國立臺灣大學法律學院法律學系

碩士論文

Department of Law

College of Law

National Taiwan University

Master's Thesis

# 制裁作為法律獨立構成理由的方式

# Sanctions as an Independent Way for Laws to Constitute Reasons

梁廷瑋

Niû Têng-úi

指導教授：莊世同、楊金穆 博士

Advisor: Shih-Tung Chuang & Chin-Mu Yang, Ph.D.

中華民國 113 年 2 月

February 2024

doi:10.6342/NTU202400318

# 國立臺灣大學碩士學位論文
# 口試委員會審定書

## 制裁作為法律獨立構成理由的方式

## Sanctions as an Independent Way for Laws to Constitute Reasons

本論文係梁廷瑋君（學號:R09A21005）在國立臺灣大學法律學系完成之碩士學位論文，於民國113年1月23日承下列考試委員審查通過及口試及格，特此證明

指導教授：

口試委員：

Hung-Ju Chen

# 謝辭

一

有些人是在死後才誕生的，有些理論也必須在全面潰敗之後才能找到新生命，或許正能說明這篇論文的來歷，但對於他的命運，我也捉摸不定了。

這篇論文歷經了多次立場與心境的波折。我的立場跟當年進來研究所時相比已經改變許多。在進研究所之前，我是一個德沃金主義者，並對法學、正義以及進步主義有著不切實際的樂觀。我曾經也沒料到，我最終是導向這樣的觀點：悲觀，並對法律、法學與人性充滿失望。或許一方面可以歸結到我的個人經歷與性格，但另一方面，也可能是處境與遭遇所導致。

作為台大法律系的外來者，在前期都是相當辛苦的。一來，台大的學生本身就有較強的競爭心與榮譽感，比外校的學生，台大的學生普遍上來說似乎都心機重了點。這種隱性的心防與競爭心態，讓台大的學生通常比外校的學生還難討論問題；二來，法學院本身就帶有獨特的氣息，是其他系所都沒有的。這四年來，我一直在思考這個獨特的氣息是什麼，這種在外校法學院僅有一絲味道的獨特氣息，在這個第一學府當中被強烈的放大了出來，幾乎窒息到我難以忍受。我認為外來者要融入這種圈子，似乎需要努力與人生智慧，但看來我是徹底的失敗了。

即便如此，我仍然倒吃甘蔗到了研究所的最後。很大一部份要感謝莊世同老師的扶持，在我幾乎脫離法學院，在各個系所跟學校流浪之際，給我了擔任基法中心助理的機會，讓我重新接觸了基法組的學弟妹，似乎又重新回歸到了這個社群。一直到現在，我都還不確定我有沒有資格獲得大家的友善。另一方面，莊老師跟韻如老師給我的工作機會，是我在獨立在台北生活的必要條件。沒有這些經濟支援，這篇論文根本不可能寫成。甚至，莊老師也在我心情低落之餘，打電話來給我支持與鼓勵，甚至還宴請大餐，讓我在這個冰冷到窒息的法學院當中，有一些溫暖的空間。雖然還是沒有改變這本論文當中因而悲觀的法學理論，但這些都是我還有餘力繼續書寫這本論文的動力。

在這空白的四年中歷經了許多碰撞、失敗與疏離，這篇論文或許可以當作我的自辯吧。這篇論文最大的敗筆，就是我沒有跟他保持好距離，讓這篇論文充滿了情緒性的色彩，但或許也是因為痛苦，我才找到全新的理論生命。尼采正是認



doi:10.6342/NTU202400318

為哲學家應該把悲觀主義當作錘子，來擊倒舊有的道德標準。一直以來法理學家對於法律的善解讀，正是我想要擊倒的對象。我一直覺得我真正的理論根源，不是來自於表面引用的分析哲學與邏輯學家，而是來自於尼采、施密特的非理性主義。

比起正義本身，我對正義的起源更感興趣。我相信所有對議題熱忱的人——特別是外校的人——都會對台大法律系作為議題先鋒有所憧憬。我也曾經是其中一員，不過，這些憧憬對於務實的法律系來說多半只是一廂情願。畢竟法學是一種實踐的技藝，比起哲學家，法律人可能更像政治家。哲學在政治的領域內只是權威的點綴，法律人的行動也必須從政黨政治與派系運作來理解，這就會形成一種弔詭——法律人需要哲學，卻在關鍵時處死蘇格拉底——就此而言，哲學家終究太過浪漫了。

## 二

這篇論文趕稿匆忙，就論述上來看，恐怕有些草率。雖然沒有達成我心目中的卓越，但至少完成了自己的一些的期許。

在理論逐漸形塑的過程中，我的兩位指導老師都幫了我很大的忙。在法理學上，莊世同老師給了非常重要的指引，他給了我 Schauer（2010）的文章，把我原先著重在規範性討論的觀點拉回了制裁理論，並且融入法理學的討論當中。在邏輯學上，楊金穆老師為我打下深厚的數理邏輯基礎，讓我有能力理解並建構邏輯系統。並且，金穆老師閱讀了這本論文的兩種早期版本：分別是關於法理學與邏輯學的，而且逐字逐句地批改，給了許多文字敘述上與結構上調整的建議。

研討會的評論人張明融、王鵬翔老師跟王一奇老師都給了這篇論文的多種前身版本不少的建議。張明融再三提醒學術文章應該要回應學術脈絡，他讓我避免了過於脫節的發想，也讓我認真思考我的論文究竟要回應什麼具體的問題。王鵬翔老師則提醒我要縮小問題意識，我處理的議題經常太廣，時常想到一個好點子就岔出去處理，但這最終只會模糊焦點。王一奇老師還特別打電話來教我學術寫作的技巧，要求我應該把我支持的論點，直接寫成命題列在緒論的第一行。這本碩論基本上可以直接作為王一奇（2015）的論文註釋，其中的一部份能獲得王一奇老師的評論實在備感榮幸。同時，也要感謝口試委員陳弘儒老師，對於本文過強的宣稱以及因果模型的技術上瑕疵做出提醒，使我能夠再三反思我的論證力度是否配得上我所宣稱的強度。但即便如此，礙於學識與書寫能力，最終這篇論文



doi:10.6342/NTU202400318

的成品恐怕還是遠遠落後於他們所要求的。

柯甯予與張明融在我理論逐漸形塑、建構的過程當中給了許多實質的意見。裡面大多數的論點都已經與他們討論許久，回應他們的回饋與批評，一直是這篇論文的進展最有成效的部分。師母林陳彩霞向我分享了許多台大法律系舊時代的往事，讓我對這種「法律系的獨特氣息」逐漸摸清了他的輪廓以及前世今生，最終還要感謝師母特地來參加我的研討會，實在讓我受寵若驚。這本論文很多的重要論述與觀點，都是來自於回應與傾聽他們的過程。

在研究所的期間，我仍然獲得了許多生活上的陪伴：劭學給了我許多工作上與學術上的機會，還兩度邀請我去居酒屋喝酒談心。我也很感念日治法院檔案解讀會議的嘉容、子涵與佩璇等人，跟大家解讀古典日文的時光是十分快樂的。宥秀和柏翰等人還特別把研究室搬來跟我同一間，讓原先孤僻的霖澤館突然變成基法大本營。方昕也不時給我飲料打氣，提醒我不要沮喪。也感謝甯予、明融在這近四年來，對於我想法上、個性上的理解與支持。期間不乏還有些人不小心成為我的情緒垃圾桶，也感謝大家包容我的陰晴不定。你們都是在這篇論文中被我聲稱拋棄、但卻成為堅實反例的共善代表。

最後，我要感謝倢伃，願意接受我從事學術的人生選擇，並且一起找到了明確的未來方向。在最終匆忙趕稿、申請學校的幾個月間，倢伃分攤事務許多。倢伃可以說是我成功的全部原因，以及對未來抱持信念的所有理由。但這交往十年來，卻有太多虧欠於她。

梁廷瑋

二〇二四年寫於士林



doi:10.6342/NTU202400318

# 摘要

本文試圖辯護一個命題：制裁是法律獨立構成理由的方式。也就是說，法律的規範性，來自於制裁。這樣的觀點最早可以溯及至約翰・奧斯丁，他將法律義務定義為「主權者的命令」。然而，在哈特出版《法律的概念》之後，制裁理論基本上成為了眾矢之的，本文透過重新註解王一奇的〈理由與提供理由的事實〉一文，重新建構一個新的制裁理論。

本文第一章批評過往的法律理論，將法律規範性的來源奠基在法律追求的理念上。這種方法忽略了法律與法學實際操作的現實，因此本章試圖論證：為何應該、以及如何採取描述性理論，並說明在德沃金的詮釋理論挑戰下，描述性理論仍得以可能的關鍵。第二章借用王一奇的基於差異製造的理由論，指出集體理由與個體理由之間存在無法跨越的三種縫隙，使得集體理由在這個社會當中穩定實踐會出現困難，因此，社會需要制裁來跨越集體理由與個體理由之間的鴻溝，而制裁本身即是個體理由。第三章評論約翰・奧斯丁，指出「主權者的命令」並不是法律獨立構成理由的方式。其中，「主權者」會破壞構成理由的「獨立性」，而「命令」並不會「構成理由」，只能「提供理由」。此外，「獨立性」與「構成理由」是分離命題的必要條件，因此主權者的命令不會是一個法實證主義的主張。第四章指出了法律作為理由的各種形式，逐步分析權威、原則、規則以及共善提供理由的方式，這些方式都將其他理由混淆進了法律理由本身。法律理由必然是一種基於制裁的個體理由，這是一種人性本惡的理由觀。

關鍵字：基於差異製造的理由論、制裁理論、約翰・奧斯丁、因果模型、集體理由、個體理由、提供理由、構成理由



doi:10.6342/NTU202400318

# Abstract

This thesis seeks to defend a proposition: sanctions are an independent way for laws to constitute reasons. That is to say, the legal normativity comes from sanctions. This point of view can be traced back to John Austin, who defined legal obligation as 'the command of the sovereign'. However, after H.L.A. Hart published 'The Concept of Law,' sanction theory largely became a target for criticism. This article seeks to reconstruct a new sanction theory by re-interpreting Linton Wang's 'Reasons and Reason-Giving Fact.'

Chapter 1 criticises past legal theories for basing the source of legal normativity on the ideal pursued by the law, which overlooks the realities of legal operations. This chapter, therefore, seeks to illustrate why and how the descriptive theory should be adopted. Chapter 2 points out, through the Difference-Making-Based Theory of Reasons proposed by Linton Wang, that there exist three gaps between individual reasons and collective reasons, making it challenging for collective reasons to practice stably in society. As a result, society requires sanctions to bridge these gaps between collective and individual reasons, with sanctions themselves constituting individual reasons. Chapter 3 critiques John Austin's view, arguing that 'the command of the sovereign' does not independently constitute legal reasons. The concept of 'sovereign' undermines the 'independence' of legal reasoning, while 'command' can only 'provide a reason' rather than 'constitute a reason.' Moreover, 'independence' and 'constituting reasons' are essential for the separation thesis; thus, the sovereign's command cannot support legal positivism. Chapter 4 explores and analyses one by one the various ways in which laws serve as reasons, including authorities, principles, rules, and the common good. All of these ways confuse other reasons with legal reasons themselves. Legal reasons must be individual reasons based on sanctions, a theory of reasons belonging to the theory of evil human nature.



Keywords: Difference-Making-Based Theory of Reasons, Sanction Theory, John Austin, Causal Models, Collective Reason, Individual reason, Providing Reasons, Constituting Reasons





doi:10.6342/NTU202400318

# 導論

本文試圖辯護一個命題：制裁是法律獨立構成理由的方式。也就是說，法律的規範性，來自於制裁。這樣的觀點最早可以溯及至約翰·奧斯丁（John Austin），他將法律義務定義為「主權者的命令」，並且，命令意味著：「若不遵守我的期許，我將給予惡害。」（頁 1）然而，在哈特（H.L.A. Hart）出版《法律的概念》之後，制裁理論基本上成為了眾矢之的。哈特區分了「受迫」（obliged）與「義務」（obligation）的差別，並指出：純粹武力的受迫，並不等於存在義務，正如同搶匪命令交出錢的時候，我並沒有實際的義務交出財產一般（頁 2）。哈特主張，法律義務若要存在，必須從法律的**內在觀點**（internal point of view）來考察。所謂的內在觀點，是指參與法律系統的人，對於法律規則存在**接受**（acceptence）的態度。而這種接受，反映在**他人不遵守規則時所產生的批判**上。如果參與者只有指引自己的行動，而不涉及對他人的批判，那麼這樣的觀點就不是接受，也不是哈特所謂的內在觀點，而只是純粹考量自己利益的壞人觀點（頁 5）。就此而言，制裁並不是法律概念的構成要素。

本文透過註解王一奇〈理由與提供理由的事實〉來解析哈特的內在觀點，並將論述的焦點關注在「理由」以及「法律作為一種理由」的性質。首先，理由是一個行動到結果的因果關係（頁 15）。這樣的因果關係，基於行動者人數的複數與單數，區分為集體理由（collective reason）與個體理由（singular reason）（頁 21-22）。集體理由因為需要多數人的協力、合作才能達成理由所追求的目標，因而**集體理由帶有批判他人的能力**（頁 23）。相反地，個體理由僅僅需要自己的行動就可以達成，因此不具備批判他人的能力。所謂的壞人觀點，即是指法律作為個體理由，只需要考量自身的「目的－手段」衡量關係，而哈特所謂接受的態度，則是指法律作為指集體理由，因此存在批判他人的效果。因此，哈特與壞人觀點之間的緊張關係可以重構為：**法律究竟是集體理由還是個體理由？**

集體理由與個體理由，都是個體在行動時會考量的兩種理由（頁 23），然而，本文指出，集體理由與個體理由之間因為存在三種縫隙，因而集體理由在這個社會上**穩定且持續的實踐**會導致困難（頁 23-28）。為了解決這個問題，社會因而設立了制裁。制裁的理由是一種個體理由，在功能上，制裁能夠完美地模擬所有規範狀態個體理由（頁 29-30），且能夠完美地模擬所有秩序態樣的集體理由（頁 31-33）。在一個大規模的社會當中，制裁有能力代替集體理由穩定且持續的實踐於



doi:10.6342/NTU202400318

社會。因此，作為維繫大規模社會秩序的法律，必然屬於一種個體理由，而這種個體理由，是基於制裁的個體理由。

再來，本文接著探討「法律作為一種理由」的形式，本文回顧奧斯丁對於法律的一般性定義：「法律是主權者的命令」（頁 34-35）。本文指出，法律理由若要成為全新且獨立的理由系統，必須遵守三種條件：構成性（constitutiveness）、思慮性（deliberativity）、獨立性（independence）（頁 50）。所謂的構成性，意味著法律必須要構成（constitute）全新的理由，而非提供（provide）一個既存的理由。在這樣的意義下，遵守命令，並不是命令本身構成了什麼理由，而是命令觸發了一個既存的理由，這個理由代表了命令者與被命令者之間的關係，因而命令只能提供理由，而不能構成理由（頁 47）。另一方面，獨立性意味著理由系統不能從屬於其他的理由系統，然而主權者作為一切制裁的來源，主權者制裁的理由，就不再是以制裁為基礎的法律理由，而是法律以外的理由，使得法律理由從屬於其他的理由系統當中（頁 40）。這三種條件，是一種漸進式的條件，並且為分離命題（The Separation Thesis）的基本要件（頁 50-51），因此一個法實證主義者若要主張分離命題，他勢必得支持以制裁——而且純粹只有制裁——作為基礎的理由系統。相反地，「主權者的命令」，卻會破壞法律作為理由的獨立性與構成性，因此不適合作為制裁理論的根基。

哈特所謂的接受的態度，實際上是把其他的理由混淆進法律理由當中。本文論證的末尾提出了法律作為理由的各種方式，指出許多遵守法律背後的理由，實際上都是在遵守法律以外的理由，包括權威、原則、規則，以及集體的計畫與共善。這些理由要嘛是集體理由，要嘛只是提供法律以外的其他理由。

本文各章大意分述如下：

## 第一章、法律理念的蠱惑

本章第一節闡述問題意識。自從哈特提出了「內在觀點」之後，對於法律規範性的探討，逐漸從被制裁的現實可能性轉移到法律追求的理念。內在觀點意味著，法體系下的社會成員對於法律規則存在接受的態度，這種接受的態度體現為：接受者以這個規則的內容作為指引自己、批判他人的準則（頁 5）。內在觀點不同於霍姆斯（O.W. Holmes）所提出的壞人觀點，壞人觀點僅僅考慮法律制裁的後果，而不會以法律的內容作為批判他人的準則。哈特指出，社會上的大多數人，並非如同霍姆斯所指出「壞人」（bad men），而更多是「困惑者」（puzzled men），渴



doi:10.6342/NTU202400318

望著社會規則的明確指引，並且真心的認同法律規定的內容（頁 5）。內在觀點構成了法律的重要部分，因而制裁理論曲解了法律實踐的本質。在哈特的理論下，時常爭執的邊緣事例（penumbrae of uncertainty）能夠獲得有效地說明，譬如憲法、國際法，以及前法律社會的規則等。這些邊緣事例都能基於哈特所說明的法律特徵，而列入法律的一環。然而，一個問題是：哈特的理論雖然成功說明了邊緣事例，然而他卻拋棄了曾經作為核心事例（core of certainty）探討的特徵：制裁。哈特的理論並沒有辦法有效說明：如果制裁對於法律的概念以及規範性並不是必要的，那麼為何所有的法體系幾乎無一例外的選擇了制裁（頁 3）？紹爾（Frederick Schauer）指出，哈特的理論過於關注法律可能的樣態，而非實際的樣態，而這樣的方案恐怕是成問題的（頁 3）。本文將法律理念視為法律規範性來源的觀點稱作法律真誠主義（Legal Sincerism）。根據這樣的觀點，法律必須真誠地接受法律的內容，這種接受體現在對他人的批判性，如果只是像壞人一般，只考慮到自己的利益，狡詐地服從法律，而不以法律的內容作為衡量他人的準則，那麼這不屬於接受的態度，也不屬於法律真誠主義的一環（頁 5）。

## 第二章、基於差異製造的理由論

理由是一個能夠「解釋」與「指引」行動的**事實**，而要達成這樣的功能，必須具備**反事實**（counterfactual）的概念，一個最常被使用的反事實概念，是反事實概念的正規解釋（canonical explanation），並且，以正規解釋理解的反事實概念，是正規的差異製造（canonical difference-making）：

**正規的差異製造**

一個事實$P$正規地製造差異$Q$，若且唯若，

（1）若$P$則$Q$，且

（2）若$\neg P$則$\neg Q$。（頁 11）

本文將這樣的差異製造，視為因果關係（causal relation）（頁 11-14）。王一奇指出，當我們用一個事實去解釋行動時，實際上是透過兩種不同的路徑去與行動產生反事實關聯，而只有一種路徑才能達到真正的解釋，這樣的解釋路徑稱作「理由事實」（reason-fact），或稱「理由」（reason）。



doi:10.6342/NTU202400318

**理由事實**

$R$是一個理由事實，若且唯若，

（1）$R$是一個行動$\varphi$到差異$\psi$的因果關係，

（2）差異$\psi$存在一個特徵$F$，且

（3）特徵$F$是期望（favourable）或不期望（unfavourable）的態度（attitude）。

（頁 20）

而另一種路徑被稱作「提供理由的事實」（reason-giving fact）：

**提供理由的事實**

$G$是一個提供理由$R: \varphi \rightarrow \psi$的事實，若且唯若，

（1）$G$導致了存在特徵$F$的差異$\psi$，且

（2）$G \neq \varphi$。（頁 15）

譬如，當我們說「你不應該喝家釀白蘭地」，這時候存在兩種解釋：

甲、家釀白蘭地含有甲醇。

乙、喝家釀白蘭地會損害你的健康。

一般而言，甲解釋了乙的存在，而乙解釋了為何「我不應該喝家釀白蘭地」，因此在王一奇的理論下：甲只是「提供理由的事實」，乙才是真正的「理由事實」。甲解釋了理由為什麼存在，而理由才真正解釋了「為什麼我應該／不應該做出這個行動」（頁 15-6）。若繪製成因果圖，則表現如下（頁 16）：

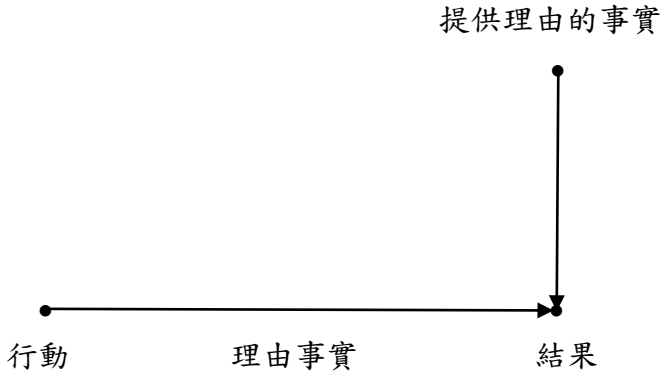


圖 2.3.1

從圖 2.3.1 來看，行動、結果與提供理由的事實形成了衝突模型（collider model），其中，提供理由的事實作為理由存在的背景條件（background conditions）（頁 16）。因此，這裡將關注放在理由本身。



doi:10.6342/NTU202400318

王一奇另一項重要的觀點，即是集體理由與個體理由的區分。集體理由是多數人的行動所導致的反事實關係，而個體理由即是單一行動所導致的反事實關係。本文歸納出集體理由具有三項基本的特點：

**集體性（Collectiveness）**：行動者的數量必須複數。
**秩序性（Orderliness）**：行動者的行動必須形成一個秩序。
**批判性（Orderliness）**：由於需要他人的配合才能達成反事實結果，因此集體理由對他人會產生批判性。（頁 22-23）

以搬石頭為例，由於石頭的重量無法一人負荷，因此要成功搬運石頭，必須仰賴多數人。並且，搬運石頭至少必須一起發力，否則石頭無法成功搬運。而當有人無法配合時，我們就會對無法配合的人產生批判。在這樣的意義下，正義、道德，以及哈特的內在觀點，都是集體理由的一環。

集體理由跟個體理由都會是我們行動的實際考慮。譬如，當我們決定禮讓行人時，我們有可能基於維護交通秩序，或者純粹避免受罰。問題是，當集體理由實踐到社會的時候，勢必要透過具體的個體來實踐，然而，由於集體理由與個體理由之間存在三種縫隙，使得集體理由**穩定且持續的實踐**會發生困難：

**解釋縫隙（Explanatory Gap）**：對於每一個個體而言，個體的單一行動都無法達成差異結果，以至於集體理由對於個體行動的解釋力不足。
**邏輯縫隙（Logical Gap）**：某些集體理由的形成，會觸發態度相反的個體理由。
**理智縫隙（Intellectual Gap）**：理性計畫者的設想，與實際個體配合的能力與意願有所差距。（頁 25-26）

為了解決這個縫隙，社會選擇了制裁作為解決方案，而基於制裁的理由本身屬於個體理由。之所以制裁能夠填補這樣的縫隙。是因為制裁能透過**交錯制裁**的機制，來模擬出集體理由的樣態。這樣的模擬具備以下的條件（頁 31）：

1. 當社群中的人不遵守規則時，存在一個個體$p_i$發動制裁。
2. 當$p_i$不發動制裁時，存在另一個個體$p_{i+1}$對$p_i$發動制裁。
3. $p_1, \dots, p_n$的制裁不能形成一個迴圈。
4. $n \geq 4$。

而這樣的交錯制裁機制，在最小的情形，可以如下圖呈現（頁 32）：



doi:10.6342/NTU202400318

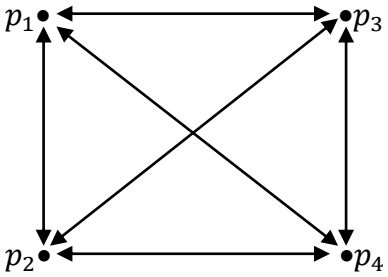


圖 2.4.1

每一個方向的箭號，都代表一個行動理由。當一個規範系統的理由機制能夠滿足圖 2.4.1 時，這個規範系統就能夠穩定且持續的實踐，並且不仰賴於其他的理由系統，此時我們就會說這個規範系統存在獨立性（independence）。而法律基本上就是這樣交錯制裁的機制。執法者透過交錯制裁，使法律能夠穩定且持續的實施於社會（頁 33）。此外，由於每個法律人對於法律的詮釋，都必須考量其他法律人對於這樣詮釋的制裁可能性，因而制裁也決定了法律概念的總體樣貌。

