

## 原初意向性與錯誤表徵

彭孟堯

### 一、緒言

傳統強人工智能(strong artificial intelligence)學者宣稱經由適當程式設計的電腦實際上具有心靈、具有理解能力、尤其具有與人類相同的認知狀態(cognitive states)。然而這個強人工智能的信念在最近十幾年已遭到來自當代心靈哲學家與認知科學家的嚴重質疑與挑戰。John Haugeland (1981)特別將近年來反對強人工智能其中的一種方式稱為「空殼策略」(hollow shell strategy)：任何一個程式設計的電腦，無論經由多麼適當的程式設計，實際上都不可能具有心靈、具有理解能力、或者具有與人類相同的認知狀態，因為它缺少了「X」(p. 31)。依John Searle (1980a, 1980b)的看法，人與強人工智能定義下的機器<sup>1</sup>的不同就在於人具有「內在意向性」(intrinsic intentionality)<sup>2</sup>。沒有一個純粹形式或語法驅動的系統能就其本身之特性

---

作者現任職於中正大學哲學研究所暨認知科學中心。

收稿日期：民國八十四年三月二十二日；接收刊登日期：民國八十四年五月二日。

兩位匿名讀者對本文多處地方提出了寶貴意見，謹此致謝。

<sup>1</sup> 這裡所指的「機器」是人工智能或計算心靈學派意義下的「機器」。Searle (1980a, p. 300)對於「機器」有相當扼要但生動的敘述。

<sup>2</sup> 這個「X」除了Searle所提的「內在意向性」之外，在目前心靈哲學界裡主要還有兩

就足以具有原初意向性；他的「華語房間論證」可說是基於他對於「內在意向性」（或「原初意向性」）與「衍生意向性」（*derivative intentionality*）的區分上。<sup>3</sup>

這個「原初意向性」與「衍生意向性」的區分直覺上的確相當具說服力。它的重要性固然可從它在批判傳統人工智能所扮演的角色看出，但同時則是它在理解「錯誤表徵」（*misrepresentation*）的現象方面具有關鍵的地位。當代心靈哲學家大致上都採取了將意向性自然化的觀點，其中一個頗具影響力的理論是早期的 Fodor, Dretske, Papineau, 以及 Millikan 的「自然目的論取向的心理表徵理論」（*natural-teleological approach to the fixation of representational content*）（以下簡稱「自然表徵論」）。<sup>4</sup> 這個理論將原初意向性理解為表徵性，而從心理語意學的角度提出了關於決定表徵內容的兩項要件：(1) 心理表徵與被表徵物之間必須有從被表徵物到心理表徵的一個律則性的因果聯結 (*symbol-to-world nomic correlation*)，以及 (2) 這個因果聯結必須滿足存活原則 (*the survival principle*)，亦即該被表徵物之被表徵可以為該認知或心理系統帶來存活的價值 (*survival conducibility*)。<sup>5</sup> 依據這兩項要件我們可以決定心理表徵的內容，而一旦表徵內容決定後，甚麼時候一個表徵的出現是錯誤的，就可以決

個理論：(1) X 就是「意識」，尤其是 D. M. Armstrong (1966) 所謂的「自我省察的意識」(p. 13)；以及 (2) X 是所謂的「關懷」(*caring*)。參 Haugeland (1981), pp. 31-34。

<sup>3</sup> Searle 對於意向性其實有更為細緻的區分：「內在意向性」、「衍生意向性」、以及「假似意向性」(*as-if intentionality*)。由於這個更為細緻的區分在本文裡不重要，將予以略過。參 Searle (1992, pp. 78-82)。Newton (1992) 認為「內在意向性」與「原初意向性」(*original intentionality*) 不同，並且詳細區分了「原初意向性」的兩種意義。他並進而詮釋 Dennett 的立場，認為 Dennett 主張雖然大自然具有的意向性實際上祇是「假似意向性」，但至少是「原初的假似意向性」(*original as-if intentionality*) (p. 19, fn. 1)。雖然我不認為「原初的假似意向性」這一觀念有任何實質意義，由於他的區分並不影響本文論證的實質內容，在本文裡，「內在意向性」與「原初意向性」兩詞將交互使用。

<sup>4</sup> 這些哲學家的觀點不盡相同，因不影響本文，將予以略過。

<sup>5</sup> 關於自然表徵論，我的博士論文 (Peng, 1994) 第二章有詳盡的闡述。

定了。因而，錯誤表徵現象的解釋有賴於預設原初意向性的存在。

然而Denial C. Dennett (1987)指出，這個對於原初意向性與衍生意向性的區分的直覺實際上並不具備充足的哲學理由支持。他甚至提出了「功能再賦與論證」(argument from functional re-ascription)與「自私基因論證」(argument from selfish genes)來否認人具有原初意向性。如果Dennett的主張是正確的，自然表徵論對於甚麼是「錯誤表徵」的理解將會出現問題。例如，一個只具衍生意向性的自動販賣機對於它自己是否表徵錯誤是無法自行決定的。但是如果正如Dennett所說，人的意向性也是衍生的，其本質與自動販賣機的意向性沒有甚麼不同，人對於他自己是否表徵錯誤也同樣是無法自行決定的。

