

# 良知的辯證

## ——康德、費希特、牟宗三的理論比較研究

彭 文 本\*

### 摘 要

這是一個東西方道德哲學的比較研究，核心問題是良知。在這篇文章中，筆者想要論證，如果不適當地限制良知的有效性範圍，則它會陷入辯證的過程而不能維持其絕對正確的地位。良知包含一種主體自我指涉（Subjektbezug）的意義，也包含一種對象的指涉（Objektbezug）意義。前一個意義是指一種主體內部的實踐理性之自我檢查，不涉及任何具體的義務內容；後者指的是一種涉及對象的具體義務內容。康德主張良知只有第一種意義，因為第二種意義不是良知的事，而是道德判斷力的事。康德認為唯有如此，才能維持良知絕對正確的地位。費希特和牟宗三則認為良知不能只有第一種空洞的形式意義，需要第二種實質的意義，而且還必須保證其絕對正確性。筆者這篇文章想要主張，康德對良知的解釋是對的，費希特和牟宗三的解釋是錯的，理由是第二義的良知無法避免陷入辯證。

關鍵詞：良知 費希特 康德 牟宗三

---

97.02.19 收稿，97.07.31 通過刊登。

\* 臺灣大學哲學系助理教授

## 前言

本文章的辯證（*dialectics*）一詞是採用康德《純粹理性批判》一書中較為貶抑的意義，這個詞在康德那裡只是一種理性邏輯的歸謬法之運用，凡是會推演出對立或矛盾的知識便是一種辯證的知識，而它是一種錯誤的知識。在理論哲學裡，康德認為這種錯誤知識之所以產生，是因為理性想要追求一種無條件的總體性知識，將原先只是關於現象知識的範疇超出經驗範圍所造成的。康德用他的三分的知識能力——感性、知性、理性——來建構整個批判哲學的知識架構。一方面知性必須面對感性所提供的直觀材料以便於形成綜合的知識；另一方面，如果知性脫離感性單獨操作，以自己的知性內容作為知識對象，這樣的知性就是理性。簡單地說，知性面對的認識對象是感性材料，理性面對的對象是知性的概念。對康德來說，前一種操作得到的才是真實的知識，後一種得到的便是會陷入辯證的假知識。相對地在實踐哲學中，康德使用類似的架構，實踐理性的定言令式（*categorical imperative*）的地位類似知性中的概念（範疇），格律（*maxim*）類似感性材料，實踐理性的自我檢查成為良知。<sup>1</sup>簡單地說，實踐理性面對的對象是格律，良知所面對的對象就是實踐理性。令人感到訝異的是，對康德來說，純粹理論理性單獨面對知性，或是理性的自我檢查會陷入辯證，而純粹實踐理性單獨自我檢查，卻不會陷入辯證。恰恰相反的是，良知作為實踐理性的自我檢查，如果進入經驗領域，實際決定哪一種格律才是義務，反而會陷入辯證。換句話說，康德認為理論理性的超越經驗的自我檢查會陷入辯證，而它與經驗結合產生的才是真知識；而實踐理性的超越經驗的自我檢查是確定的知識，而它與經驗結合產生的知識反而可能會出現錯誤。為了避免良知陷入辯證，康德不得不將良知的意義限定在純粹實踐理性的自我檢查之上。康德這個良知理論歷史上最著名的批判是黑格爾《法哲學大綱》第一三七節的自我等同的質疑（*Einwand der Tautologie*）。黑格爾認為，如果良知完全與對象割離，剩下來的不過就是主體自我等同，其中只有主體的空洞的自我斷言（*Selfbehauptung des Subjekts selbst*），這樣的良知不能指出任何具體的義務內容。而如果想要在這個空洞的自我等同加進任何義務內容，可以任意地把完全相反的內容加進良知，那麼良知便集善與惡的內容於一身，因此必然陷入辯證。

同樣的良知問題，在費希特的研究比較流行的命題是：費希特的良知不只

---

<sup>1</sup> 康德《實踐理性批判》的辯證論並不是討論良知問題，而是最高善問題。康德認為這是實踐理性追求無條件總體性的對象，但筆者認為：此處如果要以人的知識與實踐能力的架構來看，良知的能力是對比於純粹理性的。

包含康德的主體自我指涉 (Subjektbezug) 的意義，也包含對於對象的指涉 (Objektbezug)，良知的第二個意義可以代替康德回答黑格爾的質疑。<sup>2</sup>筆者認為，黑格爾對良知的批判與康德的良知理論並不衝突，而費希特所提出的對象指涉與其說是回應了黑格爾的批判，不如說反而陷入黑格爾的批判之中。

根據筆者的研究，康德的良知理論主要可以分為三個時期，第一個時期是在前批判時期，主要是他在課堂上對於包姆加登 (Baumgarten) 倫理學的批判，尤其後者的《倫理哲學》(Ethica Philosophia) 及《實踐哲學的最初根據》(Initia philosophia practica) 中。<sup>3</sup>第二個時期是出現在晚年的《道德形上學》與《純粹理性限度內的宗教》兩本著作之中。<sup>4</sup>第三個時期則出現在康德晚年的遺稿中。<sup>5</sup>康德這三個時期的良知理論應該有一個內在發展的關係，而費希特對於良知的批判的根據卻是以第二期的理論為批判重點，第一期和第三期當然因為時代因素無法為費希特所理解。筆者認為費希特對康德的批判，在康德良知理論內在發展理路來看，都已經被康德設想過。另外當代新儒家牟宗三對於康德的良知理論也採取類似費希特的批判路線。筆者想要在本篇文章中論證，費希特和牟宗三對康德的這些批判是無效的，理由是到目前為止，良知問題所能論述的有效性範圍在康德的第二時期已經窮盡了，雖然康德本人在晚年的遺稿中想要從自己第二期所設定的有效性範圍進一步突破，但是筆者認為並不成功。筆者將文章分成以下三部分：

- (1) 康德良知理論的發展：前批判期——一七九〇年的法庭模式——遺稿的最後立場
- (2) 費希特的良知理論及其辯證
- (3) 牟宗三的良知理論及其辯證

2 主要代表人物 Franz Ungler, "Zu Fichtes Theorie des Gewissens," *Wiener Jahrbuch der Philosophie* 12 (1979), 212-235。

3 康德早期倫理學講義，見：AA27.1, 351-357; AA27.2.1, 613-622; AA27.2.2, 1488-1492，以下康德文本大部分根據的是普魯士學院版 (Akademische Ausgabe)。引例：AA6 = 該版本的冊 6。

4 AA6。關於康德良知理論之二手研究，請參考：H.J. Paton, "Conscience and Kant," *Kant-Studien*, (1979), 238-251; Fumiyasu Ishihawa, "Das Gerichtshof-Modell des Gewissens," Nobert Hinske(ed.), *Kant und die Aufklärung*, 43-55, Hamburg (1992); Hoffman, T. S., "Gewissen als praktische Apperzeption. Zur Lehre vom Gewissen in Kants Ethik-Vorlesung," *Kant-Studien* (2002), 424-443。

5 AA21，筆者下文所引的文字根據該冊的編者 Gerhard Lehmann 在導論中所提，該冊是根據 1796-1803 年間的手稿 (AA21, V)，應屬康德最晚年的思想無疑。

## 一、康德良知理論的發展：

### 前批判期——一七九〇年的法庭模式——遺稿的最後立場

#### (一) 前批判期

這個時期的良知，康德使用拉丁文的 *conscientia*，這個字就其字源上本來並不一定與道德意識扯上關係。照字面上的意思，它應該被理解為「共知」(Mitwissen, con-scire)。它最原始的意思是，對一件事態 (Sachverhalt) 的認知，並且連帶對這個事態的一種第二序在反思意義上的「意識到」。這個本來是人對於事物認識上的「共知」(*conscientia*) 的現象，在啓蒙運動之後開始出現意義上的分歧。笛卡兒的意識問題只是「共知」這個概念分歧到知識論上的一個可能性變化。啓蒙運動的另一支沃爾夫 (Christian Wolff) 與包姆嘉登 (Baumgarten) 把「共知」的現象轉成道德意義上的「共知」，並用 *Gewissen* 來翻譯 *conscientia*，這時候的 *Gewissen* 比「共知」要多一個成分，*Gewissen* 不僅是「共知」，而且是知某事態為確定 (*gewiss*) 的「共知」。「良知」(*Gewissen*) 此時不僅由對於事態的認知，而且還有一個反思的「共知」，並且含有對此事態確定性的判斷等多層意義。

沃爾夫理解的良知所指涉到的事態是個別的行爲，他是這樣定義良知的：「對吾人行爲是否為對或錯的判斷被稱爲良知」(“Das Urteil von Handlungen, ob sie gut oder böse sind, wird Gewissen genennet”)。<sup>6</sup>對沃爾夫來說，良知作為道德感情是對於個別行爲直接性的判斷，而作為理性原理的道德原理並不針對個別行爲的是非判斷，而是針對前一個良知判斷的後設反思判斷，它決定前一個良知判斷的對錯：

當這個判斷是真的，那麼它便是一個正確的良知；反之，如果它是假的，便是一個錯誤的良知。例如，如果有人如此判斷：人們應該愛他的敵人，那麼他便有一個正確的良知，因為我們將在以後證明，這個判斷為真，或者它符合自然的法則。反之，如果他如此判斷：人們應該恨他的敵人，那麼他便有一個錯誤的良知，因為我們將在以後證明，這個判斷違反自然以及理性的法則。<sup>7</sup>

但是這個定義遭到康德反對，他認為：「一件行爲究竟是對或錯，是由知性來

6 Christian Wolff, *Gesammelte Werke* (1. Abt. Deutsche Schriften Bd.4), 46.

7 同上書，頁 46。

判斷，而不是良知」(Ob eine Handlung recht oder unrecht sei, darüber urteilt der Verstand, nicht das Gewissen)。<sup>8</sup>沃爾夫的弟子包姆嘉登的良知所指涉到的事態是個別行為與普遍法則之間的歸類關係，換言之，包姆嘉登將良知理解為一種判斷力。在其倫理學著作(包括《倫理哲學》(*Ethica Philosophia*)及《實踐哲學的最初根據》(*Initia philosophia practica*))中區分良知為兩種，如果歸類正確，這樣的良知叫做正確的良知(*conscientia recta*)，如果歸類錯誤，則稱為是錯誤的良知(*conscientia erronea*)，這兩者的區分可以引用盧梭的另一組語言，自然的良知(*conscientia naturalis*)與人為的良知(*conscientia artificialis*)。對包姆嘉登來說，自然的良知是先天的，永遠不可能出錯。人為良知會有錯誤的良知與正確的良知，這是因為它來自一種推論的良知(*conscientia ratiocinans*)，這種良知有時會推論正確而得到正確的良知，但有時也會推論錯誤得到錯誤的良知。包姆嘉登繼續分析在哪些情況會得到錯誤的良知，他分析出兩種：(1)邏輯上被誤導的良知(*conscientia adultera*)；(2)道德上被誤導的良知(*conscientia prave adultera*)。除此之外，包姆嘉登也從行為與良知發生的前後時間關係討論三種良知的樣式：行為前之良知(*conscientia antecedens*)、行為後的良知(*conscientia consequens*)與行為同時發生的良知(*conscientia concomitans*)。<sup>9</sup>

