第二版

當代社會研究法

質化與量化取向

W. Lawrence Neuman 著

Social Research Methods: Qualitative and Quantitative Approaches

王佳煌、潘中道、蘇文賢、江吟梓 譯

PE 學富文化事業有限公司

在這本備受推崇的著作裡，作者以全面而且均衡的方式介紹了社會研究中的質化與量化兩種研究取向，並且也強調了結合各種不同取向所能帶來的益處。本書中亦包含了數十個來自於實際研究的新範例，能夠幫助讀者瞭解各種概念以及研究方法。

「對於任何一位作者而言，要同時詳細地檢視質化以及量化研究方法，都是一項費力而艱鉅的任務。然而，本書作者卻漂亮地完成了這項任務，不但在整本書中以均衡的份量來分別討論這兩種研究方法，而且也突顯出這兩種研究方法相互重疊以及彼此歧異之處。」

——雷提克（Bethany L. Letiecq）

美國蒙大拿州立大學

如果想知道為什麼你應該購買這本新版的《當代社會研究法》，以下是五個你非買不可的好理由！

1. 在整本書裡，作者已經加強以及擴充了關於更廣泛的社會研究方法論議題的討論內容，而不只是狹隘地探討各種研究技術的應用方式而已。其中包括了整合量化以及質化取向、把社會研究置入於更大的科學事業的脈絡之中、以及解釋文化與文化信念對於進行社會科學研究的關聯性。
2. 作者更新了關於社會研究各種認識論基礎的討論內容，並且在其中加入了出現在近期文獻中的新議題以及新觀念。
3. 作者的寫作方式能夠鼓勵學生積極投入於內容當中，因為作者強調了讀者作為研究者的身分，而且也把讀者放置於一個決策制訂者的位置上。除此之外，本書中有數個章節的開頭部分都是一項研究，而且在整個章節內容中都會一再提及該項研究，以作為證明某些概念的實際例子。在內文中的各種圖形、表格、範例、補充資訊等等，則依據主要的類型來重新加以組織，而且圖形也變得更為清晰易懂。
4. 關於社會科學裡的個案研究方法，本書已經增加了相關的討論內容，而且也進行了更為詳細的說明。此部分內容目前包含了個案研究對於建立理論的重要性，並且也增加了在文獻當中所提到的關於個案研究的特殊重要性。
5. 本書第九章「實驗研究」以及第十章「調查研究」的內容，都已經大幅度地增加了最新的發展趨勢以及內容。第九章的新特色之一，是更為強調理論在實驗研究中的重要性，而第十章裡則增加了在調查研究中各種可能的誤差來源，以及在調查訪談中的最新發展，其中包括對於認知訪談的說明與解釋。

PE 學富文化事業有限公司
臺北市大安區和平東路二段118巷2弄20號
TEL/(02)2378-0358 FAX/(02)2736-9042

ISBN 978-986-6624-94-0
9 789866 624940

當代社會研究法

質化與量化取向

第二版

Social Research Methods

Qualitative and Quantitative Approaches

W. Lawrence Neuman

王佳煌、潘中道、蘇文賢、江吟梓 譯

PE 學富文化事業有限公司

國家圖書館出版品預行編目資料

當代社會研究法：質化與量化取向／ W. Lawrence Neuman 原著；王佳煌、潘中道、蘇文賢、江吟梓 譯 --二版.-- 臺北市：學富文化, 2014.1
面；　公分
譯自：Social Research Methods: Qualitative and Quantitative Approaches, 7th ed.
ISBN 978-986-6624-94-0（平裝）
1. 社會學 2. 研究方法
540.1　102020156

二版一刷 2014 年 1 月

當代社會研究法：質化與量化取向

Authorized Translation from the English language edition, entitled SOCIAL RESEARCH METHODS: QUALITATIVE AND QUANTITATIVE APPROACHES, 7th Edition, ISBN: 0205615961 by NEUMAN, W. LAWRENCE, published by Pearson Education, Inc, publishing as Pearson, Copyright © 2011.
All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.
CHINESE TRADITIONAL language edition published by PRO-ED PUBLISHING COMPANY, Copyright © 2014.

原　　著　W. Lawrence Neuman
譯　　者　王佳煌、潘中道、蘇文賢、江吟梓
發 行 人　于雪祥
發 行 所　學富文化事業有限公司
地　　址　臺北市大安區和平東路二段 118 巷 2 弄 20 號
電　　話　02-2378-0358
傳　　真　02-2736-9042
E - M A I L　proedp@ms34.hinet.net
編　　輯　蔡福根
印　　刷　上毅印刷有限公司
定　　價　830 元（不含運費）

ISBN: 978-986-6624-94-0

本書如有缺頁、破損、裝訂錯誤，請寄回更換

中文版權學富文化事業有限公司所有。未經許可不得以任何方式或形式進行複製或重製。翻（影）印必究！

PE

原序

隨著新版的面世，本書從第一版發行至今已經過了大約二十年的時間。在每一個版本中，我持續不斷地更新內容，並且進行一些小幅度的修正，而且我也相信，本書龐大而忠實的讀者們，也都很歡迎這些修正的內容。我不斷努力使本書成爲一本既全面、又可信賴的社會科學研究入門書，並讓初次接觸社會研究的學生們覺得本書不但易讀，而且實用。

本書比較偏向研究方法論，而不是研究方法。所謂的*方法*（*methods*），指的是研究設計、測量、資料蒐集、以及資料分析的技巧。而*方法論*（*methodology*）則包含了方法以及其他更多的內容。方法論所強調的是瞭解整個研究過程。與研究方法相較之下，方法論的焦點側重於一些核心的原則，並以此進行高品質的研究、以及學習對研究方法作出適切的判斷。在方法論中包含了察覺並認知下列三項事物：社會研究所處的社會—政治脈絡、社會研究在較廣大的科學社群和科學研究中所佔據的位置、以及社會研究對於社會事實與知識創造的假定。一本關於方法論的書，不能僅僅侷限於幫助學生成爲一個研究技術人員而已，還要能夠幫助學生發展成爲一個具有反思能力的社會科學家。

自從本書大約在二十年前初版以來，就在以下數個方面別具特色。首先，對於社會研究中的質化與量化取向，我都一視同仁並且全然接受。在進行社會科學研究方面，不論是質化或是量化取向，都具有其重要的價值，也都能夠促進我們的知識與理解。其次，我也討論到在社會研究中，各種不同取向背後的認識論議題。在各種不同的研究取向背後，都有支持該研究取向的基本假定，而藉由瞭解這些基本假定，我相信學生以及研究社群裡的成員都能看到，每種取向如何能夠成爲一個創造社會科學知識的基礎。而這些通常未被言明，且隱而不顯的假定，也能夠使我們看到，爲何每種取向會以目前的方式運作並且發揮作用，以及各種相異的取向如何能夠彼此互補。第三，我從近期的研究文獻中汲取出爲數眾多的範例，以闡明方法論的概念以及技術。我相信，只有在特殊而具體的研究中，方法論才能以活生生的方式展現出來。學生們學習方法論的最佳方式，就是看看那些從事知識創造的活躍社會科學家，是以何種方式把方法論應用於具體、特定的研究中。最後，統計以及量化資料分析的相關章節，仍然維持著只介紹基本概念的方式，也就是較爲著重概念性、而非計算性的議題。我認爲那些打算進行量化研究的學生，必須再另外修習一些與統計相關的課程，所以在本書中僅僅提供統計分析的基礎概念而已。

除了持續更新範例、以及在本書中融入特定技術或方法論辯論的最新進展之外，本書自從第六版以來，我也作了一些其他的改變，試圖使得本書的內容更加清晰易讀，並且讓學生更易於理解本書所提及的各種概念。我重新撰寫了數個章節，其中包括那些關於理論以及研究設計的章節。在數個章節的開頭部分，我則安排了研究範例。我之所以加入研究範例，是爲了使抽象的方法論概念變得更爲具體，且貼近

現實生活。另外，我也改變了書中內容的描述方式。現在，當我在進行描述時，我比較常談到我們身爲社會科學研究社群的成員所做的事，並且讓學生以及讀者們覺得，自己彷彿即將作出方法論方面的決定、以及即將規劃或是進行一項研究一樣。

本書的補充教材包括了以下數項：附有測驗題的教師手冊、自我測驗的題庫（My Test Test Bank）、線上課程的多媒體簡報檔案、學生版 SPSS 軟體（購買本書的讀者可以用特別的價格申請購買）、以及研究工具箱（My ResearchKit）。在研究工具裡，提供了每章的學習目標、章節摘要、閃字卡、練習測驗、以及幫助學生學習和理解的影片和活動。

在本書的這個全新版本中，我仍然執著於我一直以來的信念，也就是要讓學生們知道，社會研究是個令人興奮而且充滿生命力的過程。真實的人們在「真實的世界」裡進行研究，以促進並且增廣我們的知識範疇。社會研究是一個充滿活力、有生命力的過程，並在此過程中發現或者是創造出知識。大多數的學生只要經過適當的努力和研習，就都能夠精通或至少瞭解此種過程。雖然專業的研究者有著高度的熱忱與承諾，但是他們也都只是普通人，並且也生活在一個真實的社會、政治、以及歷史的世界裡。

最後要說明的是，本書的目標群眾仍然是以全世界的讀者爲主。身爲研究者，我們必須察覺到文化的假定、信念、以及價值觀等等會如何影響到社會研究活動。由於我們進行社會研究的目的，是去檢視發生在真實人類社會中的議題，因此，我們必須要能敏銳地察覺到，在這些社會中的種族中心主義或是國族主義的壓力，可能會以何種方式限制了社會研究在促進理解方面的潛力。

譯者潘序

社會科學研究方法是研究（讀）社會科學相關領域（科系）人士或學生必須具備的專業知識。它包含了對社會現象存在的形式及知識建構的基本哲學思考、對社會現象探究的不同取向（量化與質化）、研究的設計與倫理、研究對象的選取、研究資料的蒐集與分析、研究報告的撰寫等等。它是一套以系統性、邏輯性、組織性的方法來理解社會現象的重要知識。然而，在許多探討社會科學研究方法的教科書中，經常忽略了對社會現象哲學思考的詳細討論，並且要不就以統計計量的角度來介紹量化取向的研究方法，要不就以詮釋爲基礎的觀點來撰寫質化取向的研究方法，兩者均難以綜觀社會科學研究方法的全貌，也容易落入研究方法二元對立的困境。

W. Lawrence Neuman 是美國威斯康辛大學社會學系的教授，他所著的“Social Research Methods: Qualitative and Quantitative Approaches”，不僅對社會科學研究方法所涵蓋的重要內容進行完整的說明，且將質化及量化研究方法的取向完整且精準的整合呈現於此書中，使得讀者容易將此兩種研究取向進行對比、及融會不同研究取向之研究方法的個別長處、優點、適用條件、及其分別之限制。此書已經發行七版，可見作者爲因應社會變遷及研究方法的最新發展不斷修正此書所付出之努力，也可看出此書受到廣大讀者之閱讀與支持。

學富出版公司于雪祥先生有鑑於此書在美國學術界所受到的肯定及其樂見國內讀者能有閱讀此書之機會，特別邀請我及王佳煌老師等國內熱衷於研究方法探討之譯者，將此書翻譯成中文。書中之翻譯若有涉及學術或專有名詞、或特定文化脈絡中所使用的詞語、及人名均附上原文，以利尋找相關的參考文獻；另外，由於此書由多位譯者共同翻譯，在多次討論及校稿中也力求中文語法使用上之前後一致連貫，並希望能達到信、達、雅之目標。

在此特別感謝學富出版公司全體同仁的努力及協助，使本書能夠順利出版，也希望譯者的努力能爲台灣社會科學研究方法知識的累積帶來貢獻。譯者們均已竭盡全力完成此書之翻譯，然而書中難免仍有錯誤之處，尚祈先進與讀者們的指正與建議，是所至盼。

國立暨南國際大學
社會政策與社會工作學系助理教授

潘中道

2013/10

譯者王序

一本學術書籍的好壞，可以從形式與實質兩方面來看，兩者都與書的品質密切相關。在形式方面，首先要看作者的專業及其形象與聲譽。作者的專業可以在某種程度上確保書的品質，因爲作者的形象與聲譽是一種口碑，代表同行對作者的肯定，不致於荒腔走板，外行裝內行。

書的好壞也要看出版社。出版社的出書策劃、選書策略與行銷方式可以顯示他們對某一本書的定位與期望，以及對作者的瞭解。出版社在業界的形象與名聲也是書籍品質的間接保證。大型出版社挾其規模經濟、龐大的編輯群與顧問，以及豐厚的人脈關係，可以說服好的作者將書籍交給他們出版。中小型出版社以專業取勝，雖缺乏規模經濟，也無力砸錢作廣告，但因他們必須在競爭激烈的出版市場上求生存，除了掌握市場利基之外，選書、出書也必須精準，才能維持形象與聲譽。

當然，書的好壞也要看它的銷售狀況。關於這點，專業書籍要想洛陽紙貴，比通俗書籍還要難上千倍，因爲通俗書籍要緊跟流行浪潮，還要看作者是否爲知名人物，專業書籍或教科書以專業爲導向，市場接受度遠不如通俗書籍，作者亦多非明星或政治人物，實難與通俗書籍相比。即使是教科書，也要看適不適合教師與學生使用。許多學者滿腹經綸，下筆千言，卻令讀者不知所云者，所在多有。

從上述這些形式條件來看，本書堪稱好書。作者 W. Lawrence Neuman 是美國威斯康辛大學社會學系的教授，鑽研方法論多年，本書是他的代表作，在學術界享譽甚隆。出版社 Allyn and Bacon 與 Longman 同屬 Pearson Education 出版集團，是知名的出版商，專營教科書、參考書、教材之出版銷售，在出版界與學術界口碑卓著。再者，本書出到第四版，也顯示其內容廣受教師與學生歡迎。

形式條件只是外在的，重要的還是看書的內容。市面上許多教科書，尤其是談研究方法的教科書，最容易令人味同嚼蠟，難以下嚥。作者如果要撰寫一本好的教科書，或者學生要選擇一本好的教科書，最好能注意三個要件：

第一，生花妙筆，平實易懂。

第二，架構完整，面面俱到。

第三，介紹基本概念，適時搭配範例，以便學生迅速吸收，活學活用。

讀者當可發現，本書已經具備這三個內容要件。

第一，作者用平易的語句與社會生活中的實例，論述基本概念，語調輕鬆活潑。

第二，全書分爲五大部分，包括基礎認定、籌劃與準備、量化資料搜集與分析、質化資料搜集與分析、社會研究與報告撰寫。從頭到尾，一氣呵成。值得注意的是，第四版的編排有兩個特色：一是討論研究設計時，兼論質化與量化，二是特別將倫

理問題提到前面。第一個特色意在打破傳統研究方法將量化與質化分開來看的簡化做法。許多學者常有的誤解是質化與量化屬於不同的邏輯與思維，卻忽略兩者其實都是科學方法，以致各分畛域，彼此攻訐。第二個特色與一般研究方法的教科書不同，因爲傳統或一般的教科書常將研究倫理放在最後面，聊備一格。但事實證明，研究倫理不宜放在最後一章，研究方法不只是技術問題，更是自重與尊重他人的問題。即使學者成就傑出，著作等身，只要有輕忽研究倫理之處，即使是小差錯，也可能使一世英名與成就毀於一旦。在學術研究的歷史上，這種例子比比皆是，卻總有人忘記歷史的教訓，最明顯的就是抄襲剽竊，忽略同僚、研究助理與學生的貢獻，或者暴露敏感的、隱私的資料，傷害到研究對象。此外，在利益衝突中不思迴避，標榜純正血統，結幫成派，各據山頭，壟斷學術資源，也是極容易犯的毛病。作者將研究倫理的篇章提到前面，可謂用心良苦。

第三，本書每介紹一項方法原則或概念，隨即有範例、圖示緊密搭配，讓讀者活學活用。用電腦的術語來說，本書的撰寫、編排充分達成「使用者友善」（user-friendly）的目標。

本書譯者對研究方法有濃厚的興趣，迻譯本書時受益良多，也希望與讀者分享這本好書。讀者對本書中譯本如有任何問題與意見，歡迎與譯者、出版者聯絡、交流、討論。

元智大學社會學系助理教授

王佳煌

草于癡愚齋

2002/1/18

譯者序

譯者本身在就讀研究所時，就曾經讀過 Newman 這本知名著作中的一些章節，當時譯者對於作者深入淺出地說明各種研究方法的能力，感到印象深刻，也從中獲得許多與研究方法相關的寶貴知識。沒想到在數年之後，承蒙學富文化事業發行人于雪祥先生的抬愛，竟有機會親自翻譯本書的中文版，讓譯者深感榮幸，也讓譯者終於有機會再一次親自領略這本知名著作的原文內容。

本書在研究方法的領域中堪稱是相當知名的著作，因爲作者在書中不但詳述了各種研究方法背後的哲學基礎與進行方式，而且也引用了許多範例與小故事來加深讀者對於各種研究方法的體會與認識。因此，不論是打算撰寫論文的研究生、或是想對研究方法有更進一步認識的一般讀者，本書都是極佳的參考資料。即使讀者本身沒有撰寫論文或進行研究的需求，閱讀本書亦可充實關於研究方法的知識。此外，當今的報章雜誌中常常充斥著各種似是而非的社會研究與結果，致使社會大眾在不知不覺間受到誤導而不自知，針對此種情況，作者在書中也舉出了許多誤用研究方法的例子、以及使用各種不同研究方法的適合時機，可以幫助讀者不會輕易受到劣質的社會研究所蒙蔽而作出錯誤的判斷、甚至有能力去辨別某項研究所使用的研究方法是否恰當。由此可見，本書的用處絕對不僅止於寫論文時的參考書而已，更應該是現代知識份子都應擁有的研究方法參考書之一。

由於本書篇幅浩大、而且內容相當廣泛，在翻譯本書期間，譯者除了盡心體會作者字字句句的原意之外，也不時上網或翻查手邊的參考資料，務必力求譯文能忠實傳達作者的意思。而本書上一版由王佳煌、潘中道教授主譯的中文版，對譯者而言更是提供了莫大的幫助。王、潘教授的譯文流暢詳實，許多專有名詞的中譯也極爲貼切，因此成了譯者最佳的參考對象，若非有兩位教授先前中譯版所奠下的良好基礎，譯者恐怕無法順利完成本書的新版中譯稿。所以，譯者在此要特別向王佳煌與潘中道教授致上最誠摯的謝意。

除此之外，本書的中文版能夠順利出版，必須特別感謝學富文化公司的于先生以及所有同仁辛苦地協助本書的編輯、校稿、印刷工作等等。而譯者家人的支持與鼓勵，亦是完成本書中譯版過程中不可或缺的助力，因此譯者也在此一併致上謝意。

最後，雖然譯者在翻譯過程中力求盡善盡美，對於原文的意義也在反覆推敲、確實瞭解之後才譯爲中文，但因本書內容極爲龐大，且譯者本身才疏學淺，故闕漏之處必在所難免，希望各方先進若有發現錯誤之處，能夠不吝指教，譯者也必虛心受教，以使這本研究方法領域中的大作能造福更多的讀者。

蘇文賢、江吟梓　謹識

目錄

第1部分　基礎

第1章　為什麼要進行研究？　1

社會科學研究之外的其他選擇　3
研究的內涵：科學化的取向　11
社會研究的類別　23
研究過程的步驟　25
爲什麼要學習進行社會研究的方法？　32
結論　33

第2章　社會研究的主要類別　35

研究的運用與讀者群　37
研究的目的　53
個案內或是跨個案研究　57
單一時間點或是多重時間點　62
資料蒐集技術　65
結論　73

第3章　理論與研究　75

什麼是理論？　77
社會理論與意識型態　79
社會理論的組成部分　84
結論　120

第4章　方法論的意義　123

哲學基礎　124
三種取向　128
實證主義社會科學　130
詮釋取向社會科學　138
批判取向社會科學　148
女性主義以及後現代研究　160
結論　163

第 5 章 如何回顧文獻以及進行合乎倫理的研究 169

文獻回顧 170

社會研究中的倫理 195

結論 219

第 2 部分 規劃以及準備

第 6 章 研究設計的策略 223

三角檢證法 224

研究的質化以及量化取向 226

質化設計議題 238

量化設計議題 243

結論 261

第 7 章 質化與量化測量 267

測量的必要性 268

量化以及質化測量 270

測量的過程 271

信度與效度 279

量化測量指引 292

量表與指數 298

結論 315

第 8 章 質化與量化抽樣 319

進行抽樣的理由 319

抽樣策略 321

結論 357

第 3 部分 量化資料蒐集與分析

第 9 章 實驗研究 363

適當的技術 364

實驗簡史 367

隨機分派 371

實驗設計的邏輯 375

內在與外在效度 386

實際的考量 398
實驗研究的結果：進行比較 399
與倫理有關的議題 400
結論 402

第 10 章 調查研究 405

調查研究的歷史 406
調查研究的邏輯 409
建構問卷 411
調查的類型：優點與缺點 443
調查訪談 450
合乎倫理的調查 462
結論 463

第 11 章 非反應式研究以及次級分析 469

非反應式研究 470
內容分析 473
現存的統計數據／文件 483
調查資料的次級分析 490
推論的議題以及理論考驗 498
倫理的考量 499
結論 499

第 12 章 量化資料分析 503

資料處理 504
單變項的結果分析 507
雙變項的結果分析 517
超過兩個變項 532
推論統計 540
結論 545

第 4 部分 質化資料蒐集與分析

第 13 章 田野／實地研究 549

瞭解田野／實地研究 550
田野／實地研究訪談 590

資料品質 598
田野／實地研究中的倫理兩難困境 601
焦點團體研究 603
結論 605

第 14 章 歷史比較研究 611

歷史比較研究簡史 613
適合歷史比較研究的研究問題 613
歷史比較研究的邏輯 615
進行歷史比較研究計畫的步驟 625
歷史脈絡中的資料以及證據 630
比較研究 641
歷史比較研究中的同義議題 659
倫理 663
結論 664

第 15 章 質化資料分析 669

資料分析方法的比較 670
編碼以及概念形成 673
質化資料的分析策略 683
其他的技術 703
結論 712

第 5 部分 與他人溝通

第 16 章 撰寫研究報告以及社會研究的政治 717

研究報告 718
社會研究的政治 739
客觀性以及價值中立 756
結論 758

附錄 763

漢英索引 765

參考書目請見本書隨附光碟

第 1 章

為什麼要進行研究？

社會科學研究之外的其他選擇
研究的內涵：科學化的取向
社會研究的類別
研究過程的步驟
為什麼要學習進行社會研究的方法？
結論

社會學家所關心的，就是以一種受過訓練而有紀律的方式來瞭解社會。而此種訓練的本質是科學的。換句話說，社會學家對於自己所研究的社會現象所發現的、以及所說的一切，都必須在一個定義相當嚴謹的參考框架中進行。

——Peter Berger, *An Invitation to Sociology*, p. 16

撰寫本書的目的，是爲幫助你們瞭解社會科學家如何進行研究，然後好讓你們有能力去進行自己的研究。在這一章裡，考量了兩個主要的議題：爲什麼要學習如何進行社會研究、以及社會科學研究的內容基礎。在後續的章節裡，會逐一探究社會研究的各種不同取向，並且會檢視當你在進行研究時會使用到的原理和技術。

社會科學研究在日常生活中無所不在，而且也會影響到自己、家人、朋友、鄰居、以及同事。社會科學研究所得到的發現，會出現在廣播新聞節目中、在雜誌和報紙上、以及在許許多多的網站和部落格裡。這些研究發現涵蓋了數十種主題和領域：法律以及公共安全、學校教育、健康照護、個人與家庭關係、政治議題、商業活動、以及國際和社會趨勢等等。當我們在處理與家人、朋友、以及同事之間的關係時；在參與社區活動或是公共政策時；或是在商業領域、職場生涯、以及健康照護方面作出日常的決定時，不論是以直接或是間接的方式，我們都會用到社會科學研究的知識以及原則。社會研究並非只與大學課堂或大學教授有關而已，事實上，不論是中等學校的教師、家長、企業負責人、廣告商、經理人、行政人員、政府官員、服務業者、健康照護專家、以及其他各行各業的人們，都會用到社會研究所得到的發現以及原則。人們利用社會研究的發現與原則來養育兒童、減少犯罪、管理健康問題、銷售產品或服務、以及消化吸收新聞事件等等。因此，社會科學研究的重要性可以說是無庸置疑的。雖然偶有一些持不同論調的批評意見，但在理解一般的社會生活方面、以及每天所作出的各種決定方面，社會研究仍然具有高度的關聯性與重要性。