## 第三章、法律作為主權者的命令

奧斯丁對於法律的一般性陳述是：「法律是主權者的命令。」其中，法律是命令的一種，而命令意味著：「若不遵守我的期許，我將給予惡害。」且當我不遵守命令，因而會遭致惡害時，我就存在一個義務（頁 36-37）。因此，奧斯丁對於義務的一般陳述是：義務來自於制裁。然而，紹爾卻認為制裁實際上是區別道德義務與法律義務的關鍵，道德義務並不來源於制裁（頁 35-36）。因此，本章批評紹爾對於奧斯丁的解讀，指出奧斯丁區別道德義務與法律義務的關鍵，是在於發布命令的主體，而非是制裁。對於奧斯丁而言，正是制裁，是一切義務的來源，法律與實證道德的差別僅在於：前者是主權者對人的命令，而後者作為人對人的命令（頁 35）。

不過，在晚期的版本中，奧斯丁還另外指出了實證法與實證道德的差別：實證法的制裁是基於深思熟慮過後的行動，而實證道德的制裁則是基於情緒、性格的直接反應（頁 37）。因此，本章首先歸納出奧斯丁區分實證法與實證道德的兩個關鍵：

**主權（sovereign）：**主權作為一個領地的制裁終極來源，對內必須有最高性，對外則有獨立性。

**思慮性（deliberativity）：**制裁的設定必須經由深思熟慮，而非是性格、情緒的直接反應。（頁 37-38）



doi:10.6342/NTU202400318

除此之外，法律作為理由的方式是透過命令，因此，本章接著探討主權、命令、思慮性的三種概念，在法律理由當中能夠扮演的角色。

### 主權是必要的嗎？

本文認為，主權的存在會有害於法律的規範性。因為主權作為一切制裁的根源，那麼主權者決定制裁的理由，就不再會是法律理由。此時法律系統就會破壞掉圖 2.4.1 的圖式，也就同時破壞掉了獨立性（頁 40）。

### 命令可以構成理由嗎？

本文區分了提供理由（providing reasons）與構成理由（constituting reasons）兩種法律作為理由的方式：

**提供理由（Providing a Reason）**

一個行動$G$給予了理由$p \to q$，若且唯若：

1. $G$並不在$p \to q$的模型內，且
2. $G$與$p \to q$具有因果上的關聯。

**構成理由（Constituting a Reason）**

一個行動$G$構成了理由$p \to q$，若且唯若：

1. $G$在$p \to q$的模型內
2. $G \neq p$（頁 43）。

提供理由是透過行動與理由之間的因果聯繫，來使一段**既存的理由**進入到行動者的行為思慮當中。構成理由則是透過行動直接的創造**全新的理由**，精確來說，是讓原先認識論上的因果聯繫$p \to G$，變成實踐的反事實差異，進而創造全新的理由。根據這樣的分析，命令沒有這樣的功能，因為命令本身透過觸發（triggering），來提供一段既存的理由聯繫。譬如，當我遵守女友的命令，並不是因為女友的命令創造了新的理由，而是女友的命令觸發了一段本來就存在的理由關係：「如果我不遵守女友的命令，那麼……」是原先與女友的關係構成了理由，而非女友的命令構成了理由。法律若要成為一個不同於其他理由的新理由系統，他就必須是構成性的（Constitutiveness），而不能單單是提供一個既存的理由，否則只會是其他理由的變體（頁 47-48）。

### 分離命題的三種要件

分離命題（The Seperation Thesis），是法實證主義者的基本主張。一般而言，分離命題可以概述為：法律不會因為違反道德而喪失其效力。本文以基於差異製



doi:10.6342/NTU202400318

造的理由論，將分離命題重構為：

**分離命題**

法律理由並不以其他理由作為該理由成立的背景條件（頁 48-49）。

本文至此分析了三種要素：

**構成性（Constitutiveness）**：法律不能是提供理由，必須是構成理由，否則只是其他理由的變體。

**思慮性（Deliberativity）**：法律的構成必須：

1. 具有權力：不能有其他的制裁者共同決定制裁的成立。
2. 比較其他存在或潛在的選擇：不能只是基於情緒與性格的反應。

**獨立性（Independence）**：法律的構成必須形成 Figure 2.4.1 的圖式，否則法律理由的成立會依附在其他理由。（頁 50）

本文透過因果模型，來分析這三種概念在因果模型當中的意義，並對應到法理學當中的分離命題的概念。首先，分離命題要成立，法律理由就必須先是一個理由。因此構成性是分離命題的最基本條件。再者，思慮性要求消除一切的背景條件，使得理由的構成是基於理由，而非基於原因，思慮性的存在是區別實證法與實證道德的關鍵，這對應到了奧斯丁晚期版本的區分。最後，獨立性是讓法律理由成為一個獨立自主規範系統的關鍵要素，法律必須存在獨立性，才能讓法律的完全治理成為可能，否則，法律只會是其他理由實踐的工具，同時也會破壞分離命題。因此，這三個概念都是分立命題的必要條件，本章將分離命題的三個條件與國家治理階段對應如下（頁 52）：

| 分離命題的條件 | 國家治理階段 |
|---|---|
| 構成性 | 道德治理階段（Morality） |
| 思慮性 | 法制階段（Rule by law） |
| 獨立性 | 法治階段（Rule of law） |



doi:10.6342/NTU202400318

## 第四章、論法律作為理由的各種形式

本章開始指出，過往法理學家所論證的法律作為理由的形式，實際上都不是法律獨立構成理由的方式。

### 論權威

首先是權威。在通常的語意中，權威意指「一個知道很多知識的主體」。譬如物理學權威意味著，權威者知道很多相關的物理學知識，而醫學權威則意味著，權威者知道很多相關的醫學知識。在實踐的領域，實踐權威意味著：權威者知道很多理由（頁 58）。本文以拉茲（Joseph Raz）對於實踐權威的分析指出，我們遵守權威的理由，無關於權威指示的理由：權威指示的理由實際上觸發了我們遵守權威的理由。因此權威的理由是提供理由，而非構成理由（頁 56-58）。另外，當我們在遵守權威時，我們會依據遵守權威的理由來選擇權威，但在法律領域當中，我們實際上根本沒有這樣的選擇空間，也就是說，權威模式大多不是法律規範性的來源。在法律當中，最接近權威的模式似乎就是「行政指導」，但很顯然的，大多數的法律都不是「行政指導」，這似乎顯示了法律並不打算以這種方式提供理由（頁 58-59）。

### 論原則

原則實際上帶有幾項特性：首先，原則並不一定直接的顯示在法律條文當中（頁 61）；再者，即便顯示在條文當中，原則也不直接的指示法律效果（頁 60）；並且，原則也常帶有批判性的特質。最後一項特性顯示：原則本身屬於集體理由，而不是法律體系內的理由（頁 62）。法律人透過反覆地實踐，將原則逐漸帶入了法律體系當中，而最終構成行動導致制裁的反事實聯繫（頁 62）。就此而言，在法律人真的適用了原則之前，原則並不存在於法律理由當中（頁 63），而將原則建立起可操作的法律規則，常常是法律釋義學的根本任務（頁 65）。

### 論規則

通常，規則具有兩種意涵：立法者立下的規則與法律人實踐的規則。從法律現實主義的脈絡下當中，即是書本中的法律（law-in-book）與行動中的法律（law-in-action）。哈特的規則是在立法者意義下的，也就是說，哈特主張的是書本中的法律，並且，法律實踐者透過**接受**書本中的法律規則，形成法律的內在觀點，而這就是法律規範性的來源。所謂的接受，不同於壞人觀點，必須要將法律的內容作為批判他人的準則（頁 66）。然而，本文指出，接受作為批判他人的準則，實



doi:10.6342/NTU202400318

際上是將法律以外的集體理由，混淆進作為個體理由的法律當中。我們遵守法律的理由可以是多元的，也可以完全無關法律的理由，只是我們的行動恰好的吻合了法律。此外，書本中的法律，如同權威一般，只能夠觸發，而不能構成理由。因此法律構成理由的方式，只能透過行動中的法律——也就是實際的制裁。

另一方面，法律現實主義常常將法律推向一種極端的論點：法律規則並不存在。本文透過帕爾（Judea Pearl）對於因果模型的見解，來對規則的存在進行辯護。本文認為，法律作為一種因果關係，它的存在是**認識論**的，而非形而上學的。在因果模型當中，我們是先有理論判讀經驗資料，因而對經驗資料產生因果的理解。而在法律的領域，這種理論就是法律理論（頁 70-71）。

### 論法律真誠主義

法律真誠主義（Legal Sincerism）基本上就是主張集體理由是法律規範性的來源；相反的，本文主張的是一種「法律狡詐主義」（Legal Deviousism），也就是主張以個體理由作為法律規範性的來源（頁 72）。本文以菲尼斯（John Finnis）與莎皮羅（Scott Shapiro）來作為法律真誠主義的代表。菲尼斯主張，法理學家的任務就是要抓住法律概念的核心事例，而在法律規範性的領域當中，內在觀點的核心事例即是：指導與批判自己和他人的能力，而這種能力只有共善（Common Good）才能達成。因此，菲尼斯主張：共善即是法律規範性的來源（頁 73-74）。而莎皮羅則以法律作為一種社會計畫（Social Planning）來作為法律規範性的來源，這種社會計畫也是一種集體反事實，並且，法律是透過社會計畫來實踐道德目標（頁 74-75）。

本文主張，基於集體理由與個體理由之間的實踐縫隙，無論是共善或是集體計畫基本上難以穩定且持續的實踐。法律基於它的功能，必然是要求人民必須狡詐的對待它。霍姆斯的壞人觀點，基本上就來自這樣的隱喻（頁 76）。法律的功能若從一個穩定的共同體去考察，則容易將其他的理由混淆進法律理由當中。我們必須從一種例外去思考：當一個共同體已經覆滅，政治的激化與對立，讓不安全感與不信任感瀰漫了整個社群，人民再也找不到共通的標準作為互信與溝通的基礎。此時唯一能夠重新維繫這個共同體的方式，就是強而有力的制裁，也只有這樣，人民才會認真的考量自己的利益，從而回歸社會秩序，這也就是法律出現的理由（頁 77）。



doi:10.6342/NTU202400318

# Contents





doi:10.6342/NTU202400318

# 1. Obsession with Legal Ideal

## 1.1 So, Was Austin Wrong After All?

I aim to support a simple theory:

> *If we do not obey the law, we may be punished; therefore, we* ought to *obey the law.*

Yet, such a straightforward theory becomes controversial when articulated in philosophical terms: *legal normativity arises from sanctions*. This is a concept originally traced back to John Austin, who defined an obligation as an imperative statement backed by sanctions that would be imposed on those who disobeyed the command, and a legal obligation as an imperative statement made by a sovereign or a determinate authoritative entity. As far as Austin was concerned, sanctions were not only a source but also the only source of obligations[1]. The distinction between a legal obligation and a moral obligation lies in the issuer of the demand: the former comes from a sovereign or a determinate authoritative entity, while the latter originates from ordinary humans or God. Thus, the presence of sanctions gives the law its normative force to bind individuals, and the avoidance of punishment serves as a compelling reason for us to obey the law, underpinning a sanction-centred theory of normativity.

Austin's analysis of obligations stands out for its simplicity; however, its drawback is that it is overly simplistic. This simplistic image of obligation faced significant

[1] Some legal philosophers, such as Brian Bix (2006:6-7) and Frederick Schauer (2010: 14-16), have argued that Austin's project focuses exclusively on *legal obligation*, rather than *obligation simpliciter*. In this context, sanctions are viewed as essential components specific to legal obligations, distinguishing them from other types of obligations. Turning back to Austin's work, however, he clearly defined a command as: (All *italics* are in the original text.)

> 'If you express or intimate a wish that I shall do or forbear from some act, and if you will visit me with an evil in case I comply not with your wish, the *expression* or *intimation* of your wish is a *command*.' (Austin, 1995: 21)

and the relation between the command and the law as:

> '[T]he term *command* comprises the term *law*.' (Austin, 1995: 21)

Furthermore, Austin delineated the relationship between a command and duty, explaining:

> 'Being liable to evil from you if I comply not with a wish which you signify, I am *bound* or *obliged* by your command, or I lie under a *duty* to obey it.' (Austin, 1995: 22)

From these definitions, we can deduce that Austin's analysis centres on the general concepts of obligation and command, rather than exclusively on these concepts within the legal sphere. The project in this thesis follows a similar trajectory. Later in the thesis, I will defend the ambition of reducing the obligation to sanction and explain why the distorted version proposed by Bix and Schauer is doomed to fail.



doi:10.6342/NTU202400318

criticism when H.L.A. Hart published his magnum opus, *The Concept of Law* (1994). Hart highlighted the deficiency of the sanction-centred theory with the gunman scenario: a person threatened by a gunman for money may feel *obliged* but does not have an actual *obligation* to comply. It shows that Austin's theory conflates being *obliged* with having an *obligation*. While the former refers to a compelled action often taken reluctantly, the latter serves as a benchmark or guiding principle for puzzled individuals, allowing them to assess behaviours and critique non-compliance, including their own. Furthermore, some laws, even without sanctions, are still recognised as valid. Examples include laws that, rather than imposing sanctions, bestow powers of legislation, adjudication, and inauguration. These examples also encompass international and constitutional laws, frequently characterised as laws without sanctions. Hence, it's insufficient to lump all these diverse types of laws together under the same sanction-based model that applies to criminal law.

Hart's critique of Austin has since dominated contemporary jurisprudential thought, leading to the emergence of numerous anti-sanction theories of legal normativity. Although the diversity of these theories is so dazzling that the conflicts among them are not milder than what they are fighting against, all of these can be roughly axiomatised[2] as a proposition: the normativity of law is not something from, but rather, beyond sanctions. This proposition causes the sanction theory to be viewed as a public enemy. Neil MacCormick, for example, referred to the sanction-based theory of normativity as 'one of the perennial and persistent fallacies in legal philosophy' (MacCormick, 1973: 101). In more recent times, Scott Shapiro described such conflation of sanctions with normativity as 'myopic in that they ignore or obscure one of the motivations that people might have for obeying the law' because '[sanctions] cannot account for the intelligibility of legal practice and discourse' (Shapiro, 2006: 1158). Further, some philosophers, like Leslie Green, have even endeavoured to expel sanctions from legal systems. They argued that a system of norms depending solely on sanctions cannot be considered 'a system of law' (Green, 2008: 1049) and that sanctions are not essential to certain legal systems, such as international and constitutional laws, as well as other power-conferring statutes mentioned earlier. Evidently, the idea of law without

[2] Schauer has repeatedly referred to Hart's critique of Austin as an *axiom* of jurisprudence (Schauer, 2010: 1&7). Here is an analogy borrowed from him.



doi:10.6342/NTU202400318

sanctions has become somewhat axiomatic in the realm of legal philosophy.

Such an axiom, however, has led legal philosophy to an eccentric theory, one that is far removed from the real world where sanctions are ubiquitous. The concept of law without sanctions, as a way for Hart to deal with *penumbrae of uncertainty* about what the law is, has made actual inroads upon what was once the *core of certainty*: sanctions. A descriptive theory, in which Hart allegedly engaged, must at least account for the common phenomena of explicandum. However, there would remain an explanatory gap for Hart as to why sanctions are ubiquitous in almost all legal systems if those are neither essential to the concept of law nor genuinely important accounts of legal normativity. In Frederick Schauer's critique of Hart, he argues that it would be problematic to be over-preoccupied with 'what could be the case rather than on what is the case' (Schauer, 2010: 11) if his accounts are purely descriptive. Meanwhile, it would distort and confuse the authentic source of legal normativity to downplay the importance of sanctions. As far as Hart was concerned, social rules, rather than habits and threats, were the entities that possessed real reason-giving power, because only the former had the internal point of view, which is used for stipulating how we should act and for criticizing oneself and others when failing to conform to the rules. It is still unclear, however, how the internal point of view infuses rules with normative force, as well as the role it plays in our practical deliberations.

Some legal philosophers, such as Joseph Raz, Andrei Marmor, David Enoch, and others, endeavoured to explore the reason-giving nature of social rules. They aimed to develop a theory of normativity without sanctions, seeking to shed light on the ambiguities left by Hart. In contrast, others defended Hart's project, asserting that the internal point of view was not designed to account for the reason-giving mechanism of legal practice. Instead, its goal was to 'explain the intelligibility of the activity,' (Shapiro, 2006: 1167) something Austin failed to achieve. However, without full clarity on the deliberations of the actors in their actions, which includes the reason-giving mechanism of the law, the explanation of such activity remains elusive. Furthermore, once the reason-giving mechanism of legal practice and our practical deliberations are fully understood, we could also better examine the programmes of Joseph Raz and others in greater depth.

Taking into account the aforementioned problems with the legal theories at hand, this



doi:10.6342/NTU202400318

thesis will return to a primary question once believed to be resolved: *Was Austin really wrong after all?* Such a challenge has previously been raised in full by Schauer (2010) and will be followed and developed in this thesis. To answer this question, this thesis will delve deeper into the theory of reasons, focusing on *what is the case*, rather than *what could be the case*, of our practical deliberations. That is, pursuits of ideals like goodness, virtue, and justice, which are merely illusions and too varied to grasp, will be set aside. The focus will shift solely to:

1. How we actually deliberate on our actions, and
2. How we actually evaluate states of affairs or actions based on certain facts.

to reconstruct a theory of reasons. I will introduce a difference-making-based theory of reasons, concentrating on our actual practical deliberations, as, to support sanction-centred legal theory. This theory of reasons has been well-developed in Taiwan, first proposed by Linton Wang (2015), subsequently followed and integrated into legal philosophy by Peng-Hsiang Wang (2015), and further explored by Peng-Hsiang Wang & Yun-Chien Chang (2019).

## 1.2 Legal Sincerism

The methodology of this thesis follows the so-called 'descriptive theory', first proposed by Hart in his Postscript of *The Concept of Law* to describe his study. It was also Hart, however, who shifted the focus of legal theory towards the ideal of legal normativity. Hart posited that there are two distinct perspectives on the law and any social rule: an 'external point of view', which comes from observers who do not *accept* the rules, and an 'internal point of view', held by participants who *accept* the rules and use them as guides to conduct, to evaluate behaviours, and to critique non-compliance[3]. For Hart, merely avoiding sanctions cannot serve as a valid reason without the *acceptance* of rules even if considering the avoidance of sanctions is a kind of practical view containing belief, practice, and hermeneutic attitude (Shapiro 2006: 1158-1159). Rules serve as reasons, instead, only if participants agree to their content and view the rules as their criteria. Rather than being like the 'bad man' who plans meticulously and accounts carefully to avoid punishment, Hart argued that most people in society

[3] For the specific meaning of the internal point of view, see Shapiro (2006).



doi:10.6342/NTU202400318

resemble the 'puzzled man' who simply wonders what the law is and feels predisposed to comply with it (Hart 1994: 40). Therefore, the law is positive, well-meaning, and must be treated, once accepted, with sincerity as reasons for actions, regardless of the motivations behind their acceptance[4].

Let us refer to such a claim that laws only serve as reasons if submitted sincerely to *the content of laws* as the 'legal sincerism.' Legal sincerism posits that the law serves as a reason through its acceptance, an acceptance that manifests in *the criticism of those who do not follow the rules*. If a participant is guided solely by their own actions, without any critique of others, then such a perspective does not constitute acceptance, nor is it what Hart would describe as an internal point of view. Instead, it merely reflects the selfish perspective of a 'bad man' who considers only their own interests. This theory, however, does not exclude the possibility that the law may be treated deviously like what a bad man does, but persists firmly that it is not the way for the law to constitute reasons. So preoccupied is this claim with the ideal of legal normativity that it overlooks a pressing reality: *if legal normativity must be a sincere reason for our choice, then why are sanctions ubiquitous in the law?* This oversight undermines Hart's descriptive theory. Furthermore, the introduction of the 'internal point of view' leads to a paradoxical question: which point of view must legal philosophers adopt to accurately capture such an internal point of view in a descriptive theory?

Such paradoxicality of Hart has been pursued and critiqued relentlessly by Ronald Dworkin, who focused totally on the internal point of view and discarded the external one. For legal professionals, Dworkin argued, not only does legal subsumption remain controversial, but *the law itself* also does. To clarify such disputes, Dworkin distinguished between the proposition of law and the ground of law. The former refers to the statements and claims about what the laws command, prohibit, allow, exempt, entitle or deprive, while the latter deals with the fact determining whether a proposition of law is true or false (Dworkin, 1986: 4). Sometimes, of course, legal professionals simply disagree and argue about the proposition of law–about whether the statute really claims that people ought to do something–but agree about the ground of law–about

[4] In Hart's work, he pointed out that people's acceptance of the rules may be based on many different considerations: 'calculations of long-term interest; disinterested interest in others; an unreflecting inherited or traditional attitude; or the mere wish to do as others do.' (Hart, 1994: 203)



doi:10.6342/NTU202400318

whether the statute passed by Congress is a law. More often than not, however, legal professionals disagree and argue not only about the proposition of law but also about its ground–what constitutes applicable laws. The law is not a plain fact written in history books like codes; rather, it is flexible and can be interpreted to conform to specific political and moral principles. In other words, there is no external point of view that provides a 'meta' platform allowing us to distance ourselves from the secular world and analyse the law, much like how Archimedes used mathematics to explore the earth (Dworkin, 2006: 141). Instead, what we have are internal points of view that draw from political and moral standpoints to shape a reality that includes the law, which, in turn, influences us too. For Dworkin, the perception of *law itself* carries with it a fidelity to political and moral principles. Therefore, the *law itself* is much more positive and well-meaning, whether it serves as a reason for action or not.

The law, as far as Dworkin is concerned, is not a natural kind term with an ingrained essence that can be referred to; instead, it is something more like democracy, liberty, equality, and other political kind terms. When politicians or political philosophers define what democracy is, what they are doing is not capturing the reality of political practice, nor describing the usage of this word, but rather making a political proposal and claim: what democracy should be? When we question whether a political system, such as China's, is democratic or not, the best response should never be, 'Yes, it is, because that is how we use the word "democracy".' Such an answer would surely be disappointing[5]. In practice, then, the question is not 'what the law is' but rather 'what the law should be'. The notion of a purely descriptive legal theory is impossible.

Jurisprudence, as often described by Dworkin, plays a more active role in the practice of law: it seeks a proper interpretation. Dworkin termed such a task *philosophical*, and anyone involved in the law would participate in this philosophical endeavour. 'Lawyers are always philosophers,' Dworkin wrote, 'because jurisprudence is part of any lawyer's account of what the law is, even when the jurisprudence is undistinguished and mechanical.' (Dworkin, 1986: 380) Shapiro even called Dworkinian lawyers 'philosopher lawyers.' (Shapiro, 2011: 307) However, is it not exaggerated? Are there really that many people with the authority to interpret the law?

[5] This example is from Dworkin (2006: 283-284).



doi:10.6342/NTU202400318

Just as Dworkin asserted, in dogmatic theology, interpretation serves to make political proposals and claims. Different interpretations of doctrine are often seen as sources of reformation, revolution, or, if unsuccessful, heresy. Many Reformers, such as Martin Luther, advanced their political movements through distinct interpretations of the scriptures. For these political movements to gain traction, interpreters must either possess political prestige or hold authoritative power, and those without such prestige are left with no choice but to follow. In law schools, for example, it is unrealistic to expect law students to have the authority to interpret the law, and in a court of law, similarly, it is also obligative for judges to follow the precedent or superior court. Students who interpret it differently from their professors are likely to receive low grades, and judges who interpret arbitrarily without following precedent or superior court might face warnings, personnel transfers, or even imprisonment, as their sanctions. In law schools, only prestigious jurists, while in a court of law, only the Supreme Court, have the authority of legal interpretation. The internal point of view that Dworkin espouses seems to be a narrow perspective shared only among prominent legal scholars, Supreme Court justices, and Dworkin himself. For the majority of people, their true internal points of view are simply to abide by the rules as interpreted by these legal authorities, who are also the source of legal knowledge.

The question of who is qualified to interpret the law touches on profound institutional and political concerns[6]. When interpreting the law, we must consider other interpreters who might impose sanctions for what they deem to be incorrect interpretations. The present form of the law therefore depends on the judges' collective conception, rather than on an individual idea, of what the law ought to be. The sanctions for people's commitments, for officials' enforcements, and for judges' interpretations, form together the ecology of law. Those who ignore others' interpretations are thus just caught up in naive wishful thinking. In order to know what the law is, we must know how the judges interact with one another, and how jurists build their authority among all the others. If sanctions towards actions and interpretations can establish a stable order, much like *finding an equilibrium in a game*, then it becomes possible to describe this state from

[6] For example, Scott Shapiro argued that who is qualified to interpret the law and how much authority one can interpret is related to the arrangement of an overall legal planning. Regarding Shapiro's legal planning theory, please refer to Shapiro (2011).



doi:10.6342/NTU202400318

outside the interpretative process.