在一七七〇年之後，康德對於包姆嘉登的良知理論有兩個主要批判點：(1)康德認為良知是一種「本能」—指向自己的本能，<sup>10</sup>不只是一種判斷力。康德認為，如果良知只是像包姆嘉登說的是一種判斷力，那麼它不可能對我們的行為有所強制性，良知便不擁有一種主導主體行為的力量。(2)良知如果是判斷力，它也不單是下對錯的一種判斷(*urteilen*)，而是判決有罪(*verurteilen*)或無罪釋放(*loslassen*)，<sup>11</sup>良知因此轉出一個法庭的意識。這是一個很重要的轉變，良知這時真正面對的是實際的道德行為，而不是抽象地解決一些邏輯上找出來的困難例子，而是一個實際的人面對實際的狀況。而康德這時候也還認為良知可能是錯誤的，「當我們按照或然性來做判斷，良知可能是確定的，也可能是辯證的」(*Unser Gewissen kann sicher oder dialektisch sein, wenn wir nach*

8 《純粹理性限度內的宗教》(*Die Religion innerhalb der Grenzen der bloßen Vernunft*)，以下引用簡稱《理性宗教》，AA6, 186。

9 上述的區分出自 Alexander Gottlieb Baumgarten(ed.), *Ethica Philosophia* (1751), §§175-190, AA27.2.1, 780-784, 915-919。

10 AA27.1, 196。

11 AA27.1, 197。

Wahrscheinlichkeit urteilen)。<sup>12</sup>這時出現了一七九〇年的法庭模式的雛型，但是應用的範圍卻大不相同。《道德形上學》裡的法庭模式剛好與此時期相反，良知面對的不是一個個別實際的狀況，而是一個普遍的知性判斷力的過程。

## (二) 一七九〇的法庭模式

內在法庭的觀念其實也是來自包姆嘉登。<sup>13</sup>康德在這裡逐漸形成自己的立場，一方面他認為包姆嘉登所說的錯誤的良知根本與良知無關，而是判斷力的錯誤，另一方面，他又將法庭模式的理論位置上升一層，它不再面對實際的倫理狀況，而是對於判斷力結果之為對錯的檢查功能。判斷力的確有時會出錯，但是良知沒有出錯的可能，因為對他來說良知如知覺般出現，根本沒有對對象做任何判斷，無對錯可言：

因為對於客觀的判斷是否某些事為義務時，人們也許偶而會犯錯；但是對於主觀的判斷我是否為了那個判斷的緣故，將它與我的實踐理性（此處正確判斷的理性，richtende）做比較，我不可能犯錯，因為那樣似乎我實際上根本沒有下判斷；這種情況下既不會有錯誤，也不會有真理發生。<sup>14</sup>

康德這裡使用的論證方式和《第一批判》中的同出一轍，康德在其「先驗假相」一章中提及感官知識並無真假可言，而只有在判斷中才會出現真假，他說：「因為真理或假相並不是對象在被直覺的情況下，出現在對象之中，而是在它被思考的情況下，出現在對於它的判斷之中。因此正確地說是，感官是不會有錯的——不是因為它們總是做了正確的判斷，而是因為它們根本沒有做任何判斷」。<sup>15</sup>舉例來說：我在天空中見到一些像幽浮的對象，我的感官只是收到一些中性的知覺材料，我根據這些材料判斷我所見到的是幽浮。但是康德認為，材料本身並無真假可言，真假的可能性出現於感官材料與幽浮的觀念之間的判斷作用。同樣的，良知只是主體內部的一種直接性的知覺，根本並未對對象下任何判斷，這時候既不能說它是真，也不能說它是假。

在晚年的《純粹理性限度內的宗教》及《道德形上學》裡，康德才確立他

<sup>12</sup> AA 27.1, 199。

<sup>13</sup> 他使用的是拉丁文forum internum，AA19, 81。康德使用的是德文Gerichtshof，見：*Metaphysik der Sitten*, AA6, 438。

<sup>14</sup> 《道德形上學》，AA6, 401。

<sup>15</sup> K.d.r.V, A293 (K.d.r.V.=*Kritik der reinen Vernunft*，康德，《純粹理性批判》A版，頁293)。

這一套法庭模式理論，把良知和判斷力區分開來，判斷力是對於個別行為（客體的）的檢查，而良知則是對於上一個檢查功能（主體自身的）第二序的檢查。在《道德形上學》一書裡，康德是這樣描述良知的法庭模式：

每一個義務概念包含透過法則（作為道德的，限定吾人自由的命令）而來的客觀的強制，而它是屬於提供規則之實踐的**知性**；但是一件行為作為內在地歸類於法則之下的事件是屬於**判斷力**（*iudicium*），它作為行為歸類的主觀原理，有效地判斷（在法則之下）行為是發生或不發生；接下來便是理性的結論（判決），亦即行為與法律上的效力的連結（宣判有罪或無罪）：以上一切發生於法庭宣判（*Gericht*）之前，後者作為產生法則效用的道德人格，可稱之為法庭（*Gerichtshof; forum*）。在人心中內在地法庭的意識（在它之前他的思考互相控告或饒恕對方）叫作**良知**。<sup>16</sup>

在這個法庭內一共出現三組不同的能力，知性是判斷一個行為格律（*Maxime*）是否符合普遍法則以便分辨對錯的能力，判斷力是將一件特殊行為歸類到普遍法則或者行為格律的能力，良知則是檢查前兩組能力執行過程是否正確的能力。這三者在庫德的倫理學裡牽涉到完全不同的對象。Allen Wood在二〇〇七年出版的《康德的倫理學》（*Kantian Ethics*）一書中對這個法庭內三種不同的角色（控方、被告、法官）的詮釋裡，<sup>17</sup>另外引用的是康德《倫理學講義》的以下文字與上述引文對照：

良知的內在地法庭或許可以適切地與一個外在的法庭相比較。因此我們在自己心中發現一個原告，它如果沒有法則的話，便不能存在。而這法則並不屬於市民社會的實證性法則，而只存在於理性之中……除此之外，同時在人類心中也存在著一個辯護者，亦即**自愛**（*Eigenliebe*），它為自己脫罪，而且對於控訴提出許多反駁，另一方面原告則又提出這些反駁的駁斥。最後，我們在心中發現一個法官，他判決無罪釋放或者是有罪譴責。<sup>18</sup>

這兩段文字之間對於法庭模式最重要的詮釋上差異在於，依據《道德形上學》的文字，被告應該是指「知性與判斷力」兩者，但是依據《倫理學講義》的文字，被告似乎應指「自愛」。這裡的詮釋爭議出自「自愛」一詞的歧義性。康

16 AA6, 437-38.

17 Allen Wood, *Kantian Ethics*, (Cambridge: Cambridge University Press, 2007), 184.

18 *Vorlesung über Ethik*, AA27, 354，粗體字為筆者之強調。

德在《純粹理性限度內的宗教》一書將「自愛」理解為人的三種天生傾向（*pre disposition*）之一。這三種天生傾向分別是：(1)做為自然生命的存有者（*living being*）傾向，(2)作為自然生命（*living being*）同時又是理性存有者（*rational being*）的傾向，(3)作為既是理性存有者又是能夠負責的存有者（*responsible being*）的人格之傾向，而康德認為「自愛」屬於人類這三種天生傾向的第二種。<sup>19</sup>「自愛」對於康德來說是一種帶有理性思維的情感，他認為「自愛」又可以依據其對象之差異分成對於「善意的愛」（*love of good will, Liebe des Wohlwollens*）和「滿意的愛」（*love of pleasure, Liebe des Wohlgefallens*）。<sup>20</sup>前一種是理性的，原因在於：「一方面鑒於目的而要僅僅選擇那能夠與最大且最持久的福利共存的東西，另一方面又要為幸福的每一種要素選擇最適宜的手段」。<sup>21</sup>理性這時候是為人的自然之性好（*natural inclination, natürliche Neigung*）服務，康德認為以這個意義的自愛當作無條件的道德原理，將是產生惡的泉源。「自愛」的第二種意義是以追求一種快適感受為目的的行為，在此又可區分出兩種不一樣的意義，第一種意義涉及的是因為採取正確手段以至於達成目的所得到的滿足感，例如：「一位在投機買賣中大獲成功的商人，鑒於所採納的準則，而對自己的遠見卓識感到高興一樣」，<sup>22</sup>在這裡道德惡的起源與前一種「善意的愛」是相同的。第二種意義是將這樣的快適感受視為無條件（不受行為結果的所影響）的原理，此種情況就與前者有所不同，這時候「自愛」不能與道德法則有所衝突，因為康德認為不可能有一個關心道德的人，在行為格律（準則）與道德法則衝突的情況下，仍然產生一種無條件的自滿自足的感覺。筆者看來，這兩種「自愛」的意義其實來自上述人的第二種天生的傾向之成分——對自然生命存有者與理性存有者之混和。經由理性估量所採取的行為而使得「自然生命」或者「理性自身」得到的滿足的感受。

從上述康德對「自愛」一詞的分析出發，可以看出《倫理學講義》中談論「自愛」作為原告所可能產生的詮釋上的問題。筆者認為，這裡的「自愛」可能有兩種詮釋：或者(1)自愛是一種經由理性估量的行為所產生的自我滿足的感受，其滿足對象可能是「自然生命」或者「理性自身」，在這樣的情況下，自愛是一種合理性的愛（*rational love*），或者(2)將「自愛」採納為一種行為格律

19 參考《純粹理性限度內的宗教》，AA6, 26-27。

20 同上書，AA6, 45-46，翻譯參考：康德著，李秋零譯，《單純理性限度內的宗教》（臺北：商周出版社，2005），頁103。

21 同上書。

22 同上書。



的內容 (content of maxim)，這時候「自愛」是一種普遍的思維，不是一種情感。筆者認為Wood在這裡不可能採取第一種意義的「自愛」，理由是這樣的自愛只是理性作用的結果，它本身不可能單獨地決定一件道德行為，因此，它不是人格，也不能為行為負責，所以也不可能列為被告。基於這個理由，筆者認為第一種情感意義的「自愛」與《道德形上學》的文字是有衝突的。筆者認為Wood詮釋的「自愛」應該採取第二種作為思維內容的意義的「自愛」，這樣的意義固然與《道德形上學》的文字沒有衝突，但是採取這樣意義的「自愛」作為「被告」不足以涵蓋「被告」的其他可能身分，例如：在一些特殊的倫理兩難狀況，人們只能在不同衝突的義務選擇其中一種義務，而必須忽略其他義務時，即使所選的義務完全與「自愛」無關，這時他仍不能免於被良知法庭列為被告。例如：如果舜的父親犯了殺人罪，執法的人到舜的家裡抓他父親，若舜讓執法的人將父親抓走，那是因為他服從國家法律，而不是因為「自愛」；或者舜不讓他們抓走，那是因為孝順是他對於父親的義務，也不是因為「自愛」。因此，按照Wood的詮釋，舜這些不在「自愛」的格律下所產生的行為，應該都不會被良知法庭列為被告，但是事實上不然，這種義務衝突狀況正是良知法庭的判決介入的重要時刻。因此，「自愛」作為行為者自我辯護的能力，固然屬於「被告」的身分之一，但是它不是唯一可能的被告，我們不能忽略《道德形上學》裡所提到的「知性」和「判斷力」的能力。我們似乎可以將這三者放在同一個「道德人格」之下，將後者視為結合前三者的一個「被告」<sup>23</sup>。

以下筆者順著《道德形上學》一書的架構，說明知性、判斷力與良知在內在法庭中的功能之差異。

## 法庭模式裡的三種不同的檢查

### 1. 知性的檢查

康德認為：「一件行為究竟是對或錯，是由知性來判斷，而不是良知」，<sup>24</sup>要了解這句話的意思，我們可以問下面的問題：知性如何判斷一件行為是正確或錯誤的？康德認為，知性直接檢查的對象不是一個個別的行為，而是包含許