爲了理解社會研究與現實層面的關聯性，讓我們來看看一對養育一個三歲孩子的夫妻的例子。有一項研究（Wrigley and Derby, 2005）發現，花錢請人照顧孩子是相當安全的，但是，該項研究同時也發現到，在各種不同的兒童照顧方式中，兒童的死亡率有著極大的差異。把兒童交由托育中心來照顧，會遠比交由私人保姆照顧安全得多。另一項研究（Bridges et al., 2007）顯示出，把兒童交由托育中心來照顧，能夠明顯地提高兒童在閱讀與數學方面的分數，但是卻會對兒童的社會行爲測量結果有負面的影響（舉例而言，托育中心裡的兒童比較不會展現出與他人合作的傾向，反而會展現出侵略攻擊的傾向）。比起較爲年幼或是較爲年長的兒童，年齡約二至三歲的兒童會獲得最多的益處。除此之外，若是父母積極並且經常與兒童進行互動的話，則能夠減輕兒童因爲接受照顧時而產生的負面行爲。還有另外一項研究（Love et al., 2003）則顯示出，托育中心在品質方面的差異極大。與兒童接受照顧的時間長短、以及是否由父母或是托育中心提供照顧等因素相較之下，照顧品質的良窳對於兒童所產生的影響是比較明顯而重大的。最後，另一項研究（Sosinsky, Lord, and Zigler, 2007）則發現到，一般而言由非營利組織或是非宗教組織所經營的托育中心的品質，會優於其他類型的托育中心。根據上述這些研究的發現，這對夫妻或許會決定去尋找某種特定類型的托育中心、花時間去確認這所托育中心的照顧品質、並且特別費點心思去鼓勵他們的孩子發展一些社交方面的技巧。上述這些研究的發現不僅僅和某些特定的家長切身相關而已，對於公共政策、以及一個社區如何處理兒童照顧議題等方面，這些研究發現也都能提供一些啓發與參考。

社會科學研究能夠產出極具價值的資訊、並且能夠擴展我們對於社會的瞭解，但是社會科學研究卻並非是百分之百不會出錯的。換句話說，社會科學研究並不保證每次都能得到完美的結果、或者是提供「絕對的事實」。或許這就是爲什麼有些人不信任那些以研究爲基礎所得來的知識、或者是爲什麼有些人（其中包括一些媒體上的評論家）甚至會譏諷專業的研究者以及研究結果的原因。雖然有這些嘲弄的意見與評論，在與我們用來瞭解世界與作出決定的其他替代方法直接比較之下，研究仍然輕易地穩居上風。而這也就是爲什麼當專業人士、受過教育的人士、以及負責任的領導者想要更加瞭解社會或者是作出重大的決定時，他們仍然會始終如一地參考社會研究的方法、原則、以及發現的主要原因。

本書將會討論到社會科學研究的方法論以及研究法。雖然這兩個詞彙可能看起來像是同義詞，但是方法論的內涵較爲廣泛，而且也含括了研究法在內。*方法論*（*methodology*）的意義是指去瞭解整個研究過程，其中包括了研究的社會－組織背景脈絡、哲學假定、道德原則、以及研究得到的新知識在政治方面所會產生的影響等等。而*研究法*（*methods*）所指的，則是我們在一項研究中用來選擇個案、測量與觀察社會生活、蒐集與處理資料、分析資料、以及報告結果等方面所用到的特定技術的集合。這兩個詞彙是彼此緊密相連、並且相互依賴的。

閱讀以及進行社會研究是一件令人興奮的事：社會研究是個發現的過程，而我們則在這個過程之中學習到許多新的事物。在進行社會科學研究時，研究者必須要擁有堅持到底的耐性、正直的個性、能夠容忍曖昧不明的狀況、與他人進行互動、

並且要爲自己正在進行最高品質的研究工作而感到驕傲。除此之外，要進行社會科學研究也必須要進行邏輯性的思考、仔細而小心地遵守規範、並且一再地重覆一些步驟。在研究的過程中，我們以一種有系統的方式將理論或想法與事實結合起來，而且有時候也必須發揮自己的創意。爲了要進行研究，我們必須進行安排與規畫的工作。我們必須去選擇適合於特定問題的研究方法。我們必須永遠都以合乎倫理道德規範的方式來對待研究中的參與者。除此之外，我們還必須把我們如何進行研究的方式、以及我們從研究中所獲得的結果傳達給他人知曉。

在本章之中，我們會考量社會科學研究之外的其他選擇，以及爲什麼進行研究會較爲受到偏好。我們接著會檢視整個科學研究系統的運作方式，其中包括了進行研究的步驟，以及社會科學研究的種類。

社會科學研究之外的其他選擇

在這個部分裡，我們要來看看除了社會科學研究之外，其他四種經常被人們所使用的替代方法，而許多人會依賴這些方法來獲取知識以及作出決定：

- 個人經驗與常識
- 專家以及權威
- 來自於大眾以及媒體的訊息
- 思想信念以及價值觀

來自於個人經驗以及常識的知識

如果某件事發生在我們身上，也就是我們親眼見過、或是親身體驗過某件事，我們就很有可能會把這件事當作是事實。個人經驗或是「眼見爲憑」的知識，是一種強而有力的知識類型。不幸的是，此種知識卻有可能會誤導我們。類似於視覺上的錯覺或是幻覺，有可能會發生在我們身上。雖然某件在表面上看起來爲真的事件，實際上有可能是來自於錯覺，但是，此種個人親身且直接感受某件事的影響力量是非常強大的，因而會使得我們甚至在完全不瞭解某件事的情況之下，就輕易地被錯覺所蒙蔽。這就是爲什麼有許多人堅持相信他們個人所經驗過的事件，而不願去相信他們在閱讀一份仔細進行的研究後所得到的知識的原因，即使該份研究設計已經極力避免個人經驗所會造成的錯誤時亦然。當研究結果與個人經驗或是一般常識有所矛盾時，此種情況會特別明顯。除此之外，個人經驗所造成的錯誤會彼此互相增強。有少數人甚至會蓄意地透過宣傳活動、錯誤的言論或刻意欺騙、魔術表演的手法、政治操作、以及廣告手法來誤導其他人。

整個研究次領域都在致力於發現我們作出錯誤判斷、高估或是低估、以及犯錯的方式。以下就是一個例子：女性傾向於持續使用一些根本沒有效果的肌膚保養乳霜。除此之外，某項美容產品或是療程的效果越差，女性就越有可能會持續使用這

項產品或是療程。這些發現是來自於一項針對 300 名年紀從 27 歲到 65 歲的女性所進行的研究，這些女性都在試著使用乳霜、維他命、以及其他美容療程來使自己的外表看起來更加年輕。這些發現和我們所可能認爲的結果並不相同：這些女性竟然最忠於使用那些根本沒有作用的產品或療程！在那些認爲美容療程沒有效果的女性當中，有百分之二十七的人停止接受療程。而在認爲美容療程具有效果的女性當中，竟然有高達百分之五十五的人停止接受療程。這些研究者認爲，女性之所以會持續使用某項沒有效果的產品，是因爲當人們對自己覺得不滿意時，恐懼會比成功具有更爲強大的驅動力量。由於害怕自己會看起來比實際年齡還老，因此即使當美容產品根本毫無效果時，這些女性仍然會持續地使用這些產品。[1]

雖然這些發現我們有誤判傾向的研究讀起來相當有趣，它們卻指出了一項普遍的原則：日常生活的推理方式以及觀察結果，不但是不完美的、而且也很容易出錯。更重要的是，我們幾乎很少會馬上就注意到、或是瞭解到自己已經犯下了這類的錯誤。

來自於個人經驗的知識、常識所告訴我們的「事實」、以及日常生活的推理方式雖然有可能是正確的，但是卻都有可能把我們引離正途（請參閱資訊箱 1.1：我們認爲自己將會去做的事，與我們實際上真正所做的事）。例如，常識告訴我們，在高中校園裡發放免費的保險套將會鼓勵青少年從事性行爲，或者，實施諸如死刑之類的嚴厲懲罰將能夠遏止暴力犯罪。然而，爲數眾多的研究指出，上述兩個想法都是錯誤的。大多數人都以爲，犯罪現場目擊者的證詞是正確無誤的，但是，許多研究卻顯示出，這些證詞都是極爲不正確的。我們之中有許多人會擔心發生不幸的意外或是可怕的事件，像是墜機或是校園槍擊事件等等。然而，我們所擔心的通常是那些「不太可能發生」的事件，這是因爲我們對於某件事情發生機率的估計，通常與經過仔細研究後所得到的實際機率相去甚遠。相同地，我們也有可能會被表面的現象所誤導。許多人會因爲安全的因素而購買一輛大型的、外表看起來非常堅固的休旅車，然而，碰撞測試與車禍記錄卻顯示出，比起許多看起來較爲溫和的車款，休旅車其實是比較不安全的。[2]

不正確的「常識」所產生的錯誤認知，將會造成重大的影響。除此之外，媒體通常會重覆並且散佈這些錯誤的認知，學校以及企業會以這些錯誤認知來作出決定，而且立法者和政治人物也會以這些錯誤的認知爲基礎來推動新的法令或是政策。我們經常會在日常生活的決定中犯下以下的五種錯誤，但是研究過程則是要盡力減少或消除這些錯誤。

- 過度推論
- 選擇性觀察
- 過早下結論
- 月暈效應

資訊箱 1.1　我們認為自己將會去做的事，與我們實際上真正所做的事

社會科學家注意到一個矛盾的現象：雖然大部分的人都強烈地譴責公然的種族歧視行為，但是，露骨的種族歧視行為仍然時有所聞。為了仔細審視這個現象，川上以及同事（Kawakami, 2009）進行了一項研究。他們認為，或許人們會錯誤地判斷自己在親眼目睹種族歧視時的感受以及行為。為了檢證這個想法，他們詢問了一些非黑人學生，如果在他們面前發生了種族歧視的行為時，他們會有什麼感受、以及他們認為自己會採取什麼行動。大多數學生都預期自己會覺得非常沮喪難過。然而，當研究者刻意在這些學生面前安排種族歧視的行為讓他們看見時，大部分的學生幾乎都沒有顯示出任何痛苦或是難過的樣子。大多數學生說他們會避免去接觸那些作出粗暴的種族歧視言論的人，但是，同樣地，人們所說的話與實際上所做的事之間，又再次出現南轅北轍的情況。研究結果顯示，種族歧視依舊存在的原因之一，是因為當種族歧視的行為真的發生在面前時，許多認為自己會覺得難過或是採取行動的人，實際上卻是冷漠以對。顯而易見的是，我們並不擅長於預測當事情真的發生時，我們在此種真實情況下所將會作出的反應。

1. 當我們擁有某項可信的證據，並且假定這項證據也能夠應用到許多其他的情況時，就會產生**過度推論**（overgeneralization）的錯誤。請注意到「過度」這個詞。適度推論當然是恰當的，但是卻有其限制。我們只有在極爲謹慎小心的情況下，才可以把少量的證據推論到一個較大的情境。不幸的是，我們之中有許多人傾向於以有限的證據去推論到超過可以接受的範圍。我們經常會從已知的事實推論到未知的領域。舉個例子，假設在過去幾年以來，我總共認識了五個眼盲的朋友，而且這五個人全都是非常直率而且友善的人。那麼，我是否可以就此認爲眼盲的人都是友善的人呢？我個人所認識的這五個人，是否可以完全代表世界上所有眼盲的人呢？

2. **選擇性觀察**（selective observation）則與過度推論稍微不同。當我們特別注意到某些人或是事件、並且由此作出推論時，就會犯下選擇性觀察的錯誤。我們通常會把焦點集中於特殊的案例或是情況上面，尤其是當這些案例或是情況符合我們先入爲主的概念時。我們也會傾向於找出那些能夠證實我們已經擁有的信念的特定證據。與此同時，我們大多數人會傾向於忽視整體各種不同的案例。我們經常會摒除那些與我們的信念互相矛盾的資訊，並將其視爲可以置之不理的例外情況。舉例而言，相信那些體重過重的人比起身材瘦削的人更加直率而且友善。我的這種想法，是從父母以及媒體那兒學到的刻板印象。於是，當刻意去觀察體重過重的人時，在無意識的情況下，會特別去注意到他們的微笑以及笑聲等等。而當在觀察身材瘦削的人時，則比較會注意到他們看起來嚴肅、心煩意亂、或是生氣時的樣子。在毫不知情的情況下，會去注意到那些能夠增強先入爲主的思考方式的人們以及情況。許多研究也記錄了我們此種傾向，也就是我

過度推論　一種陳述或說明，而該陳述或說明已經超過依據某人所擁有的資料或是實徵觀察所能合理推論的範圍。

選擇性觀察　一種檢驗的過程，該過程是以一種會強化已經先入為主的想法的方式來進行，而不是以中立並且平衡的方式來進行。

們會傾向於「找出」（“seek out”）並且扭曲記憶，以使記憶更加符合我們已經擁有的想法。

3. **過早下結論**（premature closure）的錯誤，會與前兩種錯誤一起發揮作用，並且還會強化前兩種錯誤。當我們覺得我們已經找到答案，並且不再需要去傾聽、尋找資訊、或是提出問題時，就可能會犯下這個錯誤。事實上，在某個時間點上，我們的確必須停止蒐集資訊並且作出結論。不幸的是，我們大多數人都有一點懶惰、或者是有一點草率。我們只蒐集了少量的證據、或是短暫地檢視一些事件，然後就認爲我們已經想出答案了。換句話說，我們通常會去尋找證據以證實或是拒斥某個想法，而在得到少量的證據之後，就不再努力地繼續蒐集證據，並且以此少量的證據而過早作出結論。

4. 當我們相信某些人事物具有高度正向的特質、或是具有高度的名望，並且由這些人事物去進行過度推論時，就會犯下**月暈效應**（halo effect）的錯誤。換句話說，對於我們所尊敬的人事物，我們會賦予這些人事物一個光環、或是正面的名望，然後，我們會把此種光環「暈染」（“rub off”）到其他我們幾乎一無所知的事情或是人們上面。例如，閱讀了一份像是哈佛大學或是劍橋大學這類知名大學裡的學者所作的報告。假定該名作者既聰明又有天份，而且也預期這將會是一份品質極好的報告。但是，如果另一份報告是由某一所沒沒無聞大學裡的學者所寫的，那麼，就不會作出相同的假定了。由於在事先就已經形成了先入爲主的意見，因此，在閱讀個別的報告時，就無法單獨就其價值與優點來作出判斷了。再舉另一個例子：假設有某一個我們信任的名人或人物，爲一個我們幾乎一無所知的產品或是選舉中的候選人背書。於是，就可能會用我對於信任的人物所具有的正向感受，取代了我原本應該要去自行發現這些產品或候選人真正特質的工作，或者，我們也可能會以此種正向感受，作爲一種下決定時的簡便參考方式。

5. 已經有許多研究（Marks and Miller, 1987）記載了**錯誤共識**（false consensus）這種心理學上的效應。此種效應暗示著，我們其實並不擅長於分辨我們個人所想的事物、以及我們認爲大多數人所相信的事物之間的差別。簡而言之，我們傾向於認爲，其他大多數的人所抱持的看法，與我們本身的看法是相近的。此種情況並非是我們刻意去迎合並且模仿群眾所持有的觀點。更確切地來說，其實是我們大多數人會覺得，在與他人相較之下，我們本身的看法是「正常的」、或是「普通的」。雖然此種想法或許有可能是正確的，但是我們卻可能過度高估了我們的看法與他人的看法之間的契合程度。就社會事件和社會議題方面而言，許多研究顯示，我們大多數人其實都不擅長於判斷周遭的人們所持有的想法。

社會研究能夠幫助我們處理個人經驗所造成的

過早下結論　在蒐集到科學標準所要求的足夠的證據數量或深度之前，就提前作出判斷、達到結論、或是結束調查的動作。

月暈效應　一種常見的錯誤，該錯誤使人們在作出評估時，會受到某個人物、地方、或是事物的先前聲望所影響，而不是以一種中立、平等的方式來作出評估。

錯誤共識　一種人們把自己的思考方式投射到他人身上的傾向。換句話說，人們會假定其他所有人的想法和他或她自己的想法是很類似的。

錯誤。個人經驗時常會造成錯誤判斷、偏見、以及扭曲的想法，而研究的標準、規定、以及原則都是設計以減少上述這些問題的。

來自於專家以及權威的知識

我們所知道的大部分知識，很有可能是來自於我們的父母、老師、專家，以及書本、電影、電視、網際網路與其他的媒體。通常，我們之所以會接受某事爲真，是因爲某個具有專門知識、或是處於權威地位的人這麼說，或者是因爲它出現在一個具有權威性的、受人信任的來源。這種情況就是把權威當作是知識的基礎。在許多方面來說，在學習事物時，依賴專家或是權威所擁有的智慧，是一種快速、簡便、以及省事的作法。某位專家可能已經花費極多的時間去學習某項事物，而我們則可以從該名專家的經驗和努力中輕鬆地獲得益處。

然而，依賴專家其實是有其限制的，而且我們也很容易會過度高估某人的專門知識與能力。權威有可能會針對他們所知甚少的領域發表評論，而且他們也有可能會錯得很離譜。在某個領域擁有專門知識的專家，可能會把他或她所具有的真實權威延伸到一個不相關的領域。利用月暈效應，某個領域內的專家可能會以不合理的方式在一個完全不同的領域內充當權威。想想看，你是否曾經看過一些電視廣告，在這些廣告中，某個電影明星或是足球英雄正試圖要說服你去購買某項產品呢？

到底是誰決定了某個人是不是一個貨真價實的專家或是權威呢？一個人或許會在某個擁有令人印象深刻的名稱（像是「某某領域的科學研究中心」之類）的私人「智庫」（“think tank”）裡擔任「資深研究員」或是「助理學者」之類的工作。雖然有些智庫是合法的研究中心，但是卻有許多智庫是由一些財力雄厚的特殊利益團體所設立的，而這些團體的真正目的，是在從事政治方面的特定宣傳活動。由於在法律上沒有任何有關於智庫名稱與頭銜的規定，因此，任何人都能成爲某個智庫裡的「學者」。智庫透過某些「專家」對大眾媒體發表具有權威性的聲明，並且營造出一種中立而且具有知識性的假象。而這些所謂的專家可能根本就不具有真正的專門知識，或者，當他們在發表聲明時，其實是以特定的意見或意識型態爲基礎，而不是以真正的研究結果爲基礎。[3] 在本章的後面部分，你將會讀到有關於科學社群如何運作、以及科學社群如何去決定誰才是名符其實的專家的內容。

即使我們在某個特定的領域找到一些專家，他們也有可能不認同我們的看法。或許你已經聽過在健康與飲食方面許多相互矛盾以及令人困惑的研究建議。你可能會認爲，如果各種研究結果之間如此不一致，那麼進行研究根本就沒有太大的意義。然而，之所以會發生此種研究結果之間自相矛盾的情況，是因爲許多充斥在大眾媒體之間、打著「研究」或是「科學」名號的聲明或說法，根本就與科學研究完全無關。不幸的是，即使在技術上根本沒有一項真正的研究去支持某項聲明時，媒體仍然經常使用「研究」一詞。不過，科學家或是專家們也不一定總是百分之百同意彼此的意見。在許多案例中——像是最好的飲食方式、健康習慣、公共政策、或是氣候變遷等等——仍然有一些意見分歧的狀況。在本章稍後的部分，你將會讀到有關

於科學的原則以及科學社群的運作方式，而且你也會看到，作爲科學研究過程的一部分，此種意見分歧的情況是如何產生、以及如何解決的。

除了找到專家之外，重要的是，我們也應該要學習如何獨立思考、並且靠自己的力量去評估研究。一味地依賴專家或是權威，並不符合一個自由、民主社會的原則。專家在提倡一些想法時，有可能是爲了要強化他們自身的權力與地位。如果我們盲目地跟隨權威，就可能會失去自己作出決定的能力。這正是爲什麼我們要學習如何去進行研究、以及獲得研究技巧的理由之一，因爲如此一來，我們才能夠去分辨優質與劣質研究之間的差異。

來自於大眾以及媒體的訊息

除了依賴常識、個人經驗、以及專家之外，我們也可能會藉由與他人談話以及從媒體中擷取訊息的方式，試圖擴展我們的知識。雖然這是個不錯的方式，但是卻有嚴重的限制。與他人談話雖然有所助益，但是許多研究已經發現，在有關科學能力、地理知識、以及清晰與合乎邏輯的思維方式等方面，大多數人仍然是較爲不足與欠缺的。即使在像是二十一世紀的美國這樣富有、先進、以及具備高教育水準的國家裡，此種情況依舊相當普遍（請參閱本章稍後部分裡的資訊箱 1.2 科學素養裡的相關討論）。擁有使用先進科技（例如像是 iPhone、衛星定位系統、或是配備先進裝置的車子等）的能力，並不意味著我們會普遍地以合乎理性的、科學的方式來思考。一份於 2006 年針對 18 到 24 歲的年輕男女所進行的研究，發現到約有一半的人無法在美國地圖上正確地找到紐約州和俄亥俄州的位置（其中男性與女性的比例分別爲 50%與 43%），而且，雖然自從美國於 2003 年 3 月攻打伊拉克以來，新聞報導中幾乎每天都有伊拉克的相關消息，卻仍然有超過一半以上的人（大約 63%）無法在中東的地圖上找到伊拉克的位置。大部分的美國民眾對於一些早已被科學社群斥爲無稽的現象仍然深信不疑，像是幽浮（約有 31%的美國人相信）、占星術與占星學（約有 31%的美國人相信）、鬼魂與妖精（約有 51%的美國人相信）、女巫（約有 34%的美國人相信）、或是惡魔（約有 61%的美國人相信）等等。[4]

雖然美國民眾接受正式學校教育的平均水準已經上升，但是仍有許多人缺乏根據事實而來的知識、依賴不正確的資訊、或是堅持一些不合邏輯的思考方式。有些人雖然接受過學校教育，但是卻沒有學到任何知識，或是未能持續地把在學校裡所學到的知識、技能、或是思想，應用於日後的生活或是工作上的決定之中。除此之外，有許多人傾向於「人云亦云」（“follow the herd”），或是僅僅依賴大眾的意見。大眾媒體通常只會單純反映群眾的意見，但是卻沒有經過審慎的評估。正如同你早已熟知的，僅僅因爲大多數人相信某事爲真，並不必然就代表某事確實爲真。然而，我們之中有許多人只會依循「多數人所抱持的想法」，即使此種想法有可能出錯時亦然。

我們之中有許多人依賴大眾媒體（例如像是電影、電視、報紙、雜誌、以及網路來源等等）來獲取資訊。不幸的是，媒體通常會把各種不同類型的陳述混雜在一起——其中一些陳述可能有著堅實的研究基礎、而另外一些陳述則可能毫無事實根

據可言。除此之外，媒體也可能會扭曲社會議題。媒體傾向於強化一些文化的迷思，或是創造出一些「不實的宣傳」（“hype”），以使大眾認爲某個嚴重的社會問題確實存在（但是該問題實際上是不存在的）。我們可能會在媒體上聽到一個可怕的問題，然而，只要仔細地審視該問題、並且進行少量的調查或研究，我們或許就會發現，該問題其實是被嚴重地過度誇大了。