本文的目的是要駁斥Dennett對於人的原初意向性的否定，以辯護自然表徵論對於錯誤表徵現象的理解。第二段將先簡略地敘述自然表徵論者對於原初意向性的瞭解，第三及第四段分別用於省察並駁斥Dennett的兩大論證，在最後一段則是對於原初意向性的問題作一省思，以找出自然表徵論在理解原初意向性時，尚有可能遭遇的難題。

## 二、自然表徵論對於原初意向性的理解

雖然本文不擬討論Searle的「華語房間論證」，我們可以從他的討論裡瞭解他主張我們人類具有原初意向性的理由，從而探討他對於原初意向性的主張。依Searle (1980a)的看法，我們人類乃是具有某種生物結構的有機物。這個生物結構在某些條件下具有產生知覺、行爲、理解、學習等等意向現象的內在因果能力(intrinsic causal power)<sup>6</sup>；只有具有這種內在因果能力的生物有機體才有可能產生

<sup>6</sup> 關於「內在」(intrinsic)一詞除了本文的用法外至少尚有兩種含意：(1)所謂一個屬性是「內在」的意指它是一元的(monadic)屬性，而不是表達關係的(relational)屬性；(2)所謂一個屬性是「內在」的意指它是一個「必然」的屬性。這兩種用法並非此處

認知現象，才具有原初意向性(p. 299)。<sup>7</sup> Dretske (1985)也提出了一個與Searle相差不遠的看法。他認為一個具有原初意向性的系統其表徵必須具有內在的意義或內容，而內在的意義指的是「從表徵到外在世界」的聯結(symbol-to-world correlation)，不是那種依靠使用者的意圖和目的而產生的意義(p. 29)。依據Dretske的看法，我們可以說機器不具原初意向性，只具衍生意向性，就是由於機器的表徵內容有賴於外在詮釋者的決定，機器的表徵內容是由詮釋者所賦與的。而我們人類以及某些高等生物的認知表徵系統之所以具有原初意向性，則是因為我們的心裡表徵的確具有與外在世界相聯結而產生的內在意義。簡單說，自然表徵論將原初意向性理解為認知系統的表徵能力，這表徵能力則進一步理解為從被表徵物到心理表徵的律則性聯結，而不是依賴於任何外在詮釋者。

必須注意的是，上述的理論並未排除人造物具有原初意向性的可能。對Searle (1980b)而言，凡是具有至少與人腦相同因果能力的系統都可以具有原初意向性(p. 453)。Searle所反對的只是傳統強人工智能意義下的，只是語法或形式驅動的系統。這類系統因為不具有類似於人腦的因果能力，故不可能具有原初意向性。

為了有助於瞭解Dennett否認人有原初意向性的論證以及他如何從反對原初意向性的兩大論證拒斥自然表徵論對於錯誤表徵的解釋，我們首先將Searle與Dretske關於「原初意向性」的主張條列如下：

(OI-1)我們人類具有原初意向性(Searle)；

---

Searle、Dretske與Dennett討論「內在意向性」或「原初意向性」的含意，不可混淆。值得一提的是，此處所談的「內在意向性」並不帶有任何神秘不可解的色彩。如同綠葉的內在意向性就是行光合作用，它的內在意向性是可以瞭解的。認知的內在意向性就是意向性，它的內在意向性也是可以解釋的。自然表徵論就是一個試圖解釋內在意向性的理論。

<sup>7</sup> 為何只有生物有機體具備這種可以產生認知的內在因果能力呢？可惜的是Searle並未對於這個問題有進一步的說明，也沒有對於這種內在因果能力是如何產生的做進一步的解釋。

- (OI-2)人的意向性之所以是原初的，是由於人的心靈或認知機制具有內在的因果能力以產生知覺等認知現象 (Searle)；
- (OI-3)人的心靈或認知機制之所以有這種內在因果能力純是因為人是生物有機體 (Searle, Dretske)，使得人的心理表徵內容建立於表徵與外在世界的聯結，而不依靠於外在的詮釋者 (Dretske)。

從上述三項我們可以看出 Dennett 想要作的是否定 OI-1 之為真。一旦 OI-1 被否認，人的心理表徵內容就無從被決定，因而自然表徵論對於錯誤表徵的解釋也只能完全放棄。這是因為如果人的意向性也是衍生的，也是有賴於外在詮釋者賦與表徵一個內容，不僅自然表徵論將意向性自然化的工作會出現循環的問題，而且人本身也將不具任何立場去肯定人有錯誤表徵的可能。現在讓我們來檢查 Dennett 的兩個論證是否能建立他的結論。

