

# 宗教的果報觀如何促進道德行為與社會 利益：以演化賽局模型分析

張榮富

台北教育大學社會科教育學系助理教授

## 摘要

本文主旨在於以模擬的道德情緒演化賽局模型，探討宗教的「果報觀」如何促進道德行為與社會利益。首先，本文介紹法蘭克(Frank,1988)結合道德情緒與心理學「配合法則」(Herrnstein,1970)所型成的「名譽模型」。並介紹運用「名譽模型」討論「果報觀」如何透過獎懲、罪惡感和報應的機率三方面影響道德行為的近期研究(張榮富、鄭雅安，2007)。其次，根據前述研究，假設「果報觀」的信仰能增強法蘭克理論中的「名譽現象」，亦即信徒的「身份」可以成為內在德行的可靠外在「傳訊」特徵。將此假設帶入法蘭克的道德情緒演化賽局模型，模擬分析並討論果報觀在促進社會利益上所扮演的角色與價值。本文主要結論為：(1) 果報觀信仰能增加短期與長期的社會整體利益。(2) 果報觀信仰帶來的不是「社會正義」而是「社會利益」。最後本文亦以模型探討兩個相關問題：(a)為什麼在工業文明之前的富強國家多也是有果報觀信仰的國家？(b)為什麼信仰深者的「魔鬼的誘惑」(不義利得的誘惑)也將愈大？

**關鍵字：**因果報應 宗教 道德情緒 演化賽局 傳訊 社會利益

## 壹、緒論

人類的道德情緒究竟是從何而來？是先天即具有道德情緒，抑或是受後天教化而產生？從古至今皆備受爭論。天論者，主張道德是經由人類演化過程中產生的。反之，後天論者認為，道德乃是社會後天教化的結果，人皆存有慾念之心，透過後天教化的力量，才使得人們具有拒絕誘惑的品格。儘管先天論或後天論各有其理論依據，但要**更提升**人的道德行為的有效途徑則皆需透過後天的道德教化。

從歷史上來看，宗教無論對東方和西方社會的道德發展都有著深刻的影響，甚至在許多地區，地方上的道德觀念即是由宗教發展而來的。在過去社會的大部分時間裡，宗教和道德並未有明顯的分別，宗教在某些時代與某些社會中曾被放入正式的道德教育體系中，而近代在多數有宗教自由的國家中，宗教可被視為一種非正式的道德教育。

過去的相關研究多集中在正式的道德教育對道德的影響上，而少有人聚焦於非正式的道德教育其所可能發揮的影響力。宗教可說是非正式的道德教育中最重要途徑之一。很多人可能會認為宗教是藉由規範的方式強制信徒遵守道德行為，不過本文作者發現法蘭克(1988)對於道德情緒有很特別的道德先天論的見解。雖然屬於道德先天論的範圍，但可運用於道德教育領域(張榮富，2007)。因此嘗試以法蘭克道德情緒理論中的演化賽局模型為基礎，延續法蘭克的「功利主義」與「演化論」的觀點，分析宗教的「果報觀」在非正式的道德教育上所扮演之角色與其對社會整體利益的影響。或許可以為宗教與道德方面的研究提供另一種不同的思考方向。

本文第貳節將先介紹道德情緒的在先天論方面的各家說法，其次介紹法蘭克結合道德情緒理論與心理學的「配合法則」(Matching Law)(Herrnstein, 1970, 1997)而成的「名譽模型」。最後介紹張榮富、鄭雅安(2007)以法蘭克的「名譽模型」討論宗教的「果報觀」如何影響道德行為的成本效益模型。第參節則首先介紹法蘭克以承諾困境與傳訊理論為基礎的道德情緒「演化賽局模型」，說明社會中道德人口組成

的演變情形與果報觀的道德教化如何促進個人與社會的利益。其次，根據第貳節之二，張榮富及鄭雅安(2007)的結論(即宗教的「果報觀」可以促進道德行為)為推論的起點，假設「果報觀」的信仰能增強法蘭克理論中的「名譽現象」，亦即信徒的「身份」可以成為內在德行的可靠外在「傳訊」特徵，以此來分析宗教「果報觀」對整體社會利益的影響。第肆節則歸納本文推論與結論。

## 貳、宗教的果報觀對道德行為的影響

### 一、道德的先天性及其演化

根據張榮富(2007)的文獻介紹，近代演化生物學中談論道德的方向主要著力於道德感或道德情緒的先天性及其演化(e.g, Alexander, 1987; Axelrod, 1984; Boorman & Levitt, 1980; Hamilton, 1964; Trivers, 1971)。要討論道德的先天性可以從演化論談起，達爾文(Charles Darwin)在 1859 年發表的《物種起源》(The Origin of Species)中提及「自利的對象是個體本身」，對人性的自利行為建立了早期的理論基礎。然而，隨著生物學基因理論的發展，道金斯(Dawkins, 1976)修正了達爾文的想法，認為演化的基本自利單位是「基因」，而非生物個體。順著這種思維，演化論開始產生了一些對利他行為的解釋，而很重要的一個理論即是漢密爾頓(Hamilton, 1964, 1971)的近親選擇(Kin selection)理論。漢密爾頓認為：一個個體常為了基因的遺傳流布而選擇自我犧牲，以利近親的生存。然而近親選擇理論卻無法解釋非近親者之間的利他行為。群體選擇論(Group Selection)與互惠性利他(Reciprocal Altruism)理論則試圖彌補此一學術缺口，不過各有其弱點。