This thesis will not address the inherent value of the law or the ideals it pursues. In other words, I will not discuss what makes the law worthy of obedience or which concepts should be encompassed within the law's content, whether the law is viewed as an authority, a form of social planning, or an embodiment of social justice. Such discussions often entail numerous unprovable explanations, descriptions, and assumptions so much that deviate from practical realities, with legal sincerism being a prime example. A clear counterexample is this: if the law or the ideals it pursues are really so good, why can people not form a spontaneous order without the need for sanctions? This thesis therefore does not presuppose the sincerity of law-keeping, nor consider the goodness of the law, but rather explores humans' practical deliberations and the position of the law in the process of such deliberations.

In my view, the ultimate test of a descriptive theory of normativity's success is to determine if people, when practising in accordance with such a description, can foster a spontaneous and stable order. Clearly, legal sincerism would fall short in this regard. This thesis seeks to craft a descriptive theory bearing these criteria in mind, emphasising the description of our practical deliberation and the dynamics among jurists–mainly the relationships based on sanctions–and, in doing so, defining what the law is.

# 2. Difference-Making-Based Theory of Reasons

## 2.1 Reasons and Normativity

A *normative proposition* is a sentence containing a deontic modal, such as 'ought to', 'permitted', 'forbidden', or 'exempted'. This makes it different from a *factual proposition*, which simply states a *fact*. For example:

**Example 2.1.1**
(a) I ought to obey the law.
(b) I obey the law.

(a) is a normative proposition, while (b) is a factual proposition. How, then, do we understand normative propositions? Moreover, under what circumstances and



doi:10.6342/NTU202400318

conditions do we make normative claims? A state of affairs or a condition $F$ allows us to claim that we ought to $P$, if and only if the state of affairs or the condition $F$ is normative. 'Reasons', in recent discussions, have often been used for answering both of these questions: if $A$ ought to do $P$, then there is a reason $R$ for $A$ to do $P$. A fact $F$ is normative, if and only if the fact $F$ can serve as a reason for us to do $P$[7]. Thus, 'I ought to obey the law' implies 'I have reasons to obey the law', which is equivalent to saying that there exists a normative fact $F$ that provides reasons for me to obey the law. However, there remain questions as to:

1. How a normative fact justifies an action, and if this relationship can be made clear;
2. What kind of fact can provide a reason, and what are the limitations on the qualification of this fact?

In question 1, T. M. Scanlon asserted that '[p]ractical reasons for actions are facts that count *in favour of* these actions.' (T. M. Scanlon, 1998: 3, *italics* added) it must be asked, however, how do these facts count in favour of actions? Scanlon argued that it is not possible to delve deeper into the nature of this supportive relationship. In other words, '"[b]y providing a reason for it" seems to be the only answer.' (T. M. Scanlon, 1998: 17) As he said:

> ⋯I am quite willing to accept that "being a reason for" is an unanalysable, normative, hence non-natural relation. …(Scanlon, 1998: 11)

Such a statement is called *reason primitivism*. According to this statement, to provide support is tantamount to providing a reason, which is the end of the matter. John Broome, on the other hand, has posited a contrasting perspective, arguing that reason is the *explanation* for action, but what is the relationship between explanation and action?

[7] How strong is the normative claim that 'the fact $F$ can serve as a reason for us to do $P$'? When a fact $F$ presents, it necessarily implies we must have a reason to do $P$. However, whether or not we act on that reason is our choice, and even if we decide to do P, other reasons might influence us. In essence, being eligible to serve as a reason does not necessarily compel action. According to John Broome's (2007: 163-164) distinction between strong and weak normativity, weak normativity means that 'Necessarily, if rationality requires $N$ to $P$, there is a reason for $N$ to $P$' (Broome, 2004: 164, symbols changed); 'the fact $F$ can serve as a reason for us to do $P$' has thus only a weak degree of normativity, but this thesis tries to capture all degrees of normativity by claiming weak normativity.



doi:10.6342/NTU202400318

Broome has provided an ambiguous explanation:

> A perfect reason therefore need not be a unique canonical reason. Suppose you ought not to drink home-made grappa because it damages your health. The fact that home-made grappa damages your health explains why you ought not to drink it, so it is a perfect reason for you not to drink it. Another explanation of why you ought not to drink home-made grappa is that it contains methyl alcohol. This is not a rival explanation; it is consistent with the first. So a perfect reason for you not to drink homemade grappa is that it contains methyl alcohol. Now we have two distinct perfect reasons for you not to drink home-made grappa. This is confusing, but only because the individuation of explanations is confusing. We need not fuss about it. (Broome, 2004: 35)

Besides the fact that Broome has not offered a clear definition of 'explanation', he has also contended that there can be more than one explanation for a normative proposition. In the mentioned statement, Broome has suggested that the normative proposition 'You ought not to drink home-made grappa' can be explained in two ways: 'Home-made grappa damages your health' and 'Home-made grappa contains methyl alcohol'. These two explanations, which are consistent with each other, are presented side by side as compelling reasons to support the normative proposition. However, Linton Wang (2015, pp. 109-110) critiques this, asserting that these two facts relate to the normative proposition in entirely different manners, and only one offers a genuine explanation. Without a clear understanding of 'explanation', these two types of facts could easily be conflated.

Clarifying the meaning of 'explanation' relates to the nature of the fact itself, that is, to question 2: *what kind of fact can provide a reason, and what are the limitations on the qualification of this fact?* Returning back to Example 2.1.1, it is clear that factual proposition (a) 'I obey the law' cannot justify normative proposition (b) 'I ought to obey the law', for to do so would be to violate *Hume's law* 'a normative proposition cannot be derived from factual proposition'. However, when we justify a normative proposition, such as in Broome's case, we always use a factual proposition to support a normative proposition. In Example 2.1.1, however, why can't (b) support (a)? Perhaps it is not so much a matter of 'a normative proposition cannot be derived from factual proposition' as it is a matter of 'a fact's necessary quality for an explanation'. Going



doi:10.6342/NTU202400318

back to the example provided by Broome, if the reason for 'you ought not to drink home-made grappa' is that 'you do not drink home-made grappa', it would not be a good reason. But 'home-made grappa damages your health' and 'Home-made grappa contains methyl alcohol', although Linton Wang thinks they are different in nature, are at least acceptable reasons. So, where is the acceptable intuition? What is the difference in nature between these three facts? This is the first issue that must be dealt with before I clarify what is the reason.

## 2.2 What is the Explanation?

According to Linton Wang, an explanation is achieved by pointing out the relevant difference-making. In general terms, when we explain a fact, explanans $P$ explains explanandum $Q$ if and only if $P$ 'causes' $Q$. The so-called cause, to define the concept in the broadest and most commonly used way, is the definition of the *counterfactual*. Linton Wang uses 'canonical difference-making' to capture the simplest and most central counterfactual concept:

**Definition 2.2.1 (Canonical Difference-Making)**
A fact $P$ canonically makes difference $Q$, if and only if
(1) if $P$ then $Q$, and
(2) if $\neg P$ then $\neg Q$

The difference must be understood through the lens of *counterfactual* imaginings. Specifically, we need to compare the outcomes of two scenarios: the first involves an outcome $Q$ that stems from a situation $P$, which conforms to reality; the second involves an outcome $\neg Q$ resulting from a situation $\neg P$, which is contrary to reality. This disparity between $Q$ and $\neg Q$ constitutes the *difference*. Nonetheless, Linton Wang posits that not all difference-making relies on counterfactual properties. He provides the following two examples (Linton Wang, 2015: 116):

**Example 2.2.1 (Causal Difference-Making)**
While we agree that the power plant's electricity supply *causally makes a difference* to the building having power, it does not mean that if the power plant stops supplying electricity, the building will be without power. This is because the building might have a backup power system.



doi:10.6342/NTU202400318

**Example 2.2.2 (Disposition to difference-making)**

While we agree that being bitten by a venomous snake *disposes to make a difference* to the bitten individual dying, it does not mean that one will necessarily die from a venomous snake bite. This is because those bitten can be administered anti-venom serum.

But what exactly are 'causal difference-making' and 'disposition to difference-making'? While Linton Wang does not offer an exact definition of the two in his works, it is evident from the cases presented that 'causal difference-making' lacks the necessity of counterfactuals, whereas 'disposition to difference-making' lacks the sufficiency of counterfactuals. Linton Wang categorises these two types as 'genuine difference-making'. This form of difference-making possesses more intricate conditions compared to regular difference-making, leading to numerous exceptions in counterfactual hypotheses. C. B. Martin (1994: 7) has contended that counterfactual conditionals cannot be employed to account for ambiguous dispositional associations in natural language, and several scholars have since endorsed this view, dubbing the argument that counterfactual conditionals are intrinsically linked to dispositional associations a 'conditional fallacy'[8]. Consequently, Linton Wang prefers to label 'substantive difference-making' as 'conditional knowledge' instead of 'counterfactual knowledge' on the one hand, yet he posits that explanations are feasible as they must pinpoint the relevant difference-making on the other hand. The predominant method to elucidate difference-making is through counterfactual dependency, potentially explaining his hesitancy. While this paper does not delve into this segment of the discourse due to its primary focus and length constraints, for subsequent debates, I will attempt to present a counterfactual-dependent interpretation of genuine difference-making, drawing from Judea Pearl's causal model.

Pearl contends that when we analyse clauses denoting causal relations, such as '..., so...' and 'to cause', we encounter several challenges. Firstly, these clauses suggest a rule-like necessity between two events, but simultaneously, we recognise their inherent contingency. We might assert, for instance, 'reckless driving causes accidents' or state 'you will fail the course owing to your laziness' (Pearl, 2009: 1), fully aware that the antecedents generally make the consequences more probable, rather than definitively

[8] For a discussion of the conditional fallacy, see Bonevac et al (2006).



doi:10.6342/NTU202400318

certain. Secondly, these clauses possess so many exceptions that they struggle to function effectively within a logic anchored in necessity. This is largely because our natural language cannot embody all contextual circumstances and background conditions, or speakers even remain oblivious to them (Pearl, 2009: 2). Given this complexity, Pearl initially favours Bayesian probability over strict logic, emphasising data correlations. Nevertheless, he asserts that causal analysis grounded in probability remains tenable. Pearl categorises (2009: 29) our cognitive approach to causality into three hierarchical levels: *prediction*, *intervention*, and *counterfactual*. 'Prediction' involves examining the data correlation between an antecedent and its consequent, and deducing the conditional probability that should the antecedent occur, the consequent will follow. 'Intervention' requires the active alteration of the antecedent's value, with predictions made about the likelihood of a specific consequent value emerging. 'Counterfactual', on the other hand, demands a retrospective look at the point an antecedent in a conditional phrase appeared, modifying this antecedent to deviate from the actual events, and then forecasting the conditional probability of a subsequent alteration in value, following conventional developmental trajectories.

If we formalise these three stages of cognition, at the stage of *prediction* we get:

$$P(X|Y) = \frac{P(X,Y)}{P(Y)} \tag{2.1}$$

which is interpreted in English as 'Given that we know $Y$, the probability of $X$ is…' Subsequently, at the stage of *intervention*, Pearl (2009: 70) introduces a new operator $do(X)$, which is called the *do-operator*, to express the modification of value $X$. With this operator, we use:

$$P(X = x|do(Y = y)) \tag{2.2}$$

to compute 'If we *change* the value of $Y$ to $y$, what *will be* the probability that $X$ takes the value $x$?', where lowercase letters indicate specific values of their corresponding uppercase variables. According to Pearl (2009: 73-83), the do-operator can be reduced to the level of *prediction* under some specific circumstances, which he calls *identifiability*. Finally, at the stage of *counterfactual*, we obtain both of the following formulae:



doi:10.6342/NTU202400318

$$PN = P\left(X_{Y=y'} = x' \middle| Y = y, X = x\right) \quad (2.3)$$

and

$$PS = P\left(X_{Y=y} = x \middle| Y = y', X = x'\right) \quad (2.4)$$

in which $X_{Y=y'} = x'$ is referred to as $P(X = x'|do(Y = y'))$ and $X_{Y=y} = x$ is referred to as $P(X = x|do(Y = y))$. In the above two formulae, $PN$ represents the *Probability of Necessity* from $Y$ to $X$, which is interpreted in English as 'Now, we know $X$ takes the value $x$ while $Y$ takes the value $y$, but if we *had changed* change the value of $Y$ to $y'$, what *would have been* the probability that $X$ takes the value $x'$?' $PS$, on the other hand, represents the *Probability of Sufficiency* from $Y$ to $X$ and is interpreted as 'Now, we know $X$ takes the value $x'$ while $Y$ takes the value $y'$, but if we *had changed* the value of $Y$ to $y$, what *would have been* the probability that $X$ takes the value $x$?' To encapsulate the aforementioned three stages of cognition, we can envision a ladder-like hierarchy, where high-level calculations are both built upon and encompass low-level calculations. Consequently, counterfactuals emerge as the most crucial core in causal models.

In Example 2.2.1, there is a high probability of sufficiency but a low probability of necessity for the supply of electricity from a power plant that leads to the availability of electricity in a building. In contrast, in Example 2.2.2, there is a low probability of sufficiency but a high probability of necessity for death resulting from a fatal snake bite. Even so, we can still identify counterfactual dependencies in both cases, but these counterfactual dependencies have different types and degrees of quantitative relationships.

If we assume that Pearl's approach is feasible, then Linton Wang's point of view can be defended to some extent, and we can also derive theoretical support from Pearl's approach. Therefore, in my future discussions, I will not distinguish between the various types of difference-making but will refer to them as causal relations, where the concept of causality is in line with Pearl's theory. We can conclude that counterfactuals are crucial for explaining facts. Next, I will show why counterfactual dependency is a necessary notion for a reason explanation, which is a matter of factual qualification.



doi:10.6342/NTU202400318

## 2.3 The Reason Explanation

### 2.3.1 Explanatory Inversion

In contrast to the explanations of factual propositions, which seek to understand why an event occurred, explanations of normative propositions necessitate a kind of explanatory inversion. Specifically, they involve a search for what outcomes arise from the regulated facts (Linton Wang, 2015: 112). Using Broome's quote as an illustration, when I advise 'you shouldn't drink home-made grappa', a given reason is that 'drinking home-made grappa can seriously damage your health'. This reason highlights the outcomes of the regulated action. Conversely, it also provides the counterfactual assumption that 'if you hadn't drunk home-made grappa, then you wouldn't be seriously harming your health'. This explanation points to a relevant difference-making, and therefore serves as a reason why 'you shouldn't drink home-made grappa'. In addition, Broome pointed out that another perfect reason is that 'home-made grappa contains methyl alcohol'. However, this fact does not imply or suggest any counterfactual outcomes. As a result, it does not offer an actual explanation of the regulated facts. Linton Wang opines that this is not a 'reason-fact' but rather a 'reason-giving fact'. The reason-giving fact does not explain the normative proposition directly; instead, it elucidates how the difference in outcomes materialises. Linton Wang defines 'reason-fact' and 'reason-giving fact' as follows (2015, p. 111):

**Definition 2.3.1 (Reason-fact)**

$R$ is a *reason-fact* for one to $\varphi$, if and only if

(1) $R$ is a causal relation from $\varphi$ to difference $\psi$, and

(2) Difference $\psi$ has feature $F$.

**Definition 2.3.2 (Reason-giving fact)**

$G$ is a *reason-giving fact* for one to $\varphi$, if and only if

(1) $G$ is a cause of difference $\psi$ having feature $F$, and

(2) $G \neq \varphi$.

In Broome's case, the statement 'home-made grappa contains methyl alcohol' essentially explains 'why drinking home-made grappa can seriously damage your health'. Put another way, it offers a causal explanation for the outcome, rather than a



doi:10.6342/NTU202400318

reason explanation for the action. However, it is intuitively perceived as explaining the action because the explanation for the outcome hints at a causal path from the regulated fact to the outcome. As a result, the reason-giving fact suggests the presence of a reason. Nevertheless, the reason-giving fact is ineffective in elucidating actions and normative propositions when the context is ambiguous, as demonstrated in the subsequent example:

**Example 2.3.1**

Alison: 'Why are you here?'

Timothy: 'Because you're here!'

In Example 2.3.1, Timothy's response serves as a reason-giving fact, alluding to the presence of an action-produced difference connected to 'Alison is here'. However, due to the limited contextual understanding of Timothy's response, we are uncertain about the potential consequences if Timothy decides to join Alison. It's possible that both Alison and Timothy were vying for a limited item, with only one remaining. If Timothy had not approached Alison, the item might have been claimed. Alternatively, Alison might be somewhere she should not, and if Timothy had not intervened, she could have faced consequences, and so forth. In instances where the context is unclear, the distinction between 'reason' and 'reason-giving fact' becomes evident. Moreover, the feasibility of using the facts of a causal path as reasons is contingent upon the establishment of the background conditions of said path. Specifically, whether its reason-giving facts hold, which Linton Wang (2015: 122) terms the *circumstance sensitivity of reason-giving fact*. The causal graph in Figure 2.3.1 below illustrates the relationship between '*action*', '*reason-fact*', and '*reason-giving fact*':

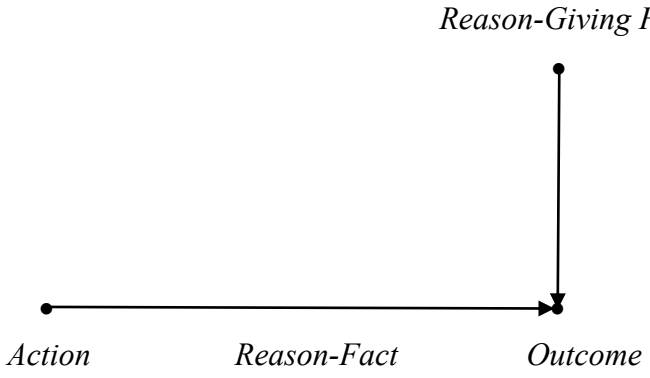


Figure 2.3.1



doi:10.6342/NTU202400318

In Figure 2.3.1, the action, outcome, and reason-giving fact together form what is termed a 'collider model'[9]. Within this model, both the action and the reason-giving fact jointly precipitate the outcome. As a result, the reason-giving fact establishes a background condition, determining whether the action can instigate a difference. Furthermore, in instances where the outcome of the action is known, a pseudo-relationship emerges between the action and the reason-giving fact: both consistently manifest and vanish in tandem. This dynamic elucidates our usual inclination to view the reason-giving fact as the motive behind the action. Additionally, Linton Wang's observation highlighting the distinction between the explanations of normative and non-normative facts also unveils *the dual nature of action explanation*. Action explanations proceed in two distinct directions: firstly, the *causal explanation*, which delves into the physiological precursors of actions–such as organic diseases or physiological mechanisms–and secondly, the *reason explanation*, focusing on the purposes or intents behind actions.

Both causal and reason explanations appeal to counterfactuals as the foundation of their arguments. Even the reason-giving fact implies a counterfactual trajectory inherent in the reasons. Consequently, facts devoid of counterfactual considerations do not inherently qualify as explanations. Revisiting Example 2.1.1, why can't 'I obey the law' serve as a reason for 'I should obey the law'? This is because 'I obey the law' simply describes the current state of affairs articulated in the normative proposition; it does not indicate the consequences that might arise if the law were disregarded. A compelling reason aiming to uphold the status quo would point out the bad consequences of not maintaining said status quo, which is also inherently counterfactual. Turning back to Broome's case, neither 'you don't drink home-made grappa' nor 'home-made grappa contains methyl alcohol' directly indicate the counterfactual outcomes of the action. However, the latter at least implies the presence of a reason, whereas the former provides no such implication. As a result, 'you don't drink home-made grappa' lacks the necessary quality for an explanation.

Reasons serve as an explanatory inversion, highlighting their significant properties in terms of deliberation and action. Agents can leverage this to decide whether to perform an action to either prompt or prevent a particular outcome. Linton Wang (2015,

[9] For the collider model, see Pearl (2009: 16-17).



doi:10.6342/NTU202400318

p. 114) borrows the term *practical deliberative usefulness* from DeRose (2010: 25) to describe this concept:

**Definition 2.3.3 (Practical Deliberative Usefulness)**
We refer to a fact as *practical deliberative usefulness*, if and only if we can use the fact to deliberate whether to make the antecedent true as a way of promoting or resisting the consequent being made true.

In addition to explaining the action, the Practical Deliberative Usefulness of reasons reveals a secondary function: guiding the action. Explaining the action and guiding the action entail different metaphysical implications. Explaining the action is a past-oriented assumption, necessitating a review of past actions, imagining different action options from reality, and evaluating potential outcomes. Conversely, guiding the action is a future-oriented idea. It involves an interventional assessment, namely, disregarding the causes underpinning the action and focusing solely on its consequences. From a Bayesian probability perspective, the explanation of action aligns with the counterfactual level of (2.3)-(2.4), whereas guidance pertains to the interventional level of (2.2). However, when an action is considered as an intervention, and the underlying reasons are eliminated, the potential confounder between action and result is also removed. Therefore, a relationship as per Definition 2.2.1 between action and result is established. It can be thus asserted that, whether guiding or explaining action, reliance on the concept of counterfactuals is essential.

Owing to the Practical Deliberative Usefulness, reasons can be positive, verifiable, and shared. Similarly to how we adopt a scientific approach in searching for the causes of events, the quest for reasons behind actions can also be scientific. By utilising our own data and methods, we can generate or stimulate reasons to achieve the same result, a process known as reproducibility in science. However, Definition 2.3.1 still falls short regarding practical deliberative usefulness, a topic we will explore in the next section.

### 2.3.2 The Interpretation of Attitude

Although Linton Wang distinguishes between causal and reason explanations, he fails to address how viewing reasons solely as objective causal paths can overlook the crucial intentional orientation behind those reasons. Consider the following example:



doi:10.6342/NTU202400318

**Example 2.3.2**

If you don't use contraception during sexual activity, there is a risk of becoming pregnant.

Example 2.3.2 illustrates a causal path where 'sexual intercourse without contraception leads to pregnancy'. According to practical deliberative usefulness, this fact should inform the decision about whether an action should be taken. However, does this causal pathway lead to the conclusion that 'contraception should be used' or 'contraception should not be used'? Alone, this causal path does not lead to a normative conclusion. If one of the partners does not want to get pregnant, then the causal path becomes a reason why 'contraception should be used'; but if both partners want to get pregnant, then the causal path becomes a reason why 'contraception should not be used'. In other words, the attitudes we hold towards the outcomes are also important in explaining behaviour. It is these attitudes that transform a reasoned action into an intentional action.

Furthermore, the lack of an attitude of purpose does not perfectly explain the action. Consider the following case:

**Example 2.3.3**

(1) Because Timothy left the pesticide on the kitchen table, Alison drank it.
(2) Timothy did not want Alison to drink the pesticide.

In the events of Example 2.3.3, we obtain a chain of causality:

(3) If Timothy leaves the pesticide on the kitchen table, there is a risk that someone, like Alison, might drink it.

What explains Timothy leaving the pesticide on the table? Referring to (1) and (2), it seems that the incident in Example 2.3.3 was an accident. But can the causal path in (3) explain why Timothy left the pesticide on the table? Usually, we would consider (3) invalid because a successful explanation of (3) implies that Timothy 'wanted' Alison to drink the pesticide, which contradicts the premise of (2). Alternatively, Timothy may not have been aware of the causal chain, or he might have thought the chances of it happening were so low that he disregarded it. In either case, (3) is unlikely to serve as a reason. That is, the subjective state of the agent must be taken into account when explaining the actions or justifying the normative propositions. Therefore, I have



doi:10.6342/NTU202400318

extended definition 2.3.1

**Definition 2.3.4**

$R$ is a *reason-fact* for an agent to $\varphi$, if and only if

(1) $R$ is a causal relation from $\varphi$ to difference $\psi$, and

(2) Difference $\psi$ has feature $F$.

(3) The agent has a favourable/unfavourable attitude towards feature $F$.

However, the introduction of (3) raises new questions: What categories do favourable and unfavourable apply? Furthermore, what attributes, when linked to attitudes, imbue these reasons with normative force? The first question concerns the categories of these attributes. Each area of evaluation has a corresponding dichotomous criterion, letting actions within a certain category ultimately serve this purpose. In morality, this dichotomy is good and evil; in aesthetics, beautiful and ugly; and in economics, advantageous and disadvantageous. Therefore, in which domain of evaluation does the dichotomy of favourable and unfavourable apply? The second question delves into the sources of the validity of reasons: what is the metaphysical basis for the normative validity of these reasons? Linton Wang applies his 'difference-making-based theory of reasons' to the three dominant views of reasoning: desire-based, value-based, and subjectivity-based theory of reasons. It becomes apparent that Wang is not proposing a new theory of reasons to be placed alongside these three, but rather proposing a general analysis of reasons. Nevertheless, if attitude is an essential explanatory component of the difference-based theory of reasons, does its inclusion compromise its neutrality? This paper does not intend to delve into the substantive controversies of the theory of reasons. However, any concrete definition of attitudes may challenge the neutrality of the difference-making-based theory of reasons. Therefore, we should consider attitudes as a 'way of categorisation': attitudes are merely a set of various kinds of difference-making, rather than a distinct or additional entity. I categorise the various kinds of difference-making as follows:

**Definition 2.3.5 (Attitude)**

**Favourable** We consider feature $F$ belonging to difference $\psi$ caused by $\varphi$ to be *favourable* $F = e$, if and only if $\varphi$ is obligatory.

