負擔必然導致競爭理論的不可共量」，就如同堅信「觀察是純淨中立的「所予」（given）可以充分驗證或反駁特定理論」，兩者都犯同樣獨斷的錯誤。

人們的觀察具有某些「預設」（presupposition），這些「預設」是當時背景下認為可靠的知識，不必然是一個嚴格意義的「理論」。部分實用論者以「整體論」（holism）觀點主張，既然觀察皆預設了研究者的背景知識，表示觀察句都遭到「污染」，實驗結果必定受到歪曲，那麼研究者各自的經驗陳述皆可做為支持不同理論的「證據」。這也意味，理論與觀察之間的驗證，有極大的彈性空間，兩方都允許做出修正，以適應既有的公認觀點。這結果將導致，沒有任何觀察上的證據可以充分決定一個理論。迪昂與蒯因曾在此問題上提出看法，我將之區別成三個論題（theses）並逐一討論批評。<sup>38</sup>

迪昂在其名著《物理學理論的目的與結構》（*The Aim and Structure of Physical Theory*）挑戰實驗觀察可充分決定理論的看法。他提出：

**TH1**：一個現象允許無限多種描述方式，沒有一個實驗可以確定地驗證或否認特定理論，不存在所謂「判決性實驗」（crucial experiment）。

他說：

在兩個互為矛盾的幾何學定理之間，沒有第三個判斷存在的空間：若一個為假，另一個必然為真。物理學中的兩個假設總是能夠構成這樣嚴格的兩難嗎？我們總是敢於斷定沒有其他可設想的假設了嗎？（……）與

---

<sup>38</sup> 關於迪昂與蒯因的觀點有許多分類方式，最通常的是將之統稱為「迪昂-蒯因論題」（Duhem-Quine thesis），見 Harding (ed.) *Can Theories Be Refuted? Essays on the Duhem-Quine Thesis* (1976)。當然，依各人論證進路（approach）的不同，也有其他分類方式，可參見：D. Gillies' "The Duhem thesis and the Quine Thesis" in Martin & Cover (1998)，Greenwood (1990)，Jardine (1986 ch.6)，Kitcher (1993 pp.247-256)，Lakatos (1978 pp.96-101)，Laudan's "Demystifying Underdetermination" in Laudan (1996)，Newton-Smith's "The Underdetermination of Theory by Data" in Hilpinen (1980)。我的分類並沒有跟隨這些文獻的分類方式，而且也不絕對是互斥的。



幾何學家使用的歸謬法不同，實驗矛盾沒有能力把物理學的假設轉變為無可反駁的真理；為了授予它這種能力，它就必須完整徹底地列舉包含確定之現象群的不同假設；但物理學家從未有把握，他窮盡了所有可設想的假定。物理學理論的真理不是像擲硬幣那樣由正反面來決定。  
(Duhem 1991 pp.189-190)

**TH2**：科學理論是與其他輔助理論一起受檢驗，這個整體不能片面地（piecemeal）切割以檢驗其確定性。當理論與實驗不一致，方法論規則無法肯定或否定其一。

他說：

物理學家從來也無法將一個孤立的假設交付實驗檢驗，而只能將一整群（a whole group）假設交付實驗檢驗。當實驗與他的預測不一致時，他所得知的是，構成這整群假設中至少有一個是不能接受的，應該受到修改。但是這個實驗並沒有指明，哪一個假設必須更改。（Duhem 1991 p.187）

迪昂抱持科學理論的工具主義（instrumentalism），他認為，科學理論的發展不能證明是更接近真理（true），其目的在於「保全現象」（save the phenomena）。理論旨在描述現象，無所謂可「更深入」解釋自然界本質，也不能說哪個理論更為真實，它們純粹是人的概念、人的發明。從工具主義出發，迪昂主張科學理論是整體不可分割。科學家以實驗來檢驗理論，並非純粹是理論與實驗兩方的對決。事實上，一邊是一組理論的系統，另一邊是實驗觀察的系統，兩個體系都包括在一個「所有接受的知識整體」中。當理論與實驗結果不符時，實驗結果不能取消哪個特定理論。實際上，科學家是運用所有他接受的知識來進行科學活動，實驗一方同樣代表一組理論，**實驗**若沒有**理論**作為工具，便不能控制或詮釋特定的觀察數據。如果科學家懷疑某一個理論，他必須設計實驗來檢驗該理論所預測的現象是否出現，此時他不僅利用該理論來構思實驗，也包括他接受的所有知識。因此，當實驗的結果與該理論的預測現象不符合時，不能完全怪罪理論，只能斷言該科學家所接受的整體知識不一致。