群體選擇論認為個體的選擇有「利群」的偏好天性，也就是說自利是以「群體」為單位。然而演化學者 Trivers (1985)針對此假設提出了「群體選擇的謬論」(The Group Selection Fallacy)，認為當個體間的利群程度不完全一致時，少數的較自私的利群者，會比其他利群者有最佳的生存機會與繁衍更多的後代，如此循環演化的結果，一個原本以「利群」個體所組成的群體最後將完全質變為充滿著自私的個體。

互惠性利他理論則認為，生物個體之所以不惜降低自己的生存競爭力，幫助另一個與自己毫無血緣關係的個體，是因為它們期待日後得到回報，以獲取更大的收益。利他在此理論當中變成是一種變相自利的投資期待，互惠性利他被 Wilson (1975) 稱為「軟性利他」(Soft-core Altruism)，但這並非一般人所認為的真正的利他道德行爲。

在上述理論之後，Trivers (1985) 提出了中介性情緒(Mediating Emotions)的概念，中介性情緒(包括如同情心、憐憫心、報復心、罪惡感等)可以幫助我們解釋為什麼很多時候我們會對無基因關係者仍會採取利他行爲。但 Trivers 的理論仍被經濟學者法蘭克(Frank, 1988)指出了其中的缺陷。Trivers 的理論無法證明這些情緒如何能對個人(利他者)在**物質利益的獲得**上有所幫助。假如利他者因道德情緒而有利他行爲但其物質利益的報酬小於經濟學上的理性自利者，則根據演化的原則：**物質資源較少者子孫將較少**，在繁衍子孫的競爭中，有此道德情緒的個體如何得以競爭得過極大化自利的個體呢？或至少不被完全淘汰呢？法蘭克試圖解答此一疑惑，在其名著 *Passions within Reason* (Frank, 1988) 中，結合了心理學、演化論與賽局理論，建立了道德情緒理論。首先，法蘭克以「承諾困境」來說明道德情緒如何在演化上有利於人的生存與物質利益的獲得，並且其物質利益的報酬可能大於理性自利者。以下是法蘭克的理論見解<sup>2</sup>：

經濟學所謂「理性選擇」假設中對人性有兩個最基本定：(1) 人性是自利的(Self-interest)；(2) 人在作選擇時企圖極大化(Maximizes)其利益。法蘭克認為「道德情緒」，例如恥辱 (Contempt)、憎惡 (Disgust)、忌妒(Envy)、羞愧 (Shame)、和罪惡感 (Guilt) 等，能使人對抗那些源自於物質報酬的理性計算，使人**不企圖**完全極大化其利益。

**不企圖**完全極大化自己利益有什麼物質利益上的好處呢？何以**不企圖**完全極大化自己利益的人在演化中能有較多子孫呢？首先，法

<sup>1</sup> 即類以於 Trivers (1985) 的「中介性情緒」。

<sup>2</sup> 對法蘭克理論的更詳細的介紹，請參看張榮富(2007) 請參看張榮富(2007) 『從道德演化的「賽局模型」與「名譽模型」談道德教育—Frank 道德情緒理論之應用』一文。

蘭克在其「承諾模型」中提出，現實生活中我們常常會遇到承諾的問題(Commitment problem)，這類問題有一個共通的特徵，即**當一個人的利益決定於一個「捆綁的承諾」(Binding Commitment)**(例如，**如果你這樣做我就一定會不計代價地那麼做**)，而此一「捆綁的承諾」的內容卻是**違背個人的極大化利益**。法蘭克認為「道德情緒」在演化上的第一個意義為「解決承諾問題而獲益」。換言之，在承諾的問題的人際互動的情境中，做出一個違背個人極大化利益的「捆綁的承諾」，其結果反而會比追求極大化自利來的更為有利。追求極大化的理性自利者在此卻碰到了困境，他做不出這種承諾，縱使說出了也不易取信於人。在此，道德情緒就扮演著一個解決承諾問題的策略性角色。有道德情緒者較可能會做出損害自己極大化利益的(傻)事，而且別人也傾向相信他會做出這種事情。當有道德情緒者在「承諾的困境」中成功的創造出**違背個人極大化利益的「可信的威脅」**或「可信的承諾」時，他會因此獲得比不能做出這種**可信的威脅或承諾**的人(追求極大化自利的理性者)更大的物質益利<sup>3</sup>。因此懷有道德情緒的人類遠祖先的物質資源不會比較少，根據演化的原則(物質資源較少者子孫將較少)，其子孫也不會比較少，所以此道德情緒的基因得以有機會在演化的競爭中流布。

其次，法蘭克引用了心理學者 Richard Herrnstein(1970, 1997)的配合法則(matching law) 發展出「名譽模型」，來解釋為什麼欺騙行為容易產生？又為什麼有人可以不欺騙？所謂配合法則，可以用一個有趣的鴿子實驗來說明。實驗者為鴿子設置了兩種按鈕提供飼料，A 組設定甲按鈕按下經 30 秒後會出現小顆的飼料，而乙按鈕按下經 40 秒後會出現大顆的飼料，兩種按鈕出現飼料的時間只差距 10 秒，在此設計之下，鴿子大多選擇了乙按鈕，即只需多等 10 秒，得到大顆的飼料。B 組實驗設定甲按鈕無須等待，只要按下就立刻會有小顆的飼料出現，而乙按鈕也只須等 10 秒，就有大顆的飼料掉下來。實驗結果卻發現，此情況下鴿子多數選擇立即掉下小顆飼料的甲按鈕。同樣