<sup>23</sup> 這裡筆者接受匿名審查人的建議，依據 Wood 書中提到的一個觀點，即「原告」、「被告」、「法官」三者：「雖然他們是同一個自然人格 (natural person)，但是說他們是不同的道德人格 (moral person) 也不為過」 (*Kantian Ethics*, 184)，「自愛」、「知性」、「判斷力」共同屬於 Wood 所說那個作為「被告」的道德人格。

<sup>24</sup> 見注8。

多個別行為或者行為規則 (Regel) 的格律 (Maxime) :<sup>25</sup> 即我檢查我所採用的 (主觀) 格律是否總是能夠成為 (客觀) 的普遍法則, 這就是康德有名的定言令式 (kategorischer Imperativ)。根據他的解釋, 定言令式作為道德的最高原理便是知性用來檢查行為對錯的標準, 「這兒很容易說明, 好比它〔指一般人之理性〕手上擁有一個羅盤, 在眾多出現的事件中非常清楚的分辨, 何者是對, 何者是錯, 是符合或違反義務, ……」,<sup>26</sup> 這是定言令式所包含的原理之一, 即分辨原理 (principium diiudicationis)。<sup>27</sup> 這個分辨對錯的檢查又分成兩個步驟: 形式的 (或消極的) 和實質的 (或積極的) 檢查, 這兩組檢查同樣的重要, 黑格爾對康德的批判便是建立在對康德的片面的理解上, 即把定言令式僅僅用第一個形式意義來理解<sup>28</sup>。我們只要看《道德形上學的基礎》裡的定言令式的第二式的解釋或者《道德形上學》第二部分的導論VI的文本, 便可以斷定黑格爾並沒有完全理解康德:

義務概念與法則直接的發生關連 (即使我抽離所有作為它的實質內容之目的概念), 這就像在定言令式裡義務的形式原理: 「**依此而行, 亦即使你的行為之格律能夠變成普遍法則**」, 這是很明顯的; 不過這在倫理學中的法則是被當作你自己意志的法則, 而不可作為他人意志的一般意志的法則, 後者提供的將是法律裡的義務, 而這並不屬於倫理學的範圍。——格律在此被視為如此主觀的原理, 它自己僅能藉檢驗 (qualifizieren) 成為普遍的立法; 這也只能是一個消極的原理 (一個無法被駁回的一般性法則)。——但是如何可能有一個為行為格律 (而有) 的法則呢? 唯獨目的概念而同時是屬於倫理學的義務概念才構成法則對於行為的格律的根據, 這是以主觀 (人人都有) 的目的置於客觀 (人人應該做) 的目的之下 (的方式)。

25 這部分請參考《實踐理性批判》第一節關於格律的定義, AA5, 19 或者《道德形上學》第二部分〈倫理學的形上學之最初根據〉的導論VI, AA6, 389; 或見《道德形上學的基礎》, AA4, 421, 注釋。

26 《道德形上學的基礎》(Grundlegung zur Metaphysik der Sitten), AA4, 404。

27 定言令式包含的另一原理為執行原理 (principium executionis), 關於兩原理的關係, 請參考: G. Patzig: “‘principium diiudicationis’ und ‘principium executionis’: Übertranszendentalpragmatische Begründungssätze für Verhaltensformen,” *Handlungstheorie und Transzendentalphilosophie*, Hrsg. v. G. Prauss, Frankfurt, 1986, 204-218.

28 這個說法, 請參考: Allen W. Wood, *Kant's Ethical Thought*, Cambridge University Press, 1999, Preface xiii。

「你應該把這事或那事（例如，他人的幸福）變成你的目的」，這個命令才是針對意念的材質（一個對象）。由於如果行為者不同時設想一個目的（作為意念的材質），也不可能有自由的行為，因此如果存在著同時作為義務的目的，那麼作為目的的手段之行為格律必包含使他自己升格成普遍立法的條件；反之，同時作為義務的目的能夠把這樣的格律變成普遍法則，這其中對格律來說，它只要能與普遍法則互相不衝突就夠了。<sup>29</sup>

知性對於一個格律是否符合定言令式的第一步形式意義的檢查，是把前者置於一個普遍化的測試（Universierungstest）之下，康德這裡的普遍化必須和以下兩種的普遍的意義區分開來：(1)它和所謂金律（golden-rule），即「己所不欲，勿施於人」不同，康德認為，它是可由定言令式導出的一個較次要的格律<sup>30</sup>；(2)它和當代倫理學家黑爾（R. M. Hare）所講的道德語言具有可普遍化性（Universalizability）亦不同，黑爾所指的可普遍化性是對於倫理狀況而言：「更重要的解答是，所有的普遍論者被要求——作出一個道德判斷就是說，如果另一個人處於相類似的情況，就必須對他的狀況作出相同的判斷」。<sup>31</sup>事實上筆者認為，定言令式和上述二者最大的不同是它背後方法論的預設，它用的是歸謬法，這方法我們可以簡述如下：為了證明A，我們把它的反面-A當作前提，當-A被普遍化時，如果發生矛盾，那麼表示-A是錯的，也就是說-A的相反A是對的。用康德自己的例子，為了證明「人不能自殺」（相當於上述等待被證明的A），康德是由這個格律的反面-A出發，即「如果生命繼續延續下去時所將面臨的不幸，多過它帶來的愉快，則我出於我愛，『縮短我的生命』當作我的法則」。<sup>32</sup>我們發現如果-A被普遍化，必然導致整個自然系統的毀滅，也就是說-A自我矛盾，因此依照歸謬法推論，A必然是正確的。康德這個方法可以

29 《道德形上學》，AA6, 388-89。粗體字出自筆者。

30 見《道德形上學的基礎》第二節的注釋，AA4, 430。康德認為：「己所不欲，勿施於人」不能用來檢查人是否可以自殺，也不適用於法官判決當成他的格律

31 R. M. Hare, *Freedom and Reason*, Oxford University Press, New York, 1963, 48-49. 該書裡黑爾亦嘗試把他自己的可普遍化性原理與金律及康德的定言令式區分開來，見該書34-35&139。關於這樣倫理狀況的可普遍性的討論，也見於費希特的《倫理學系統》，見：GA, I/5, 212; SW, IV, 234（GA=Fichtes Gesamtausgabe der Bayerischen Akademie der Wissenschaft, Hrsg.von Reinhard Lauth. Stuttgart-Bad Cannstadt 1965; SW=Fichtes Sämtliche Werke, Hrsg. von Immanuel Hermann Fichte, Berlin 1965）。

32 AA4, 422。譯文出自：康德著，李明輝譯，《道德底形上學之基礎》（臺北：聯經，1990），頁44。

在一個理性存有者自身之內單獨地自我檢查，而不必涉及其他理性存有者（如在金律裡）或者實際的倫理狀況（如在黑爾的理論裡）。

上述只是康德倫理學裡知性如何檢查格律對錯的第一個（形式）的檢查，依康德檢查的方法，我們可以用以下原則繼續分辨格律的對錯：

(1)未通過普遍化測試的格律(-A)，一定是錯的；而它的反面 A，一定是對的。

(2)通過普遍化測試的格律不一定是對的，它必須再進一步通個第二（實質）的檢查的步驟。

事實上，第二個實質的檢查這個步驟在《道德形上學》裡被視為同等重要，因為根據第一個形式檢查，我們大部分只能得到消極內容的道德法則，例如：「我不能自殺」、「我不能撒謊」、「我不能偷竊」……。要得到實質內容的道德法則：「我應當作這件事或那件事」，則我們必須進行第二個步驟的檢查，這便牽涉到定言令式的第二式：目的原則：「你應當這樣行動，總是把人當作一目的，絕不可只當作一工具來看待」。不過我們發現，在《道德基礎》與《實踐理性批判》卻又各自有不同的著重點：前書中康德主張這兩個形式在意義上是相同的：「因此，〔對每一個有理性者（你自己或他人）如此行動，即他們在你的格律中同時被視為目的自身〕這項原則基本上等同於〔依據這樣一項格律而行，即它本身同時含有對每個有理性者而言的普遍有效性〕這項原則」，<sup>33</sup>但另一方面，在道德行為的判斷上，康德認為最好採行普遍法則的形式，因為它在方法上是較為嚴格的。<sup>34</sup>在《實踐理性批判》一書裡，康德則幾乎不談定言令式的第二式和第三式，而只著重它的普遍形式。但是到了《道德形上學》裡，康德又是兩者並重（參考上面長引文）。

筆者主張在第一組形式檢查裡，當一個格律 A 因為發生矛盾被判定為錯誤時，我們並不能因此推斷 A 的反面-A 是對的，換言之，第一組檢查只能消極地排除錯誤，不能積極地肯定其排除者的反面是對的；反過來說，僅僅由那通過檢查的格律 B，我們只能推論 B 的反面-B 因為其矛盾是錯的，但不能因此肯定 B 是對的。試舉一例：考慮「我是否應該禁欲」這個格律，如果我們把它普遍化，即讓每一個人願意做禁欲的行為，結果可能和自殺一樣，整個人類因此消失於地球，但是這樣的話人們不可能做出禁欲這樣的格律的決定，因此禁欲這個格律自我矛盾。從禁欲普遍化得到的自我矛盾並不能讓我們推論出禁欲的反面（假設是縱欲）是對的，因為依照康德，這違反人類作為理性的存有者的

33 AA4, 437-38, 李明輝譯，《道德底形上學之基礎》，頁64。

34 同上書，AA4, 436-37；中譯，頁62-63。

作為目的自身的天職。因此完全禁止性的慾望固然違反義務，但是滿足此慾望必須限在一個符合目的自身原理的範圍內，而不能單純地否定「禁欲」的格律，直接認定它的反面是對的。

## 2. 判斷力的檢查

在上一階段裡，知性的檢查發生在格律和道德法則（即兩組普遍規則）之間的關係，而這一階段判斷力（康德在此特指實踐的判斷力，*praktische Urteilskraft*）的檢查則發生在個別行為與道德法則之間。「現在（判斷）是否一個在感性裡可能的行為的確是可以放在某個規則之下的一個個例，這是屬於實踐判斷力的工作，在規則裡普遍（抽象地）被說出的東西是經由它被實際地應用到一個行為上」。<sup>35</sup>此處關涉到的實際上是一種歸類的能力，舉一個簡單的例子：在一七九一年的某一天，去科尼斯堡的費希特身無分文，因此決定向康德開口借錢，借錢這件行為依照康德的實踐判斷力可能放在以下兩個不同的行為的格律：(1)「幫助需要你幫助的人，因為你有可能也需要他的幫助」；(2)「我如果明知朋友還不起錢，又把錢借給他，是陷他於道德上的惡；因為這樣會使他為了脫困而給別人假承諾」。如果是依照格律(1)，那麼康德必須借錢給費希特；如果是依照格律(2)，那麼康德不能借錢給他。康德最後並沒有借錢給費希特，也就是他依照格律(2)來做，但是他並沒有因此不幫助費希特，因為他後來幫費希特推薦給書商，讓後者可以出版自己的著作以取得稿費度過難關。