**TH1** 挑戰「判決性實驗」的存在。判決性實驗意謂，一個實驗的結果可充分地反駁待檢驗理論，理論能夠被實驗證明是錯的。波柏認為，藉助實驗觀察的幫

助，嚴格地檢驗並淘汰理論是科學增長的特性。科學發展必須依賴判決性實驗，經由一連串實驗推翻舊理論，科學知識才能朝向真理邁進。波柏曾對 TH1 作出回應，但在某些地方誤解了迪昂的論點。他說：「迪昂在他那著名的對判決性實驗的批判中，成功的表明判決性實驗不能確立（establish）一個理論。他未能表明這些實驗不能反駁（refute）一個理論。」<sup>39</sup>不過，根據迪昂的觀點，科學理論作為人類思想產物，是整體而不可分割的。從整體觀點出發，科學理論不僅不能驗證，也不能否證。在此，TH1 與 TH2 成為相互環扣的問題：既然理論是不可分割的，當實驗結果與理論不一致時，我們應該把錯誤怪罪於哪一部份？最根本的反對論點是，科學實踐沒有如數學定理那樣簡單的邏輯結構，實驗結果的複雜現象，如何能化約成邏輯的「否定後見因而否定前件式」排除一整群理論？

TH1 的成立基礎來自於**自然科學與數學**的差異。數學只是一個命題系統的連結，而自然科學還需要一些輔助概念的運用，例如物理學概念「力」、「場」等，這些概念指涉的是數學系統以外的東西，由科學家的決定所產生，牽涉心理與經驗的應用，不單只是純數學的因素。但**概念**描述因人而異，如此就會產生許多種描述現象的方式。科學理論由**數學命題**和描述現象的**概念術語**所組成，所以，科學研究的推論不是純粹的數學運算。若將觀察結果當作從某個理論推導出來的矛盾結論，以此來證明該理論是錯的，這種歸謬法的確定性只有在數學命題中才有。因此，否定的判決性實驗所憑藉的歸謬法，在科學研究中並不適用。

假定科學研究的推論程序可以運用歸謬法，那麼必須假定實驗結果（觀察）是真實無疑的，「否定後件因而否定前件式」的否證箭頭才能起作用。但此假定面臨 TH2 的困難：實驗結果也是可疑、會出錯的。因為就算實驗結果受到認可，也不能證明實驗背後所預設的理論為真，或聲稱該理論比被否證的理論更接近真理。科學理論的選擇不是非真即假的判斷；實驗背後預設的理論，同樣無法做非真即假的判斷。任何一個理論，技術上幾乎有無限多組的競爭者，若選擇並確定其一，那麼就必須證明其他**所有**可選擇的理論為假，但事實上我們不可能窮盡所有可選擇理論，更不必說要否證它們，因此科學理論都是不充分決定的。

---

<sup>39</sup> 見 Popper (1996 p.112, n.26) 波柏在“Three Views Concerning Human Knowledge”中批評迪昂，波柏認為，如果將兩個競爭的理論同其各自的「背景知識」（事實上，有時兩個背景知識有重疊部分，甚至是相同的），拿來與判決性實驗對決，我們必定可以比較出哪一個理論正遭遇危險。而且我們並不斷言僅僅反駁哪一個理論，而是該理論包括其背景知識都遭到反駁。(1996 p.112)

我同意理論選擇的判斷或推論，沒有如數學推論那種確定性；但不能因為沒有這種確定性，就斷言理論是不充分決定的。「理論是不充分決定」若僅意謂理論不是永恆絕對的真，而是嘗試性可錯的，那麼幾乎所有哲學家都會同意；但若意謂我們因此就沒有公認判準可選擇某個理論，那麼就有待商榷了。TH1是反對：科學推論應用數學的歸謬法，理論與實驗的對決有絕對確定性。不過，TH1或許邏輯上說的通，但在科學實踐上幾乎沒有這種困擾，亦即，我們總是可以認定幾個最有可能的假設著手進行研究。TH1假定「同一現象有無限多種描述方式」，只是理論上可能，事實上不可能存在「無限多種描述」。若要求科學家必須窮究所有可能假設和證據才能作出判斷，根本辦不到。基切爾（Philip Kitcher）舉恐龍滅絕為例。恐龍滅絕原因的假設允許有無限多個，可以假設是氣候變化的因素，或太空隕石撞擊地球，甚至遭到外星人消滅、造物主的懲罰、恐龍族群自我毀滅等等，無論它們表面上看多荒謬，只要是可設想而不矛盾的情況皆可羅列。但實際上，科學家所要找出恐龍滅絕的原因，是集中在少數幾個研究社群認為**認知上重要**的假設，例如氣候、傳染病、以及行星撞擊等。<sup>40</sup>科學家將注意力集中於這些少數幾個假設，原因在於實際研究中，首要認定哪些特定過程或事件是自然狀況中發生，而不是其他非自然因素。簡言之，科學理論選擇對於證據的認定，只考慮**相對最有可能**發生的情況，而不是考慮**所有可能**的情況。