### 三、功能再賦與論證

功能再賦與論證顧名思義指的是對於系統機制的功能有可以重新賦與的可能，藉著這個可能性可以重新決定某個表徵的內容，進而消除原先認定為錯誤表徵的現象。功能再賦與的可能性在人造物 (artifacts) 的情形裡尤其明顯，所以讓我們先從人造物的例子著手。

以汽車的油表為例，如果油表的指針指著 E (空的)，但實際上油箱是滿的，我們就可說這個油表錯誤地表徵了油箱的內容，這是因為油表的功能就是在於記錄油箱的內容。現在假想另一種狀況：如果油表的指針指著 F (滿的)，油箱的確是滿的，而且沒有其它故障發生，但是油箱內裝的是水。試問，這次油表有沒有錯誤地表徵油箱的內容？

這個問題乍看之下不難回答。這個儀器既然是油表，當然它的功能是要測知油箱裡的油量，如果裝的是水，當然它是一個錯誤的

表徵。可是這個回答並沒有排除我們重新指定或賦與這個儀器功能的可能。這個儀器的功能之所以是要測知箱中的油量而不是水量，是因為使用者決定要用它來測油量，如果使用者決定要使用相同的儀器來測水量，也未嘗不可。甚至於如果我們重新賦與它的功能，將它定位在測知箱中的液體量，既不是油量也不是水量，前面我們所說這個儀器是錯誤地表徵了箱中的內容，就不再為真。功能的再賦與將「錯誤表徵」消除了。

我認為在談類似油表這類人造物的例子時，其實不必問該人造物本身是否有出現「錯誤表徵」的現象。這是由於這類人造物的意向性完全依賴於使用者的詮釋或目的。我們可以說發生錯誤的不是儀器本身，不論箱中裝的是油、水、還是啤酒，儀器與箱中物的含量多寡只有純物理的關係，說不上有正確或錯誤表徵的情況。一個儀器的表徵能力來自於使用者，撇開機械故障的情形不論，所謂正確與錯誤的表徵問題要追溯到詮釋者的意向性上。因此「功能的再賦與論證」之用於被造物的意向性問題上是相當無趣的用法。但是這類用法雖然無趣，卻點出了Dennett論證的主旨：一個系統的機制功能的再賦與如果是可能的，這個系統的意向性或表徵性就不會是原初的，而只可能是衍生的，因此，它的錯誤表徵的現象也將會因為對於它的機制功能的再賦與而有可能被消除。

讓我們來看看這個論證在生物有機體的錯誤表徵的情形。無庸置疑地，我們處理「功能再賦與論證」之用於人造物錯誤表徵的方式並不能轉用到生物有機體的錯誤表徵現象上。生物有機體的認知機制所具的功能並不依靠於我們人類的詮釋、意圖、或目的。我們可以改變一個儀器的功能，將一個原本用於測知油量的儀表拿去測水槽中的水量，但是我們無法改變青蛙的覓食機制，從它測知蒼蠅、蚊子是否出現的功能改為測知BB彈是否出現的功能，因為測知並捕食BB彈並不是青蛙的生物功能。對於生物有機體的功能的賦與並不依靠於人的意向性。

不過，我們或許可以從不同的方向來考慮生物功能的再賦與。Dretske (1986) 的深海細菌的例子指出了這個可能。<sup>8</sup> 這個例子大意如下：某類深海細菌具有一種探測磁場方向的機制，使它能游向地磁北。如果我們在地磁的反方向放置一隻磁棒，它會游向磁棒而朝與地磁北相反的方向。這時我們說它錯誤地表徵磁棒的方向為地磁北。我們根據甚麼來說這深海細菌出現了錯誤表徵的現象呢？有沒有可能重新賦與它的機制功能為探測地理北呢？甚至，為何不說它的機制功能只是探測磁場的方向而已？如果可以重新賦與它的機制功能只是為探測磁場的方向，原來我們所說它之游向磁棒是一個錯誤表徵的現象就不再成立了。這個例子並不構成對於自然表徵論的威脅。依據自然表徵論，深海細菌的地磁探測機制的功能既不是探測地磁北或地理北，也不是探測磁場方向，而是探測無氧環境的所在。這是由於這類細菌無法存活在含氧量過高的環境裡，而恰好在深海細菌的演化過程中，探測地磁北的方向可以「指示」深海細菌無氧環境的所在。這是自然表徵論裡的第二要件「存活原則」的運用。