「憤怒駕駛」的例子

美國民眾已經聽說過許多駕駛人因爲在開車時對前車感到不耐煩、或是因爲被其他車輛超車而心生不滿所引起的憤怒情緒。《新聞週刊》（*Newsweek*）雜誌、《時代》（*Time*）雜誌、以及許多大型城市裡的報紙，都曾以此種現象作爲頭條新聞而加以報導。一些主要的政府官員已經爲此舉辦過多場的公聽會，而且聯邦政府也特別撥出數百萬美元的經費給執法部門以及交通部門，希望能夠減少此種情況。現在，更有一位加州的心理學家專門研究此種失序的狀況，並且還出現在數個大型的電視節目中參與討論這個問題。

「憤怒駕駛」（“road rage”）這個詞彙最早出現於 1988 年，而截至 1997 年爲止，每年有超過 4,000 篇討論此種狀況的文章出現在印刷媒體上面。雖然媒體大量關注「激進而有侵略性的駕駛行爲」以及「因爲對前車不滿而產生的怒氣」等等狀況，卻沒有任何與「憤怒駕駛」有關的科學證據。這個詞彙並未有明確的定義，因而可以被用來指稱任何與駕駛人有關的行爲，像是從車內向他人咆哮、用粗魯的手勢來表示不滿、故意開車把自行車騎士逼到路邊、緊跟著前車行駛、以及甚至是對於汽車修理費用的帳單感到生氣等等。然而，所有關於車輛相撞以及意外事故的資料都顯示出，當駕駛人在開車時普遍感覺到不滿或憤怒時，各種交通意外的發生次數卻反而下降了！

相反地，真正的狀況是，我們之所以會認爲此種憤怒駕駛的情況非常普遍，其實是因爲媒體的大肆報導。在聽過或讀過關於憤怒駕駛的報導、並且對此種行爲形成特定的標籤之後，人們就會開始去注意到粗魯的駕駛行爲，並且會作出選擇性的觀察。直到此種行爲被適當地加以研究之前，我們無法明確地知道真正的情況爲何，但是，此種行爲的數量似乎並未有任何改變。或許，到了最後我們可能會發現，此種認爲憤怒的駕駛在全國都很普遍的想法，只是一種廣泛流傳的迷思而已，而此種迷思的幕後推手，正是大眾媒體中的報導。

「假期浩劫」的例子

報紙以及電視中經常充滿了各種可怕的警告，告訴我們在假期之中會發生許多的交通意外。因此，在美國境內，七月四日獨立紀念日的那個週末假日，就會被媒體報導爲非常致命的時期，因爲每年平均會有 161 人在此時喪命。然而，事實上假日期間並不比其他期間更加危險，而且甚至可能還更加安全一些！怎麼會這樣呢？在與其他週末假期進行仔細的比較、並且考量到交通流量變大的情況之下，假日期間的交通意外發生率並沒有顯著的不同。因此，上述的那些警告，其實是許多提倡

交通安全的團體，刻意在媒體中大肆宣揚並且扭曲統計數據，以提醒人們在開車時要更加小心的一個策略而已。

教訓

像是憤怒駕駛以及假期浩劫這些例子，並不是特殊的情況。在許多的社會議題裡，扭曲以及不實的陳述可以說是屢見不鮮。有一些所謂的「問題提倡者」（"problem promoters"），尤其是在廣播媒體中，會刻意突顯一些戲劇化的誇張案例、或是選擇性地使用統計資訊，以吸引社會大眾的注意力，並且激起大眾去關心某個社會問題。媒體報導的真正問題，並不是在於其內容錯誤，而是在於會誤導民眾。媒體報導最爲擅長的地方，是在於說服大眾相信某個想法或觀念，而不是在於完整、仔細地呈現出某個問題的整體樣貌。換句話說，如果我們過於依賴大眾媒體的報導來瞭解整個社會、重大趨勢、或是嚴重的問題等等，那麼我們就有可能會輕易地被媒體給牽著鼻子走了（Best, 2001; Fumento, 1998; and Wald, 2004）。

許多研究已經證明，在電影、電視、以及雜誌中所呈現出來的貧窮、犯罪、以及其他社會大眾關心的議題，並未能正確無誤地描繪出真實的社會情況。當編劇或作者爲了電視節目或是電影劇本而創造或是「改編」真實生活時，他們通常會扭曲現實狀況。然而，這些編劇或作者很少是刻意這麼做的。更正確地來說，他們只是在重覆一些他們挑選出來的錯誤資訊，而且他們的主要目的只是爲了要娛樂觀眾而已。舉例而言，在 400 部描繪精神病學治療過程的電影中，只有大約 5 部左右的內容是正確的。同樣地，在媒體中所報導的美國回教徒人口數目，也比那些以嚴謹的科學方法所得到的數目高了二到三倍左右。各種新聞雜誌的圖片顯示，在美國所有的窮人當中，非裔美國人約佔了百分之六十二；而在電視新聞中，這個比例則爲百分之六十五。然而，實際上，在美國所有種族的貧窮人口當中，非裔美國人僅佔約百分之二十九而已。我們在電視新聞以及雜誌圖片中所看到的資訊，會強烈地形塑了我們對於社會議題的觀點與看法。媒體中的扭曲資訊意味著，如果我們過度依賴媒體來獲得對於社會現狀的知識，我們通常會得到不正確的知識。[5]

除了提供知識以及娛樂我們之外，媒體也提供了一個公開討論的場所，讓各種不同的利益彼此互相競爭，以贏得大眾的支持。支持或是反對某項議題的團體或個人，會發起各種公關活動，並且利用媒體來形塑公眾的想法。正如同稍早已經提到過的，一些爲特定利益護航的智庫，有時候會請來一些僞「專家」在媒體中討論一些特定的主題。除此之外，在近幾年來，置入性行銷影片（video news releases, VNR）——又被稱爲「僞造的電視新聞」（"fake TV news"）——的數量大幅地增加。所謂的置入性行銷影片，指的就是某個大型公司或是特定利益團體付費生產出製作精美的影片，而該影片看起來就像是一則獨立製作出來的新聞報導一樣。在置入性行銷影片中，某個男演員或是女演員會扮演獨立記者的角色，而這名「記者」則會報導一些表面上看起來似乎是中立的資訊或是新聞。事實上，這此資訊或是報導的目的是公共關係或是提倡某種觀念或想法。大多數的電視台在播出置入性行銷影片時，

都不會告知觀眾該段影片的真正來源。如此一來，在電影上的一則新聞報導，就有可能是一種複雜的宣傳手法，其目的則在於影響我們對於某一主題或是產品的觀點或看法。因此，在我們把大眾媒體當作一個權威之前，必須要更加謹慎小心才行。[6]

許多認真熱心的科學作者、以及克盡本份的新聞記者，正在努力爲社會大眾提供正確的、以研究爲基礎的資訊。然而，由於大眾媒體所傳遞的錯誤訊息不但數量龐大、而且無孔不入，他們的努力可能會相形見絀。如同你在本章稍後部分所會看到的，在瞭解與得知各種研究方面，大眾媒體並非是最佳的來源。相反地，我們所應該倚賴的，是科學社群的溝通系統。任何具有一些研究知識的人、以及任何願意花時間去探索的人，都能接觸並使用這套系統，而且不須花費任何的費用。

來自於意識型態信念以及價值觀的知識

雖然社會科學研究有許多優點，而且也易於接近使用，有些管理者和決策者仍然會有意識地排斥社會科學研究，並且反而去提倡與捍衛那些以他們的政治、宗教、或是意識型態的信念爲基礎而採取的行動。例如，在 2001 年時，美國的聯邦政府開始提供資金給「以信仰爲基礎」的社會計畫。即使有眾多的研究都質疑這些計畫的實際效果，這些計畫卻仍然取代了由研究所支持的計畫。與此同時，在政府機構服務的一些有見識的科學家，則被政治任命的特定人士所取代，而這些人士通常是某些意識型態的熱衷支持者。即使是那些備受推崇的研究發現與結果，只要與特定的意識型態觀點相互抵觸，就會從官方的健康或是環境公共資訊中被移除。[7]

曾經有一陣子，美國政府中的重要官員都在宣揚一些反科學的信念。當時小布希總統身旁的一位重要助手甚至宣稱要拒絕「以事實爲基礎的社群」，也就是指那些「相信只要對可辨識的事實進行審慎研究，就可以找到解決方案」的人（Suskind, 2004）。在第三章中，我們將會仔細討論各種社會理論，而在該章中，你將會再次看到有關意識型態的議題。

如果你想看看，除了社會研究之外的其他選擇會如何去解釋社會生活的某個面向，可以去參閱表 1.1 所列出的例子。

研究的內涵：科學化的取向

在「以事實爲基礎的社群」（“reality-based community”）中，社會科學研究具有相當重要的地位。社會研究有賴研究者小心謹慎地研究在社會事實裡的經驗、事件、以及真相。雖然社會研究能夠幫助我們回答與社會世界有關的問題，社會研究也會提出新的問題，而且或許還會改變我們看待世界的方式。社會研究所依賴的是科學的過程與證據，而且迥異於隨興的觀察、常識的推理方式、以及其他評估證據的方法，其中包括了純粹的邏輯—理性推理方式（數學的、或是哲學上的證明）或是法律—司法程序等等。我們接下來會在進行社會科學研究的脈絡下來看看何謂*科學*（*science*）。

表 1.1 社會研究之外的其他選擇所提出的解釋

範例議題：女性比男性更可能會去清洗衣物

個人經驗與常識：就我的經驗而言，男性比較不像女性一樣那麼關心衣著和外表的問題，因此，女性喜愛清洗衣物是有其道理的。在我和我朋友的成長過程中，家裡的衣服都是我媽媽和朋友們的媽媽在洗的，而且女性朋友也都會幫她們的男朋友洗衣服，但是，我周遭的男性友人卻從未自己洗過衣服。

專家以及權威：專家們認為，從兒童時期開始，女性就被教導去製作、選擇、修補、以及清洗衣物，因為這是女性關心外表、以及照顧兒童和其他家人的一種表現。因此，女生比男性更可能會去清洗衣物的原因，是來自於她們在兒童時所受到的教導。

大眾以及媒體訊息：在電影以及電視廣告裡，經常把女性描繪為喜愛清洗衣物並且樂在其中，而男性則厭惡洗衣服、並且還會弄得一團混亂。因此，女性必須去洗衣服，因為她們不但樂在其中，而且也擅長於洗衣服。這正是我們隨處都可以看到的景象以及每個人都認同的看法。

意識型態的信念：在合乎常態以及自然的勞力分工裡，女性本來就應該要打理家務、照顧兒童、並且盡到家庭主婦的本份，其中包括了煮飯、打掃房子、以及清洗衣物等等。

科學

當大部分的人聽到「科學」這個詞時，腦海中最先浮現出來的畫面，大概會是一個放滿了試管、電子儀器、以及顯微鏡的實驗室，奇特的太空船，或是一群身穿白色外袍的人們。這些外表的特徵都只是科學的一部分而已。自然科學與生物科學——像是生物學、化學、物理、以及動物學等——所探討的是實體與物質的世界（例如，像是岩石、植物、化合物、天體、肌肉、血液、以及電力等等）。由於這些自然科學是新科技的基礎，並且也是民眾相當關注的焦點，因此，當大部分的人聽到「科學」這個詞的時候，就會先想到這些自然科學。

至於社會—文化科學（例如像是人類學、經濟學、人文地理學、心理學、政治學、以及社會學等等）所牽涉到的則是去仔細研究人類的社會—文化生活，像是人們的信仰、行爲、人際關係、互動、以及制度習俗等等。正如同我們會把自然科學以及生物科學的知識應用在相關的、而且較爲實用的領域（像是農業、航空、工程學、醫學、以及藥物學等）一樣，我們也會把社會科學的知識應用於相關實用領域中較爲實際的事務上（像是諮商、刑事司法、教育、管理、行銷、公共行政、公共衛生、社會工作、以及都市計畫等等）。

有些人把社會科學稱爲「軟性科學」（“soft science”）。不過，這倒不是因爲這些學科或領域缺乏嚴謹性或精確性，而是因爲這些學科所探討的主題內容——人類的社會生活——極爲容易產生變化、難以觀察、並且也不容易進行精確的測量。一門科學所探討的主題內容（舉例而言，像是人們的態度、原生質（protoplasm）、或是銀河系等），決定了該門科學所使用的技術以及儀器（例如，調查、顯微鏡、或是望遠鏡等等）。我們將在第四章中更進一步地討論自然科學以及社會科學之間的關係。

科學是一項人類的發明。今日科學的起源，源自於大約四百年前一次思想方面的重大變化。更精確地來說，今日科學肇始於西歐歷史上的理性時代或啓蒙時代（Enlightenment period），大約是介於十七世紀到十八世紀之間。啓蒙時代引進了一

波新的思想潮流，其中包括合乎邏輯的推理方式、對物質世界進行仔細而謹慎的觀察、對於人類進步的信念、以及對於傳統的宗教與政治信條的質疑。此種新的思想潮流建立在過往的知識之上，並且由對於自然世界的研究開始，然後擴展到對於社會世界的研究。之後，工業革命掀起天翻地覆的社會改革運動，並因而使得科學思想廣爲流傳。然而，科學以及相關應用學科的進展並非主動發生的，在其進展過程中也不時穿插著許多個別研究者的勝利與掙扎。除此之外，科學的發展也受到許多重大社會事件的影響，像是戰爭、經濟蕭條、政府政策、以及公共支持的轉變等等。

在科學的推理方法充分發展而且廣爲流傳之前，人們所依賴的是非科學的方法。這些方法包括了我們在之前所討論過其他方法，以及一些現代社會較少使用的方法（例如，像是神諭、神秘主義、巫術、占星術、以及聖靈之類的方法）。雖然這些系統在現代世界仍然持續存在著，但是科學已經被社會大眾廣泛地接受了。今日，我們仍然會使用非科學的方法，來探究一些被認爲是超出科學範疇的主題（舉例而言，像是宗教、藝術、文學形式、以及哲學等等）。

科學（*science*）一詞同時是指一套產生知識的系統，以及由該系統所產生的知識。在過去數個世紀以來，科學一直在逐步發展當中，而且目前也仍然持續以緩慢的速度向前進展。科學結合了下列的各種事物：關於這個世界的假定；累積的知識；通往知識的方向；以及許多明確的程序、技術、以及工具。科學系統最爲具體而且可見的形式，就是作爲一種社會機構，也就是所謂的科學社群（請參閱本章這部分稍後的討論）。

科學所產生的知識會被組織成爲理論，並且建立在實徵資料的基礎上。讓我們來檢視一下三個關鍵性的詞彙：*理論*、*資料*、以及*實徵*。許多人把理論和意見、毫無根據的信仰、或是胡亂猜測等詞彙混爲一談。「雖然科學家都知道，所謂理論就是指一個具有充分證據的意見……一般大眾卻認爲，理論『不過就是理論而已』，不見得會比其他意見或想法更具有說服力」（Yankelovich, 2003: 8）。現在，我們可以把**社會理論**（social theory）定義爲一個具有一致性的系統，該系統是由一些在邏輯上連貫、而且互相關聯的概念所組成的，其目的在於濃縮並且組織知識。你將會在第三章裡面讀到有關於各種不同類型的社會理論。你可以把理論想像爲一張地圖，這張地圖可以幫助我們更清楚地想像這個世界的複雜性、看出事物之間的關聯、以及解釋事物發生的原因。我們利用資料來決定某個理論是否爲真並且應該被保留下來、或是某個理論是否有錯誤而必須加以修正或是被完全摒棄。**資料**（data）所指的是實徵證據或是資訊，而這些證據或是資訊則是依據科學的原則或是程序所仔細蒐集而來的。**實徵的**（empirical）所指的是可以依據人類感官經驗（像是觸覺、視覺、聽覺、嗅覺、以及味覺

社會理論　由互相關聯的概念所組成的系統，該系統可以濃縮並且組織關於社會世界以及其運作原理的知識。

資料　包括了數字（量化的）以及非數字（質化的）的資訊與證據，這些資訊和證據都是依據規定或是既定程序所仔細蒐集而來的。

實徵的　用來描述那些我們能夠透過人類感官（例如：觸覺、視覺、聽覺、嗅覺、味覺）直接觀察或體驗的事物、或是利用一些能夠擴展感官的技術，間接地觀察或體驗的事物。

等等）而獲得的證據或是觀察結果。科學家無法利用他們的感官來直接觀察世界的某些面向（例如：智力、態度、意見、情緒、權力、權威、夸克（quark）、太空中的黑洞、力場（force field）、或是重力等等）。然而，科學家們已經創造出專門的工具以及技術，讓他們得以用間接的方式來觀察或是測量這些面向。

資料或是實徵觀察可以是*量化的*（*quantitative*）（也就是用數字來加以精確表達的）、或是*質化的*（*qualitative*）（也就是用文字、影像、或是物體來加以表達的）。在本書稍後的部分裡，你將會看到我們可以如何去測量社會世界的各種面向，以產生量化或是質化的資料。

偽科學、垃圾科學、以及「真正的」科學

數個世紀以來，科學在全世界都獲得廣泛的尊重以及接受。然而，許多民眾仍然缺乏科學素養（請參閱資訊箱 1.2 科學素養），或是分不清楚真正的科學和**偽科學**（pseudoscience）之間的區別。*pseudo* 這個字首是個希臘字，意思是指「謬誤的」或是「僞造的」。我們可以看到僞科學大量出現在電視、雜誌、電影、報紙、以及高度宣傳的特殊研討會或專題討論會之中。有些人會把科學外表的裝飾（像是專業術語、外觀複雜花俏的機器、複雜的公式與統計、以及白色的實驗室長袍），和一些科學的事實與迷思、幻想、或是希望交織混雜在一起，並用以宣稱他們發現了「奇蹟藥物」、「新的神奇療法」、「革命性的學習方案」、「外星人造訪地球的證據」、或是「新世紀的精神能量」等等。在僞科學領域裡的專家，可能擁有高等的學位，但是其學位卻通常是屬於不相干的領域、或者是由沒沒無聞或名不見經傳的學校所頒發的。

除了專家之外，雜誌或書籍也會提供通俗化的、或是「大眾化的」社會科學。有些通俗化的社會科學是由合法正統的社會研究者所撰寫的，其目的在於把一些科學的概念傳達給一般社會大眾，因此其內容雖然較爲淺白易懂，但概念卻仍然是正確的。然而，對於一個非專業人士而言，雖然其他通俗化的社會科學表面上看起來像是合法而正統的社會科學，但是實際上卻是扭曲、或者誤用了社會科學。這些作者撰寫這些書籍的目的，是爲了要僞裝成社會科學，以宣揚一個特定的政治立場或是社會立場，但是他們其實根本就不符合科學社群的標準。例如，關於女性性行爲的海蒂報告書（Hite Report），就是一份有嚴重缺陷的研究，而該研究是由一個嚴重扭曲實際社會關係的非科學家所進行的。雖然這份報告書有嚴重的問題，這本書仍然成爲一本暢銷書，而且也在電視上的談話節目中和報紙上引起廣泛而熱烈的討論。同樣的情況也發生在《鐘形曲線》（*The Bell Curve*）這本書上，因爲該書宣稱非裔美國人在智力方面較爲低劣。[8]不幸的是，在電視上或是廣播中大打廣告、在報紙文章中被大量引用、或是在書店裡販售的許多書籍，都有可能充斥著未經證實的意見、個人信念、或是有嚴重缺陷的研究。對於一個未加以提防的消費者而言，很容易就會被這些書籍誤導，或是把這些不正確或充滿個

偽科學　以大量術語以及科學外表所包裝而成的概念或資訊，其目的在於尋求人們的接受與支持，但是此種概念或資訊並未依循科學方法所要求的系統嚴謹性或是科學標準。

資訊箱 1.2　科學素養

在超過半個世紀的時間裡，一些居於領導地位的教育學者、商業領袖、以及政策制訂者都一直在強調人們必須要擁有量化以及科學的能力，才能在這個複雜多變的世界裡執行專業的任務，以及在日常生活中作出各種適當的決定。量化能力（quantitative literacy）又稱作數字能力（numeracy），是指利用數字或是其他數學觀念來進行推理或是思考的能力。一個擁有量化能力的人，能夠以數量－空間的方式來進行思考，並且能夠應用此種思考方式來解決問題。他們瞭解應該如何以計算和測量的方式來蒐集資料，以及如何以直線圖和曲線圖等各種圖表和表格來呈現資料。缺乏此種量化能力的狀況，則被稱作**數盲**（innumeracy）（Paulson, 1990）。**科學素養**（scientific literacy）所指的是以下數種能力：理解科學知識的能力；應用科學觀念、原則、以及理論的能力；使用科學過程來解決問題以及作出決定的能力；以及在與人互動時能夠反映出核心科學價值的能力（Laugksch, 2000: 76）。由聯合國的經濟合作暨發展組織（Organisation for Economic Co-operation and Development, OECD）所規畫的國際學生能力評量計畫（Programme for International Student Assessment, PISA）曾經進行過多項研究以探討學生瞭解科學的方式。該計畫對於科學素養的定義如下（PISA, 2006: 23）：

- 具備科學知識，並且能利用科學知識去發現問題、獲得新知識、解釋科學現象、而且能對科學相關議題提出以證據為基礎的結論。
- 能夠瞭解科學作為一種人類知識與探究形式的獨特特徵。
- 能夠意識到科學與科技如何影響到我們的物質環境、智識環境、以及文化環境。
- 樂於作為一名具有反思能力的公民，並積極投入於科學相關議題以及推廣科學概念。

那些缺乏量化能力以及科學素養的人，將會輕易地接受偽科學，並且在作出判斷時犯下錯誤。數盲會導致記者去報導不正確的新聞，也會造成讀者或觀眾缺乏足夠的批判能力去評估這些報導。數盲的人們將會作出差勁的金融投資決策，並且經常在賭博與相關活動方面損失金錢，因為他們不瞭解基本的數學概念。那些缺乏數字能力或是科學素養的人，將會拙於評估風險。此外，他們也比較沒有機會成為在勞動市場上成長快速、收入較高的技術—管理專業人才。

你或許會認為，在這個科技進步、極度現代化的社會裡，你根本不可能會成為上述的那種人。然而，人們有可能在使用現代科技（像是電腦、行動電話、iPod 隨身聽、飛機、以及其他最新的科技產品等等）的同時，卻還保有科學出現之前的老舊想法，或者仰賴巫術或是超自然的信仰來解釋一些事件以及作出決定。擁有

數盲　缺乏量化與數字能力的情況，也就是缺乏以數字或是其他數學概念來進行推理思考的能力。

科學素養　具有瞭解以及應用科學知識、觀念、原則、以及理論來解決問題的能力，並且能夠依據科學推理方式來作出決定，或是能夠以反映出科學社群的核心價值觀的方式來與他人進行互動。

使用先進科技的能力，並不必然意味著人們會以理性、科學的方式來進行思考。

只有大約百分之 25 到 28 的美國成人有資格被稱作是具有科學素養的人。整體而言，在其他的先進國家裡，具有一般科學素養的成人所佔的比例也和美國大致相同。然而，在為中學生所舉行的國際數學和科學競賽裡，美國在所有參賽國家中的排名卻落在大約第二十名左右。一項於 2002 到 2003 年針對美國以及九個歐洲國家所進行的跨國研究，證實了在與其他先進國家相較之下，支持演化理論的美國成人比例是敬陪末座的，即使在 2009 年，其比例也僅有百分之 32 而已。在 2007 年六月時，一項由《今日美國》（USA Today）報紙與蓋洛普所合作的民意調查發現，有百分之 37 的美國人拒絕接受關於演化的科學理論，而且有百分之 56 的人支持宗教所提供的解釋。一項於 2007 年三月所進行的民意調查則發現，有百分之 39 的人所抱持的看法，與全世界科學社群的意見完全相左。也就是說，這些人竟然錯誤地認為目前的科學證據並不支持演化論。一份由美國佩尤人群研究中心（Pew Research Center for the People）於 2006 年所進行的調查也發現，有超過半數以上的美國人，認為公立學校在教導科學議題時，應該要教導宗教方面的觀點，而且還認為應該要強制全國的公立學校都要採取此種作法。一份 2006 年的蓋洛普調查也發現到，超過半數以上的美國人相信人類並非是由演化而來的（Polling Report, 2007）。科學家一般而言都同意全球正在暖化的事實，而且有百分之 84 的科學家認為，地球暖化背後的原因，正是人類的各種活動，像是燃燒化石燃料等等。然而，只有約百分之 49 的民眾認同這樣的看法。超過九成以上的科學家贊同在研究中使用動物、以及進行幹細胞的研究。相較之下，只有略為高於五成的民眾贊同這樣的作法（Pew Research Center for the People and the Press, 2009）。