TH1 的確鬆動了絕對的、無時間性的方法論。波柏期待科學發展，是以一連串的判決性實驗，否證一個個舊理論的「猜想與反駁」過程。但實際上，「判決性實驗」難以充分、立即地認定。如拉卡托斯所言，「判決性實驗」的頭銜，必須經歷一段很長的時間才賦予。<sup>41</sup>既然如此，那麼捨棄判決性實驗的認定，我們一樣可說明科學合理進步的特性。科學發展最一般而顯著的特徵，不在尋求及辨識判決性實驗，而在實驗的證據比較偏重哪一個理論。科學家於研究過程中挑選出的理論，並非永遠絕對真或完美的理論，而是在現有競爭理論中相對比較「偏重」的理論。

TH2 主張知識的整體論，亦即，每個理論檢驗的程序，都是對研究者接受的所有理論的檢驗，任何理論和觀察都預設了研究者自身認可的理論。但實際上，任何科學家不可能許諾一種「極端的」整體論，亦即，科學知識是一個完全無法分割的整體，以至於他完全無法區別哪一個是待檢驗的假設、哪一個是實驗觀察。TH2 是來自邏輯學的觀點，而不是得自於科學變遷或實際案例的考察。或許

---

<sup>40</sup> Kitcher (1993 p.248)

<sup>41</sup> Lakatos (1978 p.4, p.111)

打擊到傳統理性論的邏輯重建模式，但沒有威脅以經驗（科學史）研究為前提的理性論觀點。若堅持「以整個科學為前提」的整體論而非局部的整體論，等於放棄了「簡單性」的要求，認識世界的願望將成為不可能。知識若是一個無法分割的整體，必定比分割成片段（例如，將檢驗程序簡化成特定步驟，將科學推論歸納出幾個方法論規則）更為分雜凌亂。約束極端整體論，而允許一定程度的化約論（reductionism）是必要的。我們要瞭解紛雜的自然現象，必然要盡量化約或分類成幾個共通範疇以供理解。因此，TH2 僅能聲稱觀察與理論群不一致時，科學家無「判定程序」可判斷哪個地方該修正。

就算「知識」在某種意義上是連結為一體的，由也此無法進一步推論：科學家所做的每個假設或預測都必須以一個「整體」為前提。因為此「整體」，不能毫無無限制。在實際的科學研究中，科學家或科學社群總是在某些限定的時空條件下，探究某些可區分的假設。我舉一個實例來說明。巴斯德（Louis Pasteur）研究發酵理論，他主張：發酵所需的微生物（酵母菌），並非自然衍生，而是來自空氣中；空氣中的微生物落在牛奶或水果等有機物，繁殖並將糖分解為酒精和二氧化碳。巴斯德的主張與當時盛行的「自然衍生」（spontaneous generation）學說相衝突。自然衍生論認為，生物可以自然出現並繁殖，例如腐肉生蛆、胃液生條蟲、乳酪或麥粒放入甕中會生老鼠等。尤其是在十七世紀末顯微鏡發明後，顯微鏡下的微生物繁殖情況，更讓當時多數科學家相信自然衍生論。從事後觀點看，我們現在都知道發酵作用所需的微生物是來自空氣，巴斯德成功地駁斥了當時的自然衍生學說。但有一個問題卻是巴斯德沒有也不能回答的：世界上第一個生物是怎樣產生的？<sup>42</sup>以整體論觀點來看自然衍生論，自然衍生論等於斷言世界上第一個生物是自然產生的，就這點而言，巴斯德事實上沒有也不能完全反駁自然衍生論，因為他無法證明世界上「第一個生物」是來自何方。我要強調的是，「無限制」的整體論觀點在科學探究上無法辯護，它在邏輯上可說得通，但在實際上卻不可能。