## 3. 良知的檢查

如上所言，「良知的檢查」只是對於前兩種「知性檢查」與「判斷力檢查」的第二序檢查，換言之，良知首先檢查「知性檢查」的兩個步驟是否完全被執行，例如：在第一個步驟的普遍性檢查，康德顯然知道自己如果完全按照上述的第(2)個格律來做，那麼他無法照顧到第(1)個格律的普遍性要求。因此，如果康德只是單純因為格律(2)而拒絕費希特的借錢行為，而忽略格律(1)，那麼「良知」此時便會出現要求具有普遍性的格律(1)應當被重視，理由不在於格律(1)的實際內容，而是在普遍性檢查的標準有沒有被遵守。按照康德的理論，如果在行為決斷的當下，格律(1)的普遍性沒有被確實遵守，事後「良知」會不停地提醒行為者，一直到這個普遍性被某種程度補救為止。

「知性檢查」的第二步驟是實質的目的性檢查，這個檢查應該是前一個檢查被確實執行之後的進一步之檢查。如前所述，普遍性檢查所得到的多半是消極性的格律，禁止人做某些行為；而目的性的檢查所得到的則是積極性的格律，

35 《實踐理性批判》，AA5, 67。

肯定某些特定內容為義務。用康德上文所使用的例子：增進他人幸福為人應該實行之積極義務，例如：醫生應該盡其所能醫治病患，或者護士應該盡其所能照顧病患。當 Sars 流行時，臺北市和平醫院的醫生或護士可不可以因為害怕被傳染而離開自己的工作崗位？離開工作崗位在「目的檢查」上是不符合義務的，因為違反「增進他人幸福」的義務；而這件事也通不過「普遍性的檢查」，因為我們無法立一個法則，允許每一個人都可以擅自離開他的工作崗位。良知對於這樣的行為會在事後進行譴責，譴責的不是「擅離工作崗位」這件事，而是對於違反「實質目的」的檢查這件事。

判斷力的檢查也經常會出錯，例如：我某日在路上遇見有慈善義工募捐，把我僅剩的一百元捐出去，我把這件行為歸類到以下的規則：「在能力範圍內我應該盡量幫助窮人」；這時我也有可能考慮遇到詐騙集團，因此沒有捐出一毛錢，這件行為是根據以下的規則：「我不應該把錢捐給詐騙集團以助長欺騙的歪風」。這時候我也許把真正的義工當成詐騙集團不捐半毛錢，也有可能把詐騙集團當成真正的義工把錢全部捐獻，在這兩種狀況裡，我的實踐判斷力都發生錯誤的歸類。良知一方面要求行為者必須百分之百確定歸類關係的正確，如果無法百分之百確定，良知會警告行為者不要對不確定的事採取行動。

筆者以上簡略的分析無法窮盡康德這一個理論模式的可能性，但是至少說明良知是位於「知性與判斷力」之上，主要檢查的部分包括「普遍性」、「實質目的」、「歸類」所出現的錯誤並加以警告，而事後對於因為此兩者所產生的錯誤提起訴訟。這時「知性」、「判斷力」或甚至「自愛」扮演法庭中「被告」的角色，「良知」扮演「原告」的角色。我們可以從康德他晚期的哲學的發展裡，良知出現底下的三種意義看出良知這樣的角色：

- (1)「良知是這樣的意識，它對於自己而言就是義務自身」（“Das Gewissen ist ein Bewußtsein, das für sich selbst Pflicht”）。<sup>36</sup>
- (2)「人們也可以這樣定義良知：它是自我指向自己的道德判斷力」（“Man könnte das Gewissen auch so definieren: es ist die sich selbst richtende moralische Urteilskraft”）。<sup>37</sup>
- (3) 在人心中的內在法庭的意識（在它之前他的思考互相控告或饒恕對方）叫作良知（Das Bewußtsein eines inneren Gerichtshofes im Menschen）。<sup>38</sup>

36 《純粹理性限度內的宗教》，AA6, 185。

37 《純粹理性限度內的宗教》，AA6, 186。

38 《道德形上學》，AA6, 438。

這三個定義共同的特徵便是如上所述，在其中個別行為作為對象被抽離。在第一個定義裡，我們可以把良知簡單地理解為「義務的意識自身」，它不是對於某種特定的行為是否為義務這樣的意識，而是使某種行為成為吾人義務之條件的那個意識，在這裡的「義務意識」不是指向行為的，而是指向判斷力自己。我們可以與上述第二個定義連上關係，它也可以簡單地表述為「自我指向的判斷力」，這時候義務概念被抽離，只剩判斷力在主體內的自我與自身的關係，或者說是判斷力的自我檢查。那個指向對象的判斷力，康德稱之為「實踐的判斷力」，而那個指向自身的判斷力，稱之為「良知」，這是上述第二個定義的主要含意。按照這個定義，良知是完全出自理性和自身的關係，康德在《純粹理性限度內的宗教》一書中學了一個很好的例子：

舉例來說，有一個審判異教徒的審判官，他固執地墨守自己的規章性信仰的唯一性，甚至到了在必要時以身殉道的地步。他要對一個被指控犯有不信罪的所謂異教徒（在其他狀況下是一個好的公民）作出判決。現在，我要問：如果他判決這個異教徒死刑，那麼，人們是否能說他是按照自己的（儘管是誤入歧途的）良知作出判決的呢？還是不論他是弄錯了還是有意做出了不義的行為，都可以指摘他完全喪失良知？<sup>39</sup>

康德認為：這裡判異教徒死刑的審判官如果只是根據宗教上歷史文獻所記載的教條內容作判決，無論是出自他完全虔誠的信仰或者是有意作出不義的判決，都是違背良知的。他進一步指出，如果現在我們將那些審判官判決所根據的法條的權威去除，讓他們捫心自問：你敢不敢下這樣的賭注，如果你在這個死刑的判決上出錯，你願意被永遠地詛咒（ewig verdammt）。康德說，這時候的宗教審判官也會不寒而慄（zittern）。我們可以說：康德的良知是一種完全內在的東西，支持良知的判決除了理性自身之外，不能有其他任何外在的權威作為其立論之根據。

但是良知如果按照上面的第三個「內在法庭」的定義，由於法庭的基本結構會使良知多出一些內容：原先我們指出，在法庭內的被告和原告是同一個人分扮兩個角色，一個是現實的人格，他接手對於對象的「知性與判斷力」的檢查功能；另一個人格檢查前兩種功能是否確實執行。而法庭內的法官由誰來扮演是相當有問題的地方。康德在正文裡主張：「把法官和由良知起訴的被告表象為同一個人，這樣的想法，對於法庭而言是一個很不一致的表象方式，因為

39 AA6, 186; 譯文參考：李零秋譯，《單純理性限度內的宗教》，頁252。

這樣會使被告每次都敗訴」。<sup>40</sup>因此，康德主張「被告和原告是同一人」，而法官是人類以外的第三者，他接手良知的另一種仲裁功能。關於這個法庭的描述，康德在《道德形上學》的一個注釋裡解釋得較清楚：

這個雙重人格，在其中，良知在自我控訴與指向自身的人類必須這樣的設想：這個雙重的自我一方面必須顛慄於它自己所信任的法庭的限制，另一方面它手上又握有來自自我天賦的權威（給予）的法官的權柄，（此人格）需要一個解說以便理性不致於陷入自我矛盾。——我，作為原告又是被告，依舊是同一個人；但是作為以自由為出發點的道德上立法的主體，在此立法裡人們聽命於自己給出的法則，在實踐的意義下來說，他可被視為不同於天生具有理性的感性存有者（按：即不同於人類的存有者）——因為沒有任何理論（可以說明）由智性世界過渡到感性世界的因果關係——，而這個特別的差異正是在人類中描繪其特徵的學科（高級的和低級的）之差異。第一個是原告，他對立於被准許在法律上有權自我辯護的被告。（被告與辯護）的行動終結後，作為有權柄者的內在法官宣布對於道德行為結果是幸運或不幸的判決；以這個資格，我們無法憑藉自己的理性進一步追究他（作為世界的主宰者）（為何有此）權力，而只能對這個無條件的判定（Jubeo）或否決（Veto）表示崇敬之意。<sup>41</sup>

這個內在法庭的過程是這樣：當行為者從事某個別由知性提供的行為內容，良知這時候開始它作為法官的職責，如果行為主體執行的符合內在的理想人格，那麼法官會做出贊成的判決；反之，如果行為主體做出與理想人格不符的行為，這時理性裡發生自我矛盾的現象，法官會做出否定的判決，以避免矛盾的現象。這裡法官的角色是由上帝來擔任。因此「自我指涉的判斷力」與「內在法庭」兩個定義之間並不完全一致，前者只涉及理性與自身的關係，最後的仲裁者是理性自己，而後者除了包含理性與自身關係之外，還包括另一個最後的仲裁者——上帝。在下文中，筆者嘗試主張，康德的第一個定義要比第二個定義理論上較站得住腳。

康德這個時期的理論是一種主觀的形式主義，良知一方面只是主體面對自身判斷力所做的判決，雖然不會出錯，卻只有主觀的有效性，意即，個人的良知之判斷不能替代，也不能指導他人的良知之判斷；但另一方面，良知不能給

40 AA6, 439.

41 《道德形上學》，AA6，439。



出任何客觀的義務的實質內容，因為這樣的實質內容是知性用普遍性的檢查決定的。

### (三) 康德遺稿中的最後立場

在法庭模式中，康德一方面主張良知純粹是理性內部的自我關係，但另一方面又認為這個法庭內部最後的仲裁者不能是理性自己，而是一個超越的理想人格——一種神的聲音，這代表康德在宗教上的一種妥協，但是當關涉到人的實踐理性的自由時，良知這樣宗教的傾向與人的自律是衝突的。這種衝突一直都構成康德道德哲學上的一種張力，在《第二批判》裡對於上帝存在作為實踐理性設準就是這樣張力的表現之一，定言令式就是這種張力之間的支撐者。在康德未出版的遺稿中，良知忽然取代了之前定言令式的角色，而且出現一個令人意外的轉變，康德將神放到人的實踐理性的內部，而以良知作為兩者之間的中介。康德本人並沒有將這樣的理論像在《第二批判》的辯證論裡一樣完全展開出來，但是在遺稿中，康德明白指出，在該書中的這樣的理論只是主觀地證成，不是客觀的證成。<sup>42</sup>筆者認為，康德晚年的這個轉變具有重大的意義，良知由一個邊緣的角色變成一個道德上重要的角色，康德這樣說：

實踐理性有無條件和有條件的義務之分。前者的創立者是神。——因此，神不是一個處於我之外的實體，而只是一個在我之中的道德關係……良知那折磨人的責難是在實踐理性中神的聲音。

(Unterschied der unbedingten und bedingten Pflicht der praktischen Vernunft. Jener ihr Urheber ist Gott.—Gott ist also keine ausser mir befindliche Substanz sondern bloss ein moralisch Verhaeltnis in Mir....Der marternde Vorwurf des Gewissens ist die Stimme Gottes in der praktischen Vernunft)<sup>43</sup>

從康德晚年遺稿片段的記錄中至少可以看出，康德想要從自律倫理學與一個超越的神之存在的衝突中另外找到一個解決之道，最後他的想法也許可以歸結出兩點：第一，康德已經不贊成傳統基督教對神的理解，第二，將神內在化於純粹實踐理性之中，並將良知當作內在神的聲音，這樣的見解也將良知注入新的理論可能性，以及它有可能具備實質義的良知，而且它可以給出一些無條件的義務。這是一個很大膽的看法。但是能否證成這樣的理論，需要進一步的論述。