**Unfavourable** We consider feature $F$ belonging to difference $\psi$ caused by $\varphi$ to be *unfavourable* $F = s$, if and only if $\varphi$ is impermissible.



doi:10.6342/NTU202400318

According to Definition 2.3.5, an attitude is one of the features $F$ resulting from difference $\psi$ ($F = e/F = s$), and attitudes labelled as favourable or unfavourable are abstract, neutral categorisations. Therefore, in response to the first question, an attitude can manifest as a difference in moral, aesthetic, or economic criteria. Consequently, attitudes such as 'good', 'beautiful', and 'advantageous' are categorised as 'favourable', while attitudes such as 'evil', 'ugly', and 'disadvantageous' are categorised as 'unfavourable'. Regarding the second question, attitudes can align with psychological motives, satisfy certain values, or conform to certain good dispositions of humans. As we delve deeper unceasingly into the reasons behind these attitudes, we eventually reach an 'ultimate reason' that defies further questioning. The result of the difference highlighted by this ultimate reason will only be the distinction between attitudes that are favourable or unfavourable.

Returning to Example 2.3.2, if one partner does not wish to become pregnant, the underlying reason might be economic hardship. The reason behind avoiding financial difficulties might be to prevent a lack of food, and the reason for not wanting to be deprived of food might be to avoid misery. If we continue to ask, 'Why not wish to be miserable?', we find that such a question becomes difficult to answer. At this point, we can no longer identify a verifiable difference, and the question transitions into a philosophical realm. Philosophers have provided various answers to this, such as the violation of human dignity, the natural human inclination towards happiness and avoidance of pain, and so on. Regardless of the ultimate difference, we can at least be certain that there exist differences that make certain behaviours worth pursuing or avoiding–and, crucially, that the explanation of reasons is incomplete without considering attitudes.

## 2.4 Individual Reasons and Collective Reasons

### 2.4.1 The Gap Between Individual reasons and Collective Reasons

With the above difference-making-based theory of reasons, Linton Wang also distinguishes between individual reasons and collective reasons for action. Individual reasons for action focus on the outcomes of differences caused by individual actors, while collective reasons for action focus on the outcomes of differences caused by institutional arrangements–that is, by the orderly actions of a group. Linton Wang uses



doi:10.6342/NTU202400318

the example of game theory to illustrate how collective action can lead to differences in outcomes:

**Example 2.4.1**

Mother Alison bought the chocolate cake, which is the favourite food of her sons, Tim and Tom. However, it bothered her that the two sons fight over the cake every time because Alison is unable to cut the cake into two pieces of exactly the same size, and the two brothers always fight over the larger piece. (Linton Wang, 2015: 134)

In game theory, the solution to this kind of problem is that one child cuts the cake first, and then the other child chooses a piece first. This arrangement is because the one who cuts the cake, knowing they will choose last, will strive to cut it evenly to avoid losing out. For the difference-making-based theory of reasons, game theory provides a causal path; by adopting this scheme, each child will get a piece of the same size, and Alison holds a 'favourable' attitude towards the outcome of 'each child getting the same size piece of cake.' Thus, the game theory scheme itself provides a reason for Alison to adopt this strategy. However, the existence of this causal path depends on the context. For example, if Tim and Tom prefer competing over sharing, they may not be compatible with the game-theoretic solution. Or, they may lack sufficient understanding to carry out the solution; therefore, this causal chain would not hold at all, and Alison would not have a reason to adopt the solution.

Collective reasons for action can range from being as small as a game theory model to being as large as a set of political systems or ideologies. To summarise, there are three important characteristics to consider:

**Collectiveness**

Collective reasons for action must rely on the simultaneous or successive actions of *several* individuals to lead to a difference, whereas a individual action alone would not result in such a difference. Therefore, it requires that the smallest set of people in the collective capable of leading to a difference be compatible with it and able to put it into practice.

**Orderliness**

To achieve a particular counterfactual difference, everyone's actions must adhere to



doi:10.6342/NTU202400318

a specific order and must not be chaotic. Even in the simplest cases–such as everyone pushing a rock together–each person must be coordinated to exert force at a specific moment in time to achieve the objective of moving the rock.

**Criticalness**

Since the achievement of a collective counterfactual outcome requires the cooperation of others, highlighting others' significance. This significance can be elucidated through the following logical sequence:

(1) The collective outcome $t$ is deemed favourable.
(2) My action $p$ combined with $Q$'s action $q$ collectively causes the outcome $t$.
(3) Consequently, both actions $p$ and $q$ are obligatory.

This, therefore, explains why collective reasons for action–such as morality, political ideologies, and social systems–become a criterion for criticising others: because the critic necessarily relies on others to achieve their favourable outcome.

In our practical deliberation, both individual and collective reasons may be taken into account. For example, when we ride a motorbike and see a pedestrian crossing the road, there are two kinds of considerations regarding whether we should give way or not: one is the consideration of individual reason, namely, that giving way to pedestrians may cost us our travelling time; the other is the consideration of collective reason, namely, that giving way to pedestrians can create good traffic order. In our daily actions, we may be guided and evaluated according to both reasons. A thorny question, however, is: *can such rationales, when put into operation in society, operate stably*? A crucial aspect of this conundrum is whether individuals can act in accordance with the collective order planned by the rational planner. As far as I am concerned, there exist at least three gaps in an individual's capacity to act in accordance with collective reasons for action:

**Explanatory Gap**

For any single participant, their solitary action is incapable of causing a collective counterfactual outcome; thus, the collective reason falls short of providing an adequate explanation at the individual level.

The Explanatory Gap is most frequently exemplified in scholarly discourse by the



doi:10.6342/NTU202400318

paradox of voting. This paradox highlights the exceedingly slim probability that an individual rational voter's ballot will alter the election result. Consequently, the reason for casting a vote with the aim of societal transformation appears insufficient. Furthermore, considering additional burdens associated with voting, such as the expense and effort of travelling to the polling station, it becomes apparent that, in the majority of instances, there is scarcely any compelling reason for a rational citizen to partake in voting.

Beerbohm contends that ascribing collective responsibility solely on the basis of individual difference-making infers the absence of a counterfactual link between each person's actions and the resultant outcome. Consider, for instance, a referendum in a city of 9,999 inhabitants, where a mere 5,000 votes are required to ratify a bill desired by the populace. Under these circumstances, the vote of any single citizen lacks a *sufficient* counterfactual relationship with the bill's enactment. Moreover, once the tally of affirmative votes surpasses 5,001, each additional vote becomes redundant; in other words, none of these votes individually maintain a *necessary* counterfactual relationship with the bill's enactment.

To address this problem, Beerbohm adopts the 'mode of complicity' as a means to grapple with issues related to collective counterfactual outcomes (Beerbohm, 2012: 63-64). In considering the responsibilities associated with complicity, it is imperative to view the 'affair' as a unit achieved collectively, and to assess the participants of the 'affair' based on the counterfactual outcomes it engenders. Thus, an individual need not have a counterfactual relation to either the genesis or the resultant state of the affair. As long as an individual is a participant in the affair, they are obliged to accept responsibility for its consequent outcomes. For example, an individual who supports a political party in an election inherently assumes public responsibility for the outcomes of that party's decisions. We have formulated the mode of complicity into the following principle of collective responsibility:

**Definition 2.4.1 (Principle of Collective Responsibility)**

Let the collective affair $P$ be a consistent set of actions. If $M$'s action $m$ belongs to $P$, and collective affair $P$ causes difference $Q$, then $M$'s action $m$ has causal effect on $Q$.



doi:10.6342/NTU202400318

The principle of Collective Responsibility can derive the rules of individual-collective transformation proposed by Linton Wang (2015:136). Wang elucidates these transformation rules to demonstrate the process by which collective rationales are converted into individual ones. Specifically, he examines the deontic terms 'prohibited' and 'permitted', exploring their respective applications:

**Theorem 2.4.1 (Transformation Rule of Individual-Collective Prohibited Reasons)** For any individual $P$, $R$ is a reason why all persons are prohibited from doing $\varphi$. If and only if, $R$ is a reason why individual $P$ is prohibited from doing $\varphi$.

**Theorem 2.4.2 (Transformation Rule of Individual-Collective Permitted Reasons)** For any individual $P$, $R$ is a reason why all persons are permitted to do $\varphi$. If and only if, $R$ is a reason why individual $P$ is permitted to do $\varphi$.

Building on the two aforementioned principles, Peng-Hsiang Wang (2015: 348) proposes the rule of individual-collective transformation to address the deontic term 'ought':

**Theorem 2.4.3 (Transformation Rule of Individual-Collective Obligatory Reasons)** For any individual $P$, $R$ is a reason why all persons ought to do $\varphi$, if and only if, $R$ is a reason why individual $P$ ought to do $\varphi$.

Theorems 2.4.1 through 2.4.3 also explain the source of the *criticalness* of collective reason: any action or inaction may be consistent with a particular collective state of affairs, which may or may not lead to a specific counterfactual outcome. That is why we always criticise others on the basis of collective reason.

Nevertheless, the Explanatory Gap still highlights a blatant fact. When individuals engage in collective actions that do not directly lead to a counterfactual outcome, it diminishes their sense of personal responsibility. For example, in a firing squad execution, distributing the responsibility among multiple shooters lessens each individual's sense of guilt. Similarly, when we buy meat instead of slaughtering an animal ourselves, our guilt is reduced. This phenomenon makes it challenging to establish or eliminate a collective affair unless there is a successful effort to foster a sense of communal consciousness.



doi:10.6342/NTU202400318

**Logical Gap**

The formation of certain collective reasons triggers individual reasons with opposite attitudes.

The logical gap illustrates that the establishment of a collective affair has the potential to be a background condition for a particular individual reason that contradicts collective one. We might consider the following two reasons:

**Example 2.4.4**

(1) Printing a lot of counterfeit money will make me wealthy.
(2) If everyone prints a lot of counterfeit money, it will lead to the depreciation of the currency.

In Example 2.4.4, (1) is individual reasons for action and (2) is collective reasons for action. Given (1), I would have a perfect reason for printing a lot of counterfeit money. Nevertheless, to achieve the outcome of becoming wealthy, it is imperative that the currency retains a certain level of purchasing power; therefore, in light of (2), I must also advocate that people should not print a lot of counterfeit money and that the monetary system should be respected. It becomes apparent that individual reason (1) for obligation implies collective reason (2) for prohibition. According to Theorem 2.4.1, collective reason (2) for prohibition can infer the individual reason 'I should not print a lot of counterfeit money'. There consequently arises a contradiction: the individual reason suggests I should print a significant quantity of counterfeit money, but the collective reason dictates that I should not do so.

Typically, a Kantian argument would hold that individual reason (1) lacks validity, as its universalisation into collective reasons for action would result in a contradiction with (2), per the individual-collective transformation principle of reasons. Indeed, individual reason (1) merely represents a private reason that lacks universality. Nonetheless, we cannot entirely dismiss the existence of (1), for there are undoubtedly individuals in society who act upon this rationale. To deny this reason is to fail in accounting for such actions, or worse, to provide an erroneous explanation for them.

Logical gaps render it impossible for certain ethical systems to cultivate a spontaneous and stable order. For instance, in corporate ethics, if most corporations



doi:10.6342/NTU202400318

comply with ethical standards without coercive intervention, this compliance will paradoxically benefit non-compliant corporations, leading to the eventual demise of those that adhere to ethical norms.

**Intellectual Gap**

There is a gap between the rational planner's assumptions and the actual individual's ability and willingness to match them.

In considering collective reasons for action, the acceptance and implementation of a certain plan by participants often depend on their own envisagement of the plan's outcomes and their perceptions of other participants' implementation capacities, such as willingness to cooperate, level of trust, and capability to contribute, amongst other factors. Particularly, collective reasons conflict with pluralism–our community, it seems, is unable to maintain a consistent attitude towards the collective creation of a counterfactual difference. Consider Example 2.4.1: Tim, who enjoys competing, holds an 'unfavourable' attitude towards equal cake distribution, while Tom, typically on the losing end, holds a 'favourable' attitude towards it. Alison, believing that competition is unfair and puts Tom at a constant disadvantage, holds a 'favourable' attitude towards equal distribution. The question then arises: how do we find a collective goal that is favourable for all members, thereby making them execute the collective scheme spontaneously? Linton Wang, therefore, questioned whether there truly exists a reason to implement an action plan conceived from the standpoint of a rational participant. While the outcome of a plan predicated on collective reasons is contemplated rationally, a rationally devised action plan may be oblivious to the actual circumstances; thus, there exists an intellectual gap between the outcome envisaged by a rational participant and the individual consideration of accepting the plan's outcome (Linton Wang, 2015: 135).

To encapsulate, the conflict between collective reasons and individual reasons occurs when:

(1) collective action reduces or eliminates the responsibility of individual participation,

(2) the collective reasons for action may engender logically related yet opposite attitudes embedded by individual reasons for action, with the former serving as a background condition for the latter,



doi:10.6342/NTU202400318

(3) different individuals do not share the same attitude towards the differences engendered by collective reasons for action, and

(4) rational planners often lack adequate knowledge of background conditions.

When running in society, these factors contribute to the uncertainty of whether collective reasons for action alone can establish a spontaneous order, except in a small-scale commonwealth.

### 2.4.2 Sanctions as a Way to Fill the Gap

In order to bridge such a gap, the collective reasons for action must be, in enforcement, replaced effectively with individual reasons for action, and the most effective way to achieve this is through sanctions. Consider the example in Example 2.4.1, where Tim and Tom do not accept the game theory solution, or they fail to understand it sufficiently to implement it. Consequently, Alison establishes a few rules:

**Rule 1:** Should Tim fail to cut the cake first, he shall receive a tap on the palm.

**Rule 2:** Should Tom fail to take the cake first, he shall receive a tap on the palm.

Although Tim and Tom neither agree with nor understand the game-theoretic solution, they are compelled to abide by its rules out of fear of being tapped on the palm. Consequently, they must take into account the individual reasons for action, resulting in these reasons superseding the collective ones. It can be argued that the sanction ensures the stability of the collective counterfactual causal chain, which in turn facilitates the effective implementation of a policy and a public idea, thus establishing a command pattern that ranges from superior planning to subordinate sanctioning. While superior planning may involve collective reasons for action, subordinate sanctioning operates on individual reasons for action. Conceptually, however, superior planning is deemed unnecessary because the reasons for sanctions are *complete*, which is to say they are *functionally complete*.

**Definition 2.4.2 (Completeness)**

A reason form $R$ is *complete*, if and only if, all collective reasons and individual reasons for action can be replaced by the reason form $R$.

Due to the normative gap between the individual reason and the collective reason for



doi:10.6342/NTU202400318

action, no collective reason is complete. In the case of the individual reason, we will be able to recognise it as complete if it satisfies the following definition:

**Theorem 2.4.4 (Completeness of Individual reasons)**

*A individual reason is complete*, if and only if

(1) It leads to attitudes that are uniformly favourable or unfavourable among all individuals.

(2) It is capable of supplanting the collective order planned by all collective reasons and achieves the same counterfactual outcome.

Theorem 2.4.4 is a logical extension of Definition 2.4.2. Given that all reasons for action are categorised as either individual or collective, a individual reason, if complete, must be capable of functionally substituting for both individual and collective reasons for action. Point (1) addresses the substitutability of individual reasons for action, positing that a individual reason can only replace all individual reasons if it uniformly induces either favourable or unfavourable attitudes in everyone. The most direct method to achieve this uniformity is through sanctioning. By its very nature, sanctioning involves the creation of an artificially unfavourable outcome; hence, if an artificially created difference is favourable, it is, by default, not included in the definition of sanctions.

Logically, a sanction is a set of all unfavourable differences. Given an unfavourable set $S = \{s_i | i \in \mathbb{N}\}$, we denote $F \in S$ as $F = s$.

**Theorem 2.4.5 (Sanction-Based Normative States)**

**Obligatory** $\varphi$ is *obligatory* if and only if the outcome $\psi$ caused by $\neg\varphi$ incurs a sanction $F = s$.

**Permissible** $\varphi$ is *permissible* if and only if the outcome $\psi$ caused by $\varphi$ does not incur a sanction $F = \neg s$.

**Impermissible** $\varphi$ is *impermissible* if and only if, the outcome $\psi$ caused by $\varphi$ incurs a sanction $F = s$.

**Omissible** $\varphi$ is *omissible*, if and only if the outcome $\psi$ caused by $\neg\varphi$ does not incur a sanction $F = \neg s$.

Theorem 2.4.5 follows from Definition 2.3.5, making the sanction capable of deriving



doi:10.6342/NTU202400318

all possible normative states. We can thus use sanctions to justify all kinds of normative propositions.

One problem is that rewards can also achieve functional completeness. We define reward as the creation of an artificially favourable outcome and denote $F \in E$ as $F = e$ given a favourable set $E = \{e_i | i \in \mathbb{N}\}$; therefore, we can also derive the following theorem from Definition 2.3.5:

**Theorem 2.4.6 (Reward-Based Normative States)**

**Obligatory** $\varphi$ is *obligatory* if and only if the outcome $\psi$ caused by $\varphi$ incurs a reward $F = e$.

**Permissible** $\varphi$ is *permissible* if and only if the outcome $\psi$ caused by $\neg\varphi$ does not incur a reward $F = \neg e$.

**Impermissible** $\varphi$ is *impermissible* if and only if, the outcome $\psi$ caused by $\neg\varphi$ incurs a reward $F = e$.

**Omissible** $\varphi$ is *omissible*, if and only if the outcome $\psi$ caused by $\varphi$ does not incur a reward $F = \neg e$.

Do rewards, however, really have the same binding power as sanctions? At least, John Austin denies it, as he put it:

> If a law hold[s] out a *reward*, as an inducement to do some act, an eventual *right* is conferred, and not an *obligation* imposed, upon those who shall act accordingly: The *imperative* part of the law being addressed or directed to the party whom it requires to *render* the reward (Austin, 1995: 24)

For Austin, the reward-based obligation is not a real obligation, but rather, at best, a right. The question of what constitutes rights and obligations is a difficult one and is beyond the scope of this discussion. At least, reward-based normative states are so much *weaker* than sanction-based normative states that the former are insufficient to imitate the latter, but not vice versa.

Point (2) addresses the substitutability of collective reasons for action, and the concern here is the substitutability of sanctions for collective reasons. To illustrate this point, we must consider the following theorem:



doi:10.6342/NTU202400318

**Theorem 2.4.7 (Completeness of the Reason for Sanctions)**

The reasons for sanctions form a *complete normative system* $R = \{r_1, \dots, r_m\}$, where any reason $r_i \in R$ if and only if, given a set of members $P = \{p_1, \dots, p_n\}$ subject to the system, where $n \geq 4$, for any member $p_i \in P$, and let $p_i' \in P$ but $p_i' \neq p_i$, if $p_i$ does not initiate a sanction on one who break the rule $r_i$, then $p_i'$ will initiate a sanction on $p_i$.

Theorem 2.4.7 posits that the substitutability of reasons for sanctions in lieu of collective reasons for action can be secured by ensuring a specified number of members from the set engage with the sanctioning process. In this way, the reasons for sanctions are sufficient to establish a framework for spontaneous collective order, thereby ruling out the room for an explanation of collective reasons for actions.

Given a minimal normative system $R_s$ that satisfies Theorem 2.4.7, and let $R_p$ be the set of reasons for sanctions that is *isomorphic* to a particular collective reason for action $R_P$, then $R_s \cup R_p$ constitutes a *minimal normative system with completeness* for $R_P$. We can perhaps liken Theorem 2.4.7 to a battery, in that a biomimetic machine, analogous to a living creature, equipped with a battery, operates just like the creature, and that a set of reasons for sanctions, isomorphic to a collective reason, equipped with Theorem 2.4.7, functions just like that collective reason.

**Proof of Theorem 2.4.7.**

To prove Theorem 2.4.7, we may employ the *process of elimination*. Let the set of members $P = \{p_1, \dots, p_n\}$ subject to the system $R = \{r_1, \dots, r_m\}$ be variables applied to the terminology of kinship (e.g., *parents*, *children*). When $p_1$ is able to sanction $p_2$, then $p_1$ is a parent and $p_2$ is children, and so on.

*Case 1*: There exists a parent within the normative system who is not sanctioned by another. In this case, should there be a parent within the normative system who is not sanctioned by another, then the reasons for sanction in the systems must be deemed incomplete. This is because the reason behind the initial parent's sanctioning of the children must not be related to the sanctions themselves.

*Case 2*: The chains of sanctions are connected in a circular diagram. In this case, we must consider the following lemma:



doi:10.6342/NTU202400318

**Lemma 2.4.1 (The Prohibition of The Sanctioning Cycle)**

There exists a set of members $P = \{p_1, \dots, p_n\}$, let $p_i \in P$. When $p_i$ does not issue a sanction, $p_{i+1}$ issues a sanction against $p_i$, forming the normative system $R = \{r_1, \dots, r_m\}$. If $p_{i+1} = p_i$, then there is no normativity for $R$.

According to Lemma 2.4.1, the normativity of the entire system is compromised when the causal sanction chains form a cycle. This is because, for any member $p_i \in P$, their sanction on $p_{i-1}$ is a prerequisite for receiving a sanction from $p_{i+1}$. Therefore, if sanctions are the sole basis for normativity within the system, they effectively neutralise each other. This neutralisation remains constant regardless of the severity of the sanctions.[10]

*Case 3*: The chains of sanctions intertwine with each other, making us able to find any sanctioning pair $p_i \to p_j$, denoted as $p_i$ sanctions $p_j$, within the set of members $P = \{p_1, \dots, p_n\}$. In this case, each member of a set is capable of sanctioning all the other members within that set, thus creating an interwoven sanction system, as illustrated in Figure 2.4.1:

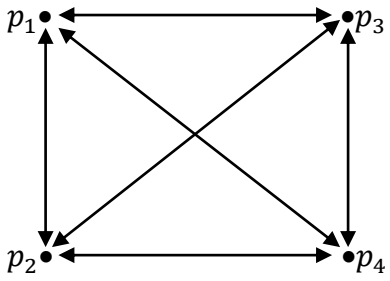


Figure 2.4.1

where each of these double arrows represents an interactive sanctioning relationship. However, in the case where $P$ has only two members, it clearly falls under *Case 2*. Subsequently, when $P$ has only three members, three types of sanctioning relationships emerge:

[10] Consider the simplest case: a set $P$ containing two members $P = \{p_1, p_2\}$. If $p_1$ does not issue a sanction, then $p_2$ will issue a sanction against $p_1$ with a magnitude 1. Conversely, if $p_2$ does not issue a sanction, then $p_1$ will sanction $p_2$ with a magnitude 2. In the absence of other reasons for initiating sanctions, it would be rational for $p_2$, to choose not to sanction $p_1$ to avoid receiving a sanction himself, even if the harm he faces is less than that faced by $p_2$. The same reasoning applies to $p_1$. Without further complicating factors, the normative system $R$ would thus collapse.



doi:10.6342/NTU202400318

**Chain:** $a \rightarrow b \rightarrow c$

**Fork:** $a \leftarrow b \rightarrow c$

**Collider:** $a \rightarrow b \leftarrow c$

In the case of the *chain*, it falls under *Case 2* if $c \rightarrow a$, or under *Case 1* if $a \rightarrow c$. For the *fork* and the *collision* cases, both fall under *Case 1*, whether $c \rightarrow a$ or $a \rightarrow c$. Therefore, this interwinding is not applicable to sets with fewer than four members. Conversely, if a set contains more than four members and the sanctions are interwound, we describe the sanctioning system as complete. □

To confirm a normative system's completeness, it is not necessary to verify that every member can sanction all others. According to Theorem 2.4.7, if there exists a minimal complete set $R$ within the system $N$, then $N$ is deemed complete. This mirrors societies governed by the rule of law, where only some individuals, such as police, prosecutors, and judges, actively enforce laws, and the majority do not engage in legal sanctions. Nevertheless, the rule of law remains constant. The stability of such a normative system, as confirmed by Theorem 2.4.7, stems from an interwoven sanctions mechanism. This mechanism ensures that there are consequences for rule-breakers, extending even to those who do not actively issue sanctions. Consequently, this system is self-sustaining and distinct from systems that rely on one-time sanctions.