因此，TH2 僅能維持一種較弱的形式，那麼似乎也只是許諾 TH1 而已。亦即，「知識」不是無限度地被視為**整體一塊**不可切割，理論受到實驗的檢驗，只意味沒有判決性實驗可以否證特定理論，就如同自然衍生論並沒有遭到巴斯德完

---

<sup>42</sup> 這問題所要面對的困難在當時顯而易見。當巴斯德決定投入「自然衍生」理論的研究時，他的老師畢歐（J. J. Biot）極力反對並勸阻巴斯德不要做這種實驗；而他另一位老師，當時的法國教育部長也是極著名化學家杜馬斯（J. B. Dumas）也勸告巴斯德說：「我不建議任何人在這個問題上太過深思。」Vallery-Radot（1923 pp.88-89）

全否證一樣。就科學上解決問題的要求而言，科學家可以透過「有限度地支解整體」來理解自然。雖然如此可能無法完理解自然現象整體（例如世界上第一個生物從何而來），但至少可以解決在特定時間和空間中的謎題。例如十九世紀中葉以前，人們雖然知道製作葡萄酒和醋需要酵母，但並不知道其作用為何，甚至認為發酵現象與有機物無關，因此釀造葡萄酒和醋時，經常發生意想不到的失敗。巴斯德堅信，要解決這類釀造失敗的問題，必須先解決酵母是從何而來的問題。亦即，若要進行發酵的研究，必須先解答自然衍生論在某些限制下的難題。巴斯德的看法是：

發酵不再是那些孤立和神秘而不可解釋的現象之一。它是營養物在特定條件下發生的一個獨特的生命過程，不同於所有尋常生命（動植物）的那些生活特徵。（……）我們甚至能夠設想發酵特性可能屬於每一個有機體，屬於每一種動物或植物細胞，在消化和排泄的化學生命（chemico-vital）作用的單獨條件下，發酵必定短暫地發生於細胞，時間或長或短。（1879 pp.45-46, 底線字為筆者所加）

將實驗限定在一定的時空和特定的條件下，可觀察到每個有機物或生物體並非自然衍生，必須由其他個體所產生，而發酵作用就是這些生命過程所產生的特殊現象。

這種有限度地支解知識整體，就是「局部整體論」的意涵之一，<sup>43</sup>即知識是局部整體地交付實驗檢驗。局部整體論加上知識都是會出錯的「可錯論」（fallibilism）觀點，可以避開 TH2 的困難。以波柏為例，他主張所有科學理論都是可錯的，理論和觀察都容許出錯的可能。科學理論只是我們對這世界所做有系統的假設或「猜想」（conjectures），這些猜想雖不能證實，但經過嚴格的檢驗可以決定是否放棄。科學家在檢驗一個理論時，總是運用到他所知的「背景知識」。他將這些長久以來認可的東西當作是沒問題的，不是因為他認為背景知識永遠為真，而是因理論的檢驗不能從渾沌開始。背景知識當作沒問題只是暫時

---

<sup>43</sup> 類似這種「局部整體論」，蒯因稱此為「溫和的（moderate）整體論」或「相對的（relative）整體論」，即僅堅持每一個科學述句都具有其各自可分離的經驗內容。（Quine 1981 p.71）整體論事實上在知識論問題上面臨許多困難，若科學述句只有共同作為一個理論才蘊含其觀察結果，那麼範圍必須是多大？能夠無限大嗎？它必須被視為這世界綜合理論的整個科學嗎？蒯因承認：「科學既不是非連續的也不是單一整體。它以不同的方式相連接，而其連接點在不同程度上是鬆散的。」「整體論原則上是整個科學這一種說法沒有什麼意義，儘管可以用合法的方式為其辯護。」Quine（1975 pp.314-5）

的，它的任何部分在任何時候都有可能被取代。長期認可的理論暫時確定後，才能將反駁的箭頭指向比較新而可疑的理論。因為我們不將長期接受的理論當作確定無誤的，所以當實驗結果與理論預測不一致時，雖然這沒有告訴我們該放棄哪一個部分，但我們可以暫時先肯定那些長期以來認可的理論，而懷疑那些更大膽、內容更廣泛的新假設。當然，遭懷疑是錯誤的理論有可能將來又被提出討論，而重新拿出來檢驗，因為也許那時會發現應該懷疑的是其他輔助理論或實驗結果。這些程序是一連串的猜想與檢驗的「動態」過程。在某種程度上，波柏也同意迪昂觀點，他說：「事實上，不曾有決定性的反證（conclusive disproof）被提出，因為總是有可能說實驗結果是不可靠的，或主張理論與實驗結果存在的差異未必是真的，它們將會隨著我們知識的進步而消失。」<sup>44</sup>在這個意義上，「時間」是重要的因素，「靜態的」TH2 並不會造成科學家「動態的」判斷上底障礙。