<sup>3</sup> 詳細推論與例子，請參看張榮富(2007)『從道德演化的「賽局模型」與「名譽模型」談道德教育—Frank 道德情緒理論之應用』一文。

的差異也在以人爲受測者的選擇實驗中上得到證實。多數人會偏好選擇在 31 天後獲得 120 美元，而非在 28 天後獲得 100 美元。然而當報酬的時間改變時，卻會選擇立即獲得 100 美元而非在 3 天後獲得 120 美元。

上述實驗得出了配合法則中的重要原則，即「**報酬的吸引力**」與「**延遲實現的時間**」成反比，而且**立即的報酬(或成本)會在心中被超比例的大幅放大**。如果報酬是可以立即獲得的，則此報酬的吸引力將會在心中被大幅放大，反之，若報酬實現的時間延遲，則隨時間的增長，報酬於人心中所造成的影響力將會被嚴重的折舊(縮小)。在前面兩個實驗中我們都可以發現，如果報酬的獲得是在一段時間甚至數天之後，人和鴿子會願意多延長一點等待的時間來獲取更高額的報酬。然而，如果報酬是立即可得的，那麼雖然立即獲得的報酬會比一段時間後少一些，但是因爲當下立即可得的報酬於心中所產生的吸引力會被**超比例**的放大，因此人和鴿子都傾向於去選擇**立即的報酬**。

法蘭克認爲配合法則正好可以用來解釋人的道德行爲，他以欺騙的例子來做說明。欺騙行爲本身可因欺騙而帶來**立即的報酬**，其所需付出的代價則是，若被他人發現後的壞名譽或刑罰。然而，因爲欺騙行爲的報酬往往是能**立即獲得的**，它在人的心中會受配合法則作用而大幅放大其影響力。欺騙後若被發現所需付出的壞名譽或刑罰則是發生在較久遠的未來。則隨時間的增長，成本於人心中所造成的影響力將會被配合法則嚴重的折舊(縮小)。「大報酬」減「小成本」的結果將使得欺騙行爲在當下心中似乎變得有利，騙子因此易傾向選擇從事欺騙行爲。

法蘭克認爲配合法則的影響並非不能被突破，人並非總是眼前放大的利益的奴隸。相反的，我們可以發現人很明顯的有相當的自制力去抗拒眼前的誘惑。如果行爲機制受到被大幅放大的**立即報酬**所影響，那麼最簡單的反制那個**立即報酬**的誘惑的方法便是，當時(同時)有另一個感覺(情緒)把行爲往相反方向拉。在上述欺騙的例子中，罪惡感即是這種情緒。罪惡感是道德情緒的一種。特別的是，這種情緒是在做行爲選擇(欺騙)的同時就會立即出現。它會隨著每個人的道德

感程度不同而有不同大小的影響力，使人感受到歉疚、不安或惶恐等不舒服的感覺。這樣的感覺因為是隨著選擇(欺騙)行為的同時就會立即出現，因此不會受到實現時間延遲所縮小。相反的，它是立即的情緒成本，將會受配合法則的作用而被放大。罪惡感的心理成本與欺騙行為原先即具有的壞名譽或刑罰的成本結合起來，將可能足以對抗被放大的欺騙報酬，使欺騙行為必須付出的代價反而可能大於獲利，因此有罪惡感的人比較會傾向於選擇不去從事欺騙行為。據此，法蘭克認為道德情緒在演化上的第二個意義為：道德情緒能抑制眼前被放大的報酬的誘惑，使得個人在長期中能建立「名譽」，而個人的名譽將會帶來與他人合作的機會。根據類似國富論(亞當斯密，1723-1790)的觀點，能與人分工合作者的利益大於單打獨鬥者或與人勾心鬥角者的利益。有道德名譽者將得到他人的信任與有更多的合作機會，因此懷有此一道德情緒的人類遠祖的物質資源不會比較少(演化的原則：物質資源較少者子孫將較少)，所以此道德情緒的基因得以有機會在演化的競爭中流布。

## 二、果報觀對道德行為的影響方式

從張榮富(2007)曾運用法蘭克的「承諾模型」與「名譽模型」來討論德行(Virtue)層面的道德教育。其文最後提及以法蘭克的「名譽模型」觀之，宗教之所以能增強人的德行(品格)，並非只是簡單的天國與地獄的獎懲問題，良心與罪惡感似乎一直是宗教的核心之一。張榮富及鄭雅安(2007)<sup>4</sup>延續此一觀點，運用配合法則與法蘭克的「名譽模型」，分析宗教的「果報觀」如何影響人的道德行為。此論文借由「名譽模型」與心理學上的「配合法則」(Herrnstein, 1970, 1997)推論得出：宗教的「果報觀」主要可藉由『獎懲』、『不道德行為成本的發生機率(報應的機率)』與『罪惡感』三方面來影響人類的行為選擇。以下是他們的理論見解：