42 AA21, 147.

43 AA21, 149.

## 二、費希特對於康德第二期的良知理論之批判 ——良知的平行模式

費希特對於康德的良知理論提出兩個批評，第一個批評是在《倫理學原理》一書的第十五節所提出，認為康德把知性的普遍性判斷與良知的主觀判斷分割開來是不適當的，因為單靠知性的普遍性檢查建立義務內容的客觀性會出現兩個問題：首先，普遍性不是倫理事件本身的規定，最後決定一件行為是否符合普遍法則依舊是主體自己；其次，判斷力的檢查過程是一個循環論證，不能理論上解釋一個行為如何可能絕對正確。

費希特在一封一七九五年給萊因厚德（K.L. Reinhold）的書信裡提到第一個批評：

一個特定的格準=A與一個對於理性存有者=B而言具有普遍性之述詞，是互相衝突的：康德敘說與證明這一點。我回答他：這的確沒錯，但是與我何干，因為是什麼使我能夠把A關連到B上去呢？我要將那個格律獨自保存；而當它變成普遍有效的，那我便知道我當然破壞了遊戲規則；但是為什麼我應該從一個特定的領域把一個格律變成我的（格律）而在下列條件下被認為是普遍的？這點康德沒有回答。<sup>44</sup>

費希特認為康德的那一套知性檢查的功能在實際的道德決斷並不可行，因為這時候我們只是認識到每一個個別的可普遍化的格律，至於在此時此刻如何挑選出剛好適合這個狀況的普遍格律，光是利用知性的普遍檢查是無濟於事的。在同一封信裡，費希特進一步主張：當我們對於格律進行普遍性的檢查時，我們必須注意，普遍性並不是格律自身所擁有的屬性，而是進行檢查的主體賦於格律的規定，因此道德法則的普遍性追根究底，必須回溯到主體自己（cf. GA, I/5, 212）。其次，這樣的檢查會陷入一個循環論證裡。當我們將一個特殊的行為置於一個普遍化測試（Universalisierungstest）之下意思是說把它放在一整個系統裡，以便看它是否在後者的每一個狀況都能通過檢查（bestehen），或者用費希特的表達：看看我當前所信念（Überzeugung）<sup>45</sup>為義務的行為與系統內一

44 GA, III/2, 385.

45 信念（Überzeugung, convince）對於費希特來說，是一種道德決斷時所出現的一種心靈狀態，是對於當下特定行為的一種肯定，它來自一種更高的「一致的感覺」（Gefühl der Übereinstimmung），見下文的敘述。

切可能的信念互相不衝突。換言之，康德是想要把個別行為的合法性在一個普遍的系統裡得到理論的證成（justification），費希特認為這行不通，因為「我的信念的系統自己除了透過我當前的這個信念之外，並沒有別的方式被給予我」。<sup>46</sup>我們訴諸系統來證成的個別行為，結果又返回原先的出發點——對個別行為的信念，因此判斷力的檢查永遠在系統和個別行為的兩端循環而得不到一個絕對確定的點；另一方面，有可能因為當前個別的信念是錯誤的，雖然它也許和整個系統不衝突，但卻可能導致整個系統的錯誤。因此費希特認為光靠知性或判斷力的檢查是不足以找到一個絕對正確的行為。因此費希特主張，康德法庭模式裡的良知不介入判斷力（或者是知性）的決定是錯的，費希特的做法是把良知置於和判斷力同一個理論層次，也介入個別行為對錯的判斷，因此筆者稱這個模式為一種「平行模式」。

費希特對康德良知理論的第二個批評出現在《自己對於基礎哲學的沉思集》及《倫理學系統》。他認為康德忽略自己在《第三批判》所發展的反思判斷理論<sup>47</sup>的成果，康德只將反思判斷用在美學及目的論太狹隘了，它同樣可以用在道德的判斷上。根據康德，反思判斷力是一種從個別對象出發尋找普遍原理的一種判斷力，如果在這個尋找過程中個別對象和普遍原理互相一致，便會出現一種美感，這時候我們可以從美感在反思判斷力所產生的過程中去逆推它與個別對象及普遍原理的關係。費希特認為良知如果也可以解釋成類似於美感在反思判斷力的產物，那麼當良知出現的時候，我們一樣可以用類比的方式去推演與它發生關聯的兩端，這時候良知不只是康德所說的那種內在的聲音，同時它也與行為作為個別對象發生關聯。在此我們發現康德系統內良知作為內在法庭的意識以及在行為前後出現的道德情感被費希特總括成一個良知概念——它是一個直接性的感覺（unmittelbares Gefühl）。費氏主張良知除了像康德說的具有主體指涉（Subjektbezug）之外，還有一個對於對象（行為）的客體指涉

46 GA, I/5, 153; SW, IV, 164.

47 反思判斷是康德《第三批判》的重要理論，判斷力是一種歸類的的能力，歸類是將一個個別的事物放置在普遍原則之下的能力。這裡有兩種方式，如果是普遍的原則先給予，然後判斷某個個別事物是否屬於這個普遍規則管轄，康德稱這種為規定判斷力（bestimmende Urteilskraft）；如果相反，個別的事物先給予，由此尋求普遍的原則以置於其下，這種方式稱之為反思判斷力（reflektierende Urteilskraft），例如：從蘋果打到頭這個個別事例中，牛頓找到萬有引力定律這個普遍規律，這便是反思判斷力的作用。費希特將人的道德判斷力歸屬於「反思判斷力」的一種應用，「良知」是一種道德情感，它與美感一樣涉及「個例」與「普遍原則」兩者。不同的是，良知所涉及的「普遍者」只有一個，即絕對自我的自由活動。

(Objektbezug)，指出某件行為是符合義務的，否則判斷力永遠在一個循環裡移動。費希特由這兩個指涉歸結出兩個對於良知的不同定義：

(1)主體指涉：良知是對於吾人較高的本性及絕對自由的意識 (das Bewußtsein unserer höheren Natur und absoluten Freiheit)。<sup>48</sup>

(2)客體指涉：良知是對於特定義務的直接意識 (das unmittelbare Bewußtsein der bestimmten Pflicht)。<sup>49</sup>

我們可以簡單把上述主體指涉之下的良知稱之為形式意義的，因為它們並未指出任何特定的道德內容，就這個形式意義來說，費希特的良知內容是與康德一致的，只需要去除法庭內的上帝角色以回到自我內在性，這個意義就與康德的良知內容完全重疊；而客體指涉的則是實質意義的，它指出某種行為在某個倫理情況是一種義務內容，這部分是康德所反對的，因為他認為實質內容的確定不是良知的事，而是道德判斷力的事。費希特認為這種實質的義務內容必須以良知的形式意義為條件，亦即沒有形式條件的客體指涉失去其理論根據。在良知方面，形式意義在主體上呈現的不是直接對於自我活動的意識，而是具有某一特殊性質的心靈狀態 (Gemütszustande)，費希特稱之為信念 (Überzeugung)，這個信念狀態是因為判斷力在一個行為表象與自我活動「一致」 (Übereinstimmung) 的結果，因此它是間接地指涉到那個絕對活動。信念狀態因為是一個「一致」的狀態，它的出現必然牽涉的兩端，它的一端是上述的形式條件，即它與自我活動的關係；而它的另外一端便牽涉到良知的實質意義，即它與某倫理狀況下的一個特定行為發生關連，即它指出某一件事是義務。這裡費希特用一個例子來說明：假設現在某一個行為主體 = A 處於現實世界的某一個倫理情況，在他和絕對自由 = X 之間假設有一段距離，那麼我可以由 A 與 X 之間畫一條直線 = L，「我們可以把這個『直線』系列稱之為有限的理性存有者的倫理規定」。<sup>50</sup> 行為主體和自我的絕對同一性兩者構成兩個極端的點，它們之間構成判斷力活動的範圍，A 便是藉著判斷力尋找適合此倫理狀況的特定內容之義務，費希特認為當判斷力剛好找到一個在 L 上某一點 P，而 P 又距離 X 最近時，主體此時出現一個「一致」的感覺，這個 P 便是我們在這個倫理狀況要找的義務的實質材料，而指向這個 P 的這個直接性感覺，就是良知的實質意

48 GA, I/5, 138; SW, IV, 147. 良知這時候指向一個普遍原理，即人人都是具有「絕對自由」這樣的身分的存在者。

49 GA, I/5, 161; SW, IV, 173. 良知這時候指向一個特殊的個例，即某個特性的內容為當下之義務，例如：乍見孺子將入於井，救人便成為當下的義務。

50 GA, I/5, 141; SW, IV, 150.

義的內容，費希特稱之為「對於特定義務的直接性意識」。

費希特分析出一個良知的決斷（Gewissensentscheidung）具有一個複雜的結構關係，首先是「信念」的一種心靈狀態，它對費希特來說不是屬於實踐能力，而是一種理論能力上的產物，一種無法再被懷疑的感覺，這種感覺一方面終結反思判斷力的進行過程，另一方面連接到實踐能力之上而與一個特定的行為內容相連。我們可以繼續追問，這樣知識能力上信念的感覺是從何而來？費希特認為信念這種介於知識能力與實踐能力的關係，事實上是出自另外一種感覺，即上述的一致感覺（Gefühl der Übereinstimmung），這個感覺就是反思判斷力的產物，一種介於個別行為與絕對自由之間協同的結果。但是如果我們進一步再將反思判斷力出現的實質內容、信念的感覺、一致的感覺全部都抽離，那麼良知的決斷中還剩下什麼東西？費希特認為還剩一個良知的形式條件，行為主體自我與自我的關係，或者稱為一種內在性（Innerlichkeit）的東西，這裡便是康德良知概念的核心。在費希特的思想中，這個良知的形式條件雖然是良知概念很核心的內容，卻不是全部內容，因為對他來說，一個良知的決斷的形成有四個成分，缺一不可：一個特定的行為內容、信念的感覺、一致的感覺、主體的內在性。沒有特定的行為內容，良知只是自我與自我等同的關係；沒有信念與一致的感覺，反思判斷力是盲的，則它所取得的實在性無從得到適當的位置；沒有主體的內在性，永遠不能確定良知的決斷是不是主體自身的決斷。因此，唯有在這四個要素在同一個決斷出現，才能形成良知的決斷。

此處來到費希特的良知理論中的證成步驟，根據什麼理由可以證明良知所指涉的特定義務內容具有一種客觀的有效性，費希特認為：「這樣一種信念置於與原初的自我之間的一種和諧關係之中。但是這個原初的自我提升於一切時間和一切變化之上；因此在這個統一之中經驗自我超克一切時間的變化，而置自身於絕對不變的境地。」<sup>51</sup> 這個論證需要說明三點：第一，之前所談的主體內在性或者康德良知概念的主體指涉便是上述引文中的絕對自我；<sup>52</sup> 第二，絕對自我是超乎時間的變化；第三，良知確實是經驗自我與絕對自我的和諧關係。前兩點我們只能在費希特的整個系統中得到證實，而此處費希特僅就良知與絕

51 SW, IV, 169.

52 絕對自我是費希特體系的第一原理，根據筆者理解，這是指一種超乎經驗中的個體自我之上的形而上實體性之自我，在「個體自我」之中主體與客體永遠處於分離狀態，因為任何經驗意識活動必須預設這樣的分離狀態；但是「絕對自我」處於一種主客合一之狀態，所以它不可能出現在任何經驗意識之中，但費希特卻認為這是一切經驗意識的基礎。

對自我之間的關係進一步引申：「因此所有我們得以作為根據的完整知識，即那些對象的終極目的之知識，在信念的感覺沒有包含在事物的終極目的中的洞見出現之前，它不會被良知所批准，而且這些知識便是那些指導道德行為的知識。」<sup>53</sup>事物的終極目的之知識，對於一切具備理性的存有者都是相同的，因此良知對於任何理性存有者總是相同的。