但是，當科學理論持續發展，新證據陸續出現，科學家的信念也會隨之改變，可能會對接受的「所有知識」做出修正。蒯因延續整體論觀點，認為「知識」如同一張信念網，人們可以修改信念網中任何部分以適應頑強的經驗。這形成「經驗證據不充分決定理論」（underdetermination of theory by empirical evidence）的論題：

**TH3：**一個經驗證據技術上可由無限多組前提導出，相互競爭理論都能透過修改各自理論中的任何部份以符合經驗證據，因而沒有一個理論可以由經驗證據唯一且充分地決定。

蒯因認為，科學理論是與其他接受觀點共同作為一個理論系列，才能蘊含其觀察的結果。因此，理論不是孤立地接受觀察的驗證或否證，而是作為一個整體來面對「經驗的法庭」。當理論與觀察相衝突，「可以通過對整個系統的各個部分做任何可選擇的修改來適應一個頑強的經驗」。<sup>45</sup>只要使信念網內部保持一致，我們可以堅持其中任何一個述句。蒯因說：

我們可以說，一個理論公式（theory formulation）蘊含的那些觀察定言句（observation categoricals）構成其經驗內容，因為只有觀察定言句才將理論與觀察連結起來。如果兩個理論公式蘊含所有相同的觀察定言句，那麼它們是經驗等價的。（……）兩個理論公式（……）可以是經

---

<sup>44</sup> Popper (1968 p.50)

<sup>45</sup> Quine (1980 p.44)

驗上等價而不是邏輯上等價。(1981 p.28)

蒯因所謂「觀察定言句」意指這類句式：「每當如此…，就那樣… (Whenever this, that)」，一個理論公式即蘊含一組這樣句式的觀察句。<sup>46</sup>我們以第谷與克普勒的例子，簡單說明蒯因的論述。「每當拂曉時，太陽從東邊升起」，此觀察定言句在第谷與克普勒時代，同時受到兩種理論公式所蘊含：「地中日動說」與「日中地動說」。兩個理論都可以得到「每當拂曉時，太陽從東邊升起」的觀察句，因而兩個理論可以是經驗等價（就此「拂曉」的觀察定言句而言），但邏輯不等價（系統的幾何架構不同）。假設當時出現新的經驗證據，不論對「地中日動說」或「日中地動說」有利，雙方支持者都可透過修正他所接受的所有知識（包括新證據），使其信念網保持一致。若這種觀點成立，顯然會產生許多競爭理論皆蘊含相同的觀察述句，而彼此卻邏輯上不一致的情況。易言之，一個觀察證據（結論）邏輯上可以由許多不同的理論（無數組前提）導出，該證據相對各理論而言是經驗等價的。<sup>47</sup>蒯因由 TH3 進一步批判傳統理性論的「驗證」與「否證」觀點。他以整體論觀點將理論與經驗之間的界線模糊，主張任何觀察都不足以充分決定一個理論。「整體論論題提供不充分決定論題信用 (credence)。如果面對不利的觀察，我們總可以在理論的各種適當修正之間自由選擇，那我們可以假定，所有可能的觀察不足以唯一地決定理論」。<sup>48</sup>

但要指出，蒯因是走邏輯實證論的老路，然後打擊邏輯實證論。也就是說，TH3 是「從邏輯觀點看理論選擇」，而非「從經驗觀點看理論選擇」。TH3「邏輯上」說得通，但沒有實際探究科學家面臨理論選擇的情況。首先，TH3 已經預設 TH1 與 TH2，即：同一現象可有無限多種描述方式，及知識是整體地為人們接受。如前所述，一個經驗證據邏輯上可以由無限多組前提導出，但事實上，知識探究首要關心的是「較有可能的」前提，而不是考慮「所有可能的」前提。恐龍滅絕的原因邏輯上有無限多種（只要可想像的），人們無須也無法一個個檢驗所有可能原因。其次，TH3 預設兩種不充分決定性概念。一種是休謨式的「經驗」不充分決定性，亦即，理論無法被過去的經驗證據所決定，因為未來的經驗證據可能與之衝突。第二種不充分決定性是心理、社會的。若允許每個人可更改所接受的知識以符合任何新證據，那可能表示我們各有不同的「信念」和「經驗」。就