Here, I am drawing an analogy between the sanctions model and law, which seems fairly intuitive. Analogizing law to a model of sanctions is understandable, given that sanctions–despite some counterexamples–are typically characteristic features of law. Such analogies have a long history, first appearing with Hobbes in 1651 and later popularised by John Austin in 1832. In fact, this thesis's account of legal obligations closely resembles Austin's. We share the view that legal obligations arise from the application of legal rules, which are conditional relationships backed by sanctions. However, this thesis takes a step further by describing them as causal relationships. Besides, there are also some legal concepts from which I differ somewhat from Austin. I will compare them in the next chapter.



doi:10.6342/NTU202400318

# 3. Law as Command of the Sovereign

## 3.1 Sanction or Sovereignty: Which Defines the Law?

As a pioneer both in analytic jurisprudence and legal positivism, Austin's theory is undoubtedly groundbreaking; through the method of positivism, Austin attempted to detach jurisprudence from the fields of philosophy and political science and make it an independent discipline of study. According to the original meaning of positivism, the construction of philosophical theories should be based on empirically verifiable facts. Therefore, legal positivism means that the research topic of jurisprudence should focus on legal phenomena that can be verified empirically, rather than metaphysical, political, or moral ideas. The *political superiors* are at the centre of Austin's focus, and Austin argues that the subject matter of jurisprudence must be the rules laid down by the political superiors, because such rules are what jurisprudence can really empirically investigate. In *The Province of Jurisprudence Determined* (1995), Austin begins by stating:

> The matter of jurisprudence is positive law: law, simply and strictly so called: or law set by political superiors to political inferiors. (Austin, 1995: 18)

Austin crawls through the various uses of the word 'law' in English. In terms of its literal meaning, 'law' refers to 'a rule laid down for the guidance of an intelligent being by an intelligent being having power over him' (Austin, 1995: 18). In this sense, law includes:

1. Law set by God to men
2. Law set by men to men

The distinction between the two aforementioned categories is based on the nature of the intelligent beings involved. The first kind is called the *law of God*, or the so-called *natural law*. In the state of nature before the establishment of man-made systems, all general laws of justice that can be discovered through reason are considered to be the law of God, such as the natural rights of equality and freedom, and so on. Different philosophers often have different views on what the connotation of the law of God is. The second kind, on the other hand, is called *human law*, which is a system of



doi:10.6342/NTU202400318

institutional rules created by intention after the establishment of man-made systems, and can be roughly subdivided into *positive law* and *positive morality*. The distinction between positive law and positive morality is also made on the basis of the nature of the intelligent beings involved. The intelligent being here is the political superior: rules made by the political superior are called *positive law*, while rules made by the non-political superior are called *positive morality*. The term 'positive' in *positive morality* is intended to distinguish it from the law of God, and to refer to moral rules that are empirically observed rather than rationally proven. In addition, the word 'law' is often extended to mean 'law of nature', which refers to the rules of natural science, but this is not the original meaning of the word 'law', it is merely a metaphorical derivation (Austin, 1995: 19). Based on the above analysis, there are five types of 'law':

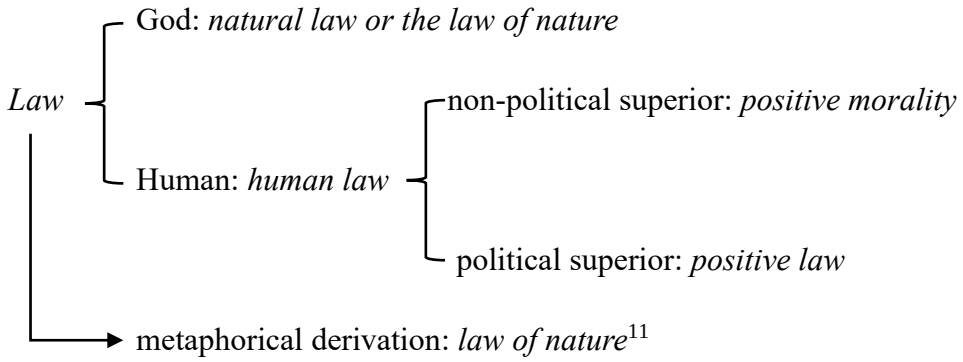


Figure 3.1.1

Apart from metaphorical derivation, the other four types of 'law' are differentiated by their subject of command. It is therefore the *subject* that makes law different from the 'morality' and the 'natural obligation'. According to Schauer (2010: 14-15), however, it is the presence of sanctions, rather than subject, that makes legal obligations different from the other types of obligations.

For Schauer, he persisted that Hart's criticism of Austin is inadequate. Hart proposed an internal point of view to counter Austin's concept of reducing legal obligations to mere empirical facts. He used the 'gunman scenario' to illustrate that our conception of

[11] In Austin's linguistic analysis, 'natural law' and 'law of nature' are, in fact, synonymous. These terms are indeed used interchangeably when referring to a system of law that is based on values intrinsic to human nature, deduced and applied independently of positive law. In a scientific context, however, the phrase 'law of nature' can also refer to the rule followed by natural phenomena. Therefore, it is used here to illustrate this point.



doi:10.6342/NTU202400318

‘obligations’ differs significantly from being sanctioned–they cannot be simply reduced to the empirical facts under threat. However, in fact, Hart’s internal viewpoint also struggled to identify the root cause: why do the actual practices of government officials create normative attitudes that guide their own behaviour and form the criteria for critiquing others? The essential issue appears to be that Hart’s perspective on legal obligations is more *deontological*, while legal philosophers like Hobbes, Austin, and Scandinavian scholars tend to hold a more *teleological* perspective. However, both of them can indeed serve as sources of obligation. Schauer noted, aligning with Kant, that normative propositions can stem from either *hypothetical* or *categorical* sources (Schauer, 2010: 15). Not only do principles, maxims, and canons act as vehicles for creating obligations, but legal rules, which are hypothetical relations between prohibited actions and legal sanctions, also serve this purpose. Thus, considering sanctions as a source of obligation is not counterintuitive. ‘An “obligation” [i]s simply the noun form of the verb “oblige”’ (Schauer, 2010: 13) On the contrary, if we do not regard sanctions as a source of legal obligations, we cannot at all explain why legal obligations are different from other types of obligations, which is, as Schauer pointed out, Austin’s sole intended contribution (Schauer, 2010: 16).

Returning to Austin’s text, however, we can see that Schauer’s reading is problematic. It is the *subject*–which Austin called *sovereign*–rather than the presence of sanctions, that makes legal obligations different from the other types of obligations. Sanctions, in all norms, should be omnipresent. Turning back to Austin’s work, however, he clearly defined a command as: (All *italics* are in the original text.)

> ‘If you express or intimate a wish that I shall do or forbear from some act, and if you will visit me with an evil in case I comply not with your wish, the *expression* or *intimation* of your wish is a *command*.’ (Austin, 1995: 21)

and the relation between the command and the law as:

> ‘[T]he term *command* comprises the term *law*.’ (Austin, 1995: 21)

Furthermore, Austin delineated the relationship between a command and duty, explaining:

> ‘Being liable to evil from you if I comply not with a wish which you signify, I am



doi:10.6342/NTU202400318

*bound* or *obliged* by your command, or I lie under a *duty* to obey it.' (Austin, 1995: 22)

From these definitions, we can deduce that Austin's analysis centres on the general concepts of obligation and command, rather than exclusively on these concepts within the legal sphere. Positive law differs from positive morality in that the former is a command of the sovereign, whereas the latter is a command of men. For Austin, ostracism, excommunication, and degradation of reputation or trust would all count as part of a general sanction. In a later version[12], Austin (1995: 20) further distinguishes positive law from positive morality by stating that positive law is an *artificially* enacted rule, while positive morality arises not from an artificial setup but rather from opinion or sentiment held or felt by men.

If we delve deeper into causal models, Austin's distinction in his later version can be understood in the following terms: The sanctions in positive law are actions emanating from a segment of a group with concentrated power, grounded in deliberate decisions. Sanctions in positive morality, in contrast, are simply acts stemming from individuals with dispersed power, based on opinions, sentiments, and dispositions that occur naturally in the environment. In causal models, there is an intrinsic difference between the 'action' and the 'act': the 'action' operates at an intervention level, while the 'act' is limited to the prediction level. This means the action establishing the sanction can eliminate complex background conditions and independently generate new reasons. This capability to create reasons independently and uniquely marks the fundamental distinction between positive law and positive morality.

Based on the preceding analysis, it is clear that Austin identifies two key methods to differentiate between positive law and positive morality: sovereignty and deliberativity in imposing sanctions.

**Sovereignty**

[12] The Province of Jurisprudence Determined underwent five editions from 1832 to 1885. However, the third (1869), fourth (1873), and fifth (1885) editions were not published as individual books. Instead, they were included in Robert Campbell's editions of Austin's Lectures on Jurisprudence. The 1995 edition, edited and published by Wilfrid E. Rumble, which is cited here, encompasses all editions of the book. The numerous passages that Campbell added to the text are marked with square brackets []. For the history of the addition, see Rumble (1995: xxv). The later edition referred to here specifically concerns the passages within these square brackets.



doi:10.6342/NTU202400318

According to Austin, sovereignty possesses two crucial attributes: *supremacy* and *independence*. Supremacy implies that within any community, there is an ultimate power, either an individual or an organisation, that holds control over all other entities. Independence indicates that this sovereign typically issues commands autonomously, without the customary obligation to adhere to the orders of others.

**Deliberativity**

Deliberativity, in this context, encompasses two essential elements: *power* and *comparativeness*. The power implies that the entity imposing the sanction must have substantial capacity to enforce an absolute sanction. The term 'absolute' signifies that the sanction's effectiveness does not depend on background conditions which are independent of the sanctioner, giving the sanctioner complete decision-making power. Comparativeness requires the sanction-imposing entity to take into account other existing or possible outcomes, rather than react directly to visible outcomes.

In terms of the difference-making-based theory of reasons, however, are both concepts necessary to explain the independence of laws that constitute reasons? Previous discourses in legal philosophy have frequently grappled with the concept of 'sovereignty'. This concept is particularly contentious because it exists outside the comprehensive framework of the law, yet significantly influences its very foundation. In the following section, I intend to delve into the critical inquiry of whether 'sovereignty' is an indispensable element for the notion of legal independence.

## 3.2 Is Sovereignty Necessary?

For Austin, a sovereign must exist within a legal society as the highest authority, accustomed not to obeying others but to issuing orders independently. Dworkin, however, criticises this view, arguing that in today's complex societies, political control does not emanate from a single entity. Instead, it arises from a multitude of compromises, cooperations, and alliances. He contends that the controlling power of law gradually emerges from community interactions, making it impossible to pinpoint a single political ruler or identify their will (Dworkin, 1967: 18-19). It is thus doubtful that the notion of sovereignty conforms to positivism.

Furthermore, if one supports legal positivism based on the difference-making-based



doi:10.6342/NTU202400318

theory of reasons, the existence of the sovereign becomes superfluous. Austin, as a legal positivist, primarily asserts that 'the law retains its validity even when it violates norms outside of legal boundaries.' In other words, legal positivism holds that the law's ability to constitute reasons should be independent. This independence is defined in the following theorem:

**Theorem 3.2.1 (Independence)**

A normative system $N$ is considered *independent*, if and only if, its necessary quality for an explanation is not subordinate to any other normative system $N'$.

Theorem 3.2.1 actually contradicts Austin's argument, primarily in the context of sovereignty: the existence of sovereignty acts as a break in legal independence. Austin's theory is best described as a command model, emphasising a hierarchy between superior planning and subordinate executing. Nevertheless, this model is arguably redundant. Theorem 2.4.7 challenges this view by asserting that sanctions are inherently complete. It suggests that the interplay of mutual sanctions among law enforcers sufficiently explains the concept of superior planning. For legislators, sanctions manifest in various forms. These include the nullification of legislative actions due to non-compliance with established processes or the compulsory enactment of legislation resulting from legislative inaction. For the judiciary, sanctions may arise from failures to adjudicate correctly, distortions of legal interpretations, or arbitrary judgments leading to repercussions such as punishment, sentencing, or impacts on professional assessments.

This framework is further reinforced by the concept of interactive sanctioning. This mechanism ensures that the reason behind every sanction initiated by a sanctioner is intrinsically linked to the sanction itself, thereby eliminating the need for sovereignty. Consequently, from this understanding, we can infer Theorem 3.2.1 from Theorem 2.4.7. This suggests that a normative system fulfilling the criterion of *completeness* also inherently possesses the characteristic of *independence*. This independence is crucial, as it underscores the self-sustaining nature of the legal framework on the basis of sanctions.

Theorem 3.2.1 can be stated more clearly through the difference-making-based theory of reasons, which gives independence a technical definition:



doi:10.6342/NTU202400318

**Theorem 3.2.2 (Independence of reasons)**

A set of reasons $R$ possesses the independence of reasons, if and only if any variable in $R$ is completely determined by other variables in $R$.

According to Austin, positive law is the command of the sovereign backed by sanctions. Consequently, the scope of $R$ encompasses the causal network linking actions to sanctions and including the sovereign as a member of $R$. Subsequently, we can follow Austin's hypothesis to prove that the setting of sovereignty is harmful to the independence of the law. We first define the following variables:

$C$ is denoted as 'the sovereign,' and $C = c$ is denoted as 'the sovereign establishes a sanction.'

$Y$ is denoted as 'The set of those who initiate sanctions,' and $Y = y$ is denoted as 'They initiate sanctions.'

$X$ is denoted as 'The set of those who do not initiate sanctions,' and $X = x$ is denoted as 'They act.'

Here, lowercase letters $\{x, y, c\}$ are used to represent the specific values of the corresponding variables $\{X = x, Y = y, C = c\}$, and

$R$ is denoted as the causal network $x \to y$ and $C \in R$.

**Proof** Theorem 3.2.2 $\nvDash C \in R$

Suppose $C \in R$. In line with Austin's interpretation of the sovereign, the sovereign is identified as the originator of all sanctions. Therefore, we position the sovereign as a child of all variables. As such, the causal path from actions to sanctions, denoted as $x \to y$, ultimately converges back to the terminal child $C$, thereby constituting the sequence $x \to y \to c$.

Considering the sovereign as the sole descendant of all variables, should the sovereign institute sanctions predicated upon reason $R'\colon c \to c'$, this reason $R'$ must necessarily lie outside the scope of $R$, which disrupts the independence established by Theorem 3.2.2. Hence, it follows that Theorem 3.2.2 does not entail $C \in R$. □

Therefore, the setting of sovereignty undermines the independence of legal normativity and makes legal normativity subordinate to other normative systems.



doi:10.6342/NTU202400318

Accepting legal positivism, which posits that law has the quality of constituting reasons independently, implies that the concept of sovereignty is not only superfluous but also detrimental, and even contravenes the essence of 'positivism'.

After excluding sovereignty, deliberativity emerges as the sole indicator differentiating law from positive morality. Within the causal model, an action $\varphi$ with deliberativity operates at the intervention level, symbolised as $do(\varphi)$. Deliberativity in law comprises two elements: power and comparativeness. Concerning power, since deliberativity functions at the intervention level, the intervening action $do(\varphi)$ that establishes the sanction cannot have parents. In other words, the sanctioner should have complete control over it. Regarding comparativeness, as per Pearl (2009: 108), actions with deliberativity require careful evaluation of existing or potential outcomes before making a decision. This means such actions are not predictable based on stimuli or motivations, as they are not stimulus-driven but based on considered deliberation of various outcomes.

In summary, while the causal pathways of empirical law are deliberately designed, empirical morality emerges organically from social interactions within a networked environment. Positive law's sanctioning actions can bypass complex interaction networks and independently forge new justifications. In contrast, positive morality's sanctioning acts result from communication and interaction. These two kinds of constituting reasons, figuratively speaking, are akin to the difference between synthesising ethanol through chemical hydration and producing it by fermenting starch.

## 3.3 Can Imperatives Constitute Reasons?

The notion of 'command' is another key point of interest in Austin's most famous proposition that 'the law is the command of the sovereign'. According to Austin, the command is to express or intimate a wish that one ought to or not to do some act, and to impose an evil against one who complies not with the wish (Austin, 1995: 21). However, Austin seems to have over-extended the concept of command. Hart contends that the term 'command' typically conveys a greater emphasis on the notion of authority rather than the capacity to inflict harm. In essence, the act of commanding does not necessarily require the imposition of sanctions. A commander's authority alone can be sufficient to elicit obedience from the person receiving the command (Hart, 1994: 20).



doi:10.6342/NTU202400318

The word ‘command’ closely implies authority, making it align too closely with the concept of law. It is thus not a sound analysis to understand law merely in terms of command. Hart argues that the conceptual overlap between law and command renders it inappropriate to define law as the ‘command.’ Instead, Hart used the term ‘imperative’ to describe a wish enforced by sanctions. However, whether the concept of law is framed in terms of a ‘command’ or an ‘imperative’, I believe both approaches risk leading to a certain kind of confusion. A fundamental question arises: Can the Imperative Serve as the Reason?

Here, we likewise adopt the term ‘imperative;’ however, this paper contends that ‘imperative’ and ‘coercive force exerting sanctions’ should be considered separately. The ‘imperative’ used here denotes sentences with an imperative mood, including ‘command,’ ‘order,’ ‘request,’ ‘plea,’ and ‘warning.’ Syntactically, reasons take a very different form from commands. Reasons are typically structured as ‘if... then...’ conditional sentences, allowing for variability in person and tense. In contrast, commands are generally limited to the present tense and are often directed towards the second person. Pragmatically, reasons, as facts, tend to be presented as ‘declarative sentences’ and have truth values. Consider Broome’s example: ‘home-made grappa damages your health’ may be false to such an extent that it does not explain the normative proposition ‘you should not drink home-made grappa’. Conversely, commands are more dynamic. They resemble ‘actions’ more than ‘propositions’ and thus do not possess truth values but rather matters of appropriateness or inappropriateness.

Separating the imperative from the reason enables us to more effectively inquire *whether the imperative can constitute a reason*. The phrase ‘constitute a reason’ is used here instead of the more common ‘provide a reason’ due to its intrinsic difference from the latter. ‘Provide a reason’ implies that the reason already exists; in other words, in nature and society, there is an established causal chain from action $\varphi$ to difference $\psi$, and the reason-giver merely prompts the agent to act by altering the agent’s cognition or by changing the background conditions of the action. Conversely, ‘constitute a reason’ suggests that the reason did not exist previously; in other words, in the nature and society, there is no pre-existing causal chain from action $\varphi$ to difference $\psi$, and the reason-giver creates this causal chain through the establishment of sanctions or



doi:10.6342/NTU202400318

rewards. In short, provided reason represents a *naturally occurring* causal chain, while a constituted reason signifies an *artificially created* one. More technically, 'providing a reason' and 'constituting a reason' are defined as follows:

**Definition 3.3.1 (Providing a Reason)**

An action $G$ *provides a reason* for $\varphi$ if and only if a reason for $\varphi$ is a causal path $R: \varphi \rightarrow \psi$, where

(1) $G$ is not in the causal path $R: \varphi \rightarrow \psi$, and

(2) There is a causal path from $G$ to the variable in $R$.

**Definition 3.3.2 (Constituting a Reason)**

An action $G$ *constitutes a reason* for $\varphi$ if and only if a reason for $\varphi$ is a causal path $R: \varphi \rightarrow \psi$, where

(1) $G$ is in the causal path $R: \varphi \rightarrow \psi$, and

(2) $G \neq \varphi$

In the definition above, if $G =$*to issue an imperative*, then how the imperative serves as a reason will depend on the relationships within the causal model. Let us examine 'providing a reason' first. Providing a reason has a more diverse and complex pattern than constituting a reason. David Enoch (2011) suggests three ways of providing a reason:

**Definition 3.3.3 (Epistemic Reason-Giving)**

The way of giving reasons is considered *epistemic*, if and only if there is a causal path from reason-giving $G$ to the receiver's action $\varphi$.

The first way is epistemic reason-giving. According to Definition 3.3.1, the reason given in this way already exists, and the provider simply informs about the existence of the reason. In other words, the act of informing by the provider is entirely independent of the reason's existence. It merely enables the person receiving the information to experience a change in perception, realising that there is a valid reason for a specific action. For instance, when someone says 'Don't smoke' and provides the reason that 'smoking is harmful to your health,' this act of giving a reason is considered epistemic. However, the harmfulness of smoking is a pre-existing reason to avoid it, and this reason remains independent of whether the state informs us about it or not. Definition



doi:10.6342/NTU202400318

3.3.3 is presented as a following causal graph:

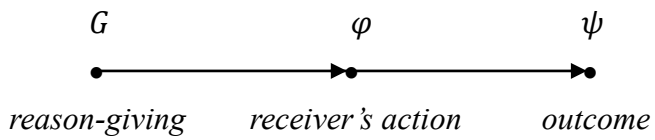


Figure 3.3.1

where the act of giving a reason $G$ causes the receiver's action $\varphi$, leading to the outcome as a difference $\psi$.

**Definition 3.3.4 (Triggering Reason-Giving)**

The way of giving reasons is considered *triggering*, if and only if the reason-giver *intervenes* and changes the value of variable $G$ in background condition $B: G \rightarrow \psi$.

The second way is triggering reason-giving. The reason, being a causal relation $R: \varphi \rightarrow \psi$, is contingent upon background conditions $B: G \rightarrow \psi$. Therefore, whether it forms a real counterfactual path or not depends on the values of these background conditions. By altering variable $G$ in $B$, the provider activates the counterfactual reason for the receiver, creating a genuine counterfactual path. For instance, consider the desire to save money as a reason not to smoke. If cigarette prices are low, there is less incentive not to smoke. However, if the government increases cigarette taxes, thereby raising its prices, it activates the counterfactual pathway 'smoking→excessive spending'. This gives civilians an additional reason to avoid smoking. Definition 3.3.4 is presented as a following causal graph:

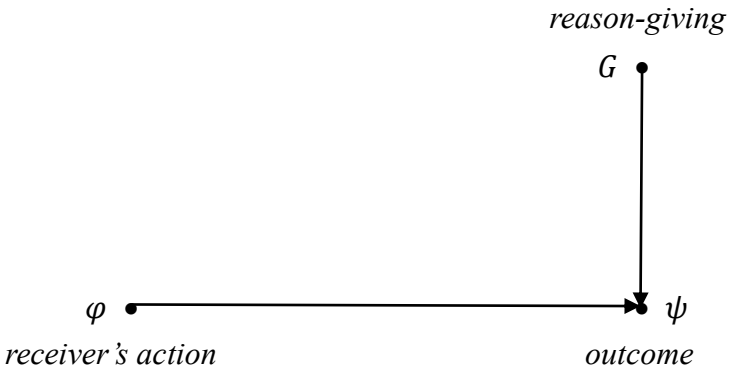


Figure 3.3.2

where the reason $R: \varphi \rightarrow \psi$ is dependent on background condition $B: G \rightarrow \psi$.



doi:10.6342/NTU202400318

**Definition 3.3.5 (Robust Reason-Giving)**

The way of giving reasons is considered *robust*, if and only if the reason-giver issues an imperative with the intention of making the receiver recognise it and act according to its content.

The last way is robust reason-giving. In this way, the provider issues a reason through an imperative, aiming to make the respondent adopt the provider's intention as a reason for action. As per Enoch's analysis, three elements must be satisfied for robust reason-giving to be considered valid:

(1) $A$ intends to give $B$ a reason to $\varphi$, and $A$ communicates this intention to $B$,
(2) $A$ intends $B$ to recognise this intention, and
(3) $A$ intends B's given reason to $\varphi$ to depend in an appropriate way on $B$'s recognition of $A$'s communicated intention to give $B$ a reason to $\varphi$. (David Enoch, 2011: 13)

In other words, robust reason-giving is a method of providing reasons through imperatives. For instance, when a family member says, 'Don't smoke,' their command does not reference an existing reason but instead creates a new one: *don't smoke because the family member forbids it*. According to Enoch, robust reason-giving is essentially a variant of triggering reason-giving. The rationale behind considering another person's request as a reason for us to act accordingly is that we already possess a reason to do so (Enoch, 2011: 8-9). For example, a special relationship with the requester leads to benefits when complying with their request, such as boosting the virtuous relationship, or to forfeits when non-compliance, such as breaking it. In essence, the other person's request triggers a counterfactual difference between 'complying with the request resulting in some advantage' and 'not complying resulting in some disadvantage'; therefore, as triggering reason-giving, the reason for $\varphi$ is dependent on the request, or other kinds of imperative. However, the difference between robust and triggering reason-giving lies in the focus of robust reason-giving on recognising the other party's intention and using it as a motive for action, whereas triggering reason-giving merely alters unintended background conditions. Definition 3.3.5 is presented as a following causal graph:



doi:10.6342/NTU202400318

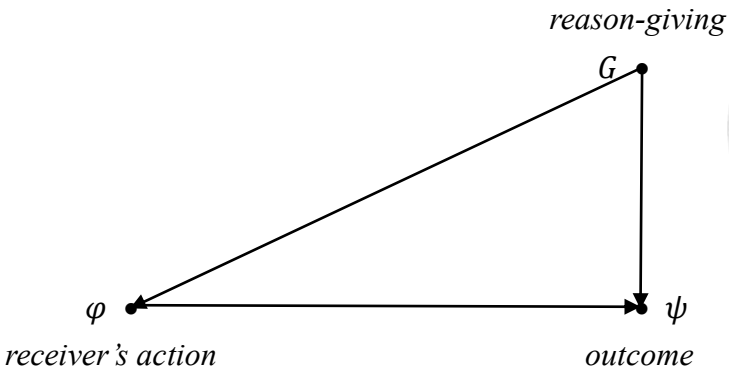


Figure 3.3.3

where the reason $R: \varphi \rightarrow \psi$ is dependent on background condition $B: G \rightarrow \psi$, and $E: G \rightarrow \varphi$ represents an epistemic causal path, meaning the reason-giver makes receiver recognise his/her intention and act according to its content.