我們現在將這個深海細菌的例子以較抽象的方式說明如下：假設某個生物機制的功能可以瞭解為  $f_1$ ，因此它的表徵內容為  $F$ ；由任何非  $F$  引發的表徵都是錯誤的。但是在不與其它的功能賦與相衝突的情形下，如果這個機制的功能也可以賦與為  $f_2$ ，而使得它的表徵內容變成  $G$ ，此時原本應該是錯誤表徵的現象可以經由功能的

---

<sup>8</sup> 生物功能再描述的可能性不限於類似細菌這類微生物，本段討論的要點亦適用於高等生物的功能，雖然再描述高等生物的功能實際上並不容易。本段僅藉 Dretske 的深海細菌來例示這個可能性。值得一提的是，究竟像深海細菌這類極其簡單的生物是否具有原初意向性，是頗令人懷疑的。Dretske (1980) 的主張似乎可以推出任何系統都具有意向性。但是 Searle 已經強調要具有原初意向性至少該系統必須具有類比於人腦的因果能力。Fodor (1986) 亦指出在生物層級中有一界線區分具有表徵能力的系統與不具表徵能力的系統，而類似深海細菌的簡單生物並不具有其他區分這兩類系統的屬性。這裡我傾向於 Searle 與 Fodor 的主張，不過這個問題將不影響本文的討論。

再賦與而消除，由G引發的表徵將不再是錯誤的。這是Dennett功能再賦與論證的本意所在。但是自然表徵論決定表徵內容的一項要件是存活原則。如果該生物之表徵F而不是G才能使它產生有助存活的行為時，它的機制功能應該理解為 $f_1$ 而不是 $f_2$ ，它的表徵內容應該理解為F而不是G。<sup>9</sup> 依據存活原則，Dennett的功能再賦與論證並不足以拒斥自然表徵論對於錯誤表徵現象的解釋。

到目前為止，所有的反例都局限在只滿足自然表徵論的第一要件而已。因此自然表徵論的支持者很自然地可以用第二要件，也就是存活原則，來辯駁這些反例。然而，Fodor(1984)指出，即使自然表徵論有這兩項決定表徵內容的要件，仍然有可能有一個以上的方式可以滿足自然表徵論的要求，因而仍然有可能有一個以上的物體為某個心理表徵所表徵(p. 49)。按照自然表徵論的主張，心理表徵的表徵內容最後乃是取決於存活原則，但是我們可以設想存活原則同時被兩個功能賦與所滿足，進而導出該心理表徵內容是無法決定的結論。這個情形可以稱為「目的論的同等性」(teleological equivalence)。Fodor(1990a)提出了這個例子：

正如同我們對於青蛙的某個機制為何是測知蒼蠅的出現與否，可以提出適當的自然表徵論的解釋，我們也可以提出它的機制是要測知「會飛的小黑點」的出現。只要在青蛙的自然環境裡絕大多數的會飛的小黑點都是蒼蠅，這個賦與仍然符合存活原則。<sup>10</sup> 依據第一種賦與（蒼蠅測知機制），青蛙如果捕食BB彈，它便是錯誤地將BB彈表徵為蒼蠅。但是依據第二種賦與（小黑點測知機制），青蛙如果捕食BB彈，它並未出現任何錯誤，因為BB彈和蒼蠅一樣都符合小黑點的賦與。此時，如果我們真要說有任何錯誤發生，我們只能說出錯的不是青蛙，而是青蛙所處的自然環境出了問題：BB彈「不該」出現在青蛙的自然生態環境裡。這裡，「蒼蠅測

<sup>9</sup> 這個觀點來自於Dretske(1988), p. 84與Papineau(1987), pp. 71-72。

<sup>10</sup> 這裡我們忽略如蚊子等其他青蛙可吃食的小昆蟲。



知」與「小黑點測知」兩個賦與都符合自然目的論的要求。青蛙的機制究竟是表徵蒼蠅，還是表徵會飛的小黑點呢？顯然，自然表徵論不能解決「目的論的同等性」這個問題。這個目的論的同等性為自然表徵論帶來了「區域性的知識論上的不確定性」(local epistemic indeterminacy)問題。

或許有人認為Fodor (1990b)「反對稱依靠性」(asymmetric dependency)的觀念可以解決這個難題。我們可以依據Fodor，而將青蛙的那個機制瞭解為具有測知小黑點是否出現的功能，而不是蒼蠅是否出現的功能，即使這兩種賦與在目的論上都是同等的。根據這個想法，既然可以中斷青蛙的某個表徵與蒼蠅的聯結而不影響該表徵與「會飛的小黑點」的聯結，換言之，既然存在某個可能世界，在其中該表徵與蒼蠅的聯結是反對稱地依靠於該表徵與「會飛的小黑點」的聯結，我們似乎可以說青蛙的機制具有的是測知會飛的小黑點是否出現的功能，而不是測知蒼蠅是否出現的功能。