在美國，由於有一些官員試圖在公立學校之內強制推行把宗教信仰當成科學的作法，因而使得演化論的議題已經變得極端政治化，除此之外，在一般的科學一量化思考方式、以及其他的科學觀念方面，美國人的表現也相當差強人意。雖然大多數人都曾接受過 X 光的照射，但是卻只有大約百分之十的美國民眾知道放射線是什麼，而且還有百分之二十的民眾認為太陽繞著地球轉－這個錯誤的觀念早已在十七世紀時就已經被科學所揚棄了！（"Scientific Savvy? In U.S., Not Much," Dean, New York Times, August 30, 2005）你或許會認為大學生的情況可能會比較好一點。然而，研究卻發現，許多大學生使用的是不合邏輯的「巫術」思考方式，而不是以科學為基礎的思考方式。有大量的大學生相信，巫毒教的巫術力量是使人們生病的真正原因，而且許多大學生球迷們甚至相信，當他們在看著電視上的籃球比賽時，他們的思想能夠影響比賽的結果（Pronin, Wegner, McCarthy, and Rodriguez, 2006）。

垃圾科學　一種被用來批評科學研究的公關術語。當某項研究所得到的發現，與某些利益團體的主張不同時，即使該項研究是以科學而且適當的方法所進行的，亦會被公關公司用這個詞彙來加以批評與駁斥。

人意見的書籍，和合法正當的社會科學混為一談。

或許你已經聽過**垃圾科學**（junk science）這個詞彙。公關公司在 1980 年代時創造出這個詞彙，以作為詆毀真正科學證據的一種策略。這些公司使用這個詞彙來攻擊一些特定的研究發現，因為這些研究發現在法庭上被當作證據，以證明一些有錢有勢

的大型企業對大眾所造成的傷害或欺騙。在新聞稿或是公開的聲明裡，這些公關公司刻意操弄其中的語言，以把*垃圾*科學與*合理的*（*sound*）科學（也就是支持這些公司立場的科學）作強烈的對比。事實上，所謂*合理的*科學和*垃圾*科學只是修辭的手法而已，而且也是不正確的說法。更重要的是，這兩種研究的品質、方法論、或是正確性，可能並沒有太大的差異。公關人員把「垃圾科學」這個詞彙用在任何一個他們反對的研究上面，不論這些研究有多麼正確或是嚴謹，同時，他們也把「合理的科學」這個詞彙應用於任何一個他們可以用來挑戰對手的研究上面，不論這些研究有多麼大的缺陷或問題。例如，菸草工業就曾經利用垃圾科學這個詞彙當作一種手法，以攻擊、批評那些探討二手煙所造成的傷害的研究，並且還花費數百萬美元，公開地駁斥抽煙對於健康的危害效果。[9] 其目的在於混淆陪審團以及社會大眾的視聽，並且創造出一種錯誤印象，讓大眾誤以為科學家缺乏前後一致的研究證據。與偽科學或是垃圾科學相比之下，真正的科學則是來自於科學社群的觀點、運作方式、以及成果（請參閱下一個部分）。

科學社群

科學社群為科學帶來生命與活力，換句話說，科學社群維持了科學的假定、態度、以及技術。**科學社群**（scientific community）是由人群、組織、以及角色所建構而成的社會團體，而且也是一組共同運作的規範、行為、以及態度。科學社群並非是存在於一個實體位置的地理社群，其中的成員也不一定都彼此互相認識，但是其成員卻會經常地與其他成員進行溝通以及互動。更確切地來說，科學社群是一個由專業人員所形成的鬆散集合體，而這些專業人員彼此間則分享著共同的科學訓練、倫理守則、價值觀、技術、以及生涯規劃等等。[10]

科學社群的組織方式，就像是一連串的同心圓一樣，其中的各個環節或是各個階層是以研究者的生產力和投入程度為基礎來加以劃分的。在其核心部分是一小群具有高度生產力、高度創造力、並且充滿熱情的科學領導人。依據其生涯階段、以及對於知識貢獻度的不同，隨著時間流逝，這一群科學家會慢慢地進入或是退出這個核心的部分。而處於科學社群最邊緣、或是最外環的部分的，則是數百萬名的科學實務工作者、臨床實務研究者、以及技術人員。這些人會經常地使用並且應用由社群核心人士所率先發展出來而且持續改良的知識、原理、以及技術。在科學社群外環部分辛勤耕耘的專業人士們，也會發展出某種層次的專門技術，並且也會經常使用各種不同的科學研究原則以及技術。然而，他們所擁有的科學知識，可能不像在科學社群中間階層或是核心部分的科學家那麼地深入。除此之外，在較外環部分的科學社群成員們，通常比較少從事一些能夠促進整體科學界發展的活動（舉例而言，像是生產重大的新知識之類的活動）。儘管如此，任何一個使用科學方法以及科學結果的人，不論是處於科學社群的核心部分、中間階層、或是在較外環的部分，都能

科學社群　一群人們的集合體，其中的成員共享一種由態度、信念、以及規則所形成的系統，該系統維持了科學知識的生產以及進展。

藉由瞭解科學社群的運作方式以及關鍵原則而獲得益處。

科學社群的邊界範圍以及成員資格是相當模糊的，而且其定義也相當鬆散。科學社群並沒有所謂的會員卡、或是主要的會員名冊。就某些方面而言，擁有科學領域內的博士學位，被視爲是一種非正式的「入會許可證」。博士學位代表了比碩士學位更高一級的學位，並且也讓擁有此學位的人具備了進行獨立研究的能力。科學社群中有一些成員並未擁有博士學位，而且許多擁有博士學位的人則是擔任一些不必進行任何研究工作的職位。他們把所有的心力都投注於教學、行政、諮詢、臨床實務、顧問、或是與更廣大的群眾分享知識等工作上面。事實上，大約一半擁有博士學位的人，並未把積極從事研究工作當成自己的終身志業。

科學社群的核心部分是由一群研究者所組成，這些研究者可能是全職的、或者是經常在從事研究，而且通常會有助理協助他們。許多的研究助理都是研究生。從事研究助理的工作，或多或少就像是在當學徒一樣，而這也是學習科學研究細節的最佳方法。大多數的核心成員都是任職於學院、大學、或是研究機構裡面。有些成員則是爲政府、非營利組織、或是私人企業組織工作，像是勞工統計局（Bureau of Labor Statistics）、美國國家民意研究中心（National Opinion Research Center）、以及美國蘭德公司（Rand Corporation）等等。這些研究單位是由大約 200 個大型的研究大學或機構所設立的，而且座落於大約十來個先進的工業化國家之內。雖然科學社群在地理上是散佈於世界各地的，但是其成員通常會以小型群組的方式共同合作，並且也會定期地彼此互通聲息。科學社群廣納民意，而其中的任何一位成員也都能對社群作出貢獻。關鍵的原則之一，就是要與社群中的其他成員分享自己的研究發現以及技術（像是新的知識等等）。隨著時間經過，科學社群會以公正的評估爲基礎，發展出對於新知識的重要性或是價值的共識。由於幾乎每天都會有新知識產生，此種產出以及評估新知識的過程因而具有高度的動態性質。

我們其實無從得知到底科學社群的確切大小爲何。截至 2006 年爲止，在美國整體的勞動人口當中，約有百分之三是受雇於科學或是工程學的領域（U.S. Census, 2008: Table 790）。在科學社群中的基本單位，是一個學術領域、或者是學科（例如，社會學、生物學、心理學等）。學術領域之間或許會有些重疊的部分，但是卻能夠讓我們在人數方面有較爲清楚的概念。在美國境內約有 11,000 位人類學家、16,000 位社會學家、以及 15,000 位政治學家，而且其中大多數人都擁有博士學位。但是，和相關的技術－專業領域裡的從業人員相較之下，這些數字就猶如小巫見大巫一般：美國境內有大約 180,000 名建築師、950,000 名律師、以及 820,000 名醫師。每一年，大約有 600 人獲得社會學博士學位，約有 15,000 人獲得醫學學位，而且約有 38,000 人獲得法律學位。

前面曾經提過，在所有取得科學領域裡的博士學位的人當中，只有大約一半的人會把積極從事研究工作當成是一生的志業。在整個職業生涯當中，就算是一個相當活躍積極的研究者，或許也只能完成兩項到十項研究而已。只有少數的研究者具有高度的生產力、並且能夠從事相當多的研究，尤其是那些具有高度影響力、並且被廣爲閱讀的研究。在任何一個時間點上，全世界或許會有一百個研究者，正在積

極地針對某個學科裡的一個特定主題進行研究（例如，像是關於離婚問題、或者是死刑的研究）。[11] 由他們的研究所產生的新知識，可能會影響到全世界數百萬人的生活，而且其影響力可能還會持續數個世代之久。對於科學社群裡的積極參與者、或者是新研究發現的應用者而言，此種創造知識的過程會使他們個人覺得努力沒有白費、並且也相當令人興奮。

科學社群的規範以及價值觀

在所有的人類社群裡，人們的行爲都會受到社會規範所約束。在多年的學校教育過程中、以及在與其他人的頻繁互動中，研究者學習並且內化了專業規範以及價值觀。這些規範與價值觀會彼此互相強化，並且對社會科學家所扮演的獨特角色有所貢獻。專業規範所表達的是適當行爲的理想典範，然而，在現實生活當中，理想卻不一定總是能夠完美地運作。研究者們也是人，也都會有偏見、自我意識、野心、以及個人的生活。上述這些因素可能會影響到少數的研究者，使得他們做出違反科學社群規範的錯誤行爲。[12]

科學社群並非在一個與「真實世界」完全隔絕的真空裡運作，因此，科學社群會受到社會力量、政治力量、以及經濟力量的影響（在第三章和第十六章裡，對這些議題會有更進一步的討論）。但是，科學社群的規範以及價值觀，仍然能夠讓我們瞭解科學社群以及涵蓋範圍更大的研究活動的運作方式。此外，這些規範以及價值觀也能提供指引，讓我們知道該如何以適當的方式進行研究，並且也能夠提供良好研究的原則。

五項基本的**科學社群規範**（norms of the scientific community）（請參閱重點重點摘要 1.1：科學社群的規範）不同於其他社會機構（例如，商業機構、政府部門、法律機構等等）裡的規範，並且也傾向於把專業的研究者獨立出來。例如，科學家必須遵守*普遍性*的規範，也就是說，即使某位傑出、具有創意的研究者的個人習慣相當怪異、或者是外表不修邊幅，科學家們仍然應該要去欣賞該名研究者。此外，依據*系統性質疑*的規範，科學家可能會與其他人進行激烈的論戰，甚至會「撕破」一份細心準備的研究報告。科學家通常都非常開放，並且樂於傾聽新的想法，不論這些想法在一開始的時候可能會有多麼古怪時亦然。而根據*公正無私*的規範，科學家可能會表現出超然或是客觀的態度。他們會把研究結果，包括自己的研究所得到的結果，都視爲是暫時性的，應該要接受外部審查以及批評。他們希望其他的社會科學家能夠閱讀他們的研究報告，並且作出回應。對於開放性的高度信仰，已經導致許多社會科學家反對所有形式的審查制度。此種信念與*研究成果共有主義*的規範是相當一致的，也就是樂於和他人分享新的知識、而不將其視爲個人所擁有的事物，而這就像在我們所共享的湯裡面加入一種新的原料一樣。然而，此種作法並非總是行得通，尤其是當共有主義和利益動機之間有所衝突的時候。例如，當在煙草業、製藥業、以及電腦晶片產業裡工作的科學家要發表其研究發現時，通

科學社群的規範　影響科學家從事研究時的一些非正式的規則、原則、以及價值觀。

重點摘要 1.1

科學社群的規範

1. **普遍性（universalism）** 不論是誰進行研究（例如，老或少、男或女），以及不論研究是在何處所進行的（例如，美國、法國、哈佛大學、或是一所不知名的大學），都只能以科學價值為基礎來對該研究進行評判。
2. **系統性質疑（organized skepticism）** 科學家不應該以一種毫無考慮、未加批判的方式來接受任何新的想法或是證據。他們應該挑戰並且質疑所有的證據，而且要仔仔細細地審查每一項研究。進行批判的目的，並不是要去攻擊個人，而是要確保在研究中所使用的方法都能夠經得起嚴密而且詳細的檢驗。
3. **公正無私（disinterestedness）** 對於意外的觀察結果以及新的想法，科學家必須保持中立、公正、接納、以及開放的態度。科學家不應該死板地堅守某個特定的想法或是觀點。他們應該要接受、甚至是去尋找一些與其立場相左的證據，並且應該誠實地接受所有來自於高品質的研究所得到的發現。
4. **研究成果共有主義（communalism）** 科學知識必須與他人一起分享，因為科學知識是屬於所有人的。創造科學知識是一項公共的活動，而且研究發現是公共資產，可供大眾取得並使用。研究的進行方法必須詳細地加以描述。要等到其他研究者已經審閱過、並且以一種特殊的形式與方法公諸於社會大眾時，新的知識才能被正式地接受。
5. **誠實（honesty）** 雖然這是一種普遍的文化規範，但是在科學研究裡卻特別地強調此項規範。科學家強烈要求所有的研究都必須誠實，而不誠實或者是欺騙的作法，在科學研究中是極大的禁忌。

常會被企業高層阻止、或是嚴重地拖延，因爲對於企業高層而言，利益動機的考量是遠高於共有主義規範的。[13] 除此之外，在從事研究以及發表研究報告時，科學家也會極度要求*誠實*。如果有任何人在研究中作出欺騙的行爲時，科學家會在道德立場上感到相當憤怒。

科學的方法以及態度

你很可能已經聽說過所謂的科學方法，而且可能會覺得好奇，到底科學方法如何能夠與此處的討論結合在一起。事實上，科學方法並不是單一的一件事物，而是科學社群所使用的各種觀念、規則、技術、以及方法的集合體。科學方法源自於科學社群內部所形成的共識。真正重要的，是要能夠確實地去瞭解科學的傾向或是態度，而不是把焦點過於放在「科學方法」上面。科學社群相當重視專業技術，以擁有創造力爲榮，並且強調高品質的標準，以及努力工作的精神。誠如葛林奈爾（Grinnell）（1987: 125）所說的：

> 大多數人所學到的都是「科學方法」，但是卻沒有學到科學態度。雖然「科學方法」是一種理想的構念，但是科學態度才是人們觀察世界的方法。從事科學工作包括了許多方法，然而，必須要得到科學界整體的接受，這些方法才能成為所謂的科學方法。

科學態度傾向於同時強調精確性與邏輯性、採取長期的觀點、具有彈性以及開放的心胸、以及樂於和大眾一起分享資訊（請參閱 Yankelovich, 2003）。相比之下，非科學的思考方式，對於追尋更高的精確性或是嚴謹性會覺得很不耐煩，對於當前的特定議題會想要得到立即的答案，而且也傾向於佔有一切事物、不願意和他人自由地分享所有資訊。

科學的期刊文章

你們可能已經看過某一本學術期刊裡面的文章。當科學社群創造出新知識時，這項資訊會出現在學術期刊或是專業書籍裡〔這些刊物被稱爲*專題研究報告*（*research monographs*）〕。你們將在第五章裡讀到更多關於學術期刊的詳細討論。大部分新的研究發現在第一次發表時，通常會以**學術期刊文章**（scholarly journal articles）的形式出現。這些文章就是科學家彼此之間相互溝通、以及散播研究結果的正式方法。這些文章也是廣受討論的「知識爆炸」（"explosion of knowledge"）的一部分。在單一學科領域內，可能就有 50 到 300 種這樣的學術期刊。每份期刊可能每隔一個月或二個月就會出刊一期，而在每一期之內則可能刊載五篇到二十五篇文章。例如，在社會學期刊內具有領導地位的《美國社會學評論》（*American Sociological Review*），每一年都會刊出大約 65 篇文章。對於研究過程和科學社群而言，學術期刊文章是極爲重要的，但是此種重要性卻不見得總是能夠被清楚地瞭解。[14]

讓我們來看看，當一個社會科學家完成了一份研究時，接下來會發生什麼事。首先，這名科學家會先以一種特定的格式，寫出一份描述該項研究以及研究結果的研究報告。通常，這名科學家會在專業學會的會議裡，像是「美國社會學學會」（American Sociological Association）或是「社會問題研究學會」（Society for the Study of Social Problems）等等，針對該份研究報告進行大約 20 分鐘左右的口頭報告。這名科學家會對著數十位的社會科學家以及學生作出口頭的摘要，並且回答在場聽眾所提出來的問題。接著，這名科學家可能會寄送該份研究報告的影本給一些其他的研究者，以獲得評論以及建議。最後，這名研究者會把研究報告寄送給一份學術期刊的編輯，像是《社會力量》（*Social Forces*）或是《社會科學季刊》（*Social Science Quarterly*）等等。一般而言，每一位期刊的編輯都是備受敬重的研究者，並且是由其他科學家所選出來監督管理該份期刊的。期刊編輯會先移除研究報告的封面，因爲這是唯一會顯示出作者名字的地方，然後，編輯會把該份報告寄送給數個審查者，以進行**匿名審查**（blind review）。這些審查者本身都是社會科學家，並且也在相同的主題領域裡進行過研究。此種審查過程之所以被稱爲「匿名」（"blind"），是因爲審查者並不知道該項研

學術期刊文章　在一份專門的刊物上所發表的文章，而該份刊物的主要讀者群是科學社群裡的成員。學術期刊文章是在科學社群內散佈新想法以及新發現的一項工具。

匿名審查　一種判斷研究報告的價值與特性的過程，在此過程中，進行審查的研究者事先並不知道某項研究的作者是誰，而且該項研究的作者事先也不知道審查者的身分。

究的作者是誰，而且研究報告的作者也不知道審查者是誰。此種作法可以強化普遍性的規範，因爲審查者會單獨依據該項研究的價值以及特質來進行評判。換句話說，他們會依據該份研究的清晰性、符合高標準的研究方法論的程度、以及對於知識的原創性貢獻度等標準，來評估該項研究。接著，審查者會把他們的評估結果寄回給編輯，而編輯則要決定是要拒絕該份研究報告、請作者進行修改、或是接受該份報告並且刊載在期刊上。

幾乎所有的學術領域都會以同儕研究者來審查研究報告或是出版品，但是並非所有的領域都會使用匿名審查的過程。在像是社學會、心理學、以及政治學這些領域裡，幾乎所有的學術期刊都會進行匿名審查，而且通常都會有三位或是更多位審查者。相比之下，在像是生物學、歷史學、以及經濟學這些領域裡，則會使用混合的審查過程，換句話說，審查者有時候會事先知道作者的身分，而且有時候只有一到二位學者來審查一份研究。在動用到較多審查者的匿名審查過程裡，整個審查過程會較爲曠日費時，而且研究最後被接受的比例也會因此較低。[15] 匿名審查是一種非常謹慎的方法，用以確保品質控制，而其目的則是促進科學社群內的系統性質疑與普遍性的規範。

有些學術期刊的讀者群非常廣泛，而且也受到高度的推崇。這些學術期刊所收到的論文，也多於它們所能刊出的數量。例如，主要的社會科學期刊，像是《美國經濟學評論》（*American Economic Review*）、《美國社會學評論》（*American Sociological Review*）、《美國政治科學評論》（*American Political Science Review*）、以及《社會問題》（*Social Problems*）等等，在它們所收到的全部論文裡，大約只會接受百分之 10 到 15 的文章。即使是那些知名度較低的期刊，一般而言也會拒絕大約一半的文章。研究者的文章獲得出版，代表了其研究暫時地被科學社群所接受。至於出版一本書籍，則牽涉到稍微不同的的審查過程，其中還包括了成本和銷售的考量，但是，書籍的接受比例通常會比期刊還要來得低。[16]

一般而言，在書報攤上所看到的流行雜誌，其文章的作者都可以得到稿酬，但是，當科學家的文章在學術期刊上被刊登出來時，他們是沒有稿費可拿的。事實上，爲了要使其投稿的文章能夠順利地進入審查程序，作者通常會被要求支付一小筆費用，以補貼行政管理方面的成本。社會科學家希望透過學術期刊的方式，讓同儕們（也就是其他的科學家與研究者）能夠得知並且接觸到他們的研究。同樣地，審查者在審查研究報告時，亦未支領任何費用。他們之所以接受這份工作，是因爲他們認爲這是身爲科學社群的一份子所應盡的責任。對於那些能夠在最重要的學術期刊上發表許多文章的研究者，社學科群的成員會給予極高的尊崇，因爲這些獲得刊登的文章，能夠證實這些研究者具有高度的技巧，並且是促進科學社群主要目標的領導人物，亦即，他們對於科學知識的累積作出了貢獻。

社會科學家主要是透過在學術期刊上發表研究的方式，以得到同儕尊敬、在科學社群內獲得榮譽、以及建立身爲技巧純熟的研究者的聲望。當一名科學家得到較多來自於同儕（例如，像是其他學識廣博的社會科學家）的尊敬時，他就越能夠往學術社群的中心位置移動。在期刊上發表研究、以及隨之而來的同儕的尊敬，也能

幫助一名社會科學家取得更多的研究經費、得到教職、擁有一批頂尖的學生跟著進行研究、改善工作環境、得到利潤豐厚的工作機會、以及提高本身的薪資等等。[17]

即使從未在學術期刊上發表過任何一篇文章，也有可能會讀到一些學術期刊上的文章。學術期刊文章是科學研究系統裡相當重要的一個部分。大多數新的科學知識都會首先發表在學術期刊上。積極、勤奮的社會科學家以及大學教師會定期地閱讀學術期刊，以得知正在產生的新知識，以及被採用的研究方法。

科學作為一種轉化的過程

在研究過程中，社會科學家會應用各式各樣的科學方法，以把想法、直覺臆測、以及問題〔有時候被稱作*假設*（*hypotheses*）〕等等轉化爲新的知識（我們將在第五章裡進一步討論有關假設的問題）。因此，社會科學的研究過程基本上會把我們的想法、理論、臆測、或者是問題，轉化成爲帶有真正價值的「成品」（“finished product”），也就是新的知識。這些新的知識能夠增進我們對於社會世界及其運作方式的理解。此外，新的知識也有可能會被用來幫助我們解決問題，或是進一步擴展未來的知識以及理解。

許多初次接觸社會研究的人，都會覺得茫然、不知所措，而且也會覺得進行一項研究超出了他們的能力範圍。進行社會研究必須要擁有分析推理的能力、複雜的技術性技能、長時間全神貫注的能力、以及付出極多的時間。然而，只要投入時間、多加練習、以及接受教育，大多數的大學生都會發現他們自己能夠精熟進行研究的基本原則。學習進行社會研究，其實與學習許多其他的活動沒有什麼不同。一開始一定是從較小、較簡單的部分做起，然後不斷反覆地練習，並且從自己的經驗和錯誤中學習。漸漸地，將會看到自己的進步，並且有能力進一步從事較大、較爲複雜的研究。除了吸收並理解科學態度之外，也必須學習如何以及何時去應用特殊的研究技巧。在讀完這本書之後，應該能夠掌握社會科學研究的方法以及方法論，並且也有足夠的能力去進行一些研究。

社會研究的類別

你們或許會認爲，社會科學研究意味著進行一項調查或是實驗，而且可能還會用到各種圖表以及表格之類的高級統計學。或者，也可能會認爲，社會科學研究所涉及到的，就是在一些自然的環境背景中，像是咖啡館、家庭聚會、或是教室裡，去仔細觀察人們所進行的各種日常活動。上述這二種看法其實都只有部分正確而已。有些社會科學研究的確會牽涉到量化資料（也就是以數字形式呈現的資料），但是其他的研究則會使用與統計無關的質化資料（也就是非數字形式的資料）。