---

<sup>4</sup> 張榮富、鄭雅安(2007.6.6)。以配合法則與道德情緒理論分析宗教「果報觀」對道德行為的影響。發表於『第四屆(教育)大學社會科教育相關系所研究生論文發表會』，國立台北教育大學，台北，中華民國。

該文中假設人類行為的理性選擇(從經濟學的角度來看)可視為一種「成本效益分析」,亦即依據「行為報酬」扣除「行為成本」之後所得的「實得報酬」的大小來做選擇時的判斷,若行為報酬大於行為成本(實得報酬大於0),則理性的人會選擇從事該行為;反之,若行為報酬小於行為成本(實得報酬小於0),則理性的人將會放棄從事該行為。

然而,當配合法則介入人類的理性選擇時,卻很容易引導出不道德的行為。因為大部分不道德的行為都能帶來立即的報酬,立即報酬受配合法則作用將在心中被大幅放大,而從事不道德行為所需付出的成本代價則往往發生在較久遠的未來,會因延遲實現而被縮小,「大報酬」減「小成本」的結果使得不道德行為在心中似乎是有利的(易感覺到實得報酬大於零),因此容易引導出不道德的行為。

理性行為模式加上配合法則容易引導出不道德的行為,這是人類心理上的弱點。然而,當宗教的「果報觀」作為非正式的道德教育介入時,其對行為最顯著的影響有三方面。首先,「果報觀」中的獎懲(如佛教中的「善有善報,惡有惡報」,基督教的末日審判等)扮演了一個不道德行為成本的角色。但此「遠期性懲罰成本」可能是發生在較遠的未來(或甚至是來世的報應),因此仍受配合法則作用在心中被縮小其影響力,使得獎懲所提供的行為成本不見得足以抑制從事不道德行為所獲得被放大的報酬的誘惑。

其次,由於當人們信仰宗教果報觀時,在人們心中相信發生因果報應的「機率」也會大幅增加。而且信仰愈虔誠,我們愈相信會發生因果報應。因此大幅增加的「報應機率」乘以(被配合法則縮小的)「遠期性懲罰成本」,將使得行為的總成本大幅地被提升。

最後,「果報觀」除了藉由有形的獎懲來約束人的行為外,也會透過心靈教化的方式增加人心中為惡時的罪惡感。罪惡感是一種道德情緒方面的心理成本。當人心念一動想要為惡時,罪惡感是一種無可避免地會立即出現在心中的成本,因此配合法則並不能去縮小它。當一個人的罪惡感愈強烈,它所形成的「罪惡感成本」將愈大,這又使得不道德行為的成本明顯增加。



總結而論，透過信仰「果報觀」所帶來的獎懲、報應機率和罪惡感三方面的改變，將使得不道德行為的成本在人心中大幅度的增加<sup>5</sup>。在其他條件都相同的情況下，一個虔誠的信徒若面臨不道德行為的選擇時，行為的實得報酬易傾向於負數。因此，我們可以發現社會上愈深信果報觀者通常比較不會去從事他心中所認知的不道德行為。

## 參、宗教的果報觀對社會利益的影響

由上節的介紹中可知，透過信仰「果報觀」中所帶來的獎懲、罪惡感和報應的機率三方面的改變，使得不道德行為的行為成本在人心中大幅度的增加。本節將於此一論點的基礎上，運用法蘭克道德情緒先天論中的演化賽局模型，分析「果報觀」的信仰所帶來的社會利益。

### 一、法蘭克道德情緒演化賽局模型

在上述第貳節之二中，法蘭克的承諾模型雖然成功的解決了 Trivers 中介性情緒理論的困境。不過法蘭克還繼續追問：如果有道德情緒者報酬較高而且子孫也較多，為什麼演化的結果是世界上還會有這麼多不道德的人存在？為什麼他們沒有在演化的洪流中完全被淘汰呢？法蘭克認為道德情緒雖能幫助我們解決承諾的問題，但這是有傳訊 (signaling) 方面的前提，即：個體要有能力發出訊號讓他人清楚地察覺到他的情緒，而且也要能正確地辨認出別人釋放出的情緒訊號。試想，如果無道德情緒者(騙子)能偽裝成有道德情緒者(誠實者)，而且有道德情緒者不能視破其偽裝，則無道德者將獲得更大的

<sup>5</sup> 以欺騙為例，行為者的行為模型如下列式子：

$$(1 - P)B - P \cdot (C_1 + C_2) - C_3 = \pi$$

$C_1$  為從事欺騙可能需付出的成本， $C_2$  則為宗教「果報觀」引發的獎懲成本， $C_1$  和  $C_2$  兩個行為成本皆受到實現時間的延遲而在心中被折舊縮小。 $B$  為欺騙的立即報酬，因受配合法則作用將在人的心中被大幅放大。 $P$  為欺騙報應發生的機率，因受「果報觀」的影響而放大。 $C_3$  為「果報觀」衍生的立即的罪惡感成本，因配合法則作用將在心中被大幅放大。 $\pi$  為實得報酬。[引自張榮富、鄭雅安(2007)之式(6)]