The manner in which the imperative provides reasons is multifaceted; it may do so epistemically, triggeringly, or robustly. In the case of epistemic reason-giving, imperatives serve more as 'reminders' or 'warnings' about the actual situation. For instance, the reason I command 'Don't open the door!' is that there are wild creatures outside, and opening the door would allow them to enter and harm you. Therefore, you already have a reason not to open the door, and my command simply informs you of this reason, particularly if it reveals a specific context.

Robust reason-giving is a common method through which imperatives provide reasons. When I ask you to 'open the door, please,' your compliance with my request could be motivated by various reasons. For instance, if you are my employee, non-compliance might risk your job. If you are my friend, you might comply simply as a favour. Alternatively, you may be aware of a wild animal outside and understand that opening the door could put me in danger, thus your compliance might be to kill me, implicitly absolving yourself of responsibility. In other words, the imperative itself is not a *reason-fact*; it is, instead, a *reason-giving fact*.

In the case of triggering reason-giving, the role of imperatives in providing reasons to actors is contingent. The actor coincidentally complies with the command due to his or her own triggered reason. Consider the following example:

**Example 3.3.1**



doi:10.6342/NTU202400318

The teacher commanded the whole class to leave the classroom. Tim, lost in a daze, failed to notice the teacher's command. However, upon seeing his best friend Tom exiting the classroom, he too left in order to chat with him.

In Example 3.3.1, Tim fails to notice the teacher's instruction and thus lacks the epistemic causal path $E: G \rightarrow \varphi$. However, because the teacher's instruction triggers an opportunity to chat with Tom, Tim consequently leaves the classroom, inadvertently complying with the teacher's instruction.

Next, let us examine 'constituting a reason'. For an action $C$ to constitute a reason, it must create either favourable or unfavourable differences $\psi$ in response to the reason-reciever's action $\varphi$. In essence, the causal path $E': \varphi \rightarrow C$ would be *epistemic*, and action $G$ transforms the epistemic causal path into one that involves action. When $C \cap \psi \neq \phi$, action $C$ itself is a difference, while $C \cap \psi = \phi$ implies action $C$ creates a difference-making action $M: C \rightarrow \psi$. This is depicted in Figure 3.3.4, where the dotted arrows indicate the uncertain relationship between $C$ and $\psi$:

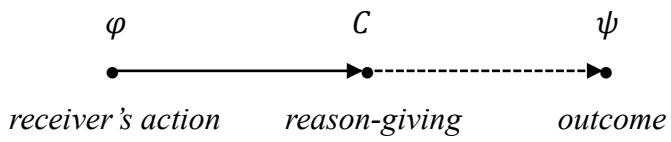


Figure 3.3.4

If $C =$ *to issue an imperative*, then an imperative can indeed *constitute* a reason, particularly in scenarios where it is employed to order a sanction. For example, when the individual under threat fails to pay the ransom and the kidnapper then orders his associates to eliminate the hostage, this order from the kidnapper does *constitute* a reason. If an imperative constitutes a reason, it can probably only be used on such an occasion. However, we still can ask why the kidnapper's associates obey the kidnapper's imperative, and perhaps this must come back to how the imperative *provides* reasons: perhaps the imperative informs existing reasons, triggers new reasons, or perhaps the kidnapper constitutes new reasons through threats–which is no longer within the scope of the imperative. In other words, *the imperative never really constitutes a reason*; the imperative is only the *reason-giving fact*.

The above analysis shows a clearer understanding of the nature of imperatives,



doi:10.6342/NTU202400318

enabling us to distinguish between imperatives and threats. Sometimes, conditional statements can indeed function as commands. For example, when a robber says 'I'll shoot if you don't give me the money,' his intention is not to state a fact but to issue a threatening command. Nevertheless, this does not imply that such a scenario captures the essence of all imperatives. Moreover, even within imperatives involving threats, 'imperatives' and 'threats' can be distinctly separated: imperatives still serve as a reason-giving fact, informing the threatened individual of the command's content through $E: G \to \varphi$, and triggering the evil outcome through $B: G \to \psi$, while the actual 'reason-fact' is the 'threat' itself.

Constituting a reason is the artificial generation of a new reason, whilst providing a reason is presenting an already established one. If we want to support legal positivism, we should likewise support the notion that the law involves constituting reasons. Otherwise, the legal system is simply a variant of other normative systems. In this regard, the imperative should not be the subject of legal positivist concern. However, this should not be misconstrued to imply that constituting reasons is the sole manner in which the law serves as reasons. In fact, there are numerous ways, a topic I shall explore in the next chapter.

## 3.4 The Three Conditions of Separation Thesis

The separation thesis forms the bedrock of legal positivism, whilst its antithesis is the connection thesis. Nonetheless, legal philosophers vary markedly in their perspectives on these two theses. Hart advocates a tempered version of the separation thesis. In his Postscript, he articulates:

> According to my theory, the existence and content of the law can be identified by reference to the social sources of law (e.g. legislation, judicial decision, social customs) without reference to morality except where the law thus identified has itself incorporated moral criteria for the identification of the law. (Hart, 1994: 269)

In the framework of difference-making-based theory of reasons, Hart's statement can be clearly expressed as the following proposition:

**Definition 3.4.1 (Inclusive Legal Positivism)**

A legal theory is *inclusive legal positivism*, if and only if the validity of legal



doi:10.6342/NTU202400318

normativity advocated by this theory does *not necessarily* have moral reasons as background conditions.

Certain theorists, notably Raz, champion a robust version of the separation thesis. Raz contends that 'the existence and content of every law is fully determined by social sources' (Raz, 1979: 46), asserting that these sources are mere facts, devoid of any moral imperative. To concisely encapsulate Raz's assertion, it might be articulated in the framework of difference-making-based theory of reasons thus:

**Definition 3.4.2 (Exclusive Legal Positivism)**
A legal theory is *exclusive legal positivism*, if and only if the validity of legal normativity advocated by this theory *necessarily does not* have moral reasons as background conditions.

Here, we will expand the concept of the separation thesis, where the 'legal reason' is separated not only from 'moral reason' but also includes other non-legal reasons such as 'religious reason', 'economic reason', or 'political reason'. This is because what constitutes a 'moral reason' is a matter to be defined by moral philosophy, which is not suitable for discussion in the study of legal philosophy. Furthermore, I also believe that legal positivists would not argue for a connection between legal reasons and the aforementioned reasons. Therefore, I have revised Definition 3.4.1 and Definition 3.4.2 as follows:

**Definition 3.4.3**
A legal theory is *inclusive legal positivism*, if and only if the validity of legal normativity advocated by this theory does *not necessarily* have all reasons other than those of the law as background conditions.

**Definition 3.4.4**
A legal theory is *exclusive legal positivism*, if and only if the validity of legal normativity advocated by this theory *necessarily does not* have all reasons other than those of the law as background conditions.

Definition 3.4.3 corresponds to Definition 3.4.1, while Definition 3.4.4 corresponds to Definition 3.4.2. We do not need to dwell on whether inclusive legal positivism or exclusive legal positivism represents true legal positivism. Our focus here is on the



doi:10.6342/NTU202400318

conditions of the separation thesis. Consequently, inclusive legal positivism can be adopted as the baseline criterion of the separation thesis, utilising it to scrutinise the specific conditions of this thesis. In this context, the connection thesis can be defined as follows:

**Definition 3.4.5 (The Connection Thesis)**
A legal theory is *not legal positivism*, if and only if the validity of legal normativity advocated by this theory *necessarily* has all reasons other than those of the law as background conditions.

Looking back at the constitution of legal reasons, for the law to constitute reasons independently, it must fulfil the various conditions previously outlined:

**Constitutiveness**
The law must constitute reasons, implying such reasons have not existed previously within the natural and social context. It is an act of reason-constitution, transmuting what was initially an epistemological causal path into a metaphysical one.

**Deliberativity**
Two prerequisites must be satisfied to constitute a legal reason: the sanction's effectiveness does not depend on background conditions which are independent of the sanctioner, giving the sanctioner complete decision-making power, and the sanction-imposing entity should take into account other existing or possible outcomes, rather than react directly to visible outcomes.

**Independence**
A constituted reason $R_i$, which is a member of set of independent reasons $R$, does not *causally* depend on any member of a different set of reasons $R'$ outside of $R$ when $R_i$ is used as a reason for action.

These three conditions form a hierarchical sequence: *Constitutiveness*, *Deliberativity*, and *Independence*, wherein the superior tiers are predicated upon the foundational ones, with constitutiveness being the most fundamental. A constituted reason is not one that occurs naturally; rather, it occurs artificially. Should the law impart normativity via reason-giving, then its normativity is essentially a variant of other forms of reasons – that is, of $R: \varphi \rightarrow \psi$ as depicted in Figure 3.3.1 to Figure 3.3.3. Absent reason $R$, the



doi:10.6342/NTU202400318

law is inherently non-normative; in other words, reason-giving fulfils the criteria set out in Definition 3.4.5, making it fail to meet that of Definition 3.4.3, which represents the rudimentary standard for the separation thesis.

Deliberativity necessitates that reason-constituting action eliminates all background conditions external from the model, ensuring that such action is solely dependent on the epistemic causal path $E': \varphi \to C$ illustrated in Figure 3.3.4. The crux of deliberativity lies in its distinction from both positive law and positive morality. The latter may also be formulated through interpersonal mutual sanctions, yet these are contingent upon numerous social and psychological factors. Conversely, sanctions instituted by the government possess sufficient power to make determinations independently of these contextual conditions. From a causal model perspective, deliberativity aims to preclude situations akin to those described in Definition 3.2.1, thereby ensuring that sanctions and reasons constituted by sanctions do not causally depend on external background conditions within the model.

The condition of deliberativity is less relevant to Definition 3.4.3 to Definition 3.4.5, whose purpose is to render the action of constituting reasons adequately artificial, thereby distinctly differentiating them from naturally constituted reasons. Furthermore, owing to the absence of any external background conditions, the counterfactual path constituted by sanctions which meet deliberativity is pure enough so that it can explain and guide the action perfectly.

Independence requires that the law be a closed logical system. When a reason-constituting action eliminates external background conditions beyond the model it has, based on the dual nature of action explanation, it becomes an action purely based on reasons. Subsequently, it is necessary to examine whether the reason for the reason-constituting action belongs to the normative system of the law itself. Should the reason-constituting action be rooted in a normative system distinct from that of the law, then the establishment of the law's normative framework becomes subordinate to this external normative system. It is necessary that, in essence, there exists a reason other than that of the law for its normativity as a background condition, which does not conform to Definition 3.4.3.

In essence, positive law and positive morality were frequently indistinguishable in



doi:10.6342/NTU202400318

primitive societies owing to both inherent constitutiveness. Subsequently, as the 'power to decide' on sanctioning became increasingly centralised within a specific group, the naturalness of positive morality gradually shifted towards a state of arbitrariness, thereby heralding the onset of deliberativity. At this point, positive morality evolves into a new form of reason, distinct from its previous form. At this stage, however, the law lacks independent normativity, and the so-called legal state is merely a strategic, provisional resolution of the political superior. Consequently, society at this point is governed not by the rule of law, but merely by law. Ultimately, when certain societal members establish the interwoven sanction system delineated in Theorem 2.4.7, the law acquires an independent normativity, heralding the emergence of a legal society governed by the rule of law.

We can organise the relationship between the three levels of the separation thesis and the three stages of governance as follows:

| **Conditions of Separation Thesis** | **Stages of Governance** |
|---|---|
| Constitutiveness | Morality |
| Deliberativity | Rule by law |
| Independence | Rule of law |

Whilst I noted that positive law and positive morality were frequently indistinguishable in primitive societies, we must avoid oversimplifying the Stages of Governance as being either progressive or regressive, or naively assuming that the morality of governance is absent in contemporary societies. Conceptually, these three states of governance are hierarchically ordered, yet in practice, they are observable across various societal segments. In communities where equality among members prevails, the stage of governance is morality; where a political superior exists, the stage of governance is the rule by law; where political superior checks and balances on each other, the stage of governance would be the rule of law.

The Stages of Governance are not related to H.L.A. Hart's distinction between



doi:10.6342/NTU202400318

primary and secondary rules. It is a misconception to think that the governance method of morality lacks 'secondary rules.' For instance, in a sizeable association without a central leader, the members may enact the rules of the association, including provisions for their amendment. Moreover, certain authorities are designated to interpret these rules' meanings. Nonetheless, in the absence of sanctions, the stage of governance is still Morality. This is because our acceptance and compliance with these rules and authorities stems from various kinds of reasons, which are inherently sensitive to contexts. As Hart put it:

> In fact, their allegiance to the system may be based on many different considerations: calculations of long-term interest; disinterested interest in others; an unreflecting inherited or traditional attitude; or the mere wish to do as others do. (Hart, 1994: 203)

If these reasons for accepting and complying with the rules and authorities are lost, then the normativity of these rules would cease to exist. The acceptance and observance of rules and authority are, by nature, a temporary phenomenon. This implies that 'secondary rules', in and of themselves, are inept at adapting to the entirety of social contexts and do not effectively make up the inherent shortcomings of primitive social norms, because these secondary rules are, inherently, replete with defects.

That is why every society will eventually resort to sanctions. In the few paragraphs in which Hart explains why sanctions exist, he explains the emergence of sanctions in this way:

> It need hardly be said that in few legal systems are judicial powers confined to authoritative determinations of the fact of violation of the primary rules. Most systems have, after some delay, seen the advantages of further centralisation of social pressure; and have partially prohibited the use of physical punishments or violent self help by private individuals. Instead they have supplemented the primary rules of obligation by further secondary rules, specifying or at least limiting the penalties for violation, and have conferred upon judges, where they have ascertained the fact of violation, the exclusive power to direct the application of penalties by other officials. These secondary rules provide the centralised official 'sanctions' of the system. (Hart, 1994: 97-98)



doi:10.6342/NTU202400318

Nevertheless, Hart merely ascribes the existence of sanctions in society to 'benefits', overlooking their intrinsic necessity to the legal system itself–a system which requires sanctions to achieve sufficient stability for its maintenance. Only sanctions, too, can satisfy three conditions of the separate thesis. Therefore, sanctions, as an independent way for laws to constitute reasons.

# 4. On the Various Ways in Which Laws Serve as Reasons

## 4.1 On Authority

Austin's general assertion categorises laws or rules as a type of command. Though it may be contentious to liken law to a command[13], for the purpose of thoroughly examining Austin's conception of law, it is worthwhile to consider 'rules' tentatively as a form of command. Let us then delve into the exploration of how rules, when viewed as commands, function as reasons. According to Enoch, the manner in which the law offers reasons, akin to the imperative, is multifaceted; it may do so *epistemically*, *triggeringly*, or *robustly*. No matter what kind of providing reason-giving, however, legal rules do not fundamentally depend on sanctions to uphold their normativity. Administrative guidance is the typical case, which is the laws in Taiwan existing without associated sanctions, as defined in Article 165 of the Administrative Procedure Act, which states:

> **Article 165, Administrative Procedure Act**
> The term 'administrative guidance' used in this Act denotes the act of an administrative authority within the scope of its duties or the functions under its

[13] Contrary to Austin, Kelsen did not perceive law as a direct command to those under political superiors. Firstly, the law functions not as a direct order, but as an indirect system of guidance: it operates by instructing officials to impose sanctions under certain conditions, thereby directing people's actions. Secondly, commands depend intrinsically on the issuer's intent, yet in the legislative context, legal obligations often disconnect from the wills of individual legislators. Kelsen pointed out that, in legislative contexts, legal obligations frequently detach from the will of individual legislators. This detachment becomes evident in legislative processes where all individual legislators may not agree to the final political compromise. Some may vote perfunctorily, others may vote for conformity, and still others may be uninformed about their voting choices. Even when all legislators responsible for enacting laws are no longer present, rendering inquiries into the original intent of the legislation impossible, the laws themselves continue to function. Kelsen posits that the law could at best be described as a 'depsychologised command'. This does not imply, however, that when I am obliged to do something, it necessarily means that this obligation is based on any specific individual's actual will (Kelsen, 1941: 55-56).



doi:10.6342/NTU202400318

control to urge specific persons to perform particular acts or omission of act, by way without mandatory force in law, for the purpose of achieving specific administrative objectives.

The way that administrative guidance provides reasons has three ways mentioned above. The first way is epistemic. The government, as an intellectual authority, provides the pre-existing reasons through notification, without creating or triggering new reasons. In other words, the reason for action already exists within the world, and the government's act of informing merely alters the actor's state of awareness. The second way is robust. Consider complex professional fields where individuals might lack knowledge on appropriate actions or require the discipline of collective effort to fulfil their objectives. In such scenarios, governmental administrative guidance provides authoritative direction, guiding individuals in their decision-making. Here, the authority's directives are robust: individuals recognise and comply with these directives, thereby realising the goals that they have reasons to achieve. The third way is triggering. In this way, individuals may not directly receive instructions from the government. Instead, the government's directives alter the social context, indirectly influencing individual actions. This indirect effect on personal behaviour will not be explored in detail here. I will focus on epistemic and robust reason-giving.

In the realm of administrative guidance, it is sometimes difficult to distinguish between epistemic reason-giving and robust reason-giving, as the government, in both instances, assumes the role of an authority guiding individuals' actions. Nevertheless, these two concepts maintain distinct core principles. Epistemic reason-giving explicitly indicates the reason, allowing individuals to analyse and judge it, whereas robust reason-giving steers individuals solely through imperatives, leaving no room for them to evaluate the reasons behind the decision. In other words, it is the faith in authority that constitutes the reason for action, rather than the situations themselves.

The concept of administrative guidance closely aligns with Raz's theory of authority, which proposes that the motivations behind our actions can be categorised into first-order and second-order reasons. First-order reasons are those that necessitate a specific action or inaction; in short, they are just reasons for actions. On the other hand, second-order reasons relate to why we should follow or disregard a particular first-order reason. Raz argued that the rule, as a kind of authority, constitutes both a first-order cause of



doi:10.6342/NTU202400318

action and a second-order reason. The rules offer practical guidance for action, while simultaneously excluding any reasons that may conflict with the directives of the rule.

Why can the rule serve as a second-order reason, thereby excluding any conflicting reasons with its directives? Raz puts forward two arguments to support this view[14]:

**The Argument from Efficiency**

Rules are perceived as a *labour-saving and time-saving device*. In their absence, we would need to individually assess the merits and demerits of reasons for every decision, a process both time-consuming and laborious. Rules simplify this by their exclusionary nature, negating the need for such extensive evaluation, and thereby facilitating swift decision-making. This exclusionary nature of rules obviates the labour of detailed consideration, aiding in more rapid decision-making.

**The Argument from Authority**

Rules are regarded as commands from an authority. Since this authority is assumed to have thoroughly deliberated the relevant first-order reasons, adhering to its directives likely aligns our actions with outcomes that are in keeping with these first-order reasons. Generally, due to the practical expertise of the authority, an individual who accepts and follows these commands is more likely to be in accord with the first-order reasons than if they attempted to act directly upon these reasons themselves.

In Raz's theory, 'authority' holds significant importance. It is authority, rather than power, that forms the foundation of the law's exclusionary force. Likewise, it is from the legal authority's recognised and respected position, rather than its coercive power, that it derives the influence, guiding compliance through its authoritative status. Indeed, this does not imply that the authority's command is inherently correct, or true, but rather it highlights the 'motive' behind human inclination to trust in authority. In the context of the difference-making-based theory of reasons, such a motive can be explained by 'reasons' and 'causal graph.'

Robust reason-giving demonstrates how reasons are imparted by authority. Consider Figure 3.3.3. The causal path $R: \varphi \rightarrow \psi$ elucidates our reason for adhering to the authority's directives, with $\psi$ representing the claimed outcome of the argument from

[14] Here are the terms used by Peng-Hsiang Wang (2008: 360).



doi:10.6342/NTU202400318

efficiency or from authority. The causal path $E: G \rightarrow \varphi$ reveals how the content of the authority's instructions influences our cognitive state. Meanwhile, the causal path $A: G \rightarrow \psi$ sheds light on the exclusionary power of the authority's commands. These commands must be robust, meaning the causal path $A: G \rightarrow \psi$ needs to be present to demonstrate the authority's directives in effecting differences. Absent this causal path, the commands become merely epistemic, impacting only cognition and leaving individuals to base their actions on personal reasons.

It is important to recognise that the reason for adhering to a rule as a practical authority, represented by the causal path $R: \varphi \rightarrow \psi$, is founded not on the intrinsic value or benefits of the rule's content, but on those derived from the act of obedience to the rule itself. As highlighted by Raz:

> Here we see clearly how rules differ from other reasons. The insightfulness and subtlety of a novel are reasons for reading it because they show why reading it is good. But the considerations which show why the rule is binding, i.e. why it is a reason for not bringing more than 3 guests, do not show that it is good not to bring more than three guests. They show that it is good to have power given to a committee. (Raz, 2009: 210)

There exists a discernible gap between the reason for obeying a rule and the value or benefit of its content. In other words, the answers to 'Is $\varphi$ good?' and 'Is the rule that ought to be $\varphi$ valid?' may differ. This discrepancy is what Raz terms the *normative gap*. He further describes the method of justifying the rule as a *content-independent justification*, emphasising that the rule's validation is not contingent on the intrinsic value or benefit of its content. Peng-Hsiang Wang (2015: 351-353) further argues that the rule's association with normativity stems from the reason that obeying the rule yields specific benefits. In answering 'why should we obey the rule', it is argued that compliance engenders certain advantages. Consequently, the 'existence of legal rules' becomes a reason-giving fact, triggering the causal path of this reason. In this context, the rule itself is not a reason-fact but a *reason-giving fact*.

Authority, as commonly understood, refers to the intellectual superiority of a person or a group of people. When we say that a person is an authority in physics, it means that he knows a lot about physics. Similarly, when we say that a person is an authority in



doi:10.6342/NTU202400318

practice, it means that he has a lot of practical knowledge. Such knowledge, according to *practical deliberative usefulness*, is a fact that can decide how to perform an action to either prompt or prevent a particular outcome. In other words, it is a reason-fact. A practical authority possesses a wealth of reason-facts and, consequently, can adeptly guide others. This guidance enables the advised individuals to act more effectively in alignment with their own reasons. From the difference-making-based theory of reasons, practical authority is defined as follows:

**Definition 4.1.1 (Practical Authority)**
A person or a group of people is a *practical authority*, and if and only if he or they know a lot of reason-facts.

Within the moral realm, such a practical authority is often characterised by the term 'wisdom,' which denotes a profound adeptness in navigating interpersonal relationships and a deep understanding of how to get along with people. Within the political realm, such a practical authority signifies a keen acumen for managing both domestic and international affairs. Within the legal realm, such a practical authority translates to a comprehensive knowledge of the law, which is a jurist, lawyer, or judge. The practical knowledge possessed by the authority is the basis for the authority's being established; nevertheless, this knowledge is an established reason rather than the actual creation of a new reason.

Occasionally, the guidance provided by an authority is epistemic, meaning that the authority simply informs others about an established reason, leaving the decision-making to them. In this case, the reason provided by the authority is the same as the reason it possesses. On the other hand, when the authority's guidance is robust, a normative gap occurs in which the reason provided by the authority is not the same as the reason the authority possesses. The reason provided by the authority is the one we trust the authority, whilst the reason the authority possesses is the one the authority is established. Therefore, the reason we obey the authority's commands can be unrelated to the benefits of their content. However, it is crucial to recognise that these reasons are pre-existent in the world, and no authority can constitute a new reason through a mere command. In other words, *an authority cannot constitute a new reason*.

Perceiving the law as an authority may partly elucidate its role in our practical



doi:10.6342/NTU202400318

considerations. This perspective underpins the utility of administrative guidance. However, a pertinent question arises: if legal reasons stem from their authority, why do statutes of administrative guidance constitute merely a fraction of the overall law, or at times, appear as exceptions? Moreover, while the law's role as an authority clarifies our compliance, it does not explain our acceptance of this authority, which precedes compliance with the authority's commands. The exclusionary force of the authority becomes established only upon our acceptance; otherwise, we continue to consider first-order reasons.