在閱讀過本書的前面數個章節之後，我們將會仔細地分析討論量化與質化資料，以及進行社會科學研究的相關取向。這兩種取向都使用了多樣的研究技術（例如，調查、訪談、民族誌等等）以蒐集並且分析實徵資料。雖然在量化與質化研究

表 1.2 量化取向以及質化取向

量化取向	質化取向
測量客觀的事實	建構社會的事實與文化的意義
焦點集中於變項	焦點集中於互動的過程以及事件
信度為關鍵因素	本真性（authenticity）為關鍵因素
不帶任何價值觀判斷	在研究中清楚明確地呈現出價值觀
理論與資料各自區分開來	理論與資料融合在一起
獨立於情境脈絡之外	受到情境所限制
許多的個案與受試者	極少的個案與受試者
統計的分析	主題式的分析
研究者採取超然的立場	研究者參與其中

資料來源：Crewsell (1994), Denzin and Lincoln (2003a), Guba and Lincoln (1994), Marvasti (2004), and Tashakkori and Teddlie (1998)。

之間確實存在著一些真正的差異，但是這兩種研究卻有許多重疊的部分。不幸的是，某種取向的支持者並非總是能夠瞭解或是欣賞另外一種取向。有些社會科學家把這兩種取向之間的差異，當作是與其他人之間的戰爭。李文（Levine, 1993: xii）把量化取向稱作是「真正的社會科學」，並且宣稱量化取向在和質化研究的對抗中已經「贏得戰爭」。另一方面，丹辛與林肯（Denzin and Lincoln, 2005: ix）兩人則主張，「在社會科學以及其他相關領域中，質化革命的擴張與影響的範圍之大，簡直可以說是令人瞠目結舌。」

雖然量化與質化這兩種取向共享一些核心的科學原則，但是兩者之間仍有非常明顯而重大的差異之處（請參閱表 1.2）。每一種取向都有其長處與限制，而且也都有其適用的主題或是議題、以及對社會生活提供真知灼見的經典研究。進行量化或是質化研究的社會科學家，會盡力去避免我們在稍早之前已經討論過的誤判以及錯誤。所有的社會科學家都會以有系統的方式蒐集資料，進行仔細的比較，並且採取批判性的思考方式。藉由瞭解兩種不同研究取向的方式，你就能夠清楚地瞭解社會科學研究的整體範圍，並且以互補的方式來運用這兩種不同的取向。

拉金（Ragin, 1994a: 92）曾經解釋過，在濃縮資料以及強化資料方面，這兩種取向如何補足彼此的缺失之處：

> 當我們把質化研究方法與量化研究方法相互對照之後，就能夠看出所有的質化研究方法所共有的幾個關鍵特徵。大部分的質化資料技術，都是濃縮資料的技術，而濃縮資料的目的則在於看見事物的全貌……相比之下，量化方法則可被視為是強化資料的技術。當資料被強化的時候，就有可能更加清楚地看到個案的關鍵面向。

最理想的研究，是結合兩種不同研究方法的研究，以同時獲得量化與質化取向的長處，然而，基於以下的一些原因，此種研究並不常出現。首先，採取混合方法的研究比較耗費時間。其次，很少有研究者能夠同時精通超過一種以上的研究方法。除此之外，每一種研究方法在指引研究過程方面都有其涇渭分明的邏輯，而把這些不同的邏輯混合在一起時，則會明顯地增加研究的複雜程度。

研究過程的步驟

步驟

在進行一項研究時，我們必須遵循一連串的步驟。然而，根據我們所依循的是量化取向或是質化取向、以及我們正在進行的社會研究型態的不同（第二章以及第四章會討論到不同型態的研究），確切的順序以及具體的步驟會有些變化。稍後你將會看到，在這兒所概述的幾個步驟可能會有些簡化、並且將實際過程理想化了，但是，這些步驟仍然可以當成一個有很有的起始點。

量化取向的社會研究

1. *選擇主題*（*Select a topic*）主題可能是一個籠統的研究領域、或者是由專業或個人興趣所衍生出來的議題。研究主題的範圍是相當廣泛的，例如像是離婚所造成的影響、造成犯罪行為的原因、無家可歸的情況所造成的衝擊、或是菁英分子如何使用媒體等等。
2. *聚焦研究問題*（*Focus the question*）光是選擇出一項主題的話，會太過於空泛而無法實際地進行研究。因此，這也使得下一個步驟顯得相當重要：我們必須把主題的範圍縮小，並聚焦在一個特定的研究問題上面，以便進行研究來加以處理。通常，研究者必須去回顧、檢閱相關的研究文獻（在第五章中將會討論到這個議題），並且發展出假設（我們將在第六章中討論到關於假設的議題），而且這個假設通常是來自於社會理論（相關討論請參閱第三章）。例如，一項廣泛的主題－造成犯罪行為的原因－可以成為一個經過聚焦之後的研究問題：就來自於東亞地區的男性青少年移民而言，那些與家鄉文化有著強烈連結、並且尚未與新社會同化的移民，是否比那些與家鄉文化連結較弱、並且已經同化的移民更有可能從事犯罪的行為？請注意到，原本一開始較為廣泛的主題，也就是造成犯罪行為的原因，已經變得比較聚焦了。我們把焦點集中在一個造成犯罪行為的特定原因上面（也就是同化的程度），並且把研究對象集中於一個特定的群體上（也就是來自於東亞地區的男性青少年移民）。
3. *設計研究*（*Design the study*）一旦我們決定好研究問題之後，就必須開始設計研究（關於設計研究的議題，將於第六章至第十一章中進行討論）。在設計一項研究時，研究者必須作出許多決定，像是個案的種類或是要挑選出來的樣本、如何

測量相關的因素、以及要使用何種研究技巧（例如，使用問卷或是進行實驗）等等。同樣地，在這個階段裡，研究者也必須要依據理論來作出決定。

4. *蒐集資料*（*Collect data*）在我們鉅細靡遺地設計好一項研究之後，就必須仔細地記錄並且確認資訊，而這些資訊都是以數字的形式呈現的。接下來，如果這些數值資料尚未轉換成可供電腦讀取辨識的格式時，我們就必須把這些數值資料先進行轉換。
5. *分析資料*（*Analyze the data*）在這個步驟裡，研究者通常必須使用電腦軟體來處理數值資料，以產生許多的圖形、表格、以及統計結果（請參閱第十二章）。這些由電腦所產生的文件資料，可以提供整體資料的濃縮圖像。
6. *解釋資料*（*Interpret the data*）在我們產生圖形、表格、以及統計結果之後，我們接著必須決定這些結果所代表的意義。我們必須檢視這些經過分析的資料、利用我們對於研究主題的瞭解、並且引用理論來回答研究問題。我們還必須考量到其他可能的解釋方式，把我們的結果與之前的研究結果進行比較，並且從中歸納出我們所得到的較廣泛的意義。
7. *發表研究結果*（*Inform others*）在這最後一個階段裡，我們會以一種特殊的格式（請參閱第十六章）撰寫一份關於這項研究的報告，並且在報告中呈現該項研究以及研究結果的描述（請參閱圖 1.1）。

接下來，讓我們來看看三個量化取向的社會研究的例子。在這三個例子當中，每一個例子都是一種量化研究的類別，而且也是本書稍後某一章節中的討論重點，其中包括了實驗（請參閱第九章）、樣本調查（請參閱第十章）、以及現有的統計數據（請參閱第十一章）。

作者以及研究名稱：羅威瑞以及同事（Lowery and colleagues, 2007），《無意識促發對於學業表現的長期影響》（*Long-Term Effects of Subliminal Priming on Academic Performance*）（譯註：「無意識促發」亦可譯爲「閾下促發」或「隱式促發」）。

1. *選擇主題*　促發效果與學業表現。
2. *聚焦研究問題*　無意識地接受了與智力相關的「促發」效果的大學生，在考試方面是否會有較佳的表現？所謂*無意識地*（*subliminally*），是指在呈現某項事物時，接受者並未清楚地意識到該項事物的存在。而當某個字、影像、或是資訊使一個人警覺到、準備好、或是「設定好」去進行後續行爲時，就是產生了「*促發*」（*priming*）效果。
3. *設計研究*　作者們進行了二項相似的實驗。在第一個實驗裡，受試者是選修統計入門課程的七十名大學生。而在第二個實驗裡，受試者則是選修社會心理學入門課程的七十八名大學生。在這兩個實驗裡，作者們在電腦螢幕的不同邊上顯示出一些字讓學生們看到。他們告訴學生們，該項研究的主題是關於他們找出這些字的能力（這其實不是真的）。其中有一半以隨機方式選出的學生，所看到的字是和智力有關的字（例如，像是精明的、聰穎的、天才、受過教育的等等）。另外

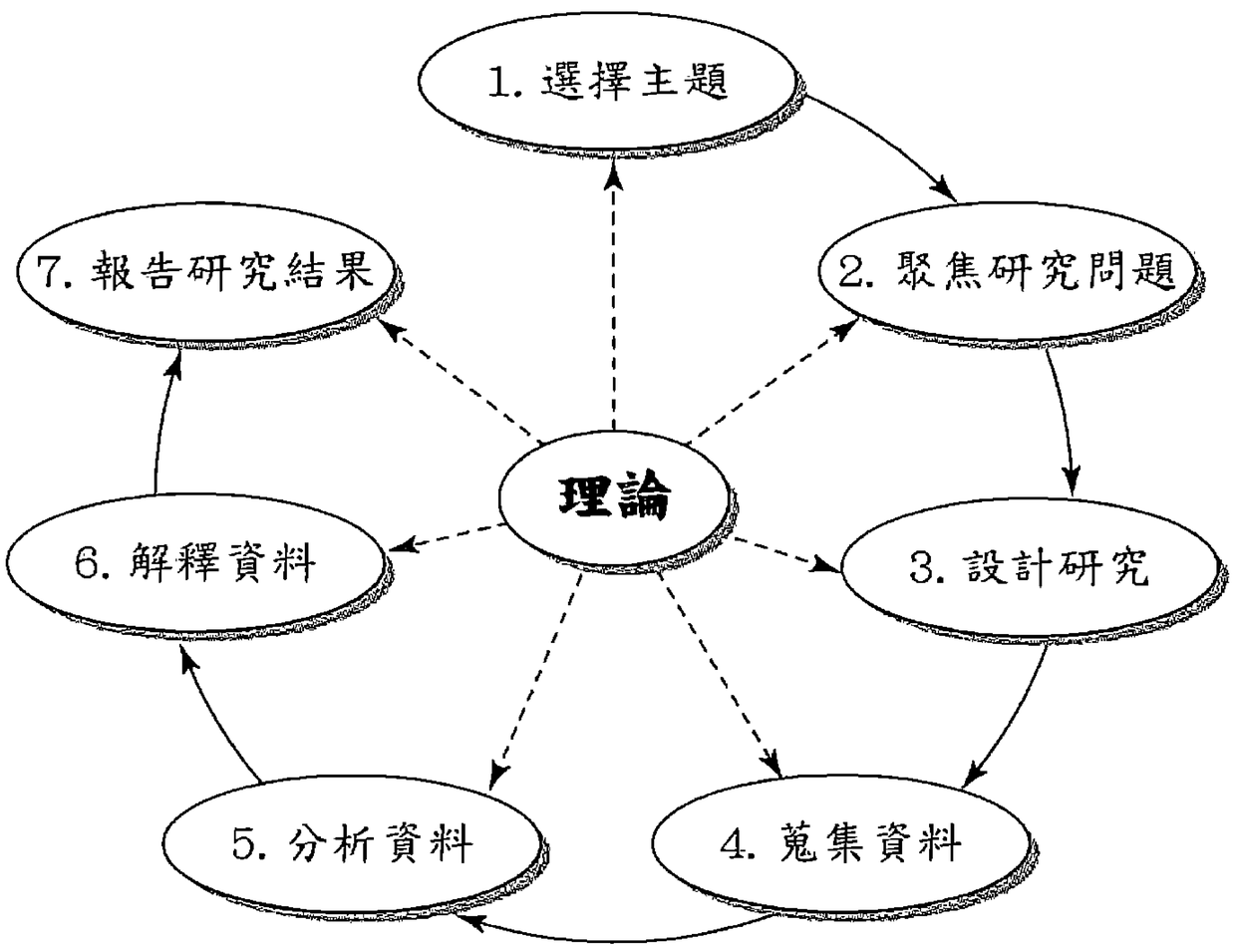


圖 1.1　量化研究過程的步驟

一半隨機選出的學生，則會看到一些不相關的字。在兩個實驗裡的學生接著參加一次練習性的測驗。而在數天之後，這些學生則在各自的課程中接受正式的測驗。

4. *蒐集資料*　在這項研究裡所要蒐集的資料，是在統計學以及社會心理學入門這兩門課程裡，學生們在練習測驗以及正式測驗中所獲得的兩個分數。
5. *分析資料*　作者們檢視了各種不同的表格，並且進行了統計考驗。
6. *解釋資料*　結果顯示，在這兩門不同的課程裡，那些曾經接觸過、或者是被智力相關的字彙所「促發」的學生，在兩次測驗裡都獲得明顯較高的分數。
7. *發表研究結果*　描述本項研究以及研究結果的文章，被刊載於《基礎以及應用社會心理學》（*Basic and Applied Social Psychology*）這本學術期刊上。

理論如何與研究結合？作者重新驗證了關於無意識促發效果的理論。他們所檢視的，是在促發事件過後數日，促發效果是否持續存在的問題。

作者以及研究名稱：艾德傑爾與崔比（Penny Edgell and Eric Tranby, 2007），《在美國境內，宗教立場在解釋種族不平等現象方面所造成的影響》（*Religious Influences on Understandings of Racial Inequality in the United States*）。

1. *選擇主題*　宗教以及種族之間的態度。
2. *聚焦研究問題*　白人新教徒的基督教次文化與信仰體系，會鼓勵、或是阻止人們對於種族不平等或是非裔美國人採取一種個人主義式的、不支持的立場？
3. 設計研究　作者們決定在 2003 年進行一項大型的、全國性的調查，其調查對象爲隨機選出的 2,081 名美國境內的成年人。
4. *蒐集資料*　在約三十分鐘的電話訪談過程中，這些隨機選出的成年人回答了許多關於社會背景、宗教信仰、對於種族不平等的解釋、以及對於非裔美國人的想法等等問題。
5. *分析資料*　作者們檢視了一些數字表格，內容是各種百分比例以及統計考驗。

6. *解釋資料*　作者發現，在接受調查的受訪者中，那些擁有強烈而保守的新教徒基督教信仰的人、以及那些最投入於宗教活動的人，在解釋黑人所受到的不平等待遇時，會側重於個人的問題（例如，個人所具有的缺點、或是缺乏動機等等），而不是結構上的問題（例如，種族歧視）。除此之外，在保守的基督教徒裡，女性的看法異於男性的看法；而受過教育者的看法，也與教育程度較低者的看法有所不同。
7. *發表研究結果*　作者們完成了關於這項研究與研究結果的報告，並且投稿到《社會問題》這份學術期刊。

理論如何與研究結合？本項研究的作者檢視了一個理論，該理論認爲白人新教徒的次文化會促使人們對於社會以及政治議題發展出特定的態度。該理論較不重視結構方面的解釋（像是歧視、政府的幫助等等），而把重點放在個人主義式的、個人應該自立自強方面的解釋。

作者以及研究名稱：麥可貝以及索柏列斯基（Rory McVeigh and Julian Sobolewski, 2007），《紅色郡，藍色郡，以及性別與種族所造成的職業隔離》（*Red Counties, Blue Counties, and Occupational Segregation by Sex and Race*）（譯註：在美國政治裡，傳統上紅色代表共和黨，而藍色代表民主黨）。

1. *選擇主題*　社會不平等以及投票行爲。
2. *聚焦研究問題*　由性別以及種族所造成的職業隔離（此種隔離現象是社會不平等的主要來源），是否會影響到人們在 2004 年美國總統大選中的投票行爲？所謂的*職業隔離*（*occupation segregation*），是指某個特定的團體（例如，某個性別、或是某個種族）幾乎獨佔性地把持了某種類型的工作。
3. *設計研究*　在政府所蒐集的、屬於各郡層級的資料中，作者發現了一些特殊的因素：種族與性別會造成對特定總統候選人的偏好、以及造成職業隔離。他們也考量了在某一個郡的勞動市場裡，那些或許會破壞或是削弱職業隔離程度的一些特色（例如，在該郡人口中的種族混合比例、女性以及非白人居民所擁有的學歷、移入該郡的流動性程度高低等等）。
4. *蒐集資料*　資料來自於美國政府對於職業、人口統計、以及投票方面的人口普查記錄。
5. *分析資料*　作者檢視了爲數眾多的相關性、圖表、以及統計考驗。
6. *解釋資料*　作者發現到，在各郡的勞動市場裡，不論是職業隔離、或是性別隔離，都會與選舉結果有所關聯。在勞動市場較爲平等、而且女性就業率較高的郡裡，民主黨的候選人會得到較多的選票。而在勞動市場高度隔離的郡裡，尤其是還存在著一些可能會影響或減弱此種隔離現象的其他狀況時，共和黨的候選人就會獲得比較多的選票。
7. *發表研究結果*　作者們撰寫了一份報告，在其中描述了這項研究以及研究結果，並將這份報告投稿到《美國社會學期刊》（*American Journal of Sociology*）這本學術期刊上。

理論如何與研究結合？作者利用種族競爭理論（ethnic competition theory）、以及分割勞動市場理論（split labor market theory），以解釋在各郡層級上的不平等現象如何影響了當地的政治氣候以及投票行爲。

質化取向的社會研究　許多採用質化取向的社會科學家，所依循的是與量化研究稍微有些不同的步驟。這些步驟也會根據研究者所使用的特定質化研究方法而有所變化。除此之外，質化取向的特色就是比較不固定，而且也比較不是線性的（所謂線性的，是指必須依照順序一個步驟接著一個步驟進行）（相關討論請參閱第五章）。

1. *確認自我以及情境脈絡*（*Acknowledge self and context*）如同進行量化研究一樣，在進行質化研究時，社會科學家也是從一項主題開始，但是，與此同時，研究者也必須先進行自我評估（self-assessment），並確定該項主題在社會—歷史脈絡中的位置。許多質化研究者會依賴個人信念、傳記、或是當前的特殊議題，以找出一個他們感興趣的、或是重要的主題。
2. *採取一個觀點*（*Adopt a perspective*）質化研究者可能會去仔細思考理論—哲學*典範*（*paradigm*）（在第四章中將會有進一步的討論），或是一邊進行研究、一邊和其他研究者持續地進行討論。這樣的作法並不是在縮小主題的範圍，反而意味著研究者其實是選擇了一個特定的研究方向，而且在此方向之內可能包含了許多潛在的問題。

3-6.*設計研究，以及蒐集、分析、解釋資料*　就如同量化研究一般，質化研究者也必須*設計研究*（請參閱第六章）、*蒐集資料*（請參閱第十三章與第十四章）、*分析資料*（請參閱第十五章）、以及*解釋資料*。與量化研究者較爲不同的是，質化研究者比較有可能會同時進行蒐集、分析、以及解釋資料的工作。這是一個不斷變動的過程，必須在各個步驟之間來來回回好幾遍。而且，研究者經常不是僅僅使用或是考驗過去的理論而已，還有可能會建立新的理論。在*解釋資料*的階段，質化研究者會創造出新的概念以及理論方面的解釋。

7. *發表研究結果*在量化和質化取向裡，這個步驟都是很類似的。不過，要再提醒一次的是，根據所使用的研究方法的不同，研究報告的格式與寫法也會有所差異（請參閱第十六章）。（請參閱圖 1.2）

接下來，讓我們一起來看看兩個質化研究的例子。每一個例子都是一種質化研究的類別，而且也是本書稍後某一章節中的討論重點，其中包括了田野研究—民族誌（請參閱第十三章），以及歷史比較研究（請參閱第十四章）。

作者以及研究名稱：范卡德胥（Sudhir Venkatesh, 2008），《當一天的幫派頭頭》（*Gang Leader for a Day*）。

1. *確認自我以及情境脈絡*　作者描述了他的個人興趣以及背景。他原本感到有興趣的對象，是內都市裡的貧窮問題，但是後來則轉移到在都市住宅計畫中的幫派上面（譯註：此處所指的都市住宅計畫，通常是指由政府出資興建的廉價住宅，以

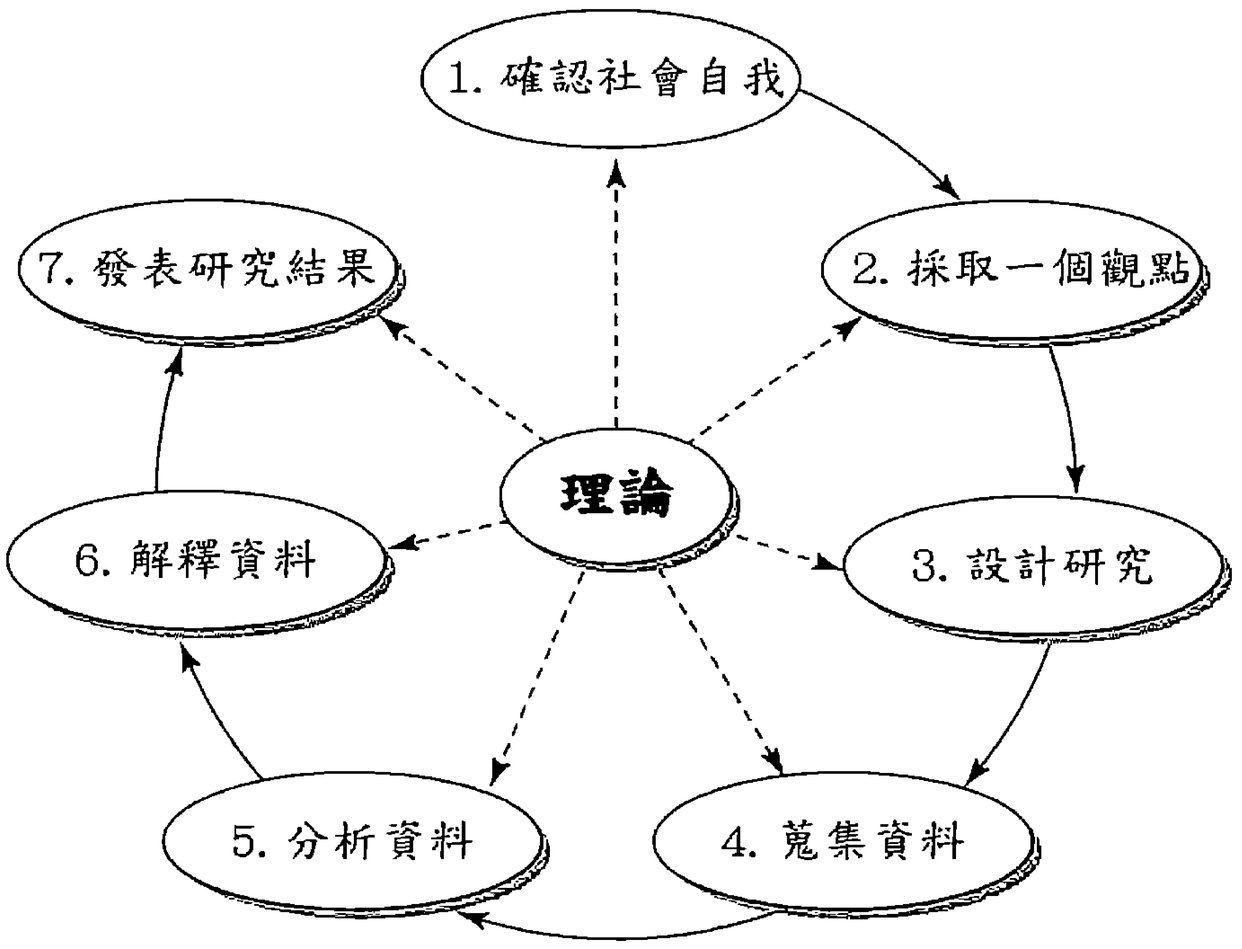


圖 1.2　質化研究過程的步驟

供低收入家庭居住）。作者也解釋了此種興趣轉移的過程。

2. *社會—文化脈絡*　本研究的實體—社會背景，是芝加哥南部地區的一項都市住宅計畫，該地鄰近於芝加哥大學，而作者當時則是芝加哥大學的研究生。在這個住宅計畫中，販賣毒品的幫派也參與了其中的運作。除此之外，住戶之間的貧窮比例非常高，而且住戶幾乎清一色全部都是黑人。