利益，這個世界將會是騙子的天下。

法蘭克以演化賽局模型來說明有道德情緒者與無道德情緒者的**並存**問題。為了解釋方便起見，法蘭克把模型中兩種可能的行為策略(有道德與無道德的行為)簡化，假設在人口中存有兩類人：合作者(Cooperators)與背叛者(Defectors)<sup>6</sup>。合作者在此例中為有道德情緒者，為誠實者，他們能克制自己在沒有可能被發現的情形下也不去欺騙別人。背叛者在此例中則是缺乏道德情緒的極大化自利的人，為欺騙者，他們是投機與偽善的騙子。當這兩類人相遇互動時(以尋找合作夥伴為例)，他們所得之物質報酬以分數計算如表一所示。表一中的報酬結構是賽局理論中典型的囚犯困境(亦即承諾的困境)的例子。當這兩類人被丟到生存競爭的環境時，每一類型人的人數演化會視模型中傳訊方面的假設而定，亦即，當其他條件不變時，辨視這兩類人的成本會影響這兩類型人的人數比例。

**表一 報酬結構**

	背叛者	合作者
背叛者	各得 2 分	合作者得 0 分 背叛者得 6 分
合作者	合作者得 0 分 背叛者得 6 分	各得 4 分

法蘭克根據現實社會的情形假設：背叛者可以偽裝成合作者，但無法完全偽裝(亦即偽裝有可能被視破)，辨視其偽裝需要付出 1 分的辨視成本。亦即，如果合作者付出了 1 分的辨視成本，則保證能與另一個合作者配對在一起；但是如果不付出辨視成本，則必定無法分辨出誰是背叛者誰是合作者。然而，合作者可視需要決定是否要支付辨視成本。

舉例來說，當原始的合作者的人口比例是 90% 時。如果不花辨

<sup>6</sup> 合作者(Cooperators)與背叛者(Defectors)為經濟學與演化賽局中的名詞，常用於標示正反兩類行為。

視成本，一個合作者與另一個合作者配在一起的機會為 90%，與一個背叛者配在一起的機會只有 10%。他的平均報酬將有  $(0.9 \times 4) + (0.1 \times 0) = 3.6$  分。這比花費辨視成本後所得到的  $4 - 1 = 3$  分報酬更高，顯然在此人口比例下，不支付辨視成本比較好。所以此時合作者會選擇不支付辨視成本。

但是，如果原始的合作者的人口比例不是 90% 而是 50% 時。如果不花費辨視成本，則現在合作者將只有 50% 機會與另一合作者合作。他的平均報酬因此只有 2 分，而花費辨視成本後所得到報酬為 3 分，顯然在這種人口比例下，花費辨視成本所得到的報酬反而較高，此時合作者將願意支付辨視成本。

事實上，這個例子隱含著一個「長期均衡點」(見圖 1)，即當合作者的人口比例是 75% 的時候，不支付辨視成本的合作者有 75% 的機會得到 4 分報酬，和 25% 的機會得到 0 分報酬，平均報酬是 3 分，也等於他已經花費辨視成本後的報酬 3 分。即在合作者的人口比例是在 75% 時，花不花費辨視成本，報酬都是一樣的。而長期均衡點的位置將隨辨視成本與報酬結構的改變而改變。

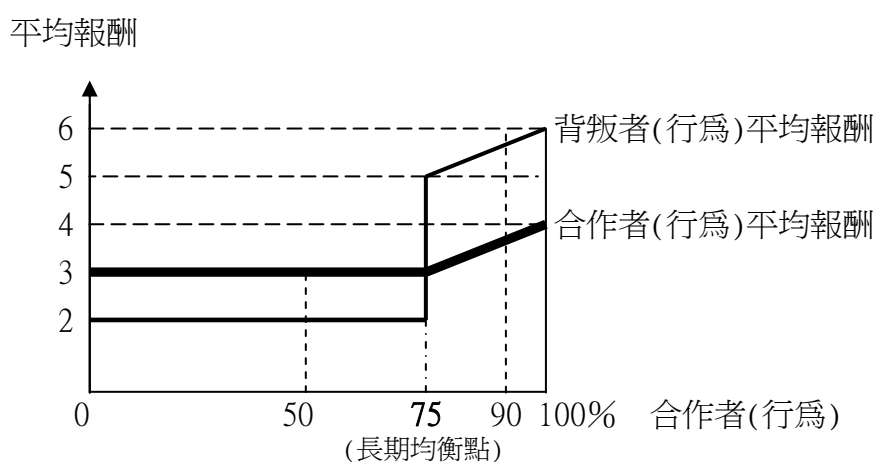


圖 1

圖 1 顯示在各種不同人口比例下，合作者與背叛者的平均報酬。當原始的(短期的)人口比例是在 75% (長期均衡點)左邊的時候，合

作者將藉由支付辨視成本而總是能找出其他合作者與之配對，得到 3 分的報酬。因為眼光銳利(付了辨視成本)的合作者不肯與背叛者合作，所以背叛者只好與其他背叛者合作，得到只有 2 分的報酬。因此，如果合作者的人口比例少於 75%時，合作者將會拿到比背叛者更高的平均報酬，這意謂他們的子孫比較多，合作者的人口比例將會持續成長至 75% (長期均衡點)。

但當原始(短期的)人口比例是大於 75%時，情況剛好相反。合作者沒有道理花費 1 分的辨視成本。然而此時背叛者的平均報酬反而會高於合作者，意謂背叛者的子孫會較多，合作者的人口比例將會下降至 75% (長期均衡點)。因此在這個例子中我們知道，短期的人口比例將在 75%這個長期均衡點左右來來回回的徘徊，而 75%為長期的穩定均衡點。當其他條件(報酬結構)不變時，演化的結果將是合作者與背叛者以 3：1 的比率並存於此社會中。