Comparing authority and law in terms of 'acceptance of authority' highlights their dissimilarity. In seeking medical or educational guidance, for example, we typically choose our doctor or teacher before adhering to their advice. Conversely, in the context of governance, we never select the government before following its instructions. Considering voting in a democracy as electing a government is misleading, as legal systems preexist democracy–we have the rules of the democratic system first, and then we choose the government according to those rules. Additionally, asserting that the laws in non-democratic countries have lost their normativity is untrue, as the laws in those countries still continue operating no matter how we evaluate these legal systems.

Furthermore, it would also be misleading to contemplate why we accept authority in the context of justifying the law as an authority, as that is not how the law works from the beginning. Instead, the crux lies in recognising that the law often does not present us with the option to deliberate on this authority. We find ourselves almost forced into accepting the authority of the law, largely due to the threat of sanctions. Most laws are intrinsically linked to sanctions through a counterfactual relationship. However, there are notable exceptions, which I intend to explore in the subsequent section.

## 4.2 On Principles

Typically, rules are connected to sanctions via a counterfactual relationship. Within the legal framework, there exist certain provisions that encompass both the constituent elements and the legal consequences within a single clause. This structure makes the counterfactual relation of the rule immediately evident from a individual provision. This characteristic is often seen in areas such as the Criminal Code and the Administrative Penalty Act of Taiwan, among other domains where express sanctions



doi:10.6342/NTU202400318

are explicitly stated. Such provisions are commonly termed as '*complete legal sentences*' (*Rechtssatz*) and represent the quintessential form of statutory law with a counterfactual structure.

Moreover, there are other statutes which exhibit this counterfactual structure, albeit dispersed across various provisions. For example, Article 345 of the Civil Code of Taiwan delineates only the constituent elements:

> **Article 345, Civil Code of Taiwan**
> A sale is a contract whereby the parties agree that one of them shall transfer to the other his rights over property and the latter shall pay a price for it.
> The contract of sale is completed when the parties have mutually agreed on the object and the price.

Article 345 only points out the constituent elements of the contract of sale, while its legal effects are distributed among sections 348 to 378. These sections are referred to as '*incomplete legal sentences*' (*Unvollständige Rechtssätze*) They still manifest the counterfactual structure within the framework of the code system.

Nonetheless, some legal provisions are fundamentally principled. The Civil Code of Taiwan has many typical examples of this:

> **Article 148(2), Civil Code of Taiwan**
> A right shall be exercised and a duty performed in good faith.
> **Article 1084(1), Civil Code of Taiwan**
> Children shall display filial respect towards their parents.

Similarly, the provisions in the Constitution's Basic Rights Clause lack directly corresponding legal effects. Even beyond the realms of positive law, numerous political, moral, or even philosophical principles, not explicitly inscribed in legislation, significantly influence legal outcomes. For example:

**Example 4.2.1 (The Sunflower Student Movement)**
In 2014, Taiwan witnessed the Sunflower Student Movement. During this period, students protesting against certain legislative proposals occupied the Legislative Yuan in Taipei City. They even attempted to take over the adjacent Executive Yuan.



doi:10.6342/NTU202400318

Following the resolution of these events, the students who had occupied the Executive Yuan were charged with offences of obstructing an officer in discharge of duties. The defendants asserted their actions constituted 'civil disobedience,' a concept not formally enshrined in law. Concerning the issue of whether the leadership of a protest, involving the occupation of the Executive Yuan, was lawful hinges on the concept of civil disobedience, the Supreme Court adjudicated as follows:

> Acts of civil disobedience represent a unique form of freedom of speech expression. The whole legal interest they aim to protect is the liberal democratic constitutional order that is about to be or has just begun to be destroyed (if the actions of the object of protest have caused systemic and serious damage to the liberal democratic constitutional order, then it falls within the scope of the right to resist). In such scenarios, the Court may invoke the emergency doctrine or apply the provisions for excessive danger-averting by analogy, thereby mitigating the violation or reducing criminal liability. (Criminal Judgment No. 3695, Supreme Court 109, Taiwan.)

Although civil disobedience is not explicitly outlined in legal statutes, it plays an important role in guiding judicial interpretation, leading to the dismissal of the prosecutor's appeal in this case. This case demonstrates that principles play an actual and crucial place in the realm of adjudication; however, how are they possible?

These principles that are not explicitly expressed within statutory provisions or do not delineate specific legal consequences seem to constitute a counterexample to the difference-making-based theory of reasons. Principles, invariably articulated as distinct from rules, can equally function as reasons. Dworkin (1978: 22-28) argued that due to the presence of principles, law cannot be exclusively interpreted through rules. Rules typically present a binary of absolute compliance or non-compliance, leading directly to a legal outcome. Exceptions to rules necessitate new rules. In contrast, principles involve gradations and numerous exceptions that do not undermine their validity or coherence. The influence of a principle is not in dictating a specific conclusion but in directing the course of legal interpretation. This implies that principles need not be formally written in the Code. Even when they are, they often remain incomplete and lack direct legal consequences, characteristic of the nature of principles. Should we accept the validity of Dworkin's assertions, may we then posit that principles themselves can constitute reasons within the legal framework?



doi:10.6342/NTU202400318

To adequately address this query, it is imperative first to elucidate the essence of a principle. Principles manifest in a multitude of forms, predominantly as normative propositions. For instance, Article 148(2) of the Civil Code decrees, 'A right shall be exercised and a duty performed in good faith,' and Article 1084(1) mandates, 'Children shall display filial respect towards their parents.' Additionally, principles may emanate from political and legal theories, such as the concept of civil disobedience or the legal state (Rechtsstaat). These principles represent a set of normative propositions.

At times, principles with legal consequences not explicitly delineated in statutes still exert a binding force upon us. In the realm of reason theory, the axiom 'if $A$ ought to do $P$, then a reason $R$ must exist for $A$ to do $P$' suggests the presence of underlying reasons for these principles. These reasons thus give us a force to consider. However, the pertinent question then becomes: Does the origin of this force stem from the law itself? Principles inherently possess criticalness, a key characteristic of collective reason; in contrast, legal reasons based on sanctions fall under the category of individual reasons. The origin of this force thus does not stem from the law itself, but rather from normative systems external to the law. Take, for instance, the principle 'Children shall display filial respect towards their parents.' This simply posits that children have a moral obligation to respect their parents, a notion deeply rooted in the positive morality of Confucian tradition. It is only when a judge confers legal effect to this principle that it metamorphoses into a 'legal reason.' For instance,

**Example 4.2.2 (Recovering Property from Unfilial Son)**
The defendant, Tom, was sued by his parent, the plaintiff, Tim, over a property dispute. Tim, diagnosed with cancer, transferred real estate to Tom under a notarised gift contract, conditional on Tom providing daily ancestral worship and monthly maintenance fees. This contract included a clause allowing Tim to reclaim the gift if Tom *failed to display filial respect towards their parents*. However, after receiving the real estate, Tom neglected these obligations, neither returning home for worship nor caring for Tim's well-being, breaching the contract's conditions. Consequently, Tim initiated legal action, arguing that Tom ought to return its ownership.

In this case, the Taoyuan District Court once pointed out in cases involving violations of Article 1084(1):



doi:10.6342/NTU202400318

> Regarding the duty of children to respect their parents: Parents have the right and obligation to protect and nurture their minor children, as explicitly stated in Items 1 and 2 of Article 1084 of the Civil Code. This recognises the familial values of children showing respect to their parents and parents providing care and education to their minor children. Therefore, children who have reached adulthood and were nurtured by their parents should fulfil their filial duties. If parents are bedridden, and their children, without any legitimate reason, fail to visit them over a long period, this neglect is considered a significant act of cruelty in a society that values filial piety, causing immense emotional pain to the parents...Consequently, adult children who do not visit their elderly or sick parents and fail to provide necessary daily care are causing great emotional distress to their parents, and such behaviour is naturally considered unfilial. (Civil Judgment No. 120, Taiwan Taoyuan District Court, 2023)

In Example 4.2.1, the judge delineates the constituent elements of filial respect, thereby stipulating its effect within the contractual framework. This interpretation imbues filial respect with a new counterfactual structure, elevating it from a mere moral principle to a legal rule.

It appears from Example 4.2.1, as well as from Example 4.2.2, that the principles are distinctly defined and their effects are specified by jurists. Until then, principles do not constitute legal reasons, as we cannot find the *counterfactual structure* in these principles. Frequently, principles do not directly result in a specific threat, but rather influence the acquisition and cessation of legal rights or obligations, indirectly precipitating it. In the realm of private law, this pertains to the right of claim, while in public law, it concerns the initiation of coercion. Here, it is unnecessary to construe sanctions in an excessively narrow manner. By definition, a sanction entails the establishment of an artificially unfavourable outcome. In terms of essence matter: nullity can be a sanction. In terms of subject matter: the entity imposing sanctions need not be the government; within tort law, sanctioning is effected by granting individuals the right to claim.

The aforementioned perspective can be extended to rules: rules do not necessarily possess a stable counterfactual structure until a judge continuously renders decisions on the legal effects ascribed to them. Numerous statutes in the statute book remain effectively dormant to this day. This could be attributed to intentional disregard or



doi:10.6342/NTU202400318

possibly obscured by interpretative techniques. Furthermore, there exists a multitude of uncertainties within the rules awaiting elucidation and clarification by legal scholars. Hart addresses this issue by pointing to the inherent ambiguity in natural language, asserting that all general terms possess a core of certainty surrounded by a penumbra of doubt regarding their references and meanings. It is not uncommon, however, for a judge to arrive at divergent decisions in situations governed by ostensibly clear legal rules, based on varying interpretations of the law. For instance:

> **Article 92(1), Civil Code of Taiwan**
> An expression of intent which is procured by fraud or by duress may be revoked by the expresser. If the fraud was done by a third party, the expression may be revoked only under the circumstances that the other party knew, or might know the affairs.
> **Article 92(2), Civil Code of Taiwan**
> The revocation of an expression of intent on the ground of fraud cannot be a valid defense against the bona fide third party.
> **Article 948, Civil Code of Taiwan**
> If, with the purpose of transferring the ownership of, or creating other rights to a personal property, a person in good faith takes the possession of such personal property, such possession shall be protected by law, even though the transferor had no right to transfer it. Provided, however, this provision shall not apply, if the transferee knows, or does not know with gross negligence, that the transferor has no right to transfer it.

Under the converse interpretation of Article 92(2), a revocation made under duress is valid even against a bona fide third party. However, Article 948 stipulates that the possession of property by a bona fide third party should be legally protected, even if the transferor lacks the right to transfer said property. Presently, the dominant interpretation, guided by the principle of *Lex specialis derogat generali*, favours upholding Article 948. Conversely, a minority opinion, considering legislative intent, suggests that the adoption of Article 948 would render the legislative purpose of Article 92(2) redundant. The question then arises: which interpretation is correct? It appears there is no definitive answer to such a query. Even a prevailing interpretation at one point in time is not immune to revision or reversal in the future.

To create and preserve the normativity of law itself, legal practitioners have



doi:10.6342/NTU202400318

developed a set of interpretative rules through the accumulation of doctrines, judgments, and precedents. This includes, for example, the application of the principle of proportionality within the Constitution, the standards for legal interpretation, and the benchmarks for sentencing. By delineating the distinctions and relative importance of various principles, lawyers are able to align them with specific legal effects, ensuring that all elements within the law function akin to a cause-and-effect system. This approach is not limited to principles alone; the method of interpreting rules must also be 'regularised' to create a set of operational guidelines. To constitute operable rules, therefore establish legal normativity, and eventually distinguish them from moral and policy judgments, is the very task of legal dogmatics and the of legal professionals.

## 4.3 On Rules

As previously outlined, the normative force of law is derived from rules rather than principles. However, the question arises: how do rules function as reasons? This section will explore two primary mechanisms through which rules serve as reasons. The first is through an attitude of acceptance, and the second, through the process of adjudication.

For H.L.A. Hart, the validity of a rule–it serves as a reason–stems from an attitude of acceptance. In 'The Concept of Law', Hart delineates our perception and comprehension of law into two perspectives: the internal point of view and the external point of view. A prevalent misinterpretation is that the internal point of view is exclusively that of the 'insider's' point of view, embodying only the participants' '*hermeneutic*' stance towards the law, as opposed to the external viewpoint, which lacks such a stance. Stephen Perry, for instance, posits that 'The general idea of the internal point of view is that an adequate jurisprudential account of law must at some point take into consideration how the practice looks to at least some of the practice's participants, from the inside.' (Perry, 1995: 99) Similarly, Gerry Postema notes, 'The law, like other similar social practices, is constituted not only by intricate patterns of behavioural interaction, but also by the beliefs, attitudes, judgments, and understandings of participants. The practice has an "inside," the "internal point of view" of participants.' (Postema, 1998:332). In this interpretation, the theory that seeks to align with the internal point of view aligns itself with philosophical behaviourism, focusing solely on stimuli and responses.



doi:10.6342/NTU202400318

Shapiro challenges this notion, arguing that the 'hermeneutic' stance alone does not suffice to define what Hart terms the internal point of view. According to Hart, the internal point of view must encompass an attitude of acceptance as well. This attitude is characterised as 'using the rules as standards for the appraisal of their own and others' behaviour.' (Hart, 1994: 98) Absent this attitude of acceptance, even Holmes's 'bad man' could be misconstrued as part of the internal point of view, despite being a subject of Hart's vehement criticism (Shapiro, 2006: 1159).

Shapiro posits that in our recognition and understanding of rules, we operate within two distinct realms: theoretical and practical. In the theoretical realm, the approach to comprehending rules, as Hart claims, falls under what he terms a descriptive theory. This theory aims to depict human actions in a manner that is neutral and devoid of value judgments. However, the method of describing empirical phenomena hinges on the stance adopted. For behaviourists, these phenomena are merely plain facts, and it is argued that one should refrain from interpretations that extend beyond the facts, as such interpretations are not empirically verifiable. Hart, on the other hand, critiques behaviourists for their excessive empiricism, leading to a misrepresentation of certain realities. Hart's portrayal not only considers the factual actions but also encompasses participants' attitudes of acceptance. In essence, Hart's descriptive theory is hermeneutic in nature: he constructs the existence of rules and employs them to explain the consistent and unified actions of individuals in empirical phenomena.

In the practical realm, our knowledge and understanding of rules must align with that of the participants, necessitating a hermeneutic interpretation of these rules. However, this does not imply an automatic acceptance of the rules. The example of the 'bad man' illustrates this distinction clearly. He exemplifies that the attitude of acceptance is fundamentally different from a hermeneutic perspective: the 'bad man' acts out of fear of punishment and complies with the law, but does not consider the law as a means of critique or guidance for himself or others. In order to clarify this, I have organised the relationship between the theoretical, practical, hermeneutic, and acceptance stances into the following table:



doi:10.6342/NTU202400318

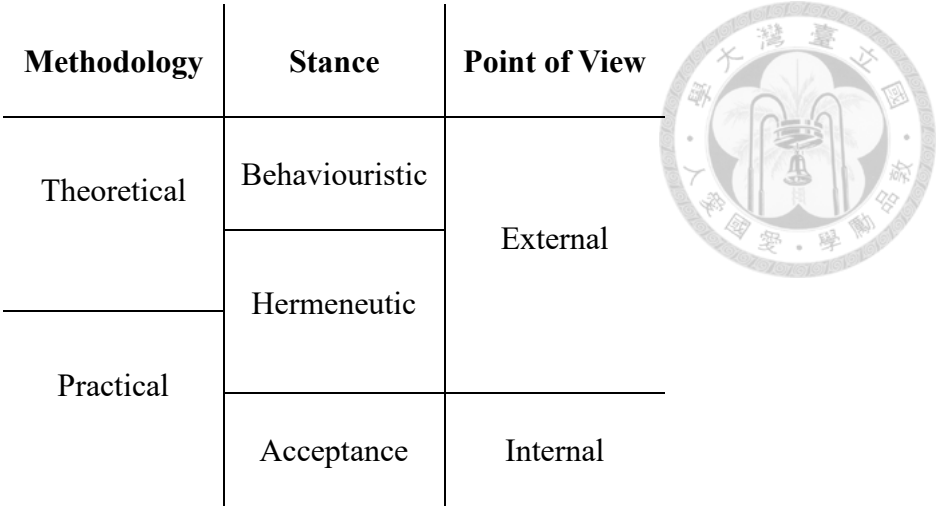


Figure 5.3.1

Figure 5.3.1 progresses through successive layers. The behaviouristic stance occupies the outermost viewpoint, characterised by Hart as the extreme external point of view. This stance is exemplified by Scandinavian Legal Realism, influenced by logical positivism, which posits that for knowledge to possess *cognitive meaning*, it must be verifiable under *the principle of verification*. A.J. Ayer (1952: 9) articulates this principle as follows:

> **The Principle of Verification**
> A proposition $P$ has cognitive meaning, if and only if $P$ is analytically or *empirically verifiable*. The term *verifiable* means that there is a way for us to know whether the proposition is true or false.

Per the principle of verification, normative propositions are deemed unverifiable, and hence, they are considered devoid of cognitive meaning. Scandinavian legal realists ascribe all legal concepts to psychological factors, reinterpreting positive legal concepts like 'rights', 'obligations', and 'property', as well as natural law concepts such as 'justice', as mere psychological emotions and reactions to constraints, because these concepts have no other verifiable facts except emotions. For instance, when we assert 'I own this pen', there is, in reality, no objectively observable factual relationship between the individual and the watch. This claim is often made in the hope of exerting psychological constraints on the behaviour of others. It is through the examination of the causal relationships that govern these reactions that genuine knowledge can be attained.



doi:10.6342/NTU202400318

The hermeneutic stance is categorised into theoretical and practical methodologies. In theoretical research, the use of a hermeneutic stance aligns with Hart-style descriptive jurisprudence, with the acceptance of rules embodying his hermeneutic approach. On the other hand, in the practical field, it is American legal realists who exclude the attitude of acceptance from practice, retaining only a hermeneutic stance. Influenced by pragmatism, American legal realism contends that the essence of law is not found in the legislative provisions established by lawmakers but in specific actions. Llewellyn, for example, asserted that 'What these officials do about disputes is, to my mind, the law itself' (Llewellyn, 1996: 3). Jerome Frank even posits that the nature of law remains unknown until a judge rules on a particular case, asserting, 'The law, therefore, consists of decisions, not of rules. If so, then whenever a judge decides a case he is making law' (Frank, 2009: 138), and 'rules, in all of this, are important to you so far as they help you see or predict what judges will do or so far as they help you get judges to do something. That is their importance.' (Llewellyn, 1996: 9). In essence, legal rules are not genuine *rules* per se but are *predictions* of legal effects, a set of probabilities tied to sanctions. Understanding the law equates to recognising the *likelihood* of judicial punishment for certain actions, rather than comprehending a definitive *reason* justifying both the judge's and our actions. As articulated by O.W. Holmes:

> If you want to know the law and nothing else, you must look at it as a bad man, who cares only for the material consequences which such knowledge enables him to predict, not as a good one, who finds his reasons for conduct, whether inside the law or outside of it, in the vaguer sanctions of conscience. (Holmes, 2007: 22)

A quintessential assertion of legal realism is the scepticism regarding whether legal provisions established by legislators represent the real law. It is divergent, however, that scholars have opinions on the actual form of 'real law'. Scandinavian Legal Realists typically contend that legal rules are non-existent, primarily on the grounds that they cannot be empirically verified. On the other hand, American legal realists generally maintain that legal rules are dynamic and not rigidly confined to codified statutes, echoing Holmes' iconic statement:

> The life of the law has not been logic, it has been experience. The felt necessities of the time, the prevalent moral and political theories, intuitions of public policy,



doi:10.6342/NTU202400318

> avowed or unconscious, even the prejudices which judges share with their fellow-men, have hade a good deal more to do than the syllogism in determining the rules by which men should be governed. The law embodies the story of a nation's development through many centuries, and it cannot be dealt with as if it contained only the axioms and corollaries of a book of mathematics. In order to know what it is, we must know what it has been, and what it tends to become. We must alternately consult history and existing theories of legislation. But the most difficult labor will be to understand the combination of the two into new products at every stage. (Holmes, 1951: 1)

It would be a mistake to interpret Holmes's critique as a narrow dismissal of the role of logic in law. Logic, in essence, forms but the foundational framework of our thought processes, as opposed to imposing constraints on legal research. Holmes's criticism is specifically aimed at the rigid confinement of legal research to positive law, rather than logic itself. This limitation steers the pursuit of legal knowledge towards conceptual dissections of law's internal framework, consequently forging a coherent and self-contained logical system.

In this discourse, I elect to commence with Holmes as the foundational reference. The rationale for this choice stems from Holmes's exemplary formulation of the prediction theory, which serves as an apt springboard for both critiquing and advancing the difference-making-based theory of reasons. Concurrently, by building upon Holmes's theory, it becomes feasible to engage with and respond to the perspectives of Scandinavian legal realists and, indeed, Hart.

Holmes draws a discerning line between the notions of 'prediction' and 'reason', positing that a 'bad man' fixates solely on the likelihood of sanctions, in contrast to a 'good man', who would delve into the reason behind the law. While Holmes's demarcation is rooted in ordinary language, there is scope to broaden the concept of 'prediction'. My contention is that the law, when viewed as 'reason', remains inextricably linked to Holmes's 'bad man theory'.

In causal models, the term 'prediction' pertains merely to discerning data relevance, without encompassing the notion of 'counterfactuals.' This, however, presents us with a predicament: data relevance alone fails to elucidate the causal nexus between an



doi:10.6342/NTU202400318

action and its consequent outcome. For instance, we might erroneously conclude that cancer patients ought to eschew chemotherapy to diminish mortality risk, basing this assumption merely on correlation. Absent an 'explanation' of data relevance, our 'predictions' regarding the repercussions of actions risk being skewed. Furthermore, as 'counterfactuals' are necessary for 'explanation', only predictive probability data cannot truly explain 'action'. Consequently, a 'bad man' cannot 'predict' sanctions based solely on data; instead, he requires 'reasons' to know whether he will be sanctioned or not.

The 'good man's' reason does not find its place within the legal framework, as the law does not construct reasons in such a manner. When the good man seeks a reason for the rule of law, his quest is essentially for a reason external to the law. This sought-after reason is not a individual reason, but rather a collective one, necessitating collective attainment. Since a collective reason relies on communal effort for realisation, it inherently bestows the capacity to critique others. The law, however, constructed on the basis of sanctions, embodies individual reasons, inherently regarding individuals through a lens that presupposes their 'bad' nature.

Pure empiricism proves unattainable[15], as interpreting data necessitates pre-existing background knowledge. If such knowledge is to elucidate facts, it must encompass insights about 'counterfactuals.' This knowledge originates from the reasoner's own pre-understanding, encompassing prior observations, life experiences, and preconceptions shaped by education and culture. Actors employ this pre-understanding to interpret data and determine their course of action. Within the realm of jurisprudence, this pre-understanding manifests as legal dogmatic theory. Hence, legal rules do not genuinely exist in the real world but rather reside within the theoretical construct devised by jurists, rendering them *epistemological* rather than metaphysical in nature. The theory espoused by Scandinavian legal realists, in fact, represents 'jurisprudence devoid of theory'. Such jurisprudence, bereft of the counterfactual element and thus unable to provide a realistic explanation of action, lacks practical deliberative usefulness.

[15] In the domain of statistics, an undue emphasis on empirical data, whilst neglecting theories for empirical interpretation, gives rise to a plethora of paradoxes, including the notable Simpson's Paradox. For further insight on this matter, one might consult Pearl's work (2009: 173-200).



doi:10.6342/NTU202400318

The *converse interpretation* demonstrates the counterfactual element of legal dogmatic theory. Let us revisit Article 92(2) of the Civil Code of Taiwan, which reveals such a rule:

> *The revocation of an expression of intent on the ground of fraud*→cannot *be a valid defense against the bona fide third party.*

According to Definition 2.2.1, we get:

> *The revocation of an expression of intent* not *on the ground of fraud*→can *be a valid defense against the bona fide third party.*

In other words, for the converse interpretation to be substantiated, legal provisions are required to encompass counterfactual elements.