3-6.*設計研究，以及蒐集、分析、解釋資料*　作者一開始本來試圖要進行量化研究，但是後來則放棄了此種研究方法。相反地，在從 1990 年到 1998 年這八年之中，他每一周都會花數天的時間到這個住宅區去觀察幫派成員和住戶，並且和他們交談。而且，他每天都會很詳盡地記錄下他所看到的、聽到的、參與的、以及思考的事物。

7. *發表研究結果*　在原本的研究結束十年之後，研究結果才出現在《當一天的幫派頭頭》這本半學術的書籍裡，不過，在這十年之間，其實作者已經發表過幾項研究與書籍，其內容都與這項研究有所關聯。

理論如何與研究結合？與許多民族誌研究相同的是，本項研究主要是描述性質的研究，因而沒有太多理論的成分。對於幫派如何爲一個當地社區提供社會組織以及服務、販賣毒品的經濟學、以及當地的窮人爲了生存下去而必須與許多其他組織進行協商的過程，作者只提供了少量的理論而已。

作者以及研究名稱：麥可卡蒙以及六名同事（Holly McCammon and six colleagues, 2008），《成爲完整的公民：美國女性的陪審團權利運動、改革的步調、以及策略性調適》（*Becoming Full Citizens: The U.S. Women's Jury Rights Campaign, The Pace of Reform, and Strategic Adaptation*）

1. *選擇主題*　女性獲得完整的公民權利。

2. *社會—文化脈絡*　在美國女性於 1920 年贏得投票的權利之後，並未得到成爲陪審團成員的權利。一直到 1975 年，最高法院才同意讓女性擁有成爲陪審團成員的權利。然而，在不同的州裡，女性獲得此項權利的時間點也有極大的差異（而且有時候會被剝奪、然後又得到此項權利）。由於這項議題得到許多女性團體的提倡與支持，因此這項議題已經被熱烈爭論了數十年之久。
3. *設計研究，以及蒐集、分析、解釋資料*　在兩年的時間之內，七位作者大部分的時間都花費在蒐集十五個州於 1910 年代至 1960 年代晚期之間的陪審團權利運動的資料。在這些州裡，他們造訪了二十二個檔案保管處（保留歷史記錄的專門圖書館）。在這十五個州的每一州裡，他們都去檢視了陪審團權利運動的組織記錄，訪問了當地報紙以及相關的雜誌，並且也閱讀了所有相關的法律以及政治文件資料（例如法院的判決結果、立法的公聽會、以及相關的法令規定等等）。除了分析每一州以及運動組織的相關細節之外，他們也審視了每一州實施女性陪審團權利所需花費的時間，並且把每一個組織以及其活動的具體特徵加以分門別類。他們所獲得的重大發現是，在那些女性最快速地獲得陪審團權利的州裡，相關組織都曾採取過一些策略性的行動。這些組織會持續不斷地調整其要求、尋求一定範圍內的政治盟友、並且改變其論點的表達方式。而在進展極爲緩慢的州裡，運動團體則是零星散佈、沒有一致的行動、或是不具有應變的彈性，而且也未能審時度勢，因而無法發揮影響力。
4. *發表研究結果*　描述本項研究及結果的研究報告，已被刊載於《美國社會學期刊》這份學術期刊裡。

理論如何與研究結合？作者們試圖要解釋爲什麼有些社會—政治運動能夠迅速地達成其政治目標，而其他的同類型運動則進展緩慢。他們以過往的社會運動理論爲基礎，並且深化了社會—政治運動中的「策略性調適」這個新的觀念。

在圖 1.1 以及圖 1.2 中所顯示的研究過程的七個步驟，其實是過於簡化的。在實際的情況裡，我們很少會在完成步驟一之後，就完全不管這個步驟、然後直接進展到步驟二，或是進展到其他步驟。研究比較像是一種相互作用的過程，而且所有的步驟都會彼此混合在一起。一個較後面的步驟，可能會促使研究者去重新思考之前的步驟。整個過程並嚴格的線性過程，而是可能會同時往數個方向流動，直到抵達某個終點爲止。研究也不會突然就在步驟七時戛然而止。這是一個持續進行的過程，而且某項研究的結束，通常會激發出新的思考方式以及新的研究問題。

這個七步驟的循環過程，只適用於單獨一個研究。每一個研究都是建立在先前研究所奠定的基礎上，並且也會對整體更大的知識體系作出貢獻。進行科學研究以及累積新知識，是一個更爲廣大的過程，而且也需要許多的研究者進行爲數眾多的研究。單一的研究者或許會同時從事多項研究，或者，數個研究者也可能會一起合作進行一項研究。同樣地，一項研究可能會產生一篇或是數篇學術文章，而且有時候在同一篇文章裡可能就報告了數個較小型的研究。

為什麼要學習進行社會研究的方法？

任職於大學、研究中心、以及政府機關裡的專業社會科學家，在進行研究時通常會有數個助理以及技術人員的協助，而他們的研究結果通常會出現在專門的學術期刊、或是大學的教科書裡。他們所進行的研究能夠擴展我們對於社會世界的瞭解，並且會間接地影響到廣泛的公共知識。你或許想要學習如何進行社會研究的原因之一，就是爲了要提昇關於社會世界的知識，並且避免那些人們在以其他的、非科學的方法創造知識時所會犯下的錯誤。

那些爲報紙、電視網、市場調查公司、學校、醫院、社會服務機構、政黨、顧問公司、政府機關、人事部門、公眾利益組織、保險公司、以及法律事務所工作的人們，也會從事社會研究。從事社會研究是他們工作的一部分，而且他們也會採用相同的社會科學研究技術。他們的研究結果主要是供內部使用，而不會與大眾分享、或是刊載於期刊上。不過，比起過去利用猜測、預感、直覺、以及個人經驗等方式來作出決定的作法，此種以研究發現爲基礎而作出的決定，比較有情報方面的根據、而且也比較不會帶有偏見（請參閱重點摘要 1.2：實務工作者與社會科學）。除了擴展知識之外，你可能想要學習如何進行社會研究的第二個原因，則是一個實務應用方面的原因：改善制訂決策的過程。

不幸的是，有一些人以及組織誤用或是濫用了社會研究，像是採用不嚴謹的研究技術、曲解研究發現、或是操縱研究以刻意符合先前已經確定的結果等等（我們將在第三章中探討更多此類事例）。除此之外，有些人相信，他們正在被過度地研究、或是認爲各種研究已經多到令他們感到不堪其擾的地步了。例如，許多人已經拒絕接受選舉期間所進行的出口民調，而且回覆調查的比例也已經下降了。對於社會研

重點摘要　1.2

實務工作者與社會科學

科學不會、也不能夠提供人們固定、絕對的「真相」(“Truth”)。這是因為科學是一種緩慢而不完整的過程，其目的在於逐漸減少不真實或是虛假的情況。科學就像是一場追尋最佳可能答案的探索之旅，而進行這場探索之旅的人，是一群獻身於科學的人，他們以一種謹慎的、系統性的、心胸開放的方式，鍥而不捨地努力工作著。許多人對於此種勤勉刻苦的步調、躊躇不前的進展、以及科學中的不確定性感到不自在，因為他們要求的是立即的、絕對的答案。許多人因此轉而求助於宗教狂熱者或是善於蠱惑人心的政客，因為他們能夠提供大量最終的、決定性的真相。對於那些必須在日常工作裡迅速作出決定的勤勞的實務工作者(例如，服務人群的工作者、專業的健康照護人士、司法審判人員、記者、或是政策分析家等等）而言，此種情況究竟意味著什麼呢？他們是否必須放棄科學思想，轉而倚賴常識、個人信念、或是政治教條呢？當然不是！他們其實也能夠使用社會科學的思考方式。而且，雖然要這麼做確實很困難，但卻還是有可能做到的。他們必須認真地盡力去找出當前能夠取得的最佳知識；採用謹慎而且獨立的推理方式；避免已知的錯誤或是謬見；並且對於那些提供完整、最終答案的教條維持謹慎小心的態度。實務工作者必須永遠對於新想法抱持著開放的態度、採用多重的資訊來源、並且經常質疑那些支持某個行動方案的證據的真實性。

究遭到誤用，人們會產生一些負面反應，而此種反應則可能會造成對整體研究的負面觀感。因此，你可能想要學習如何進行社會研究的第三個原因，則是讓自己有能力去分辨何者是合法的、有價值的研究，而何者又是僞造的或品質低劣的研究、僞科學、以及遭到誤用的研究。

結論

本章討論了何謂社會研究、研究過程的運作方式、以及是誰在進行研究。本章也描述了社會研究之外的其他選擇，也就是那些能夠得到快速的、容易的、以及實用知識的方法，然而，這些知識卻常常含有許多謬誤、錯誤的資訊、以及不正確的推理方式。本章也呈現了科學社群運作的方式、社會研究與科學活動的融合方式、以及科學規範和期刊文章爲何對於科學社群如此重要的原因。除此之外，本章也概略描述了研究的步驟。

社會科學研究的對象是人（*people*）、與人有關、並且是由人所進行的。雖然必須要注意到一些原則、規定、或是程序，但是社會研究在本質上仍是屬於人的活動。社會研究者都是和你我一樣的普通人。他們渴望著創造以及發現知識，並且認爲進行社會研究是一件非常有趣而且令人興奮的事情。他們進行研究以發現新的知識，並且瞭解社會世界。不論你是否會成爲一名專業的社會研究者、一個在工作上必須應用研究技術的人、或者只是一個運用研究結果的人，你都可以從學習關於研究過程的知識裡獲得益處。如果你能夠開始在你自己和研究過程之間建立一種個人連結的話，你的生命將因此而顯得更加豐富。

米爾斯(Mills, 1959: 196)在他所著的《社會學的想像》(*Sociological Imagination*)一書中，爲我們提出了一個彌足珍貴的建議：

> 你必須學著把你的生活經驗運用於你的智性工作之中：也就是持續不斷地檢視並且解釋自己的智性工作。就這個意義而言，技巧乃是你自己的中心，而且你會在個人的層次上參與了每一個你可能從事的智性工作成果。

問題回顧

1. 除了社會研究之外，還有哪些知識來源的其他替代選擇？
2. 爲什麼社會研究通常會比上述這些知識來源的其他選擇更爲優越？
3. 社會研究是否一定是對的？它能否解答任何的問題？試說明之。
4. 在不同的時代裡，科學與神喻如何達到類似的效果呢？
5. 何謂科學社群？其所扮演的角色爲何？
6. 科學社群的規範是什麼？這些規範具有什麼樣的影響力？
7. 一項研究若要被刊登在一份學術性的社會科學期刊上，必須經過什麼樣的過程？

8. 在從事一項研究計畫時，會涉及到哪些步驟？
9. 當我們說研究步驟並非是嚴格固定的時候，代表了什麼意思？
10.什麼樣的人會從事社會研究？他們從事社會研究的理由爲何？

注解

1. See Parker-Pope (2007) on the face cream study and related research.
2. On the limits to self-knowledge, see Wilson and Dunn (2004); on inaccurate eyewitness accounts, Wells and Olson (2003); on inaccurate risk evaluation, Gowda and Fox (2002) and Paulos (2001); on condoms in schools (Kirby et al., 1999); on SUVs, Bradsher (2002).
3. From Rampton and Stauber (2001:274–277, 305–306).
4. Results on geographic information are from *National Geographic* (2006). Results on UFOs, devils, and so forth is from Harris Poll (2003, 2005).
5. On media inaccuracy on psychiatric treatment, see Goode (2002), on the Muslim population, see Smith (2002), and on African Americans in poverty, see Gilens (1996).
6. Video News Reports are described by the Center for Media and Democracy http://www.prwatch.org/fakenews3/summary and Consumer Product Safety Commission http://www.cpsc.gov/businfo/vnrprod.html Also see Barstow and Stein (2005, March 13), "Under Bush, a New Age of Prepackaged TV News," *New York Times*; Aiello and Profitt. (2008).
7. On "faith-based" programs, see Goodstein, "Church-Based Projects Lack Data on Results," *New York Times* (April 24, 2001); Crary, "Faith Based Prisons Multiply," *USA Today* (October 14, 2007); Ferguson, et al. (2007); and Reingold et al. (2007). On restrictions of science in government, see Mooney (2005) and Union of Concerned Scientists (2004).
8. See Herrnstein and Murray (1994) and a critique in Fischer et al. (1996).
9. "Junk science" is discussed in Rampton and Stauber (2001 223).
10. For more on the scientific community, see Cole (1983), Cole, Cole, and Simon (1981), Collins (1983), Collins and Restivo (1983), Hagstrom (1965), Merton (1973), Stoner (1966), and Ziman (1968).
11. See Cappell and Guterbock (1992) and Ennis (1992) for studies of sociological specialties.
12. For more on the social role of the scientist, see Ben-David (1971), Camic (1980), and Tuma and Grimes (1981). Hagstrom (1965), Merton (1973), and Stoner (1966) discuss norms of science, and Blume (1974) and Mitroff (1974) talk about norm violation.
13. See Altman. "Drug Firm, Relenting, Allows Unflattering Study to Appear," *New York Times* (April 16, 1997); Markoff, "Dispute over Unauthorized Reviews Leaves Intel Embarrassed," *New York Times* (March 12, 1997); and Barry Meier, "Philip Morris Censored Data about Addiction," *New York Times*, (May 7, 1998).
14. Science's communication and publication system is described in Bakanic and colleagues (1987), Blau (1978), Cole (1983), Crane (1967), Gusfield (1976), Hargens (1988), Mullins (1973), Singer (1989), and Ziman (1968).
15. See Clemens and Powell (1995:446).
16. See Clemens and Powell (1995:444).
17. For more on the system of reward and stratification in science, see Cole and Cole (1973), Cole (1978), Fuchs and Turner (1986), Gaston (1978), Gustin (1973), Long (1978), Meadows (1974), and Reskin (1977).

第2章

社會研究的主要類別

研究的運用與讀者群　　單一時間點或是多重時間點
研究的目的　　資料蒐集技術
個案內或是跨個案研究　　結論

不論是社會學家、政治學家、或者是人類學家，進行學術研究的目的，都是在他們個別的領域裡設法去找出理論問題的答案。對比之下，進行應用社會研究的目的，則是善加運用資料以便作出決策。

——Herbert J. Rubin, *Applied Social Research*, pp. 6-7

在從大學畢業三年之後，提姆（Tim）和雪倫（Sharon）相約一起見面吃午餐。提姆問雪倫說：「那麼，你在社會資料調查公司擔任研究員，這份新工作做得如何呢？工作內容是什麼呢？」雪倫回答說：「現在我正在對教師進行一項橫斷式調查研究，這項研究本身是一個針對六家日間托育中心所進行的應用研究計畫的一部分。這個計畫的目的是要提供描述性資料，以應用於一項爲非營利基金會所準備的評估研究之中。」雪倫所提及的托育中心的研究計畫，觸及了社會科學研究的幾個面向。在本章裡，你將會學習到這些面向，並且會得到一份關於各種社會研究類別的「道路圖」（"road map"）。

在第一章中所呈現給你的，是一幅關於社會科學研究的簡化圖像。社會研究有著各種不同的型態與大小，而且我們也可以用數種方法來將社會研究歸類，像是實驗研究以及非實驗研究、個案研究以及跨個案研究、或是質化研究與量化研究等等。[1] 我們也可以依照五個面向（請參閱圖表 2.1）來將許多不同的研究進行組織歸類。這些面向包括了我們如何運用一項研究的發現，以及其主要的讀者群爲何；我們爲何要進行一項研究；個案的數目以及我們檢視個案的方式；我們如何探討時間的面向；以及決定要採用何種蒐集資料的技術等等。你可以把任何一項單一的研究，放在社會研究的這五個面向中的每一個面向裡去進行討論。

你將會發現到，學習並且瞭解這些面向彼此之間的相互關係，是有很有幫助的。首先，這些面向可以幫助你更易於瞭解你所聽到、或是在學術期刊上所讀到的研究報告。在你確認了一項研究的各種面向之後，你就可以快速地掌握這項研究所傳達的訊息、以及進行的方式。其次，當你在進行自己的研究時，你必須作出許多

圖表 2.1　社會研究的面向以及主要類別

研究的運用與讀者群

- 基礎研究
- 應用研究
  - 評估研究
  - 行動研究
  - 社會影響研究

研究的目的

- 探索性研究
- 描述性研究
- 解釋性研究

個案內或是跨個案

- 個案研究
- 跨個案研究

單一或是多重時間點

- 橫斷性研究
- 縱貫性研究
  - 時間序列研究
  - 同樣本多次研究
  - 世代研究
- 個案研究

資料蒐集技術

- 量化資料
  - 實驗研究
  - 調查研究
  - 非反應式研究（內容分析、次級資料分析、現存統計資料）
- 質化資料
  - 田野研究（民族誌、參與觀察）
  - 歷史比較研究

決定。你可以把這些面向想像爲你在發展出一個特定研究計畫的過程中，所會遭遇到的決策點。爲了要作出好的決定或是決策，你應該要清楚知道在每個決策點上的權衡取捨、以及各種的優點與缺點。除此之外，這些面向都是互相關聯的，而某些面向之間會容易產生連結性（例如，研究目標以及資料蒐集技術）。當你在學習這些面向時，你將會開始瞭解到，你應該要怎麼做才能夠結合各個面向，以處理重要的特定研究問題。

研究的運用與讀者群

社會研究有兩大分支，其中之一是較爲超然的「科學性的」、或是「學術性的」取向，而另外一派則是較爲務實的、實用的、以及行動的取向。不過，這並不是一種嚴格死板的分類方法。有許多研究者的研究工作兼具這兩種取向，或者，他們會在不同的生涯階段裡，從一種取向轉移到另一種取向。此兩種取向的差異之處，就在於如何運用研究發現、以及主要的讀者群爲何。

基礎研究

基礎研究（basic research）又稱作*學術研究*（*academic research*）或者是*純研究*（*pure research*），其目的在於促進對於社會世界的基礎知識。基礎研究是大多數新的科學觀念以及思考社會事件新方法的來源。科學社群正是基礎研究的主要讀者群。研究者利用基礎研究去支持或是反駁一些理論，而這些理論則是關於社會世界運作與改變的方式、各種事件背後的成因、以及爲什麼社會關係或是事件會以某種特定方式進行等等。

有些人會批評此種基礎研究的取向，並且問道：「基礎研究有什麼貢獻？」他們認爲進行基礎研究只不過是在浪費時間和金錢而已，因爲他們認爲基礎研究不但沒有任何立即性的效用、也無法解決任何一個具有急迫性的議題。雖然很多實務工作者也想要找到問題的答案，好讓他們可以在一周之內、一個月之內、或者是一年之內，就可以把這些答案轉化爲實際的行動或解決方案，但是，進行基礎研究的研究者或許必須投入數年的時間，煞費苦心地尋找問題的答案，而這些答案則可能會在將來數十年的時間之內，重新塑造我們的思考方式。許多基礎研究無法在短期間之內提供實際方面的應用價值，但是卻能夠爲知識以及廣泛的理解奠定基礎，而這些知識與理解則會對許多議題、政策領域、以及研究領域造成極大的影響。基礎研究也是所有研究者使用的一些工具（包括方法、理論、以及觀念等等）的主要來源。幾乎所有在知識方面的重大突破與影響深遠的進展，都是源自於基礎研究。基礎研究爲核心理解紮下根基，並且對於一些在研究進行當時或許仍未存在的議題，也可能有所啓發與影響。

基礎研究可能會去檢視一些在表面上似乎不切實際的議題，因爲由這些議題所產生的知識，可能在未來的數年、甚至數十年之間都無法應用於實務層面上。通常，只有當各種不同的知識在經過一段長時間的累積進展之後，我們才能夠看到實際方面的應用出現。舉例而言，在 1984 年時，當時英國萊斯特大學（University of Leicester）的遺傳學家傑佛瑞斯（Alec Jeffreys）正在從事一項基礎研究，其研究主題則是基因的演化。在他發展一項新技術的過程中，由於一項間接、偶然的副作用，他意外地學會

基礎研究　基礎研究的設計目的，是為了促進並且發展關於世界運作方式的基礎知識，並且藉由專注於「為什麼」的問題的方式，建立並測試理論性的解釋。科學社群是基礎研究的主要讀者群。

了如何製造人類 DNA 的「指紋」(“fingerprint”)，也就是每個人的 DNA 裡那些獨一無二的標記。其實，發展出這項分析技術並非是他的本意。傑佛瑞斯甚至還說過，如果他原本的目標就是要創造出 DNA 指紋的話，他可能反而永遠不會想到這項技術。到了 1990 年代中期時，該項技術的實際應用方法已經被發展出來了。時至今日，DNA 分析法已經被廣泛地運用於犯罪調查以及其他的領域中。有許多重大的突破與創新都有類似的發展過程，也就是皆源於起初看似毫無任何關聯的基礎研究。

對於像是「爲什麼會發生偏差行爲？」這類的基礎研究的問題，極少有實務工作者（例如，警察、青少年罪犯的諮商師等等）能夠看出這個問題與他們工作之間的關聯性。然而，在回答此種基礎問題時，卻能夠刺激新的思考方式。而所得到的答案，則可能會在實務工作者的工作方面，帶來革命性的變化或者是戲劇性的進展。在未能理解各種事件或是行爲背後重要成因的情況下，公共政策以及社會服務有可能會變得效能不彰或者是被誤導。同樣地，應用研究也會建立新的知識。然而，基礎研究對於知識的擴展是必要而不可或缺的。大多數的基礎研究，都是由處於科學社群中心附近的研究者所進行的。

應用研究

當我們在進行**應用研究**（applied research）時，我們會試圖去處理一個特定的、明確的議題。我們或許會提出一些解決方案， 以回應一個由僱主、當地社群、或者是一項社會運動所提出的問題。[2] 和基礎研究不同的是，我們在進行應用研究時，很少會去建立、測試、或是產生與理論之間的關聯性。大多數的應用研究都是短期的、小規模的研究，而且它們所提供的，是可以在一年之內直接派上用場的實際結果。舉例而言，假設有某一所大學的學生自治會想要減少酒精濫用的情況。因此，該自治會想要知道，如果在明年舉辦一個不提供酒精飲料的宴會，該校學生被逮到酒醉駕車的次數是否會因此而下降。在此種情況下，進行應用研究將會是最合適的方式了。

企業、政府機關、健康照護機構、社會服務機構、政治組織、以及教育機構都會進行應用研究，並且依據研究發現來作出決策。應用研究的發現影響了許多的決策。舉例而言，這些研究發現可能會促使有關當局作出進行一項新計畫的決策，以減少社會福利接受者在獲得救濟之前的等待時間。應用研究的發現也可能幫助警方決定是否要採用一項新的治安措施，以減少家暴的問題。應用研究的發現或許也可以幫助一間公司決定把產品行銷的對象鎖定爲成熟的成人，而不是青少年族群。

目前正在任職中的實務工作者（例如，教師、醫生與護士、業務人員、諮商者以及個案工作者、法官、經理人、管理者、以及市政行政官員等等）都是應用研究發現的主要讀者群。在這群廣大而且多元的讀者群當中，有許多人並不具備研究方面的背景、或是強烈的科學觀點，而這樣的情況可能會造成許多的混淆或者障礙。舉個例子來說，假定在一個訴訟過程中，法院取得了一個