## 二、宗教與道德在道德情緒演化模型中的義涵

張榮富(2007)把上述模型套用在道德教育上，得出若干延伸的論點，解釋了道德教育的增強、個人利益與社會利益三者之關係。然而，由本文第貳節可知，宗教為非正式的道德教育，而「果報觀」的信仰，能增強人的道德行為，故上述模型亦可適用於「果報觀」的分析。根據第貳節之三張榮富及鄭雅安(2007)的研究，我們把「果報觀」的信仰定義為：「借由獎懲、罪惡感和報應的機率三方面的影響，增強信徒的內在德性，使得人為的而且**短期**的改變合作者(好人)與背叛者(壞人)的人數比例，使合作者的人數增加」。在此定義下，「果報觀」的影響在模型中的涵義如下：

- (1) 如果社會本來在短期上處於長期均衡點的左方，而「果報觀」信仰使社會人口比例在短期內能調整至均衡點的右方。以一個社會的總報酬(總產出)來看，處於均衡點的左方的低果報觀(或低道德)社會的總報酬會較小，而處於均衡點的右方的高果報觀(或高道德)社會的總報酬會較大。以一個社會有 100 人為例，當社會中合作者的人口比例是 50%時，此社會由人與人合作產生的總報酬為

$(50 \times 3) + (50 \times 2) = 250$  分。而當社會中合作者的人口比例是 90% 時，此社會由人與人合作產生的總報酬為  $(90 \times 3.6) + (10 \times 5.6) = 380$  分。

- (2) 前述結果，似乎也可以解釋，為什麼工業文明之前的各個文明社會都是自古以來較重視宗教果報觀(或道德)的社會？在族群與族群、社會與社會的競爭中，高果報觀(或高道德)的族群或社會的總產出將高於低果報觀(或低道德)的族群或社會。低果報觀信仰(或低道德)的族群或社會較易在演化競爭的洪流中被淘汰。
- (3) 有趣的是，如果由個別合作者或背叛者的利益的角度來看，則情況與(1)(2)相反。由圖 1 可知，在比較沒有道德的社會中(均衡點的左方)，個別合作者因背叛者很多，所以會處處提防以免受騙上當，亦即此社會中的合作者會支付辨視成本，因此其平均報酬是 3 分，而個別背叛者此時的平均報酬是 2 分，背叛者的平均報酬確定會小於合作者。反之，在一個壞人很少的社會(均衡點的右方)，例如合作者的比例為 90% 時，合作者需支付辨視成本，故個別背叛者的平均報酬反而高於個別合作者。果報觀信仰(或道德教育)帶來的似乎是「社會利益」而非「社會正義」(好人有好報，壞人有壞報)。一個公平正義的社會的產生不能只有宗教與道德，可能是要依靠其他因素，例如：法治、文化、或是政府對社會與經濟制度的干預。
- (4) 如前文所述，合作者與背叛者此兩類人口的假設，是爲了模型的方便解釋。我們可把模型中的兩類人視爲一個人心中的兩種可能的行為策略(有道德與無道德的行為)。由此再回頭來看圖 1，演化均衡點在此代表個人採行背叛策略與合作策略的平均值。75% 代表演化的結果，一般人平均而言，有 75% 採行合作策略，而 25% 採行背叛策略。在均衡點右方的任一點表示合作行為大於平均值的人，定義爲好人(有道德的人)。由此新定義觀之，一個 90% 採行合作行為的好人，其單次採取背叛策略時的報酬會大於單次採取合作策略時的報酬。反之，在演化均衡點的左方的任一點表示合作行為小於平均值的人，定義爲壞人(無道德的人)。一個採行 40% 合作行為的壞人，其單次採取背叛策略時的報酬會小於單次

採取合作策略時的報酬。簡單的說，一個好人偶爾一次的背叛行為(欺騙)的報酬會比一個壞人一次的背叛行為(欺騙)的報酬高很多。

- (5) 承上述，如果我們把信仰深者定義為更有道德的人，把「魔鬼的誘惑」定義為背叛行為(不道德行為)的「不義利得」，則我們可以推論得出，信仰深者(更有道德的人) 偶爾一次的背叛行為(例如欺騙)的不義利得會比常人與壞人高很多。隨著信仰的加深，個人將愈有道德，同時也愈令人相信他是好人(好的名譽)，然而偶爾一次的不義行為的潛在利得也會愈大，亦即「魔鬼的誘惑」也將愈大。

### 三、當信徒的「身份」成為可靠的「傳訊」特徵時在模型中的義涵

在本文第貳節之二中，張榮富及鄭雅安(2007)的推論得出：相信宗教的果報觀者因「獎懲」的上升、心中立即「罪惡感」的高漲，和相信報應的必然發生(報應的機率=1)，故不易受眼前(在心中大幅放大的)立即利益所誘惑。承續這樣的論點，加入上文中傳訊與辨視成本的考量，本文假設：一旦成為果報觀的「信徒」則就是傳遞出了道德行為較優的訊息。亦即果報觀信仰的信徒「身份」本身可成為內在德行較優的「可靠的」外在「傳訊」特徵。