不過，以我之見，我們似乎不必使用到Fodor這個「反對稱依靠性」的觀念來解決「目的論的同等性」帶給自然表徵論的難題。對於自然表徵論來說，在出現目的論的同等性的情形裡，如果出現兩種對於機制的賦與都符合存活原則的情形，問應該選擇哪一個功能賦與其實已是多餘的了。究竟青蛙測知並捕食的到底是蒼蠅還是會飛的小黑點其實無所謂，重要的是到底青蛙捕捉的是否是有助於它們自己（或整個青蛙族群）存活的食物。從自然表徵論的立場來看，青蛙的那個機制所表徵的，不論是蒼蠅還是會飛的小黑點，其外延在目的論上是相同的(teleologically coextensive)；BB彈雖然也是會飛的小黑點，但是由於存活原則的緣故，並不落在青蛙的表徵外延裡。其實從自然表徵論的觀點來看，或許將青蛙的那個機制的功能定為測知「可食的物體」是否出現，既不是測知蒼蠅也不是會飛的小黑點是否出現，要更為恰當。

或許有人質疑雖然上述的看法至少在較低等的、認知與表徵

機制較簡單的動物的例子裡可以成立，不過一旦我們考慮較高等的、認知與表徵機制都非常複雜的動物的情形時，功能再賦與的困難將不可免。的確，原則上功能再賦與的困難在高等動物的情形裡是有可能出現的，即使實際上要舉出這樣的例子有其困難。區域性的知識論上的不確定性仍然存在。然而，上述所提的「可食的物體」的回答方式仍然適用。如果真的出現功能再賦與以及目的論的同等性這類情況時，到底採用哪一個賦與已經無所謂。但是在自然表徵論裡允許實用的因素以及知識論上的不確定性，並不對自然表徵論構成威脅。自然表徵論是針對「表徵內容是如何決定的？」這個問題，不是針對「表徵具有哪一個確切的内容？」提出解決。

功能再賦與論證是否能夠支持Dennett對於原初意向性的否認呢？討論至此，我們應該可以看出答案是否定的。至於這個論證為自然表徵論帶來的區域性的知識論上的不確定性，我想再提出兩點簡短的觀察：

(1)這裡Dennett並未以知識論的不確定性來拒斥自然表徵論所提的決定表徵內容的兩個條件，認為這兩個條件既非充分又非必要。知識論的不確定性只是用來顯示這兩個條件的使用必須相對於一個目的論的參考架構。但是這並不足以顯示我們不能決定在某些情境裡錯誤表徵是否出現。只要固定了目的論的參考架構，甚麼情況是錯誤表徵的情況是可以決定的。

(2)正如第二節所討論的，Searle與Dretske將「原初意向性」理解為一種具有因果能力的認知狀態所表現的意向性或是不依賴於外在詮釋者的意向性。然而，Dennett的功能再賦與論證既不是在否定人的認知狀態具有因果能力，也不是在主張人的意向性必須依賴於外在的詮釋者。因此，Dennett無法以功能再賦與論證拒斥自然表徵論對於錯誤表徵的解釋。

的確，我們人類自己就是自己心靈表徵內容的詮釋者，而且也是由我們人類（科學家）決定詮釋的參考架構。這是所謂的「自我

詮釋」(self interpretation)。這個自我詮釋的特徵的確為自然表徵論引進了所謂的「廣域性的知識論上的不確定性」(global epistemic indeterminacy)。「廣域性的知識論上的不確定性」之出現，不僅是由於功能再賦與的可能性仍然存在，更是由於不論各種可能的原始資料為何，原則上有可能有許多同樣適用的目的論的參考架構。簡單說，存在有許多種自我詮釋的可能。但是這個可能性的存在將不只會對自然表徵論構成威脅，事實上它會對所有的表徵理論構成威脅。這裡我們只能說，如果「廣域性的知識論上的不確定性」會構成問題，它是所有人的問題。事實上，我傾向於同意 Searle 的看法。Searle (1992)認為原初意向性是人類生物本性的一部份，而與人類如何使用它或者人類如何看待他自己或者人類選擇如何描述他自己無關(p. 79)。原初意向性只是有關人類的一個事實而已。

暫時不論人類的原初意向性是否只是一個事實或者受到自我詮釋所帶來的「廣域性的知識論上的不確定性」的影響，Dennett 甚至於並未用這個自我詮釋的可能以及廣域性的知識論上的不確定性來否定人類具有原初意向性。相反的，Dennett 所採取的策略是論證表徵內容是依賴於外來的意向性，因此人的意向性也不是原初的，而是衍生的，依賴於這個外來的意向性的。但是奇怪的是，他這裡所謂的「外來」的意向性竟然是來自於「自私基因」或「大自然」本身的意向性。這個策略似乎削弱了 Dennett 否定人類原初意向性的論點。尤其，依據 Newton (1992)的詮釋，Dennett 其實將大自然理解為具有「假似意向性」而已。如果 Dennett 的確主張大自然只具有「假似意向性」，說人的意向性是依附於大自然的意向性將是更奇怪的事。底下我們將檢討他這個論證。