應用研究 應用研究的設計目的，是為了針對某個具體的問題提出具有實用性的解決方法，或是處理臨床工作者或實務工作者們所擁有的立即、特殊的需求。

類似調查研究所得到的結果。然而，這些非科學家們（法官、陪審員、律師等）卻是以非科學的方式來評估這份調查研究的方法論以及研究發現。[3] 如此一來，他們就有可能會以錯誤的方式來解釋研究結果，並且使用一些與科學社群不同的評估標準。他們可能會接受來自於一項未符合基本科學研究的標準所得到的發現，但是卻拒絕接受來自於一項符合最高科學研究標準的研究所得到的發現。應用研究者必須把科學與技術方面的發現，轉化成一般的決策者所能瞭解的語言。這些應用研究者必須強調一項研究的設計或是發現裡的優點或是限制。

研究者可能會爲某位決策者進行一項應用研究。對於這位決策者而言，他對於該項研究的進行細節完全不感興趣，他所想要得到的，只是該項研究的重大發現的簡短摘要而已。然而，這位研究者仍然應該要準備一份完整的、詳細的研究報告。其他一些有時間與能力評估該項研究品質的人，可能會對該項研究感到興趣，或者，之後也可能會產生一些爭議。與應用研究相關的一項限制，就是應用研究不太可能會出現在那些經過同儕審查的刊物裡。在許多時候，應用研究的發現只會在有限的範圍內散佈流通，而且也只有某個組織裡的少數決策者、或者是實務工作者，才能取得這些研究發現。

由於我們會將應用研究付諸實際行動，因此，應用研究就有可能會產生爭議。這種情況已經是屢見不鮮了。例如，在 1903 年時，艾爾伍德（Ellwood）曾經進行了一項關於監獄和救濟院的應用研究，並且提出了文獻以證明這些機構具有嚴重的缺陷。結果，他的研究報告引起了社會大眾對於這些機構的不滿與憤怒。然而，與此同時，他卻被指控爲在誹謗雇用他進行研究的州政府。[4] 懷特（William Whyte, 1984）在分別進行關於奧克拉荷馬州的一間工廠以及芝加哥的數家餐廳的應用研究時，則是遇到了研究結果相互衝突的狀況。在關於奧克拉荷馬州那間工廠的應用研究中，管理部門比較感興趣的是如何去擊敗工會，而不是去學習與瞭解勞資關係。而在關於芝加哥的餐廳所進行的應用研究裡，餐廳經營者則是想要使餐飲業在表面上看起來運作良好，但是卻不願意讓任何人得知其日常經營運作的實際細節。一些企業組織所擁有的心態，與那些以研究爲導向的研究者，所擁有的心態並不相同。學習在不同的心態間進行協調與溝通，是一項必須加以培養的重要技能（Reingold, 1999）。另外一個相關的議題，則是有時候政府官員在遇到政策方面的爭議時，會刻意呼籲進行應用研究，以當作一種緩兵之計。他們這麼做的目的，只是想要轉移批評的焦點、或者是要延遲作出某項決策，直到政治方面的激烈爭議逐漸平息爲止。事實上，他們對於該項應用研究或者是研究發現，根本就沒有任何興趣可言。

應用以及基礎研究在研究方法論的取向方面並不相同（請參閱表 2.1）。在應用研究裡，研究者必須在科學嚴謹性方面作出更多的權衡與妥協，以獲得快速而可用的研究結果。儘管如此，一些不得已的妥協作法，並不能當作研究品質不佳的藉口。應用研究者必須學習如何在一個應用環境的限制中進行研究，並且要在科學的嚴謹性與實際需要之間取得平衡。要取得此種平衡，研究者不但要對研究有深入的知識與瞭解，而且也要能夠清楚地意識到在科學標準上作出妥協所可能造成的後果。

表 2.1 基礎研究與應用研究之比較

比較層面	基礎研究	應用研究
主要讀者群	科學社群（其他研究者）	實務工作者、參與者、或是管理者（非研究者）
評估者	研究同儕	實務工作者、管理者
研究者的自主性	高	低到中
研究的嚴謹程度	非常高	依情況而有所不同，中等
最高優先事項	經過驗證的事實	相關性
目的	創造新的知識	解決實際的問題
成功與否的指標	研究獲得出版，以及對於知識／科學家造成影響	直接應用於處理一個具體的問題或者是重大的議題

應用研究的三種類型 應用社會研究有將近一打以上的不同形式，而在這兒，你將會學到其中三種主要的類型：評估研究、行動研究、以及社會影響評估。

1. **評估研究**（Evaluation research）是最被廣泛使用的應用研究類型。[5] 大型的官僚組織（例如像是企業、學校、醫院、政府、大型非營利機構等等）通常會使用評估研究，以得知一個計畫、新的做事方法、行銷活動、或是政策等是否有效。換句話說，其所針對的問題就是「它有效嗎？」。甚至還有一份期刊是專門促進評估研究這個領域發展的，而該份期刊的名稱是《評估研究評論》（*Evaluation Review*）。

在 1960 年代時，由於美國聯邦政府不斷推出許多新的社會計畫，因而使得評估研究在美國迅速地擴展。大多數的研究者所採取的是實證主義取向（a positivist approach）（請參閱第四章），並且使用了成本效益分析（cost-benefit analysis）（我們將在本章稍後部分進一步討論這個問題）。而到了 1970 年代時，大多數的政府社會計畫都會要求進行評估研究，以決定這些計畫的效果如何。

評估研究的問題可能會包含下列的這些例子：法定強制拘留方案是否能夠降低婚姻暴力？反強暴計畫是否能夠減少大學男性對於女性的強制性交行爲？彈性工時計畫是否能夠增加員工的生產力？在評估研究中，我們會去衡量一項計畫、政策、或是辦事方式的有效性。此外，在評估研究中，我們可以使用數種不同的技術（例如，調查以及民族誌田野研究），但是，如果能夠直接進行實驗的話，研究結果將最具有說服力。

評估研究 應用研究的類型之一，在此種研究中，研究者試圖決定一項計畫或是政策的運作程度是否良好、或達成其既定目標的程度。

有一些實務工作者也會進行他們自己的評估研究。然而，在大部分的情況中，外部的經理人或是決策者會要求進行評估研究。外部人士有時候會對一項研究的內容加以設限，而且也可能會明確指出他們所真正感興趣的特定結果。例如，教育官員可能會要求進行一項研究，以評估國小二年級到五年

級學生在數學技能方面的進步情況。但是，他們卻會告訴研究者直接忽略其他科目、其他的學習面向、以及兒童在認知—社會發展方面的變化。

研究倫理以及政治考量的衝突，經常出現在評估研究當中。之所以會發生此種情況，是因爲人們與特定的研究發現之間有著強烈的利益關係。這些發現可能會影響到誰會被雇用、誰可以在政治方面建立民眾的支持度、或是哪一個計畫能夠獲得實施。如果有人對於研究發現感到不高興的話，他們或許會批評該研究的研究者，或是認爲這項研究不夠嚴謹、存有偏見、或是進行的方式不恰當。有些評估研究的研究者曾經經歷過要求他們操縱研究的壓力，尤其是當他們所進行的評估研究與某些極具爭議性的議題或是計畫有關時。由於存在著產生爭議的可能性，因此，應用研究者應該特別重視誠實與公開，並且要小心翼翼地堅守適切的研究程序。

雖然評估研究所具備的價值不容否認，但是評估研究仍然有其限制。評估研究很少經過嚴謹的同儕審查過程，而且其原始資料也很少公開讓他人可以取得、加以詳細檢查或者是複製。除此之外，政策制訂者也可以選擇性地使用、或者是乾脆完全忽略評估研究的報告（請參閱範例 2.1 評估研究）。許多評估研究採取非常狹猛的焦點，也就是只檢視某些被特地挑選出來的輸入或輸出結果，而不是去檢視一個計畫的整體過程，或是其整體影響性。例如，在 1996 年時，美國的社會福利計畫曾經被大幅度地改變、或是「改革」。針對新的社會福利計畫所進行的評估研究，卻只把焦點放在這些新計畫是否降低了福利計畫的案件數負擔量、以及執行新計畫的成本上面。幾乎沒有任何評估研究考量到一些未得到滿足的家庭責任、或者是兒童們逐漸升高的痛苦與不幸。爲了要幫這些新計畫進行辯護，政策制訂者以及許多政客會選擇性地利用這些證據，並且大量吹噓其正面的效益。[6]

範例　2.1

評估研究

衛松、安尼斯基維茲、以及萊特（Wysong, Aniskiewicz, and Wright, 1994）等人評估了在美國以及 42 個其他國家內的 10,000 所學校實施反毒教育（Drug Abuse Resistance Education）計畫的成效。這項計畫被廣泛地使用，擁有充分的資金援助，而且也廣受警政部門、學校行政人員、家長團體、以及許多其他人士的歡迎。透過警政人員與低年級學生的談話，反毒教育計畫試圖利用增加青少年對於毒品的知識、發展拒絕毒品的技巧、以及提高自我評價的方式，減低青少年使用毒品的情況以及可能性。這三位作者檢視了印第安那州一所高中裡的兩組高年級學生團體的情況。其中一組學生在七年級時已經參與過反毒教育計畫，而另一組的學生則從未參加過反毒教育計畫。與許多過去的研究結果一致，作者們發現到，不論是在第一次吸毒的年紀、使用毒品的頻率、或是自我評價等各個方面，在不同組別的學生之間都沒有持續性的差異。作者們認為，反毒教育計畫之所以會廣受歡迎，或許是因為該計畫在政治上所具有的象徵性意義。這個計畫在潛在的目標方面或許有其成效（例如，幫助政客、學校行政人員以及其他人在道德方面感覺良好，並且讓他們覺得自己參與了反毒的行動），然而，在正式目標方面，卻毫無成效可言（例如，降低青少年非法使用毒品的情況）。

行動研究　一種應用研究，其主要目的在於促進社會變革，或是實現一個以價值為導向的政治—社會目標。

評估研究有兩種類型，分別是形成性評估研究、以及總結性評估研究。*形成性評估研究*（*formative evaluation*）是針對一個計畫進行監控或是持續性的回饋，以達到計畫管理的目的。而*總結性評估研究*（*summative evaluation*）則是去檢閱、回顧一項計畫的最終結果。通常這兩種評估研究都是重要而且必須的。

許多組織（例如，學校、政府機構、企業等等）已經把評估研究當成其持續運作的一部分。其中的一個例子，就是由美國國防部於 1960 年代所率先採用的*設計、計畫、預算制度*（*Planning, Programming, and Budgeting System, PPBS*）。PPBS 所依據的概念，是研究者能夠利用衡量某項計畫，達成既定目標的情形，以對該項計畫進行評估。評估者把一項計畫分割爲數個構成要素，並分析每個構成要素的成本（工作人員、供應量等等）、與達成計畫目標的情況。舉例而言，有一所婦女健康中心提供了一項孕婦教育計畫，而該項計畫共有四項構成要素，包括擴大服務範圍、教育、諮商、以及轉介服務。除此之外，該項計畫還有四個主要目標：向外接觸孕婦並且提供情感方面的支持、提供關於懷孕的最新資訊、提供孕婦在懷孕期間關於健康風險以及關心事項的諮詢服務、以及轉介孕婦到健康照護部門或是家庭計畫機構。評估研究或許會檢視每一個構成要素的成本，並且衡量該項計畫達成四個主要目標的程度。例如，研究者可能會提出下列的問題：（1）在過去一年裡，有多少工時以及多少支出是用於擴大服務範圍的活動上面？（2）有多少詢問電話及諮詢服務是來自於上述的活動與努力？以及（3）在目標族群當中，有多少位女性曾接觸過、或是親自來到婦女健康中心，尋求諮商服務？[7]

2. **行動研究**（Action research）把知識視爲權力的一種形式，並且會整合新知識的獲得與對該項知識的利用，以達成某一個特定而明確的目的。在行動研究中，我們不會保持一種超然、客觀的立場，而是會去縮減存在於研究某項議題，以及從事社會—政治行動之間的落差，以對該項議題發揮影響力。各種不同類型的行動研究，其靈感來源

> 是不同的哲學立場，主要是由關於認識論與本體論的不同核心假設所驅使的，而此種情況會使各種不同立場的實務工作者，在目標以及要求上產生不同的規範與作法。然而，此種由哲學立場的差異性所造成的影響，在公開的陳述以及說明中通常不會被特別注意到，也因此造成了許多模稜兩可以及爭論不休的狀況……（*Cassell and Johnson 2006: 785-786*）

在第四章當中，我們將會再進一步探討此種哲學方面的差異性。大部分的行動研究都具有下列的五種特質：

- 在行動研究的過程中，被研究的對象都是主動而積極的參與者。
- 行動研究包含了一般大眾所熟悉的知識以及他們所關心的事物。
- 行動研究檢視了權力關係，並且記錄下社會不平等與不公正的情形。

- 行動研究的發現必須與大眾分享，以提高大眾的相關意識、並賦予大眾權力。
- 行動研究直接和社會—政治行動、以及達成社會目標緊密連結在一起。

行動研究試圖使研究參與者和研究者之間的權力關係達到平等，並且試圖避免讓研究者擁有比研究對象更大的控制權力、地位以及權威。換句話說，行動研究鼓勵平等，並且也鼓勵研究參與者直接參與整個研究過程。行動研究想要提高參與者以及社會大眾的意識，因此，出版研究文章僅僅是次要的目標而已。相反地，行動研究真正要強調的重點，則是和研究參與者以及社會大眾一起分享研究的發現，因此，行動研究在發表研究發現時所採取的形式，通常是一般性的報告或者是小冊子、對大眾媒體發表新聞稿、或是舉辦公開的會議等等。

行動研究通常會吸引那些對於某項議題（例如像是環保議題、平等議題、女性主義議題等等）具有強烈熱情的研究者。一位充滿熱誠的女性主義行動研究者，可能會將某項研究視爲同時能夠促進知識、而且也能創造出改革性別關係的社會變化。[8]如果這名女性主義行動研究者所研究的是關於性騷擾方面的議題，則其研究結果或許會是呼籲作出政策方面的變革，以減少性騷擾的發生頻率，並且與潛在的受害者密切合作，好讓她們更有能力去捍衛自己的權利。當某個小城鎮可能因爲附近要興建水壩而即將被拆遷時，行動研究者也有可能會去努力想辦法保存這個小城鎮。行動研究者也有可能會與工會幹部以及管理階層合作，一同重新設計工作流程，以避免大規模的裁員。而在發展中國家裡，行動研究者則通常會努力讓那些不識字且貧困的農民們，能夠具備基本的讀寫能力、喚醒大眾對某些問題的意識、並致力於改善當地居民的生活環境。[9]

參與式行動研究（participatory action research）是行動研究的一個分支，此種研究所強調的是，把知識創造過程加以民主化、揭露不公正的狀況、突顯社會中的不平等以及衝突、以及致力於從事集體行動以改善現狀。參與式行動研究的一個關鍵信念，就是知識乃來自於直接體驗社會—政治方面的行動主義。隨著研究參與者投入直接的行動，他們將會獲得更多訊息、得到更多權力。他們可以由此多加學習知識，也比較容易成功。

在參與式行動研究中，研究參與者在規劃研究、設計研究、以及實施研究方面採取了積極主動的角色。他們在一個合作式的過程中，與專業的研究者共同獲得研究發現。研究參與者直接參與了定義問題以及實施研究的過程。由於大多數的研究參與者並不熟悉專業的社會研究，此時，受過專業訓練的研究者就可以扮演一個諮詢者，或是共同合作者的角色，以提供必要的協助，並在研究設計、資料蒐集、以及資料分析與詮釋等各方面，提供專業知識。

一般而言，行動研究者必須在下列兩件事物之間取得平衡：其中之一是專業標準；另外一個則是爲了適應當地情境、以及參與者所具有的特定關心

參與式行動研究　行動研究的一種類型，在此種研究中，研究參與者會積極主動地幫助設計以及進行研究。此種研究強調把知識創造過程加以民主化、以及參與集體行動，而且此種研究假定，政治知識是由參與研究過程所產生的。

事物，而產生的實際限制。讓當地的參與者直接涉入與控制研究過程，意味著研究參與者能夠和研究者共同分享研究發現的所有權。如果某位研究者想要在專業期刊上發表研究結果的話，可能會讓參與者覺得，這名研究者只是想要提升自己的學術地位而已。正因爲如此，所以當研究者想要在專業團體或管道中發表研究結果之前，事先取得參與者的同意與合作，便是至關重要的事。[10]

有一些組織或是人們會因爲擁有特定價值觀或是擁護特定論點，因而反對維護研究參與者的權益。這些組織或是人們有可能會對廣爲人知的、成功的行動研究提出質疑與挑戰。爲了因應此種情況，行動研究者必須對適當的研究程序擁有深入的知識，同時也要小心翼翼地仔細以文件記載研究方法（請參閱範例 2.2 行動研究）。

3. **社會影響評估**（social impact assessment）研究主要是在進行一項計畫性的改變之前，先去評估可能會造成的社會影響與結果。[11] 一般而言，社會影響評估（SIA）研究可能會是政府部門所要求的大型環境影響報告的一部分。在美國，1969 年所頒布的《國家環境政策法》（*National Environmental Policy Act, NEPA*），就要求任何聯邦政府機構在採取「可能會對人類環境的品質造成深遠影響的行動」之前，都要先提出「環境影響報告書」（Environmental Impact Statement, EIS）（NEPA, section 102）。要提出環境影響報告書，就必須進行社會影響評估研究，而這就會牽涉到社會科學研究，除此之外，社會影響評估研究也會同時評估正面以及負面的影響。

在尋找合適的位置以興建學校、醫院、監獄、住宅開發計畫、購物中心、工廠、垃圾掩埋場、公路、機場、水庫、公園、休閒娛樂區、以及發電廠時，都必須提出環境影響報告書。如果環境影響報告書裡面包含了社會影響評估，則該份評估研究會去評估各種不同行動的可能影響或者是後果，而這些行動包括了住宅的可得性與品質、人口結構與特質（像是年齡結構、種族多樣性、收入以及教育水準等等），以及權力與權威的分配情況。社會影響評估可能會去檢視人們的態度或者是看法、家庭之間的凝聚力、以及人們的交際網絡。在環境影響報告書中所包含的社會影響評估，或許會考量對於社區資源所可能造成的各種衝擊或影響，而這些社區資源包括了健康服務、警政服務、火警服務與衛生服務、就業機會、接受教育以及休閒娛樂的機會、以及非營利組織的活躍程度等等。而對於那些居民已經建立當地歷史與文化根基的特色社區，社會影響評估也會去考量這些社區的存活與延續，所可能造成的影響或是衝擊。[12]

進行社會影響評估的研究者，通常會在一個跨領域的研究小組內工作，以衡量不同領域內的影響狀況（請參閱範例 2.3 社會影響研究）。社會影響研究者的專業組織是「國際影響評估學會」（International Association for Impact Assessment），而該學會所發行的學術期刊爲《影響評估與方案評價》（*Impact Assessment and Project Appraisal*）。

社會影響評估 一種應用研究的類別，記錄了如果一項重大的變化被引進一個社區時，在社會生活的各種不同領域內所可能產生的影響與結果。

在數十年的發展之後，社會影響評估研究的工具和效果已經相當完備，然而，此種應用研究的類

範例　2.2

行動研究

威廉斯以及其同事（Williams and associates, 2007）使用了參與式行動研究來研究加拿大薩斯卡通（Saskatoon）地區的生活品質。他們分別在 2001 年以及 2004 年從該城市的三個地區（低所得、中等所得、以及高所得）蒐集了量化以及質化資料。他們把焦點集中於三個主題之上：（1）逐漸加大的所得差距；（2）包含了低社經地位人口的社會知識轉化策略（social knowledge translation strategies）；以及（3）如何使當地的生活品質產生正向的改變。研究者們發展了一個「混合式」（"hybrid"）的研究組織，該組織同時在大學和當地社區設有根據地，而且也有共同領導者（其中一位領導者來自於大學，另一位則來自於當地社區）。在整個研究過程中，那些關心薩斯卡通地區生活品質議題的社區領導者們，都表現得相當活躍與積極，而且他們也併用了四種知識轉化策略：經常關注當地的媒體（報紙以及電視）、舉辦數場社區論壇、與數個社區組織共同創立了「薩斯卡通生活品質指導委員會」（Saskatoon Quality of Life Steering Committee）、並且雇用了一位可以擔任政策創業家（也就是積極支持執行新政策的人）的行動研究者。社區成員也參與了研究設計、資料蒐集、以及資料分析與詮釋的工作。作者們把研究發現視為社區所擁有的學習工具，可以提高社區成員的意識並且激發行動。此外，作者們利用數種方式來傳播研究結果：發表簡短的研究報告、張貼海報、並且在社區論壇中分發研究摘要以供討論之用。與其說討論研究發現是一個終點，倒不如說這只是整體過程中的一個階段，其目標在於根據社區對於這些發現的反應，進而訂立新的政策、計畫、或者是行動。

另外一個行動研究則是由奎區與其同事（Quach and associates, 2008）所進行的一項應用研究，研究對象是在加州某個郡內的指甲彩繪沙龍中的越南籍員工，而研究目的則在於「蒐集初步的描述性資訊」（p.340）。作者們注意到，在加州地區共有 35,000 家指甲彩繪沙龍，其中總共有 300,000 名指甲彩繪師，他們每天的工作時間很長，而且也一直在使用那些包含有毒與危險成分的美甲產品。在加州地區，約有百分之 59 到 80 的合格指甲彩繪師是越南籍，而且其中有百分之 95 都是女性。在 1987 年到 2002 年之間，越南籍指甲彩繪師的比例增加了十倍之多，但是關於她們的健康情況，外界卻一無所知。研究者們設計本項研究，以提高這些越南籍指甲彩繪師對於健康議題的意識，並且也藉由成立「社區諮詢委員會」（Community Advisory Committee）來監督本項研究的方式，鼓勵這些指甲彩繪師參與本項研究。本項研究的一個重要特色，是研究對象都是外來的移民，而且許多人只具備有限的英語能力（其中有百分之 99 的人是在越南出生，有超過一半以上的人已經在美國居住超過十年以上）。社區諮詢委員會是由一所當地健康中心的外派人員所領導的，其成員則包括了十個越南社區裡的許多成員（其中包括了幾名在指甲彩繪沙龍工作的員工）、健康中心裡的病患、美容指導員、曾經罹患乳癌的倖存者、以及來自於阿拉米達郡（位於美國舊金山灣區）的心理健康顧問專家。該項研究總共研究了在該郡的 74 家指甲彩繪沙龍中工作的 201 名員工。研究者使用了一份填答時間大約十分鐘左右的越南文問卷、焦點團體、以及對沙龍環境的觀察（例如，門窗以及通風口的數量等等）。該項研究詳細記錄了為數眾多的健康議題。超過一半以上的指甲彩繪沙龍員工回報說，自從她們開始在這個行業裡工作之後，就遇上了嚴重的健康問題（例如，像是眼睛紅腫發炎、頭痛、呼吸困難等等）。絕大多數的指甲彩繪沙龍的員工，都極為關心她們過度接觸工作場所裡的化學物質的問題，然而，只有為數不到一半的沙龍有裝設排放廢氣的通風口，以減少員工接觸到有毒化學物質的程度。當地社區的成員也相當積極主動地參與了研究過程中的數個階段。該項研究的作者們利用研究發現來教育當地社區中的成員，並且也發展出一些策略，以幫助減少他們在工作環境中接觸有害化學物質的機會。