<sup>46</sup> Quine (1990 p.10)

<sup>47</sup> 從邏輯觀點較容易理解這個論題。例如，我們現在有觀察證據 P，那麼要由邏輯規則推導出 P，其前提允許有無限多組：「P」或「Q, Q→P」或「~P→Q, ~Q」或「R∨S, ~S, R→P,」……等等。

<sup>48</sup> Quine (1975 p.313)。蒯因在此篇文章中強調，整體論不應與經驗證據不充分決定理論的觀點相混淆。

「經驗」而言，觀察述句依照觀察者接受的外在刺激表述出來，這個刺激「輸入」與表述「輸出」的過程，並無法保證每個人都相同。更關鍵的是，這個過程牽涉觀察者所處的社會文化背景，因此所做出的經驗陳述可能極為不同。第二種不充分決定性比比第一種更為嚴格，等於許諾：任何理論都不被**所有且可能的**經驗觀察所充分決定，而是相對其所處的社會與文化來決定。

關於 TH3 的第一種不充分決定，並不影響理論評價。大概沒有一個科哲家不贊同存在這種不充分決定性，因為科學知識都是可錯的而且動態發展，就必然會有不確定性，但此不確定性不會導致競爭理論無法比較優劣。它只威脅這種主張：單一理論與證據的驗證對決，可產生一勞永逸（*once and for all*）的決定性判斷。我們並不抱持這種主張。第二種不充分決定性，即觀察者因所處之社會背景不同，產生的理論評價困難，將在第五章討論。以下我將就 TH3 整體提出批評。

不可否認，我們有時可透過修改或替換某些輔助理論，使得待檢驗的理論適應觀察上的異例。不過，並非所有理論永遠可經由修改其中部分，使其與觀察證據相符合，並能與競爭理論等價，導致我們無法從其中判斷優劣。如本章開頭所述，我們採取一種「比較的」觀點，即理論或假設的評價永遠與其現有的競爭者相關。<sup>49</sup>依此觀點，若某個理論沒有競爭者，它可用任何方式保留以抵抗觀察上的異例，此情況無所謂「理論選擇」的問題。理論選擇意味，從兩個或以上的理論中，評價出一個較好的理論。在這種多個競爭理論的情況下，當某個理論通過檢驗，而其他競爭理論沒有通過相同的檢驗時，選擇那個通過的理論是合理的（*reasonable*）。傳統理性論抱持一種觀點，即實驗觀察與待檢驗理論的關係不是蘊含，就是相互衝突：一個實驗觀察 E 受到認可，要嘛驗證了理論 A，要嘛就是否證了與 A 相競爭的理論 B。亦即，某觀察若否證了理論 A，那就表示驗證了 A 的否定理論 B。<sup>50</sup>接受一個實驗證據，意謂決定了一個理論而同時放棄了另一理論。為應付 TH3 的挑戰，我採行「比較的」解釋：經驗證據「驗證」或「支持」理論的基本概念是相對比較的關係，競爭理論 A 與 B，若證據 E 比較支持 A 而不是 B，那麼選擇 A 是合理的。**我們無須等待判決性實驗，也不必問 E 是否充分決定 A，只要問相對於 A 與 B，E 比較支持哪一個。**理論與實驗的驗證不是兩

<sup>49</sup> 不少科哲家提到比較性觀點，但很少有人特別強調。至少出現於 Kuhn (1977a ch.8)；Lakatos (1978)；Laudan (1984)。勞登後來直接稱此比較觀點為「比較論」(comparativism)，見 Laudan (1997 p.314)。

<sup>50</sup> 波柏特別強調「驗證」與「否證」的不對稱 (asymmetry) 關係。亦即，一個觀察述句受認可，它能夠反駁一個普遍定律，但無法驗證一個普遍定律。Popper (1993 pp.181-182)



方對決，而是包含兩個或更多競爭理論與證據之間的比較關係。兩方對決的情形，必須充分決定實驗證據；但比較的觀點認為，理論選擇的判斷就看現有證據較偏重哪一個理論來決定，如此可避免不充分決定的困難。所有競爭理論在既有的種種證據之前，不可能**永遠且完全**得到等價的支持，必定可能存在或發展至一種情況：在這些競爭理論中，只有一個理論比其他競爭者，明顯更好地得到證據的支持，其他競爭理論無論如何修正都無法取得與之等價的地位。