### 三、良知的辯證

如果我們把良知理解為人的「內在法庭的意識」，尤其是良知是上帝的聲音，會有什麼困難？叔本華在他的《倫理學的兩個基本問題》一書第二卷第七章〈康德關於良知的學說〉裡提出這樣的質疑：

他（康德）說：「把一個受良心譴責的人和審判官等同起來，用這種方法描述法庭，是荒謬可笑的；因為在那種情況下，被告者一定會敗訴的。」為了闡釋這一陳述，他還加上一個非常模糊不清的注釋。他的結論是，如果我們不打算陷入矛盾，我們必須把這（在我們內部演出的司法的良心一戲劇中的）審判官想像為是與我們不同的，事實上，是另外一個人格；不但如此，而且想像為一個對人心無所不曉者，他的命令誰都要聽從，而且對執行機關的每一個目的來說，他有無限權力。這樣他經由一條平坦的路徑，從良心始而以迷信終，使後者成為前者的必然結果。<sup>54</sup>

叔本華認為將康德對於良知的解釋剝去其法庭比喻的外衣，「那末，周圍的光環連帶全部他的強制作用，也就迅速消失不見了，那裡剩下來的不過是這一事實，即有時當我們思考我們行為時，我們有些對自己的不滿，這種不滿顯然具有特殊的特點。」<sup>55</sup>叔本華主張應該將良知現象指就其自身所顯現的內容來解釋，不應該無限制擴大。康德良知理論裡的那個上帝人格之出現是一種推論，他的推論根據是法庭的三個角色——法官、原告與被告——之間的關係，但是如果將法庭這個譬喻外衣去掉，則良知是否是一種出自上帝的聲音，需要進一步的理論證據。

53 SW, IV, 171.

54 叔本華著，任立、孟慶時譯，《倫理學的兩個基本問題》（北京：商務印書館，1996），頁195。

55 同上書，頁196-97。

如果我們放棄良知上帝的超越側面，僅僅將它理解為「自我指向的判斷力」，那麼我們必須面對以下黑格爾的質疑：「這裡在道德這一形式觀點上，良知沒有這種〔義務的〕客觀內容，所以它是為己的、無限的自我肯定（Gewißheit seiner selbst），正因為如此，它同時又是這個主體自我肯定」。<sup>56</sup>如果良知是這樣完全不具備任何實質義務內容的東西，它自身會陷入辯證的過程裡，即它會由本來是善的轉向是惡的：「如果良知僅僅是形式的主觀性，那麼他正好就處於轉向惡的根源位置上，在為己存有的、認知的、決斷的自我肯定上，道德性和惡在此處有共同的根源」。<sup>57</sup>為了迴避這個善惡同源困難，黑格爾主張良知作為主觀的道德的階段（Moralität）必須過渡到客觀精神的倫理階段（Sittlichkeit）。<sup>58</sup>在費希特研究的文獻中有不少人主張，康德的良知理論無法迴避黑格爾這個善惡同源的困難，而認為費希特的良知理論剛好可以迴避這個困難，<sup>59</sup>但是這樣的主張需要證明一個前提，即費希特良知理論中的實質意義——某個倫理情況中的實質義務內容——不可能陷入善惡的辯證裡。如果情況是說，想要在良知概念中加入任何可能的實質內容都會陷入辯證關係中，那麼剛好證明黑格爾對良知問題的責難是針對費希特的，而非康德的，因為康德正好反對良知有實質內容的可能，根本沒有發生辯證的可能性。底下筆者嘗試說明，費希特的實質意義的良知無法避免陷入辯證，因此康德的形式意義之良知才是唯一可辯護的。

費希特的良知理論是剛好要在康德良知的形式意義添加一個不會出錯的實質意義，如果他是成功的，則他可以超越康德的良知理論而同時回答黑格爾的質疑。費希特也曾經面對黑格爾所提的質疑，對他來說，良知實質意義的形成完全是出自某一種獨白的過程（monologisches Verfahren），其間出現一種經驗自我與絕對自我的和諧狀態。他反問自己，如果賦予這樣私人的信念（Privatüberzeugung）一個無限制的權力，雖然依據他的系統推論良知是不可能出錯的，那麼它就可能出現一個與現在國家法律互相衝突的情況。這時不論是將私人的信念公諸於世或者付諸行動都是一種顛覆國家的行為，但是顛覆國家本身是違反義務及良知的。但是如果完全禁止私人的信念之表達，這樣便壓制個人的自由。黑格爾會主張，國家法律是高於個人的良知，因此違反國家法

56 《法哲學大綱》，第137節，Suhrkamp Verlag, Frankfurt a.M., 1986, Bd.7, S.255。

57 同上書，第139節，S.261。

58 黑格爾認為，在道德階段中裡的道德判斷是一種個人的判斷，而倫理階段裡的判斷由於是透過具體的社會規範或成文的法律給出，本身是屬於一種普遍的判斷。

59 參考註2，頁212。

律的個人良知內容是無效的。費希特認為國家法律與個人的良知之間可以找到一個環節使兩者不會出現矛盾，他認為必須成立一個學者共和國（*Republik der Gelehrten*）：

這個矛盾的綜合統一的環節將是這樣一種社群（*Gesellschaft*），其概念之中具有以下的內容：它應該部分地被限制及被規定，因此不是包含所有的不受規定的概念的人，而是包含特定被選出來，且與它們有區隔的數目之人。……它是一個社群意識的論壇，在它之前以一種無限制絕對的自由來設想和研究一切可能的事物。<sup>60</sup>

在這個學者共和國每一位成員都允許將自己的良知的內容提出來，進行討論。國家必須能夠容忍這些成員的意見，但是國家有權阻止這些意見被實踐出來。<sup>61</sup>這時候良知的考量脫離個別主體性的範圍，而進入互為主體性的對話中（*intersubjektiver Diskurs*）。互為主體性的對話不外乎要使每一個參與對話者達成一種共識（*sensus communis*），以便於當作行動決定的參考。如果這時候對話者之間完全達不到共識，表示對於同一件事不同的良知得到不同的義務內容，而這些內容之間無法取得最後的善惡判斷，表示良知的判斷如果要維持其絕對正確性，必須退回到主體的內在性——即揚棄其實質的義務內容。如果此時對話者之間取得同一個共識，那麼道德判斷力將陷入另一個循環裡，這個循環出現在共識與個人良知的信念之間，因為對話者嘗試個人的信念透過學者共和國的共識得到證成，但是這個共識之所以成為共識，已經預設每一個對話者具備這樣共識的內容才可能，因此這樣的證成又從共識中回到個別對話者身上。而根據費希特，為了突破這個循環以結束主體間的反思過程，必須假設一個直接性的感覺（按：即良知）指出這時候的義務內容。問題出在，在主體自身內部的良知之形成如果具有絕對正確性，它不可能與第二次經由互為主體性之間對話所取得的良知內容互相矛盾，因為這剛好將前一個良知的絕對確定性取消。但是個人的良知判斷與出自群體的共識的良知判斷之間不可能是完全一致的，亦即不可能維持其絕對正確性。因此，筆者認為要從自我內在性取得任何實質的義務內容，不論是假設內部的上帝，或者個人直接性感覺，或者經由互為主體性的共識取得的直接性感覺，都不能得到一種絕對正確的義務內容。良知如果要維持其絕對的正確性，康德的內在性理論是唯一的選擇。

60 SW, IV, 248.

61 SW, IV, 250.



#### 四、牟宗三的本體論模式與其對康德的批判

牟宗三先生的良知理論是完全以中國哲學為基礎發展出來的。筆者稱牟先生的理論模式為本體論模式，是因為這套理論的特點正在於它把良知視為一種形而上的本體，這是康德的理論中所見不到的，而且是最核心的。根據筆者對康德理論的分析，良知有以下的許多不同的意義，但這些都不包含本體論意義：

- (1)良知是義務的意識自身；
- (2)良知是一種來自神（超越者）的聲音；
- (3)康德又將這種判斷力界定為一種指向自我的判斷力；
- (4)良知是一個內在法庭的意識。

這些意義都為牟先生所解釋的良知理論所涵蓋，<sup>62</sup>這裡面最重要的差異是牟先生用「良心」來翻譯康德的Gewissen，以與中國哲學的良知作理論上的區分。首先，就內在法庭的意識而言，康德的「良心」和王陽明的「良知」是完全相同：「試以王陽明所說的良知為準而比論之。良知也可以比喻為一個內部法庭之意識，良知底裁判（良知之獨知）亦如一內部法官。至此為止，康德所說的良心與陽明所說的良知完全相同。」<sup>63</sup>其次，康德的「良心」只是主觀的感受條件，本身不能提供客觀的義務內容，亦不能提供道德法則的內容，用中國哲學的用語，「良心」與「天理」是分離的。但是牟宗三認為這心理分離的「良心」還不是王陽明的「良知」：「王陽明所說的良知確是心理是一者：良知之獨知（知是知非）即是良知的感受性（心靈之感受之主觀條件），但同時良知亦即是天理（法則）——良知即法則之源，法則即出於良知；良知即理性，並非良知外別有一虛懸的理性。在此顯出良心和良知底差別。」<sup>64</sup>「心理為一」與「心理分離」成為康德的良心與王陽明的良知之間區別第一個重要的標準。

用這段分判來對比康德的良心與中國哲學的良知，還不能完全顯出後者的本體義。這裡最重要的分判在康德的內在法庭理論中那個仲裁的法官是誰？根

---

62 筆者原先認為第三個意義不包含在牟先生的解釋裡，經匿名審查人的指正，牟先生在《現象與物自身》（臺北：學生書局，1975）一書的頁100-102中，談到知體明覺之「自我震動而返照其自己」。此義是牟先生引用康德《第一批判》對於智的直覺的定義：「如果該主體的直覺只是自我活動，即是說，只是智的直覺，則該主體必應只判斷自己」，同上書，頁102，參考K.d.r.V., B68。

63 《康德的道德哲學》，頁449。

64 同上書，頁449。

據康德，這個法官是上帝，不是良知。「這位理想化的人格，有權威的良心之法官，康德心目中視其為上帝，而在康德的道德哲學中，良心不會是上帝……若依陽明所說的良知，良知是檢察官，同時亦是有權威的最高審判官。」<sup>65</sup>這裡的問題關鍵在於，康德都把良知作為一種人內部的屬性來看，這個屬性自有其所關聯的實體對象，但它決不可能就等於那些它所關聯到的實體。因此整個問題分辨到此，依照中國哲學的理路，王陽明的良知理論可以含攝康德的法庭模式，最終它還包含一種西洋哲學中沒有的本體論向度。

以上的論述是牟先生以其對於康德至《道德形上學》的理論發展的總體批判。如果我們把牟先生的理論與康德晚年遺稿中的一些想法作一個對照，可以發現其實康德晚年對於道德本性的理解與牟先生所談的陽明學相去不遠，差別只是那個內在的道德行為者的超越者是神或者稱之為良知，在康德那裡，神與良知依舊不會是同一個東西，良知是內在神的聲音，不是內在神自身。而在王陽明或牟先生那裡，依然只存在一個內在的超越者——良知。<sup>66</sup>