The legal rules advocated by American legal realists are fundamentally practice-based. Viewed through a causal graph, Figure 3.3.4 illustrates legal rules derived from practices, with $\varphi$ representing the behaviour of individuals, $G$ denoting the judge's adjudication, which subsequently leads to the sanction $\psi$. Hence, the adjudication by the judge enacts a legal rule and thus *constitutes a reason*. The rules here are not reason-giving facts but *reason-facts*. Conversely, Figure 3.3.3 depicts the legal rules as proposed by Hart. In Hart's perspective, it is the legislator who enacts the rules, which in turn provide the reasons. Herein, the rules are not reason-facts; rather, they are *reason-giving facts*. For Hart, $R: \varphi \rightarrow \psi$ symbolises an acceptance of rule, with the underpinning reasons for this acceptance being multifaceted. However, these reasons are not legal reasons; rather, they are natural reasons grounded in social contexts. The legal reason, by contrast, is artificial constructs–that is, facing sanctions–which is not listed in the reasons for acceptance by Hart.

Legal reasons are inherently individual, lacking the capacity to levy criticism upon others. Yet, for Hart, why does adopting the internal point of view let individuals use legal rules as benchmarks for evaluating others' conduct? It may be, as Hart suggests, that the motives for obeying the law are diverse. At times, our compliance with the law may seem apparent, yet it is not necessarily rooted in the legal reason itself; we might act based on alternative reasons. As Hart elucidates, these could include 'calculations of long-term interest; disinterested interest in others; an unreflecting inherited or



doi:10.6342/NTU202400318

traditional attitude; or the mere wish to do as others do,' where some reasons, such as striving for the collective common good or social justice, are collective reasons and can engender a critical stance towards others. Such reasons, albeit occasionally aligning with the law, are not in themselves legal reasons. Hart astutely observes the multitude of reasons underlying our adherence to the law. Thus, at times our obedience to the law occurs simply because our actions just align with it. However, the law intrinsically provides a reason, which is sanction.

The difference-making-based theory of reasons aligns more closely with Holmes's 'bad man theory', representing a hermeneutic stance. The hermeneutic essence of this theory is rooted in the legal dogmatic interpretation of the law, encompassing both principles and rules to construct a framework for operation. Methodologically, this theory is twofold: theoretical and practical. Theoretically, it delineates our process of practical reasoning, whilst practically, it equips us with knowledge applicable to making informed decisions regarding actions or inactions.

## 4.4 On Legal Sincerism

Every political and legal concept, even when not explicitly articulated, implicitly presupposes a theory of human nature, which fundamentally bifurcates into the *Theory of Evil Human Nature* and the *Theory of Good Human Nature*. Within the realm of law, legal sincerists staunchly contend that our acceptance of, and compliance with, the law is invariably driven by a sense of sincerity. This sincerity may stem either from support for the law's content or a belief in its authority, and it is evidenced by the way we evaluate ourselves and others in light of the law's contents. To be more precise, legal sincerism is defined as follows:

**Definition 4.4.1 (Legal Sincerism)**
A legal theory is *legal sincerism*, if and only if the legal reason advocated by this theory is the *collective reason*.

On the contrary, the opposite concept is defined as follows:

**Definition 4.4.2 (Legal Deviousism)**
A legal theory is *legal deviousism*, if and only if the legal reason advocated by this theory is the *individual reason*.



doi:10.6342/NTU202400318

In various scholarly classifications, the concepts of 'evil' and 'good' manifest in diverse forms. Generally, 'evil' is construed from the standpoint of 'self-interest', often portrayed as expansive, unbridled, and encroaching upon others' interests. Such personal interests are sometimes greed and sometimes represent human vulnerabilities —insecurity, inferiority, anxiety, and fear—triggering correspondingly compensatory behaviours—self-preservation, arrogance, impulse, and aggression. As these interests lack a public dimension, they are deemed purely individual reasons. Conversely, 'good' is characterised by 'altruism' and embodies a certain public aspect. Often, *rationality*, too, is depicted as public, stemming from its presupposition as a universal human trait. Within this framework, the 'evil' is frequently equated with *irrationality*. However, both 'evil' and 'good' are rooted in reasons, albeit of different natures. Consequently, while 'evil' is understandable, but unsympathetic.

The most straightforward argument positing legal reason as a form of good stems from natural law theory. A proposition generally supported by these theorists is *the natural law thesis*.

**The Natural Law Thesis**

Law is necessarily a rational standard for conduct. (Crowe, 2016: 91)

which can be traced back to Aquinas's view that 'Law is a dictate of reason.' Finnis categorises a 'rational standard' as a form of *common good*. He posits that this common good is attainable exclusively through the mechanism of law. It is this common good, inherent in the law, that lends practical reasonableness to our actions. In other words, the presence of the common good within the law is imperative for a perfect explanation of our actions. In the absence of such a common good, our actions are inherently flawed. Consequently, a law devoid of this common good is fundamentally deficient. However, it is critical to note that a deficient law does not necessarily equate to its invalidity (Finnis, 2011 :363-366).

Finnis adopted Hart's notion of the internal point of view, positing that legal philosophers are tasked with identifying the important attributes of law from a specific practical point of view, thereby distinguishing the core of certainty and the penumbra of doubt within the concept of law. Drawing from the attributes inherent to the internal point of view, which enable the guidance and critique of both oneself and others, Finnis



doi:10.6342/NTU202400318

astutely identified the common good as the core of certainty within the law.

Post-Hartian legal positivism, on the other hand, while not explicitly endorsing the notion that the source of law is the common good, nonetheless places significant emphasis on the law's public character. Shapiro (2017), for instance, contends that the normativity of law stems from a collective plan, an argument commonly referred to as 'the planning thesis'

**Planning Thesis:**

Legal activity is an activity of social planning. (Shapiro, 2011: 195)

Contrasting with the notion of the common good, the legal plan is positive and constitutes a publicly accessible standard, in order to 'guide, organise, and monitor the behavior of individuals and groups' (Shapiro, 2011: 200). Furthermore, the legal plan itself acts as a *de facto* authority. This means that once the legal plan is established, legal practitioners are absolved from the need for further deliberation, negotiation, or bargaining (Shapiro, 2017: 8-9). Additionally, unlike Hart articulates, there is no requisite for legal actors to accept the law's content as a standard for evaluating actions. Instead, what is required is merely a 'commitment to a division of labour' within the framework of the legal project. This 'commitment' does not necessitate an authentic 'acceptance' of the content of the project but simply demands adherence to specific procedural rules. This implies that an attitude of acceptance does not justify compulsory, nor legal authority. Such a project can be disassociated from the thoughts or intentions of any specific individual and is, in essence, a mere matter of fact.

The purpose of legal planning is to facilitate collective achievements that are beyond the scope of individual endeavours. For Shapiro, such planning is a moral aim. In pre-legal societies, these aims were traditionally attained through collective actions driven by social customs, traditions, consensus, and promise. Yet, as society expands, these collective actions, due to escalating arbitrariness and the rising costs of coordinating actions, may fail to deliver the anticipated benefits. Hence, the concept of 'law as social planning' emerges as a solution to this quandary. Nonetheless, the assertion that 'the law harbours a moral objective' does not imply that 'the law must invariably fulfil this moral aim,' nor does it suggest that 'the law ought to pursue a specific moral aim.' It also does not infer that 'the law becomes null and void if it falls short of achieving its



doi:10.6342/NTU202400318

moral aim.' Instead, 'The law harbours a moral objective' merely 'makes the law *the law*'. (Shapiro, 2011:214)

Comparing these two theories, although one belongs to natural law theory and the other one belongs to legal positivism, their remarkable similarity becomes apparent. For both, the law represents a collective purpose. The term 'moral aim' is fundamentally synonymous with 'common good', and while the absence of such a collective purpose does not render the law void, it certainly subjects it to moral censure. This argument of collective purpose envisions law as a kind of collective reason, a public value meriting collective endeavour. Fundamentally, what is essentially appealed to is *the good nature of law*.

In contemplating how such a collective project aligns with individual reasons, we must revisit the divide between collective and individual reasons. For the difference-making-based theory of reason, the application of collective reasons to individuals necessitates a thorough examination of:

1. Whether collective planning will trigger a individual reason for violation of collective planning? and
2. Whether individuals in reality possess both the willingness and capability to execute the plan?

However, in a context where no entity possesses omniscience, realising this is nearly unattainable. There exists an intellectual gap between the subjective suppositions of rational planners and the actual reasons individuals might have for executing plans. It is highly probable that, in practice, the individual's reason for executing the plan may not align perfectly with the reason underpinning the collective plan. Occasionally, an individual's actions could diverge from the intended plan, thereby ultimately leading to the collective plan's shortfall.

Indeed, the practical failure of planning does not negate that the law cannot contain social planning. Embedding plans within the law can be considered a defining characteristic of what constitutes a law. The issue at hand is, however, that what makes a law a reason is not driven by 'collective reason' but rather by 'individual reason', meaning that, according to Definition 4.4.2, the law does not look forward to good



doi:10.6342/NTU202400318

human nature.

The law stems from a fundamental distrust of human nature. Within the context of traditional Chinese philosophy, the concept of law is invariably linked to sanctions, serving as a countermeasure to the *Ritual system* (禮制) that lacks punitive measures. The Ritual system prioritises collective harmony, whereas the legal system addresses individual interests. In fragmented societies, the significance of law and sanctions becomes increasingly paramount. The emergence of *fajia* (法家) was in the wars after the breakdown of the ritual system. This shift underscores the inability of the ritual system, rooted in collective reason, to sustain itself amid the collapse of the community. Human beings with different goals and interests made ancient China eager to explore a new system at that time, and the sanction of resorting to personal interests was the new solution. Legalist cancellation of the Ritual system must be interpreted from the opposition between collective reason and individual reason. The validity of collective reason wanes with the disintegration of community, supplanted by the ascendancy of state power. As Carl Schmitt aptly stated, 'all genuine political theories presupposed man to be evil' (Schmitt, 2007: 61), an equally remarkable and disquieting diagnosis also echoes in legal theory: all genuine legal theories presupposed man to be evil. Both concepts political and legal are built upon a foundation of human pessimism. During China's Warring States Period, Legalists masterfully fused these two aspects, thereby transforming law into a technique for political governance. Holmes's metaphor of the 'bad man', too, exemplifies such legal pessimism, suggesting that laws are essentially designed to govern and serve the bad men. When jurists tried their best to find an operatable and applicable legal basis for cases, they also constantly left room for the cunning of bad men. In later periods of ancient China, consequently, the court, more and less, reverted to the Ritual system again, moved away from strictly positive laws, and vigorously suppressed lawyers, to give the 'common good' more room for non-rule discretion.

As Schmitt underscored, the presumption of evil human nature is a fundamental premise in political practice. Similarly, the evil human nature is intrinsically linked to the function of law. Observing the works of legal theorists like Hart, Dworkin, Green, Finnis, and Schapiro, it becomes apparent that a greater focus on collective reasons leads to a corresponding reduction in the role of sanctions in law, and ultimately, to the



doi:10.6342/NTU202400318

diminution of the law itself. Furthermore, theories accentuating the importance of community often tend to diminish the role of the state and law in society. The function of the law, if examined in terms of a stable community, tends to confuse other reasons with legal ones. To understand the function of the law, we have no choice but to think in terms of an exception: when a community has collapsed, with political intensification and confrontation permeating the entire community, leading to a pervasive sense of insecurity and distrust, the people can no longer find common standards as the basis for mutual trust and communication. When every dialogue is misunderstood and all good intentions are doubted, society has, without a doubt, reverted to its state of nature. In such times, the only way to re-establish order within this community is through the imposition of strong sanctions. Only through such measures will the people seriously consider their own interests and return to a state of social order. This is why the law emerges.

Without sanctions, we cannot differentiate laws from other forms of class or club regulations. We can also see that there is no law in a class or club–because a harmonious community does not need laws. In reality, the closest parallel to a legal society might be a mafia, which asserts control over a territory and levies protection fees. However, such violent organisations, irrespective of whether they meet the criteria of independence as per Theorem 3.2.2, are perceived as threats because they pose a challenge to the state's capacity to constitute reasons, described by the state as illegal and a being that needs to be eliminated.

Every legal governance system commences with an initial conquest by force. However, this does not imply that sheer violence can sustain a society indefinitely. The costs incurred from violent control frequently compel the state to seek alternative methods for maintaining the community. Consequently, following a phase of aggressive conquest, the state often initiates cultural education to mould a community compatible with the state. This approach replaces the force of law with the collective reasons inherent within the community, rendering harsh sanctions unnecessary. On the contrary, a country that cannot shape its community tends to resort to the violence of sanctions. As the Chinese adage goes, 'In tumultuous times, severe punishments are employed'



doi:10.6342/NTU202400318

(亂世用重典). Escalating penalties is a typical response to social unrest[16].

## 6. A Summary of the Argument

The purpose of this paper is to revisit a theory that has been overlooked by mainstream legal philosophy: the theory of sanctions. Specifically, this article aims to advocate a proposition: that *the sanction is an independent way for laws to constitute reasons*. In this context, the terms 'Sanction', 'Independent', 'Constitute', and 'Reason' each carry distinct meanings. Firstly, 'Reason' refers to the causal path from action to outcome, with the actor having a specific attitude towards the difference in outcome. Secondly, 'Constitute a Reason' follows Definition 3.3.2, signifying the artificial creation of a previously non-existent causal path, rather than the provision of an existing one. Thirdly, 'Independent' suggests that the reasons for causation must be self-contained and not dependent on other types of causation. Lastly, 'Sanction' is defined as an artificially created unfavourable difference.

Methodologically, this article does not delve into the concept of legal pursuit, nor does it view this concept as the normative source of law. Instead, it examines the normativity of law from a practical standpoint: how we actually think and the patterns of interaction among people, particularly legal entities. The normativity discussed here is not an abstract ideal to be pursued but rather a tangible pattern of thought and interaction already functioning in reality. However, this approach may lead to challenges in defining sanctions: not everyone harbours unfavourable attitudes towards legal sanctions. Some individuals may be indifferent, while others might even view them favourably. For these people, the law lacks the normative force that it ought to possess, and thus the sanctions do not align completely with 'reality'. Nonetheless, legal sanctions cannot be determined by individual subjective feelings alone, given the diversity of attitudes towards different situations. Legal sanctions must be established

[16] In penology, this is known as deterrence theory, which has always been criticised by humanists who advocate rehabilitation, correction, and social support mechanisms. Nevertheless, the irony lies in the fact that the latter approaches are still fundamentally anchored in 'sanctions.' Criminals must be stigmatised and marginalised first before receiving societal aid, and the establishment of social support mechanisms necessitates the implementation of systemic sanctions, encompassing employment and performance frameworks. The sole 'humane' aspect in this context is that the burden of sanctions is not exclusively shouldered by the criminal, but is instead shared by society as a whole. Nevertheless, it is still essentially based on sanctions.



doi:10.6342/NTU202400318

based on differences that the majority perceives as unfavourable. Take the statement 'home-made grappa is harmful to your health'. Not everyone will normatively agree that 'you should not drink home-made grappa'. For those intent on self-harm or suicide, this statement might be the very reason they choose to drink it. Similarly, sanctions, such as the death penalty, can be an incentive for those who are bent on revenge against society and seek death. However, no matter what reason the actors obey the law, or how they obey the law in accordance with legal reasons, a generally applicable legal reasons still exists.

Theoretically, this thesis builds upon and extends Linton Wang's Difference-Making-Based Theory of Reasons. It posits that reason is an explanatory inversion of a cause and, thus, fulfils two functions. The first function is explaining the action; this is a retrospective counterfactual relationship, necessitating a reflection on past actions and imagining 'if we had done, we would have...'. The second function is guiding the action, which is a prospective counterfactual relationship, requiring us to look to the future and imagine 'if we do, we will...'. In English conditionals, the grammatical distinction between explaining the action and guiding the action is evident. However, as both satisfy Definition 2.2.1, both constitute counterfactual as well as causal relationships.

In terms of argument, this thesis answers the three questions:

1. What is the nature of reasons?
2. What kind of reason is the law?
3. How does the law serve as reasons?

The answers to these questions will constitute an argument in favour of sanctions as a source of legal normativity. This argument can be encapsulated into the following four theses:

i. The reason is a causal path from the action to its outcome.
ii. Legal reasons are individual reasons, not collective reasons.
iii. The law constitutes reasons through sanctions.

**The reason is a causal path from the action to its outcome**

Reason is a fact that can 'explain' and 'guide' actions. To fulfill this function, it must incorporate the concept of the counterfactual. A commonly used counterfactual concept



doi:10.6342/NTU202400318

is canonical difference-making. This thesis considers such difference-making to be a causal relationship. The causal relationship alone, however, cannot fully explain the normative status of an action. In addition to the causal link from action to outcome, the actor's attitude towards the outcome is also crucial for comprehensively explaining and guiding the action. Therefore, reason refers to the causal relationship between an action and its result, with the actor possessing a specific attitude towards this outcome.

**Legal reasons are individual reasons, not collective reasons**

A reason is a cause-and-effect relationship from an action to its outcome. Furthermore, such causal relationships, depending on the number of actors involved, differentiate between collective reason and individual reason. Collective reason necessitates the cooperation of many individuals to achieve the goal it pursues, thereby endowing it with the capacity to critique others. In contrast, individual reason relies solely on one's own actions for fulfilment and, therefore, lacks the capacity to critique others. Both collective and individual reasons are considerations for individuals when taking actions. However, this thesis identifies three gaps between collective and individual reasons, leading to challenges in the stable and continuous application of collective reason in society. To address this issue, society implements sanctions. The reason based on sanctions is a form of individual reason. Functionally, sanctions can perfectly replicate the individual reason in all normative states and emulate the collective reason in all states of order. In a large-scale society, sanctions have the potential to supplant collective reason in maintaining stable and continuous societal practice. Thus, as a mechanism for upholding large-scale social order, law must be categorised as a type of individual reason, specifically one based on sanctions.

In the article *Normativity of Rules* (規則的規範性, 2015), Peng-Hsiang Wang, drawing on Joseph Raz's argument, views rules as a reason-giving fact. He contends that the genuine reason for obeying rules is a collective one, which is triggered by rules and transformed into individual reasons through Theorem 2.4.3. This perspective underscores the criticalness of legal rules, aligning with what Hart describes as 'acceptance'. However, this thesis highlights that when collective reasons are implemented in society, they encounter three fundamental gaps with individual reasons –the explanatory, logical, and intellectual gaps. These gaps prevent collective reasons from forming a stable order, suggesting that they do not accurately represent how law



doi:10.6342/NTU202400318

is practised in society.

It is crucial to clarify that it is still feasible for us to act, or to criticise others, based on collective reason. For instance, while riding a motorcycle, we might yield to pedestrians to foster good traffic order, or criticise others for failing to do so. However, these are not the reasons that genuinely govern societal behaviour; they are merely our personal assumptions about collective reasons. Though such ideas might gain widespread acceptance through education, communication, and dissemination, the law does not rely on this consensus and consequently must establish sanctions. If not, there would be no necessity for laws at all. *The law stems from a fundamental distrust of human nature.*

**The law constitutes reasons through sanctions**

This thesis highlights that for the law to satisfy the separation thesis, the way it serves as a reason must fulfil three criteria: constitutiveness, deliberativity, and independence. 'Constitutiveness' implies that the law must create a new reason rather than merely providing an existing one. In this sense, obeying an order does not mean that the order itself constitutes a reason; rather, it triggers an existing reason that represents the relationship between the issuer and the recipient of the order. Thus, an order can only provide a reason, not constitute one. 'Deliberativity' requires that sanctions must exclusively determine the constitution of reasons, without the influence of other actors. Furthermore, the implementation of sanctions must be weighed against other existing or potential options, not merely based on emotional or personality-driven reactions. 'Independence' dictates that the legal system must not be subordinate to other systems of reason. However, since the sovereign is the origin of all sanctions, the reasons behind the sovereign's sanctions cease to be legal reasons based on sanctions. Instead, they become reasons external to the law, relegating legal reasons to other systems of reasoning. Utilising these three concepts, the thesis demonstrates that the law, through adjudications, in fact, establishes sanctions to constitute reasons, and that 'the command of the sovereign' is inconsistent with the separation thesis. Moreover, many reasons for obeying the law actually stem from sources other than the law itself, including authority, principles, rules, social planning, and the common good. These reasons are either collective or merely provide reasons external to the law.



doi:10.6342/NTU202400318

# Reference


Austin, J. (1995). *The province of jurisprudence determined.* Cambridge: Cambridge University Press.

Ayer, A. J. (1952). *Language, Truth, and Logic.* New York: Devor Publications.

Beerbohm, E. (2012). *In Our Name: The Ethics of Democracy.* Princeton: Princeton University Press.

Bix, B. (2006). Legal Positivism and "Explaining" Normativity and Authority. *American Philosophical Association Newsletter*, pp. 5-9.

Broome, J. (2004). Reasons. In P. P. R. Jay Wallace, *Reason and Value: Themes from the Moral Philosophy of Joseph Raz* (pp. 28-55). Oxford: Clarendon Press.

Crowe, J. (2016). Natural Law Theories. *Philosophy Compass*, pp. 91-101.

Bonevac, D.; Dever, J. & David, S. (2006). The Conditional Fallacy. *Philosophical Review*, pp. 273-316.

Dworkin, R. (1986). *Law's Empire.* Cambridge: Harvard University Press.

——. (2006). *Justice in Robes.* Cambridge: Harvard University Press.

——. (1967). The Model of Rules. *The University of Chicago Law Review 35*, pp. 14-46.

Enoch, D. (2011). Reason-giving and the law. In Green L. & Leiter B. (eds.), *Oxford Studies in Philosophy of Law*. Oxford University Press.

Frank, J. (2009). *Law and the Modern Mind.* London: Routledge.

Green, L. (2008). Positivism and the Inseparability of Law and Morals. *New York University Law Review*, pp. 1035-1058.

Hart, H. L. A. (1994). *The Concept of Law.* Oxford: Oxford University Press.

Holmes, O. W. (1951). *The Common Law.* Boston: Little, Brown and Company.

——. (2007). The Path of the Law. In Kennedy, D. & Fisher, W. W. (eds.), *The Canon of American Legal Thought* (pp. 19-44). Princeton: Princeton University Press.

Kahneman, D., & Miller, D. T. (1986). Norm Theory: Comparing reality to its alternatives. *Psychological Review, 93*, pp. 136-153.

Kelsen, H. (1941). The Pure Theory of Law and Analytical Jurisprudence. *Harvard Law Review, Vol. 55*, pp. 44-70.

Kirchmann, J. (1848). *Die Werthlosigkeit der Jurisprudenz als Wissenschaft : ein Vortrag, gehalten in der juristischen Gesellschaft zu Berlin.* Berlin: Springer, 00.00.1848.

Llewellyn, K. N. (2012). *The Bramble Bush.* New Orleans: Quid Pro Books.

MacCormick, N. (1973). Legal Obligation and the Imperative Fallacy. *Oxford* (pp. 100-129). Oxford Essay in Jurisprudence. Second Series: Clarendon.




doi:10.6342/NTU202400318

Martin, C. B. (1994). Dispositions and Conditionals. *Philosophical Quarterly 44*, pp. 1–8.

Pearl, J. (2009). *Causality 2nd. edition.* Cambridge: Cambridge University Press.

Perry, S. R. (1995). Interpretation and Methodology in Legal Theory. In A. M. ed., *Law and Interpretation* (pp. 97-136). Oxford: Oxford University Press.

Postema, G. J. (1998). Jurisprudence as Practical Philosophy. In *Legal Theory* (pp. 329-357). Cambridge: Cambridge University Press.

Raz, J. (1979). *The Authority of Law: Essays on Law and Morality.* Oxford: Oxford University Press.

——. (2009). Reasoning with Rules. In *Between Authority and Interpretation: On the Theory ofLaw and Practical Reason* (pp. 203-219). Oxford: Oxford University Press.

Scanlon, T. M. (1998). *What We Owe to Each Other.* Cambridge, Mass: Harvard University Press.

Schauer, F. (2010). Was Austin Right after All?: On the Role of Sanctions in a Theory of Law. *Ratio Juris, Vol. 23*, pp. 1-21.

Schmitt, C. (2007). *The Concept of the Political*, translated by Schwab, G. Chicago and London: The University of Chicago Press.

Shapiro, S. J. (2006). What Is the Internal Point of View. *75 Fordham. L. Rev.*, pp. 1157-1170.

——. (2011). *Legality.* Cambridge: Harvard University Press.

——. (2017). The Planning Theory of Law. *Yale Law School, Public Law Research Paper No. 600*, p. Available at SSRN: https://ssrn.com/abstract=2937990 or http://dx.doi.org/10.2139/ssrn.2937990.

王一奇（2015），理由與提供理由的事實，收於：謝世民主編，理由轉向：規範性的哲學研究，頁 106-139，臺北：臺大出版中心。

王鵬翔（2008），規則是法律推理的排它性理由嗎？收於：王鵬翔主編，2008 法律思想與社會變遷，頁 345-386，臺北市：中央研究院法律學研究所籌備處。

——（2015），規則的規範性，收於：謝世民主編，理由轉向：規範姓之哲學研究，頁 325-356，臺北：臺大出版中心。



doi:10.6342/NTU202400318