根據傳訊理論中的「高做假(偽裝)成本原則」(costly to fake principle)，偽裝成本愈高的事項，愈能用作為傳遞訊息的可靠中介物(Frank, 1988, Ch.5)。如果無內在的果報觀信仰(或道德情緒)，人們將不易克制受眼前因配合法則作用而在心中大幅放大的立即利益所誘惑，因將不易成為虔誠「信徒」或建立起道德的「聲譽」。

當信徒的「身份」本身成了內在德行較優的可靠外在「傳訊」特徵時，此即是法蘭克理論中的「名譽現象」<sup>7</sup>。因此人們在尋找合作伙伴時，即可以根據信徒的「身份」做為簡單的辨視他人是否有內在

<sup>7</sup> 參見法蘭克(Frank Robert, 1988) *Passions within Reason*, Ch. 4。

德行與是否可能採取背叛行為的依據。上述假設如果成立，則對前述演化賽局的影響將是模型中的辨視成本的下降。

當一個社會有些人信仰果報觀時，辨視好人與壞人的成本將會比本文原先模型(圖 1)中的假設低。「辨識成本」的降低，將使圖 1 中的長期均衡點右移，亦即長期而言好人的比例會「自然的」增加。例如，當辨視成本由原例中的 1 分降為 0.8 分時，均衡點會由原來的 75% 右移至 80%。亦即人口比例中有 80% 會是好人，壞人縮減為 20% (如圖 2 所示)。此時社會的總報酬(總產出)將比長期均衡點在 75% 時大。以一個社會有 100 人為例，在長期均衡點為 75% 時，總報酬為  $(75 \times 3) + (25 \times 5) = 350$  分；在長期均衡點為 80% 時，總報酬為  $(80 \times 3.2) + (20 \times 5.2) = 360$  分。

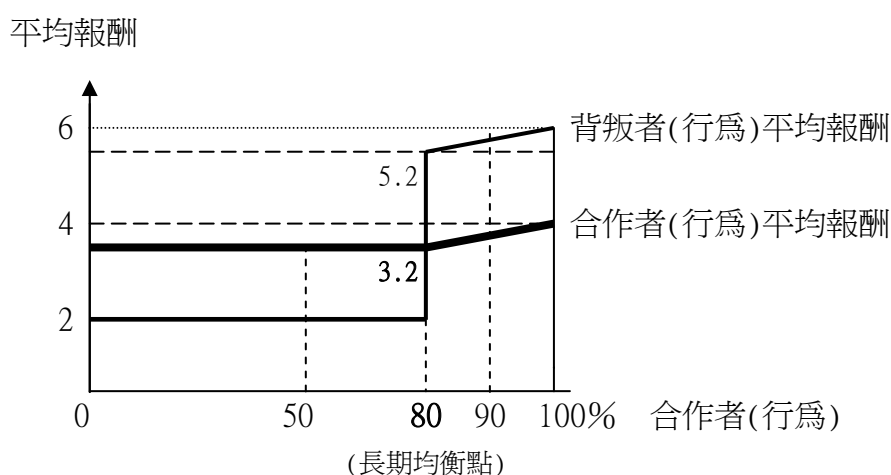


圖 2

## 肆、結論

本文使用經濟學者法蘭克(Robert H. Frank) 在 1988 年所建立的道德情緒演化賽局模型，接續張榮富(2007)與張榮富及鄭雅安(2007)二文的分析，討論宗教的果報觀在促進道德行為與社會利益上所扮演的角色與價值。本文經由賽局模型的模擬推導得出幾個主要的結論如下：

## 一、果報觀能增加短期與長期的社會整體利益

- (一) 果報觀在短期上，「借由獎懲、罪惡感和報應的機率三方面的影響，使信仰者的不道德行為成本升高，因此增強信徒的內在德性」，因此能在短期上使得模型中合作者的人數的比例增加，而社會總利益因此增加。亦即在長期均衡點不變的情況下，藉由短期上改變合作者與背叛者的人數比例，果報觀的信仰能使社會總利益增加。
- (二) 在長期上，根據傳訊理論的高偽裝成本原則，假設人們普遍認為相信果報觀者較易抗拒被配合法則放大的眼前不當利得的誘惑。信徒的「身份」即可成為內在德行較優的可靠的外在「傳訊」特徵，亦即法蘭克理論中的「名譽現象」。由於好人可以由果報觀信仰去辨識別人的好壞，這將使人們在尋找合作伙伴時的辨視成本下降。由於辨視成本的下降，在賽局模型中的長期均衡點將會右移，而社會的總利益亦會因此增加。

## 二、低宗教信仰(或低果報觀)的族群或社會較易在演化競爭中被淘汰

前述(1)(2)也可以解釋，為什麼工業文明之前的各個文明社會都是自古以來較重視宗教(或果報觀)的社會？在族群與族群、社會與社會的競爭中，當其他條件一樣時，高宗教信仰(或高果報觀)的族群或社會的總產出將高於低宗教信仰(或低果報觀)的族群或社會。低宗教信仰(或低果報觀)的族群或社會較易在演化競爭的洪流中被淘汰。