牟先生是宋明儒學傳統下，第一個嘗試用西方式的論證去證成這個良知理論的哲學家，《智的直覺與中國哲學》一書中，牟宗三首先將「心理合一」的良知當成意志自律的條件，這裡「理」被理解為康德的定言令式，而在康德那裡發布定言令式的意志即是牟宗三的心，或稱為良知或性體：「道德即依無條件的定然命令（案：即定言令式）而行之謂。發此無條件的定然命令者，康德名之為自由意志，即自發自律之意志，而在中國的儒者則名曰本心，仁體，或良知，而此即吾人之性體。」<sup>67</sup>第一，心理不能分離的理由是，一旦「理」離開「心」，意志所發布的命令便不是由自己產生，那麼自律便不可能；第二，如果作為定言令式的理本身具有一種無條件性，那麼發布理的心本身必須是無限的存在者，這個無限者既創造吾人的道德行為，也創造天地的一切，換言之，良知本身是一個創生的原理：

性體既是絕對而無限地普遍的，所以它雖特顯于人類，而卻不為人類所限，不只限于人類而為一類概念，他雖特彰顯于成吾人之道

65 同上書，頁451。

66 雖然良知指涉到內在神的聲音的這一事實可以見到康德與牟宗三兩人理論的相近性，兩人之間依舊有一些本質性的差異：(1)牟宗三指的良知是實體，康德指的良知是屬性；(2)牟宗三的良知的聲音可以決定個別的行為，康德的良知的聲音不能決定個別行為。筆者兩者都反對，良知只能是「內在性」的聲音，而內在「神」只是一種推論。

67 牟宗三，《智的直覺與中國哲學》（臺北：商務印書館，1974），頁191。

行為，而卻不為道德所限，只封于道德界而無涉于存有界。它是涵蓋乾坤，唯一切存在之源的。不但是吾人之道德行為由它而來，即一草一木，一切存在，亦皆繫屬於它而為它所統攝，因而有其存在。所以它不但是創造吾人道德行為，使吾人的道德行為純亦不已，它亦創生一切而為一切存在之源，所以它是一個「創造原則」，即表象「創造性本身」的那個創造原則，因此它是一個「體」，即形而上的絕對而無限的體，吾人以此為性，故亦曰性體。<sup>68</sup>

但是依照康德，道德法則是適用於道德世界的普遍法則，而自然法則則是適用於自然世界的普遍法則，兩者各自管轄不同的世界，牟宗三卻認為道德法則可以超越道德世界決定自然世界，理由何在？牟宗三使用的是間接證法：

今試設想性體是一有限概念，未達此絕對而無限的普遍性之境，還能有此無條件的定然命令否？性體是一有限的概念即表示本心仁體其本身即是受限制而為有限的（此不是說其本身是無限，而在具體的表現中為具體機緣所限而為特定的表現）；其本身既是受限制而為有限的，則其發布命令不能不受制約，而無條件的定然命令便不可能；復次，本心受限制而為有限的，則本身不復是本心，本心轉成習心或成心而受制於感性，即喪失其自律性。<sup>69</sup>

這裡牟宗三的間接論證是以道德法則的無條件性和自律性為前提，假設發布道德法則的「良知」僅僅限於道德世界，則它必然是有限的，但是第一，有限存在者不可能發出具有無條件性的道德的命令，否則便失去道德的無條件性，第二，有限的良知不是本心，人變成由受感性決定的習心所決定，便失去道德的自律性。筆者將牟宗三的整個論證用以下幾個步驟整理：

(1)如果承認道德法則是無條件的，則發布道德法則的存在者必定是無限的存在者。道德法則的發布者與道德法則之間存在因果關係，根據果不能大於因的原則，不可能由一個有限存在者作為原因，而得到無條件道德命令的結果。<sup>70</sup>

68 同上書，頁191。

69 《智的直覺與中國哲學》，頁191；參考：牟宗三，《心體與性體》（臺北：正中書局，1973-1975），冊2，頁223-224。

70 這裡牟宗三的確是使用因果原則在做推論，見上述引文的下一頁，即《智的直覺與中國哲學》，頁192：「又，當吾人就無條件的定然命令而說自由意志為自由自律時，此自由自律即表示只能為因，而不能為果，即只能制約別的，而不為別所制約……。如是，這第一因與發布無條件定然命令的自由意志其性質完全相同。如果

- (2)如果人能夠自律，則發布道德命令者就是他自己，否則便是他律。
- (3)現在良知是道德命令的發布者。
- (4)現在，道德法則是無條件的。
- (5)因此，良知內在於人而且必然是無限的存在者。
- (6)這個無限的存在者以其無限實體的身分創造這個世界，而且世界中的存在物無一例外，都是它創造的結果。

這看起來是一個嚴整的論證，但內部卻隱藏至少三個理論的困難：

(1)這裡最為直接的是辯神論的反駁：良知是全善的，它所創造的世界卻是有善有惡，可見良知不是全善的，或者世界不完全是它創造的。王陽明似乎也有想到這個問題，因為他的四句教說「無善無惡心之體，有善有惡意之動，知善知惡是良知，為善去惡是格物」，牟宗三說：

無善無惡心之體是就「至善是心之本體」而說。無善無惡是謂至善。然則無善無惡是「無有作好，無有作惡」之意。善惡相對的謂詞俱用不上，只是一自然之靈昭明覺亭亭當當地自持其自身，此即為心之自體實相。至善是心之本體，猶言是心之自體實相，簡言之，就是心當體自己也。此心須當下即認為是超越之本心，不是中性的氣之靈之心也。心之自體是如此，然其發動不能不受私慾氣質之阻隔或影響因而被歪曲，因此「有善有惡意之動」。其發動即得名曰意。故「意」可以說是經驗層上的。然發動的或善或惡，此心之自體及其靈昭明覺之上自己未嘗不知之，此即所謂良知。如是，這良知即超越在經驗層上的意之上而照臨之。意有善惡兩枝，而照臨此意之良知則是絕對的純一，故它是判斷底標準。<sup>71</sup>

心之體無善無惡的意思不是說，它是中性無善惡的事物，反而是一個至善，或者超乎善惡的本體，超驗層的東西，心之體發動成為意（念），因受物慾的阻隔而有善有惡，善惡出於經驗層的意。但是心之體與意的關聯是什麼？前者照臨後者，給予後者是非標準。王陽明和牟宗三的意思也許是，經驗層的事物雖然是超驗層的良知所創造，但是創造之後因意的發動出現善惡是經驗層的責任，與至善的超驗層無關。這樣可以同時保有良知的至善，以及解釋經驗世界

第一因是絕對而無限（隱指上帝言），則自由意志亦必定是絕對而無限的。天地間不能有兩個絕對而無限的實體，如是，兩者必同一」。

71 牟宗三，《從陸象山到劉蕺山》，《牟宗三先生全集》（臺北：聯經，2003），冊8，頁195。

的惡的出現。《圓善論》中牟先生談康德論惡的起源之附錄中講得較為清楚：

人之生命中的那作為「採取格言」之基本根源或終極根據的善性格（性向）或惡性格（性向）早已在人生時即現存於人，但這並不表示說這生時之生是這善性格或惡性格之原因。善或惡的性格（性向）之所以為善或惡，即那使其為善或惡的原因或理由，是源於人之採用此格言或彼格言，而此卻是存在於人之有自由選擇作用的意志中，或存在於自由之使用中。案：「有自由選擇作用的意志」是譯德文Willkür一詞，此亦可譯為「自由決意」，此是人的現實意志之作用，故有善有惡，此相當於王陽明所謂「意」，「意念」之意，「心之所發為意」之意，此是一個表示「意志之受感性影響的現實作用」的字，故英譯亦譯為volition，此是一個作用字，與設定其為「自由」，是其自由為設準而不能被證明的那個意志不同，此後者是一個實體字，在德文是Wille。凡說純粹的意志，善的意志，或自由的意志，皆是指的這個實體字，此則不可有善有惡，只能是純善。<sup>72</sup>

牟宗三用康德的「自由選擇的意志」（Willkür）一詞來解釋王陽明的「有善有惡意之動」，對於康德來說，擁有「自由選擇的意志」的人可以自由地選擇善的格律或者惡的格律，這樣選擇的原因不出於造物主，而是在「自由選擇的意志」之中，因此人必須為自己的選擇之善惡後果負責。用王陽明的文字脈絡，善的意志或是良知是一個實體，無善惡可言，而意念只是這個實體的一種作用，但是這種作用具有善惡的性質，它的「原因」不在於善的意志這個實體，而在於「意念作用自身」。但是按照這樣的看法，良知有個空白地帶是它創造原則所無法涵蓋的，那就是人的善惡選擇的自由，因為這樣才可以解釋惡的起源；但是如果這個空白地帶不是良知所管轄的，那麼它就有所封限，因此不是無限者，而是有限者。牟宗三必須在兩個理論中做選擇：或者良知不是無限者，或者它是善惡混。而良知如果進入善惡混，則陷入錯誤的辯證知識之中。

對於這個批評的可能性反駁，牟先生在《圓善論》談到根本惡問題並沒有正面的回答。筆者認為牟先生唯一深入談這個問題是在《康德的道德哲學》中對《第二批判》的譯註，他借用康德對於「現象與物自身」的區別來回答這個難題：

如果人是上帝所創造的有限存有者，則他雖是一實體物（自體物），亦必依待於上帝，此即表示他不能有自由。康德解答這難題是如此：

<sup>72</sup> 牟宗三，《圓善論》，《牟宗三先生全集》（臺北：聯經，2003），冊22，頁62-63。

上帝創造人乃至萬有是就其為物自身而造之，不是就其為現象而造之。物自身是不在時間串系中的，因此萬有雖是被造物，就人說（人亦是被造物之一），仍可說自由。因為創造之觀念只應用於他們或它們底超感觸的存在，不能應用於他們或它們底感觸的存在。<sup>73</sup>

康德認為，把上帝當作物自體存在的原因，而同時又將祂當作現象的原因是矛盾的，上帝這時又是超時空的存有物，又是時空內的存有物。要解消這個矛盾，只能將物自體的原因歸於上帝，而現象的原因歸於人：「在時間串系中並在範疇決定下的現象是不依靠於上帝，亦不為上帝所創造。說上帝是現象底創造者，這是矛盾的，亦必否決上帝之無限性與獨立性」。<sup>74</sup>牟先生認為用康德所撐開的上帝、物自體與現象的關係，可以普遍地解釋儒、道、釋三家，只是其間各自有其分際。如何用在儒家的道德理論上？良知就是那個創造物自體世界的上帝，而現象是在時空屬性這種感性條件下的產物，其中善惡的區分因為「自由的選擇」而在現象內部構成而獨立於良知的因果關係，這些是屬於人的感性身分的存有所創造，其原因不在良知之上，而是人作為感性身分的存有。

(2)筆者認為，牟先生在對於道德法則的「無條件性」解釋上，弄錯了應然與實然的關係。康德說的無條件並非指自然或現實上的原因與結果的條件關係，而是指一種「無條件的應該」（*unbedingtes Sollen*），這裡牟先生混淆了自由與自然的界限，<sup>75</sup>亦即，這個無條件性最大可能的理論範圍只由涵蓋人的行為。因此以無限實體來解釋人的自由意志犯了一個理論上的錯誤，實踐上的應然在還沒有採取行動之前現實上並不存在，無條件的命令是等待被實踐的東西，採取（或發布）這項命令的人是一個自律者，但是後者和前者不是原因和結果的關係，換言之，即使是一個有限的意志，一樣可以採取一個不是以自身作原因的無條件命令作為其行為的根據。