#### 四、自私基因論證

Dennett 的自私基因論證是 Dawkins「自私基因」理論的一個變形。Dennett (1987)主張我們人類都是「天擇的機器」(survival ma-

chine)設計，來延續我們的自私基因的存活的；自私基因由於不能快速而有效的行動，以達到它們自己的利益，因此將人類設計為「天擇機器」，人是自私基因所造的，人也是「被造物」。所以人類的意向性是從我們的自私基因的意向性衍生的(p. 298)。人的一切行爲、知覺、相信等認知都是相對於大自然進化的目的(p. 300)。Dennett 因此批評 Dretske，認爲他在解釋表徵內容時雖然看到了表徵內容是獨立於外在的詮釋者，可是他沒有看到表徵內容並未獨立於大自然的意向與目的，因而也沒有看到人的意向性最終也只不過是衍生的(p. 305)。

然而，Dennett 的自私基因論證並不具信服力。第一，他的主張基因或大自然具有（原初）意向性是錯的；第二，在 Dennett 的論證裡所提到的三個詞：「被造物」、「衍生意向性」以及「相對性」都有問題。讓我們依序探討這三個觀念：

(1)「被造物」。顯然，說人是「被造物」與說自動販賣機是「被造物」有很大的不同。我們非常清楚自動販賣機的確是被造的。它是經由人類的設計以滿足人類的需求，是人類智能的產物。但是在說我們人類自己也是被造物時，也是同樣的意思嗎？說人類是「被造物」、是由大自然「設計」的，至多只是一種比喻的說法而已，大自然或演化歷程本身並沒有智能或目的。

Dennett (1987)當然知道演化的過程並沒有預見一個未來的目的。因此，說人類是「被造物」與說自動販賣機是「被造物」是不同的。人類是「被造物」、是由大自然「設計」出來的說法只是比喻性的、擬人式的說法，但是說自動販賣機是「被造物」則不是。Dennett 應該能同意這點。他的「自私基因論證」顯然並未達到他對於人的原初意向性的否定。

或許 Dennett 的主張有較深刻的角度，或許問題的關鍵在於我們如何瞭解自然目的論。我們可以將物競天擇的演化歷程看作是一個隨機的歷程。今天仍存活的種屬之所以存活到現在純是由於

它們的祖先碰巧進化得很好而足以面對它們生存的自然環境。或許我們可以主張進化是目的導向的(goal-directed),而不是盲目的,物競天擇的演化歷程是有理性的。如果持這個看法,所謂人是大自然設計出來的,就可能不再是一個比喻的、擬人式的說法。這個有關進化的哲學問題已非本文所能處理,只有暫述到此。

(2)「衍生意向性」。類比於上述對於「被造物」的說明,同樣的,人的意向性是從大自然或自私基因的意向性衍生出來的說法也只是比喻的、擬人式的。自動販賣機的意向性之所以是衍生的,是由於我們人類賦於它一個詮釋,以滿足人類的需求。(事實上,自動販賣機有意向性這個說法本身就是比喻的、擬人的說法。)但是,主張人的意向性是衍生自大自然或我們的自私基因的意向性的說法,卻是很奇怪的事。大自然或是自私基因並沒有類比於人類設計自動販賣機的智能,基因是否能有最起碼的意向性都是一個值得懷疑的主張。對於人的意向性是從自私基因的意向性衍生出來的這一點,實在沒有足以令人信服的理由。

在第二節的討論裡,我們約略敘述了Searle和Dretske對於原初意向性的理論。任何企圖駁斥他們理論的論證都應該針對他們對於原初意向性的理解。或者要論證出人的意向性和自動販賣機的意向性本質上沒有甚麼不同,都是具相同意義的衍生意向性;或者要論證出人的表徵系統缺乏因果力,就如同Searle「華語房間論證」的作法一樣。可惜的是,Dennett的「自私基因論證」兩者都沒有做。

說我們的自私基因或是大自然具有意向性的確非常反直覺。或許Dennett真正的訴求在於我們人類的意向性,心理表徵內容的決定,必須取決於物競天擇的演化歷程。或許Dennett是主張原初意向性實質上來自於大自然演化歷程。但是這個主張早就為支持「自然目的論取向的心理表徵論」的哲學家(Dretske, Fodor, Millikan, Papineau)所認同。自然表徵論的第二要件就是「存活原則」,心理