## 三、果報觀信仰(或道德教育)帶來的不是「社會正義」而是「社會利益」

如果由個別行為者的利益的角度來看，在比較沒有道德(或果報觀信仰)的社會中(在長期均衡點左方的社會)，個別合作者因背叛者很多，所以會處處提防以免受騙上當，較會支付辨視成本，因此個別合作者的平均報酬會大於個別背叛者的平均報酬。反之，在一個壞人很少的社會(在長期均衡點右方的社會)，合作者需支付辨視成本，



故個別背叛者的平均報酬反而高於個別合作者。當其他條件不變時，果報觀信仰(或道德教育)帶來的似乎是「社會利益」而非「社會正義」(好人有好報，壞人有壞報)。一個公平正義的社會的產生不能只有宗教與道德，可能是要依靠其他因素，例如：法治、文化、或是政府對社會與經濟制度的干預。

#### 四、信仰深者「魔鬼的誘惑」也將更大

- (一) 當把模型中的兩類人口的假設重新定義為「一個人心中的兩種可能的行爲策略(背叛與合作)」時，演化的長期均衡點代表個人採行背叛策略與合作策略的平均值。好人(大於平均值的人)單次採取背叛策略時的報酬會大於單次採取合作策略，而壞人(小於平均值的人)單次採取背叛策略時的報酬會小於單次採取合作策略。亦即一個好人偶爾一次的背叛行爲(欺騙)的報酬會比一個壞人一次的背叛行爲(欺騙)的報酬高很多。
- (二) 如果把信仰深者(遠高於平均值的人)定義為更有道德的人，把「魔鬼的誘惑」定義為背叛行爲(不道德行爲)的「不義利得」。則由模型中可看出，信仰深者偶爾一次的背叛行爲(例如欺騙)的不義利得會比常人與壞人高很多。隨著信仰的加深，合作行爲的比例將會更大於平均值，個人將愈有令人信服的道德聲譽，然而偶爾一次的不義行爲的潛在利得也將隨之擴大，亦即此時「魔鬼的誘惑」也將愈大。

## 參考文獻

- 張榮富，2007，〈從道德演化的「賽局模型」與「名譽模型」談道德教育—Frank 道德情緒理論之應用〉。國立台北教育大學教育實踐與研究，20(1)，pp.119-146。
- 麥克·阿蓋爾/陳彪譯，2005，《宗教心理學導論》。中國人民大學出版社。
- 戴康生，2006，《宗教社會學》。台北縣：世界宗教博物館基金會。
- Alexander, R., 1987. *The biology of moral systems*. New York: Aldine.
- Axelrod, R., 1984, *The evolution of cooperation*. New York: Basic Books.
- Boorman, S., & Levitt, P., 1980, *The genetics of altruism*. New York: Academic Press.
- Darwin, C., 1859, *The origin of species*. Cambridge, Mass: Harvard University Press, 1966.
- Dawkins, R., 1976, *The selfish gene*. New York: Oxford University Press.
- Evans, D.、張勤譯，2005，《情感來自演化》。台北：左岸文化。(Evans, D., 2001, *Emotion: The science of sentiment*.)
- Frank, R., 1988, *Passions within reason*. New York: W.W. Norton.
- Hamilton, W., 1964, "The genetical theory of social behavior," *Journal of Theoretical Biology*, 7, 1-32.
- Hamilton, W. D., 1971, "Selection of selfish and altruistic behavior," In J. Eisenberg & W. Dutton (Eds.), *Man and beast: Comparative social behavior*. (pp. 57-91). Washington, D. C.: Smithsonian Institution Press.
- Herrnstein R., 1970, On the law of effect. *Journal of the Experimental Analysis of Behaviour*, 13, 242-66.
- Herrnstein R., 1997, *The matching law*. New York: Russel Sage Foundation.
- Trivers, R., 1971, "The evolution of reciprocal altruism," *Quarterly Review of Biology*, 46, 35-57.
- Trivers, R., 1985, *Social evolution*. Menlo Park, Calif.: Benjamin/Cummings.
- Wilson, E., 1975, *Sociobiology: The new synthesis*. Cambridge, Mass.: Belknap Press of Harvard University.

# **How religious retributive justice improve individuals' moral behavior and social benefit: analyzed by a evolutionary game theory**

**Jung-fu Chang**

Assistant Professor, Department of Social Studies Education, Taipei  
Municipal University of Education

## **Abstract**

The aim of this study is to extend Robert H. Frank's moral sentiments evolutionary game theory (1988) to discuss how religious retributive justice (Buddhism's "karma"), which means "good is rewarded with good and evil meets evil", improve individuals' moral behavior and social benefit. Firstly, Frank's "reputation model" and Herrnstein's "matching law" are introduced for discussing that religious "karma" affects individuals' moral behavior by three approaches: "rewards and punishments", "the moral sentiment of guilt" and "the possibility of punishment". Secondly, it is assumed that retributive justice is able to strengthen the "reputation phenomena". This means that the "identity" of believer himself can be a trustable signaling characteristic of inside morality. By this assumption, Frank's moral sentiments evolutionary game is re-simulated. The main findings of simulation are: (1) "karma" not only can increase short-run but also long-run social benefit. (2)

“karma” (or retributive justice) does not bring social justice, although it brings social benefit. Two questions are discussed in the end. First, why, before the industry revolution, are most rich countries also those taken retributive justice belief? Second, does the deeper the belief of “karma” , the higher the temptation from evil (the temptation from injustice benefit)?

**Keywords: virtue, moral sentiments, reputation, retributive justice, evolutionary game**