73 《康德的道德哲學》，頁324。

74 同上書，頁324。

75 有人會認為，康德《第二批判》導言的主張：「自由是道德法則的存在根據」剛好證明牟宗三與康德在道德法則的無條件性的理解並無二致，如果牟宗三有錯，康德也同樣有錯。筆者卻認為，康德的那個主張和此處牟宗三說的發布道德命令者是無限者不是在說同一件事。康德由道德法則推演出來的只是一種個人的自由，他以自由為基石分析出最高善，再由此推出最高善的保證者—無限的神。在康德那裡，自由是道德法則的存在根據，而上帝不是道德法則的條件，而是道德法則所決定的意志之必然對象的條件（《現象與物自身》，學生書局版，頁54），而牟宗三從道德法則的無條件性直接推出無限者（即康德意義下的上帝），自由和上帝之間不能是

對於這個混淆實然與應然的批評，牟先生在許多地方都有反駁，在《心體與性體》第一冊，綜論第三章，牟先生談到中國哲學道德理性三義中，是這樣談實然與應然的關係：

若依宋明儒之大宗說，道德性的天理實理是本心性體之所發。本心性體或於穆不已之道體性體是寂感真幾，是創造之源，是直貫至宇宙之生化或道德之創造。自宇宙生化之氣之跡上說以及自道德創造所引生之行為之氣之跡上說，是實然，自然，是服從自然因果律，但自創造之源上說，則是當然（案：即應然），是服從意志因果的。如是，則這種契合是很直接而自然的，不必須曲曲折折強探力索地去艱苦建構了。這個問題不是直接地就「道德底當然」與「自然底實然」這兩個自性絕對不同的概念而去搏鬥，一方面說很難，一方面說很容易。<sup>76</sup>

牟先生的意思是，康德把一切外在對象的牽連斬斷，顯現道德理性之必然性之應然，只涵蓋中國哲學道德理性的第一義，而「定然地真實的性體心體不只是人的性，不只是成就嚴整而純正的道德行為，而是直透至其形而上的宇宙論意義，而為天地之性，而為宇宙萬物底實體本體，為寂感真幾、生化之理，這是『涵蓋乾坤』句，是道德理性底第二義」。<sup>77</sup>

牟先生對中國哲學的體悟使他消弭了西方哲學家對應然與實然這個兩個世界的分離，他真正的洞見就在於，在良知或性體那裡，實然與應然是不分的，它同時給出應然的道德命令，也實然地是世界的創造者，這兩者是不衝突的。是在《心體與性體》和《智的直覺與中國哲學》之中，牟先生似乎想直接地肯定自然世界與道德世界的同一性。<sup>78</sup>但是在較晚的著作裡，他似乎轉向以價值世界來解釋良知所創造的世界（或者說是良知所創造的物自體世界等於價值的世界）：「儒家說的天道創生萬物，這也是對於天地萬物所作道德理性上的價值的解釋，並不是對於道德價值作一存有論的解釋」。<sup>79</sup>這是一個值得注意的轉變，但是這樣的轉變不是直截了當地把實然等同於應然，而是將自然界的實然理解為價值上的實然，牟先生把原本屬於自然的實然物轉變為價物而置之

---

等同關係。因此，康德並未混淆自由與自然的界限，牟宗三卻刻意打通兩者。

76 《心體與性體》，冊1，頁115-116，參考頁157-158、172；《康德道德哲學》，頁340-341。

77 《心體與性體》，冊1，頁138。

78 《心體與性體》，冊1，頁115-116，《智的直覺與中國哲學》，頁201。

79 《圓善論》，頁131。參考：《現象與物自身》，《牟宗三先生全集》，冊21，頁10-14。

於應然的道德之內部，來解消兩者的分裂，這是一個較為保守卻比較容易為之辯護的立場。

(3)這裡最為嚴重的理論問題，是牟先生從應然的道德的「無條件性」推演出給出無條件性的主體，然後把這個主體等同於一個「無限性的實體」，這個理論的混淆造成一個創造天地的良知，把內在法庭意識中的法官等同於良知；造成一個消弭自然世界，僅存價值世界的理論；造成一個什麼內容都包括的良知概念：既是判斷力，也是內在法庭意識，也是法庭內的法官，也是主觀的感受性，也是道德情感。

首先，康德分析的道德無條件性，不是一種自然現象上的因果分析，康德只是說，一件行為之所以是道德的，不是為了某種行為的結果為善，而是它本身就是善的，所謂無條件指的是不為自己之外的事物之善的緣故而為善，無條件性只是一屬性，根本不是一種實體，而這時候「目的自身」的概念很重要，一個具備目的自身的存在者，它是無條件者，但絕對不是無限者，<sup>80</sup> 否則目的王國中只會有一個成員，這不合理。其次，為了自律概念，那個給出無條件的道德法則必須是道德主體自身，這個想法沒什麼問題，但在這個給出過程中，牟宗三作了另一個錯誤的推論：「今試設想性體是一有限概念，未達此絕對而無限的普遍性之境，還能有此無條件的定然命令否？性體是一有限的概念即表示本心仁體其本身即是受限制而為有限的（此不是說其本身是無限，而在具體的表現中為具體機緣所限而為特定的表現）；其本身既是受限制而為有限的，則其發布命令不能不受制約，而無條件的定然命令便不可能。」這是一段無效論證，第一，主體給出的道德法則是「無條件的」，不是「無限的」。第二，而且就算道德法則是「無限的」，也只是理念上應然的無限，不是現實上的無限，因此，一個有限實體絕對是可以給出一個應然的、無限的道德法則，正如一個自我要求完美的人，本身不一定要事實上已經是完美的人，他卻可以用完美的概念要求自己。在此頂多再假設一個能夠實現這個無限的道德法則的能力而已，但這個無限的應然與實現它的能力合在一起，並不能得到一個現實的無限實體的結果。

80 按照康德，無條件者有兩種意義，第一種不容例外地包含一切事物的系列之整體，在這種情形下的無條件者是無限者；但是無條件者也可以指這些系列的部分，它們不處於其他條件之下（《純粹理性批判》A417=B445）。理性對知性範疇要求總體性這樣的無條件性是屬於第二種，這個情況下有三個無條件者：自由、上帝、靈魂（《純粹理性批判》，B394註釋），而道德法則的無條件性也是屬於第二種。因此，這個無條件者並不等於無限者。



## 結語

康德以其獨特的方式對良知進行分析，一方面是對於包姆嘉登的修正，另一方面受基督教傳統的影響，他理解的良知對下不能指定任何一項行為是義務（這是判斷力的工作），對上不能等同創造主的上帝的無限實體。它只能是在主體內部的一種感受性的東西。費希特在其良知理論中，除了保留康德主觀的自我指涉之外，嘗試加上一個對象指涉之意義以尋求特定之義務內容，但是如上所述，這樣的增加反而使良知陷入辯證。牟宗三更進一步，把上帝等同於良知概念，而這樣的良知概念，是牟宗三分析康德道德無條件性概念和自律的概念所得到的。誠如筆者上述的分析，牟宗三從康德道德分析的無條件性的應然推演到給出應然的主體的無條件性，然後把無條件性等同於無限性。這裡出現兩個推論的錯誤，(1)給出無條件性的主體本身不一定要是無條件者（因為這是應然的無條件性）；(2)無條件性並不等於無限性（目的自身是無條件的，但是它不是無限的）。

## 引用書目

### 一、中文

- 牟宗三，《心體與性體》三冊，臺北：正中書局，1973-1975。
- 牟宗三，《智的直覺與中國哲學》，臺北：商務印書館，1974。
- 牟宗三，《現象與物自身》，臺北：學生書局，1975。
- 牟宗三譯註，《康德的道德哲學》，臺北：學生書局，1982。
- 牟宗三，《從陸象山到劉蕺山》，《牟宗三全集》，冊 8，臺北：聯經，2003。
- 牟宗三，《圓善論》，《牟宗三全集》，冊 22，臺北：聯經，2003。
- 叔本華著，任立、孟慶時譯，《倫理學的兩個基本問題》，北京：商務印書館，1996。
- 康德著，李明輝譯，《道德底形上學之基礎》，臺北：聯經，1990。

### 二、西文

- Baumgarten, Alexander Gottlieb (ed.). *Ethica Philosophia*, 1751.
- Fichte, J.G. *Gesamtausgabe der Bayerischen Akademie der Wissenschaften*. Hrsg. von R. Lauth. Stuttgart-Bad Cannstatt: Frommann, 1962-.
- Fichte, J.G. *System der Sittenlehre nach den Prinzipien der Wissenschaftslehre*, Bd.5.
- Hare, R. M. *Freedom and Reason*, Oxford University Press, New York, 1963.
- Hegel, G.W.F. *Gesammelte Werke*. Hrsg. von E. Moldenhauer. Frankfurt a.M.: Suhrkamp, 2000.
- Hegel, G.W.F. *Grundlinien der Philosophie des Rechts*, Bd.7.
- Hoffman, T. S. "Gewissen als praktische Apperzeption. Zur Lehre vom Gewissen in Kants Ethik-Vorlesung," *Kant-Studien*(2002), 424-443.
- Ishihawa, Fumiyasu. "Das Gerichtshof-Modell des Gewissens," Nobert Hinske(ed.): *Kant und die Aufklärung*, Hamburg: Meiner, 1993, 43-55.
- Kant, I. *Gesammelte Schriften*. Hrsg. von der Preussischen Akademie der Wissenschaften. Berlin: Walter de Gruyter, 1911-.
- Kant, I. *Grundlegung zur Metaphysik der Sitten*, 1785: AA Bd.4.
- Kant, I. *Kritik der praktischen Philosophie*, 1788: AA Bd.5.
- Kant, I. *Die Religion innerhalb der Grenzen der bloßen Vernunft*, 1793: AA Bd.6.

- 
- Kant, I. *Metaphysik der Sitten*, 1797: AA Bd.6.
- Paton, H.J. "Conscience and Kant," *Kant-Studien* (1979), 238-251.
- Ungler, Franz. "Zu Fichtes Theorie des Gewissens," *Wiener Jahrbuch der Philosophie* 12 (1979), 212-235.
- Wolff, Christian. *Gesammelte Werke*. Hrsg. Jean Ecole. Hildesheim: Olms. 1. Abt. Deutsche Schriften Bd.4
- Wood, Allen. *Kant's Ethical Thought*, Cambridge: Cambridge University Press, 1999.
- Wood, Allen. *Kantian Ethics*, Cambridge: Cambridge University Press, 2007.

# Dialectics of Conscience: A Comparative Study of the Theories of Kant, Fichte and Mou Zongsan

Pong, Wen-berng\*

## Abstract

The central theme of this paper is the problem of conscience in a context of comparative study of Eastern and Western moral philosophy. I argue that the absolute rightness of conscience cannot be justified if one does not limit the sphere of its validity. Conscience can be understood in two different ways: it can either contain a subjective reference (Subjektbezug) or an objective reference (Objektbezug). Subjective reference means self-examination based on practical reason, while objective reference signifies reference to a concrete content of duty. Kant holds that conscience can only have a subjective reference, because an objective reference is not the task of conscience, but of moral judgment. Both Fichte and Mou Zongsan hold that conscience cannot just have the empty form of subjective reference, but should also contain some concrete substance without losing its absolute rightness. My conclusion is that Kant's interpretation of conscience is right. The reason is quite simple: the objective reference of conscience cannot avoid the difficulty of dialectics.

**Keywords:** conscience, Fichte, Kant, Mou Zongsan

---

\* Assistant Professor, Department of Philosophy, National Taiwan University.