的表徵能力乃是受到自然演化「型塑」的(teleologically shaped)。因此,如果Dennett「自私基因論證」的訴求與自然表徵論沒有甚麼不同,自然也不足以否定人的原初意向性,進而拒斥自然表徵論對於錯誤表徵的解釋了。

(3)「相對性」(relativization)。自然表徵論對於表徵內容的決定是相對於自然目的論的。這是因為我們曾經提過的「自我詮釋」這個特徵的存在。但是這種相對性與我們之說自動販賣機的意向性是相對於詮釋者的詮釋,以及人的意向性是相對於大自然或自私基因的意向性的說法,是截然不同的。這是由於表徵內容的決定有賴於認知機制的生物功能的決定,只要決定表徵內容的目的論的脈絡的決定方式不是進一步地依賴於外在意向或目的,這個相對性其實是可以接受的。Dennett應該論證但沒有做到的是,他沒有論證出:人的意向性之相對於自私基因的意向性,與自動販賣機的意向性之相對於人的意向性,兩者具有同樣的意義。

在以上回應Dennett以自私基因論證否認人的原初意向性的討論裡,我們指出「被造物」、「衍生意向性」、以及「相對性」這三詞在Dennett自私基因的論證裡有歧異。這三重歧異導致自私基因論證喪失了它的信服力。事實上,Dennett的自私基因論證可以從三方面來理解,而不論是哪一個方式都無法導出人類不具有原初意向性的結論。

第一,我們可以將自私基因論證理解為一種比喻、擬人的說法。在這理解下,自私基因論證或許有趣,但是與自然表徵論或原初意向性的問題毫無關連。

第二,自私基因論證不是比喻的、擬人的。大自然或自私基因的確具有意向性,而且人的意向性的確依賴於自私基因的意向性之上。然而,這個說法不僅荒謬,而且缺少任何哲學上的或科學上的理由;Dennett更未曾提出任何這方面的有利論證。意向性是意識的根本特徵,說自私基因或大自然具有意向性就是說自私基因或大

自然具有意識。我們有任何哲學理由主張不具意識但卻具有（原初）意向性的系統存在嗎？

事實上，我們可以舉一個類比於Dennett自私基因論證的類似論證，姑且稱之為「神論證」：人的意向性不是原初的而是衍生的，因為人是神創造出來的「被造物」，人的意向性必須依賴於神的意向性，人的意向性是神賦與的。可惜的是，用這個神論證來否定人具有原初意向性，只怕不是很好的策略。

第三，Dennett的自私基因論證指出的是，對於人的意向性的瞭解有賴於自然目的論（存活原則）的參考架構。但是如果Dennett的自私基因論證所要強調的只是如此，他的主張與自然表徵論沒有基本上的差異。既是如此，自私基因論證如何能否定人具有原初意向性？我們不否認人的意向性有可能是衍生的。很可能Searle與Dretske對於原初意向性的瞭解根本就是錯誤的，原初意向性不能以自然表徵論的方式來理解。但是，在Dennett未提出另一套與自然表徵論相競爭並藉之以論證人的意向性是有賴於大自然的意向性之前，我們無法接受他對於人的原初意向性的否定。

## 五、人的原初意向性被否定了嗎？——結語

在上述的討論裡，我們駁斥了Dennett否定人具有原初意向性的兩個論證。在功能再賦與論證方面，我們主張由目的論的同等性所帶來的區域性的知識論上的不確定性這一問題，對自然表徵論而言並不構成難題。這是因為在面臨目的論的同等性時，應引進實用層面的考慮，而這個實用層面的考慮對於表徵內容的決定以及錯誤表徵現象的解釋是無害的。其實，功能再賦與的可能性也有其物理上的或生理上的限制。例如，無論油表的設計有多精密，我們都不可能將它的功能重新描述為測量空氣的含量。要使一個油表能測量空氣的含量只有改變這個儀器的整個構造，但如此一來，它也已經完全是一個新的儀器，而不再是油表了。在這情形下，我們談的不

再是功能再賦與的可能性，而是不同儀器的不同功能。同理，我們對於生物有機體機制的功能再賦與也受到生物或生理學的限制，否則我們談的功能再賦與的可能性就會變成談不同生物機制的不同功能了。

另一方面我們承認廣域性的知識論上的不確定性是大難題。這個難題出自於人的認知具有「自我詮釋」的特徵。這個難題或許是任何表徵理論或意向性理論都無法解決的，或許將來有新的理論可以克服，目前只有存而不論。

至於Dennett的自私基因論證，在上面的討論裡我們已清楚指出這個論證可以從三方面來瞭解。但是其中任何一個方式都不足以幫助Dennett達到否定人有原初意向性的結論。一個類比於自私基因論證的神論證指出了自私基因論證實在不具有說服力。

綜合以上的討論，我們可以大膽駁斥Dennett對於人的原初意向性的否定，並進一步維護自然表徵論對錯誤表徵現象的解釋。



## 參考書目

- Armstrong, D.M. 1966. "The Nature of Mind." D.M. Armstrong, *The Nature of Mind and Other Essays*. Ithaca: Cornell University Press.
- Dennett, D.C. 1987. "Evolution, Error, and Intentionality." In D.C. Dennett, *The Intentional Stance*. Cambridge, MA: The MIT Press, pp. 287-322.
- Dretske, Fred I. 1988. *Explaining Behavior*. Cambridge, MA: The MIT Press.
- \_\_\_\_\_. 1986. "Misrepresentation." In W. Lycan (ed.), *Mind and Cognition*. Cambridge, MA: The MIT Press, pp. 129-143.
- \_\_\_\_\_. 1985. "Machines and the Mental." In Western Division APA Presidential Address, *Proceedings and Addresses of the APA*, vol. 59, pp. 23-33.
- \_\_\_\_\_. 1980. "The Intentionality of Cognitive States." In P.A. French, T.E. Uehling, and H.K. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. 5, pp. 281-294.
- Fodor, J.A. 1990a. "A Theory of Content, I: Problem." J.A. Fodor, *A Theory of Content and Other Essays*. Cambridge, MA: The MIT Press, pp. 51-88.
- \_\_\_\_\_. 1990b. "A Theory of Content, II: Theory." J.A. Fodor, *A Theory of Content and Other Essays*. Cambridge, MA: The MIT Press, pp. 89-136.
- \_\_\_\_\_. 1986. "Why Paramecia Don't Have Mental Representations." In P.A. French, T.E. Uehling, and H.K. Wettstein (eds.), *Midwest Studies in Philosophy*, vol. 10, pp. 3-23.
- \_\_\_\_\_. 1984. "Semantics, Wisconsin Styles." In J.A. Fodor, *A Theory of Content and Other Essays*. Cambridge, MA: The MIT Press, pp. 31-50.
- Haugeland, John. 1981. "Semantic Engines: An Introduction to Mind Design." In J. Haugeland (ed.), *Mind Design*. Cambridge, MA: The MIT Press, pp. 1-34.

Newton, N. 1992. "Dennett on Intrinsic Intentionality." *Analysis* 52(1), pp. 18-23.

Papineau, D. 1987. *Reality and Representation*. Oxford: Basil Blackwell.

Peng, M.Y. 1994. *An Examination of Fodor's Disjunction Problem and the Nature of Misrepresentation From Within Natural Teleological Theories of Intentionality*. Doctoral Dissertation. University of Iowa.

Searle, John. 1992. *The Rediscovery of the Mind*. Cambridge, MA: The MIT Press.

\_\_\_\_\_. 1980a. "Minds, Brains, and Programs." In J. Haugeland (ed.), *Mind Design*. Cambridge, MA: The MIT Press, pp. 282-306.

\_\_\_\_\_. 1980b. "Response to Comments on 'Minds, Brains, and Programs'." *The Behavioral and Brain Sciences* (3), pp. 450-456.

## Original Intentionality and Misrepresentation

*Eric M. Peng*

### Abstract

The ability of a cognitive system to misrepresent presupposes that the intentionality the system exhibits is original or intrinsic rather than derivative. This is true especially when we are talking about human intentionality, because, on the naturalistic representationalism proposed separately by Dretske, Fordor, and Papineau, the fixation of representational content is prior to what counts as a case of misrepresentation and has to be done in a non-derivative way.

Dennett, however, rejects the claim that human intentionality is original or intrinsic via his, as I call them, “Argument From Functional Re-ascription” and “Argument From Selfish Genes.” He argues that human intentionality is derivative either because of the possibility of re-ascribing functions to human intentionality is derivative either because of the possibility of re-ascribing functions to human cognitive systems or because of the reason that human beings derive their intentionality from the intentionality of our selfish genes. He thus argues that what are taken to be cases of misrepresentation might be dismissed and become cases of veridical representation.

This essay first presents what human original intentionality is by a brief survey of Searle’s and Dretske’s theories, since Dennett takes their views as his main target. His two arguments are then examined. With regard to the Argument From Functional Reascription, I argue that we stick to the survival principle (natural teleology) which is a major factor in the fixation of representational content in the naturalistic representationalism, even if we come to cases in which functional ascriptions are teleologically equivalent. With regard to the Argument From Selfish Genes, I examine three possible defences for the view that genes have intentionality from which human intentionality is derivable. I argue that none of the three defences work. I thus reject Dennett and defend the view that human intentionality is original or intrinsic and that the way the naturalistic representationalism explains misrepresentation is not affected.