範例 2.3

社會影響研究

在過去的三十年之間，許多不同形式的賭博活動、或者稱為「博奕」活動，已經大幅地擴張到美國的許多地區。在 1980 年時，僅有少數幾個州把賭博活動合法化，而且每年所產生的收益也少於一百億美金。時至今日，賭博活動在高達 48 個州裡都是合法的，而且每年所產生的收益也超過五百億美金。立法者希望在不提高稅收的情況下，尋找新的收入來源，並且也想要同時促進經濟發展。而博奕產業則允諾可以帶來新的工作機會、重新為經濟帶來活力、並且可以讓政府獲得一部分從賭博活動所收取的現金。換句話說，博奕產業讓立法者可以在不必提高稅收的情況下，創造工作機會、活絡當地經濟、並且獲得更多收入。法律方面的裁決已經承認了美國印第安族群的條約權利（treaty rights），也就是說，美國的博奕法條並不適用於印第安族群的土地上。當有人提議在美國紐約州羅契斯特市（Rochester, New York）的市中心區興建一間新的賭場時，肯特（Kent, 2004）被委託去進行一項相關的社會影響研究。就像大多數的社會影響研究報告一樣，該項研究的報告也未曾發表於學術期刊上。該份報告估計，這間新的賭場將為羅契斯特市增加 1,300 個新的工作機會。從賭場運作所徵得的稅收中，紐約州每年可以從中額外獲得二千三百萬美元的收入，而羅契斯特市則可以獲得大約一千一百萬美元的收入。為了要估計興建新賭場所造成的影響，該項研究的其中一部分是去比較來自於數個設有賭場以及未設有賭場的城市的資料，並且也考量了過往有關於賭博成癮行為的研究資料。這個部分的研究同時考量了在設有賭場的城市裡所出現的經濟利益，以及額外增加的社會成本（例如，犯罪率、賣淫行為、毒品濫用、強迫性賭徒等等）。該項研究指出，病態的賭博行為會隨著賭場的興建而增加，也會造成個人以及家庭的損失（會造成離婚率上升以及兒童虐待案件的增加）。該研究報告估計，每年因此造成的社會成本，可能高達一千萬美元。

型仍然未被充分地加以利用。此種情況乃是肇因於數個因素。首先，大多數的環境影響報告書並未要求進行社會影響評估。立法者、政策官員、或者是決策制訂者在批准或同意重大計畫之前，很少會要求先進行社會影響評估。除了非常少數的大型計畫之外，大多數的決策制訂者在改變區域劃分的規定、開發新的商業園區、訂定住宅發展計畫、或是變更運輸路線等等時，都不會系統性地考量到社會影響的問題。在決定這些議題時，都是以政治上以及經濟上的利益爲考量重點。其次，進行社會影響評估研究必須花費大量的時間與金錢。官員們對於花費過多經費在進行社會影響評估上面相當反感，而且也反對因此而拖慢了制訂決策的進度。由於官員們必須在相當短的時限內完成工作，因此，即使進行社會影響評估研究可以讓他們作出更明智的決策、並且就長遠來看可以節省金錢與日後的困擾，他們也仍然不願意進行此種研究。第三，在許多地方，整體的政治—文化氛圍對於制訂計畫這件事是持保留態度的，而且對於所謂「專家的」建議也抱持不信任的心態。此種不信任的心態通常會伴隨著對於社會科學研究知識不足的情況。結果，人們就傾向於死守著傳統的決策制訂方法不放。對於其決策所可能造成的社會影響，他們所憑藉的是自己所作的猜測，而不是以研究爲基礎所得來的知識。第四，新方案的提倡者或者是投資者通常會反對進行社會影響評估研究。這是因爲他們擔心研究發現會造成進度延宕、迫使他們作出所費不貲的變更、或是由於發現了社會大眾所真正關心的議題，而必須大幅更動整個計畫等等。最後，

就社會影響評估研究而言，官員們常常會因爲一些更爲優先的政治考量、以及受到根深蒂固的政治—經濟利益的影響，因而刻意忽視這些研究的結果。

應用研究的兩項工具　許多應用研究者會使用下列兩項工具來當作其研究的一部分：需求評估以及成本—效益分析。

需求評估（needs assessment）牽涉到廣泛地蒐集資料，以決定主要的社會需求，和這些需求的迫切性。在政府機構或是慈善機構決定要採行某項幫助人們的策略、或者是進行更深入的研究之前，通常會先進行需求評估，以當作一個預備性的步驟。然而，在一個關係複雜的社區當中，需求評估經常會變得盤根錯結、紊亂不堪，而且當我們在進行需求評估時，也可能會遭遇到下列數個不同的議題（請參閱重點摘要 2.1 需求評估的兩難困境）。

首先要處理的第一個議題，就是先針對嚴重的需求，或是排定優先處理的問題順序。一個社區或許會有十來個議題與關心的事項，像是遭受嚴重家暴的婦女、吸毒的國小兒童、那些因爲無家可歸而睡在公園裡的遊民、因爲賭馬而賠掉一大筆錢的勞工們、或者是那些在鄉村俱樂部喝太多酒，還違規酒駕的公司主管們等等。哪些議題較爲重要、並可以先進行需求評估？有時候，最顯而易見的需求，往往不見得就是最嚴重的、或者是最能夠引起大眾強烈共鳴的需求。

第二個議題，則是要確認進行需求評估時的資訊來源。舉例而言，如果有一項援助無家可歸者的計畫，而且我們決定要針對這項計畫進行需求評估時，誰最有資格提供相關的資訊呢？有一些商店業者經常抱怨那些露宿在街道上的、無家可歸的遊民，我們應不應該和這些業者討論這些遊民的需求爲何？我們應不應該去徵詢目前正在爲無家可歸的遊民提供服務的人們或機構（例如，像是社工人員、健康照顧中心、學校、爲遊民所設置的暫時安置中心、免費爲遊民提供食物的場所等等）的意見？我們是否應該去仰賴執法人員（例如，像是警方、監獄看守人員、法庭裡的官員等等）爲我們提供資訊？我們是否應該去詢問遊民們的朋友、家庭成員、以及爲他們奔走的非專業人士的意見？我們應不應該直接去詢問這些無家可歸的遊民呢？理想上，我們應該要把所有的資訊來源都含括進來，但是，要確認所有的資訊來源，不但是一件吃力不討好的工作，也可能會耗費過多的時間。

第三個議題，則是那些明顯、立即的需求，或許並未涵蓋所有其他較爲不明顯的需求、或是把這些較不明顯的需求，和長期的解決方案連結在一起。舉例而言，我們知道那些無家可歸的遊民說，他們需要可以讓他們住的房子。然而，在檢視了整個情況之後，我們認爲如果這些人有工作的話，那麼他們就可以負擔得起或是租得起房子。因此，居住的問題是由工作的需求所引起的，而工作的需求則可能是由技能的需求、得到「能夠維持生活的工資」（"living wage"）的需求、以及某些種類的產業需求所引起的。因此，要解決住屋的需求，就必須吸引特定的產業、執行新的最低工資的規定、並且提供職業訓練的機會。通常，表面上顯而易見的需求，都是根

需求評估　一種應用研究的工具，用來蒐集關於一項需求、議題、或者是關心事項的描述性資訊，其中包括了重要性、範圍、以及迫切性等等。

重點摘要 2.1

需求評估的兩難困境

1. 誰能夠定義什麼是最嚴重的議題、並且應該對該議題的需求加以評估？
2. 在瞭解某一個團體的需求時，你應該要去詢問誰的意見？
3. 你是否應該同時考量那些清楚、可見的需求，以及那些未曾言明的、隱而未顯的需求？
4. 當有許多不同領域的需求同時存在時，你應該要把哪些需求含括在評估範圍內？
5. 那些長久以來掌握權力的人會設下種種的限制，你是否應該把解決某項需求的方案，侷限於可以在上述的限制中具體實施的方案？或者，你是否應該要考量所有的可能性，即使這些可能性或許會對現狀造成破壞，也在所不惜？

植於更深層的問題與起因，而許多人對於這些更深層的問題與起因渾然不覺。例如，飲用了受到污染的水、飲食不均衡、及缺乏運動等等，都可能造成健康照護的需求大增。此種情況是否意味著，我們必須興建更多的健康照護中心？抑或是，我們真正需要的，其實是更好的水質處理方案與公共健康教育計畫？

第四個議題，則是需求評估有可能會引起政治方面的爭議。這是因爲需求評估所建議的解決方案，可能會超出當地的控制能力範圍，或者是根本不可能有實施的機會。有些握有權力的團體，可能不想讓某些社會需求被記錄下來，或是公諸於世。我們或許會得知某個城市裡有許多隱匿未報的犯罪活動，然而，如果我們把此種情形公開的話，可能會損及這個城市安全而且運作良好的形象。而此種良好的形象，可能正是當地的商會以及市政府所極力推銷的形象。一般而言，某個團體（例如像是那些花費太多金錢在賽馬場裡賭博的人）的需求，會和那些因爲創造出此種需求，而有所獲得的利益者（例如像是賽馬場的主人以及員工）的行動有所關聯。當我們在詳細記錄這些需求、並提出解決方案時，我們有可能會被困在兩個利益互相衝突的團體間，進退兩難。

應用研究的第二個工具，是由經濟學家所發展出來的**成本效益分析**（cost-benefit analysis）。成本效益分析牽涉到估計某個提議方案在未來可能產生的成本，和帶來的效益，並針對這些成本及效益，賦予貨幣上的價值。在進行成本效益分析時，我們首先必須確認所有可能產生的影響與結果，其中包括了具體而明確的結果，像是創造工作機會、形成新的產業、或者是畢業率等等；以及無形的結果，像是乾淨的空氣、政治自由、優美的風景、或是某個計畫或方案的壓力程度較低等等。接下來，我們會爲每項可能的結果，賦予貨幣上的價值，有些（例如成本）可能是負值，而有些（例如效益）則可能是正值，而有些則可能是中性的。我們接著計算每個結果發生的可能性或機率，而在最後，我們必須把所有的成本以及效益進行比較，並且決定它們是否達到平衡，抑或是互相抵消。

成本效益分析 一種由經濟學家所發展出來的應用研究工具，在進行此種分析時，某個過程的輸入與結果都會被賦予貨幣上的價值，然後，研究者接著會去檢視輸入與結果這兩者之間的平衡狀態。

雖然從表面上看來，成本效益分析似乎是一種非政治性的、理性的、以及在技術上相當嚴謹的決

策制訂策略。然而，成本效益分析仍然經常引起爭議。就如同需求評估一樣，人們對於哪些活動應該被認爲是相關的、或者是重要的，經常會有不同的看法或意見。因此，有些人可能會認爲，最重要的關心事項，應該是商業活動的穩定性與獲利性、更低的稅率、以及創造新的工作機會等等。然而，其他人則可能會認爲，最應優先處理的事項，應該是健康而乾淨的環境、開放的綠地、以及更多的藝術活動和言論自由。除此之外，對於某項可能的結果究竟應該被賦予正值或是負值，人們的意見也相當分歧。例如，對我而言，我認爲把道路拓寬是一件好事，因爲道路拓寬之後，我就可以更快速地開車到公司去上班，而且也有助於減少塞車的情況。然而，對於住在路旁的屋主而言，拓寬道路卻被視爲是一項成本，因爲在拓寬馬路時，勢必得佔去他或她前院的一部分，噪音與污染也會增加，而且房子的市場價值也會下降。在針對紐約州羅契斯特市設立新賭場（請參閱範例 2.3）所進行的社會影響評估研究中，該份研究報告權衡了經濟利益（收益、工作、稅收）與社會成本（犯罪、賭博成癮、家庭破碎、濫用毒品）之間的關係。在該份報告的陳述中，興建新賭場所帶來的經濟利益，超過了所需付出的社會成本。然而，真正的問題在於，得到經濟利益的人（也就是當地的企業老闆以及納稅人、還有那些在賭場內找到工作的人），與因爲興建賭場而付出社會成本的人（也就是那些由於強迫性賭博行爲，而造成家庭破碎的人、那些因爲濫用毒品而使得健康狀況惡化的人、與那些成爲娼妓的女性），根本就不是同一群人。

當我們在決定成本以及效益的貨幣價值時，可以使用下列的兩種方法。*權變評估*（*contingency evaluation*）是直接去詢問人們，對他們而言，某項事物的價值是多少。舉例而言，假設有一個小鎮，正在考慮是否要允許一座會造成污染的工廠在該鎮設廠。因此，我們想要估計空氣污染對於一般人的健康所造成的成本爲何。我們或許會問人們下列這個問題：「如果你雖然沒有嚴重咳嗽的情形，但是卻因爲氣喘的問題，每年必須有十天左右的時間無法工作。那麼，你認爲此種情形所造成的成本價值是多少呢？」如果在一個人口約二萬人的小鎮裡，人們對此種情形所賦予的平均價值爲 150 美金，我們就可以估計，健康狀況的權變評估或是主觀利益爲每年 150 美金乘上 20,000 人，也就是大約 300 萬美金左右。接著，我們會把這項成本拿來和在允許污染的情況下，公司因此所獲得的較高利潤、以及所創造的工作機會兩相權衡比較一番。此種估計方法的問題，就在於人們很少會給予精確的數字。除此之外，不同的人所給予的價值，通常會差了十萬八千里。對於一個窮困的人來說，咳嗽以及無法上班的價值或許是 150 美金。但是對於一個富有的人來說，同樣的情況卻可能高達 15 萬美金以上。除此之外，也存在著一些影響範圍較廣的結果。在這個例子裡，造成污染的公司將會遷移到那些居民收入較低、而且所給予的價值也較低的城鎮裡。如此一來，就會使低收入地區的生活環境更加惡化，並讓貧富之間的生活品質差距更形擴大。

利用同樣的例子，*實際成本評估*（*actual cost evaluation*）則是評估真正的醫療與工作損失所造成的成本。我們或許會去評估對於健康所可能造成的影響，並且去計算雇主在找人代替生病或失能的員工時，所可能增加的額外醫療費用與成本的總

額。假設每個人的平均醫療費用爲 150 美金，而當有員工因病請假時，雇主每聘請一個代替的臨時工，每天平均要額外支出 300 美金的費用。因此，每年治療 10,000 人的費用就相當於 150 美金乘以 10,000 人，也就是 150 萬美金。而聘請 1,000 名臨時工來工作兩天的費用，則爲 600 美金乘以 1,000 人，也就是 60 萬美金。把上述兩個金額加總之後，即可得知估計總金額約爲 210 萬美金。值得注意的是，此種方法會忽略了病患所承受的痛苦和折磨、不便之處、以及間接的成本（例如，父母親必須至少有一人請假在家陪伴生病的孩子、或者是某個兒童會因爲罹患氣喘而無法參與運動比賽等等）。爲了平衡用此種方法所計算出來的成本和效益，這間造成污染的工廠，將必須能夠賺取 210 萬美金的額外利潤。

成本效益分析所根據的假定，是我們可以指定一個貨幣價值給所有的事物（例如，像是兒童的學習、健康、愛情、快樂、人性的尊嚴、及道德方面的正直等等），而且人們會賦予類似的貨幣價值予同樣事物。然而，我們或許會質疑此種假定的正確性。除此之外，成本效益分析也有可能會引起道德和政治方面的疑慮。因爲那些承擔成本的人，有可能並不是那些得到利益的人。而且，成本效益分析的計算方式傾向於偏袒那些富裕的、收入較高的人們，因而忽視了那些貧困的、收入較低的人們。與一個低收入的人相較之下，一個高收入者的時間價值較高，因此當可以在上班的路程中節省十五分鐘的通勤時間時，高收入者就會指定較高的價值給這十五分鐘的時間。例如，一個高收入者可能會認爲，節省這十五分鐘時間的價值爲 50 元美金，但是對於一個低收入的人而言，同樣的時間，其價值卻可能只有 5 美金而已。在成本效益分析的結論中，經常會認爲讓低收入的人們承受一些不方便之處、或是打亂低收入人們的生活，是較爲符合「成本效益」的作法。

成本效益分析也傾向於隱蔽決策中的道德—政治面向的問題。舉例而言，面對一個病情嚴重、必須靠維生系統才能活下去的老年人，我們是否應該要拔掉機器的插頭？還是應該讓這個老年人繼續多活半年？讓我們先來比較一下不同作法的效益與成本爲何。假設讓這個老年人的生命額外延續半年的醫療支出爲 20 萬美金，那麼，讓這樣一個在社會上已無生產力可言的人多活半年，其效益是否有達到 20 萬美金的價值呢？除了經濟方面的考量之外，成本效益平衡的決策，也有其道德的面向。然而，只有當某些決策牽涉到一個與你有個人、以及情感的連結的某個活生生的人時（例如像是你的祖母之類的），這些決策中的道德面向才會比較清晰可見。如果決策中所牽涉到的人，是某個被認爲不相干的個體（像是在一千個醫院中的某個病患），而且是與決策制訂者沒有那麼直接、或是個人接觸的人時，道德面向的考量就會變得比較模糊，亦不那麼重要了。不過，雖然在此種情況中道德層面的問題被忽略了，但卻始終存在，且無法逃避。

超越基礎研究與應用研究的二分法 把研究劃分爲基礎研究與應用研究的二分法，事實上是過度簡化的作法。以下有三個相關的議題可以讓我們反思此種區別，且可以建立這二種研究類型之外的其他研究類型：

1. 一項研究所創造出來的知識形式

2. 能夠使用研究發現的讀者群範圍
3. 是由一名獨立的研究者或是其他人開始進行、設計、以及控制一項研究

知識形式　社會研究者會產出兩種形式的知識，分別是工具性知識以及反身性知識。雖然這兩種知識形式有重疊的部分，這兩種形式的知識卻反映了下列兩種行爲之間的區別：其中之一是中立的、公正無私的、以及以任務爲導向的行動；而另外一種則是有原則的、以價值爲依據的、並且極度投入的行爲。大部分發表在學術期刊上的研究、以及由實務工作者所進行的應用研究，都是在建立、並且擴充**工具性知識**（instrumental knowledge）。此種知識是一種手段——目的或是任務導向的知識。我們會使用此種知識來完成某件事，像是一項實際的任務、或是促進我們對於世界運作方式的瞭解。當我們延伸舊有的研究技術、或者是發明出新的研究技術，並且蒐集、驗證、連結、以及累積新的資訊時，就會創造出此種知識。工具性知識能夠促進我們的理解範圍。當我們創造出工具性知識時，我們可以避免直接去面對道德方面、或是價值導向方面的疑慮。

相比之下，**反身性知識**（reflexive knowledge）則是一種自我覺察、價值導向的知識。此種知識是有原則的、並且導向一個終極的價值或是目的本身。我們創造出反身性知識的目的，在於建立特定的道德承諾、有意識地反思知識創造的脈絡與過程、並且強調知識的各種不同意涵。當我們在創造出反身性知識時，我們可能會提出像是下列的問題：我們爲什麼、以及如何創造出這項知識？這項知識的關聯性或者重要性爲何？又對誰具有關聯性與重要性？對於其他的知識，或者像是正義、真理、公平、自由、平等之類的道德原則，這項知識所具有的意涵爲何？

研究發現的讀者群　正如同之前已經提過的，基礎研究的主要讀者群，是在科學社群裡的其他專業研究者。至於應用研究的主要讀者群，則是實務工作者以及非專業研究者。我們還可以進一步把從事實務工作的讀者群劃分爲四個類別：一般民眾、行動主義者、一般的實務工作者、以及狹義的實務工作者。每一種類別所感興趣的事物各自不同。多數的*一般民眾*（*the public*）只有廣泛而籠統的興趣，而且是在求學時期或是從大眾媒體的管道得知各種研究結果。*行動主義者*（*activists*）、社區擁護者、與行動研究中的研究參與者，對於研究結果都有直接且立即的興趣，因爲這些研究結果與他們切身的議題之間，具有高度關聯性。*一般的實務工作者*（*generalist practitioner*），像是在政府機構或是大型組織（例如像是企業、醫院、警政部門等等）裡工作的高階決策制訂者，或是政策專家等等，則想要整合實務知識，以利於他們在制訂許多目前或未來的決策時，能夠使用這些知識。相較之下，*狹義的實務工作者*（*narrow practitioner*）則是想得到特定的研究發現，以處理一個明確的、急迫的問題。

工具性知識　一種知識形式，其主要焦點限定於回答一項基礎研究，或是應用研究所提出的問題、議題、或對於某項結果的疑慮，並且具有任務導向的傾向。

反身性知識　一種知識形式，此種知識被用以廣泛地檢視基礎或是應用社會研究的假定、脈絡、以及道德價值立場，其中包括了研究過程本身，以及所創造的知識意涵。

委託研究 當一項研究是由研究者以外的人所提供資金、並且指示研究的進行方式時，就稱為委託研究。對於研究問題、研究方法、以及研究結果的呈現方式等等，進行委託研究的研究者通常沒有太多的掌控權力。

研究者自主性以及委託社會研究 在理想而美好的研究情境中，研究者應該擁有完全的自由，可以在不受任何限制的情形下去追求知識。理想中的研究者應該是獨立自主的、擁有充分的資金、並且對於研究的進行方式，擁有完全的掌控權。與此種理想情境相反的，則是研究受到許多限制，而此種情況正是受雇的研究者——雇員（hired researcher-employee）的寫照，或被稱爲**委託研究**（commissioned research）。大多數的委託研究都會對研究者的自主性設下限制，換句話說，在進行研究時，是由另外的人提供資金、指定研究問題，及研究發現的傳播範圍。其他的「附帶條件」（"strings"）則或許包括，只能檢視某些特定議題之類的限制。除此之外，研究者也可能必須面對嚴格的時間限制，以求完成研究。即使沒有時間上的限制，他們也可能被指定在研究中必須使用哪些研究工具或方法，抑或必須去聯絡、接觸哪些特定人士等等。

基礎研究與應用研究類型的擴充組合 現在，我們可以結合知識形式、讀者群、委託研究與自主研究等類型，以創造出與基礎研究、應用研究、及研究者角色相關的各種組合（請參閱表 2.2）。爲科學社群所進行的基礎研究，能夠產出反身性知識或是工具性知識，而這兩者分別是批判性研究和專業的研究。[13] 一個大型的私人基金會或是政府機構，或許會委託一名研究者去進行基礎研究，這就是基礎契約研究。有時候，研究者會擔任公共知識份子的角色，並且生產反身性知識，以促進一般的討論及公眾論辯。而在其他時候，研究者則會從委託研究或是自主研究中，產出工具性知識。此種工具性知識可能專門獻給一項特定的政策，並且能促成政策辯論。

表 2.2 基礎研究與應用研究類型的擴充組合

	知識形式		
讀者群	**反身性**	**工具性**	**工具性**
	自主的	*委託的*	*自主的*
基礎研究類型			
科學社群	基礎批判研究	基礎契約研究	基礎專業研究
應用研究類型			
一般民眾	公共知識份子	專屬政策	民主政策
參與者	公眾教育者	諮詢者	參與研究者
一般實務工作者	民主的考量	民主的契約	民主的應用研究
狹義實務工作者	專屬考量	專屬契約	專屬應用研究

若是一名研究者爲研究參與者設計了反身性研究，則該名研究者就扮演了公眾教育者的角色。當所生產的知識是工具性知識時，研究者可能會充當研究參與者的諮詢者，或是充當一個與研究參與者地位相等的參與研究者（participatory researcher）。在某些場合裡，一般的與狹義的實務工作者會創造出反身性知識，並將此種知識應用於辯論中，或是關於各種議題跟決策選擇的考量之中。在大部分的情況裡，實務工作者的焦點都是工具性知識。有時候一般的實務工作者會創造出知識，並利用此種知識來促成開放的、民主的決策。而在其他時候，實務工作者則會聚焦於一個特定的議題，該議題可能不具有太多的適用性、研究發現的散佈也受到許多限制。[14] 在此情況中，一個外界的團體或雇主可能會委託研究者進行研究，或者，研究者也可以自主地去進行此種研究，並創造出知識。

研究的目的

我們可能會爲了許多形形色色的理由而進行研究，像是老闆的命令、課堂上的作業、自己本身的好奇心、或是受到身邊朋友的鼓勵等等。進行研究的理由的數目之多，幾乎是可以說有多少個研究者，就有多少種理由。我們可以把研究的目的分成三個組別：探索新的主題、描述社會現象、或是解釋某事發生的原因。[15] 一項研究可能會有多重的目的（例如同時進行探索以及描述），但是通常都會有一個最顯著的目的（請參閱重點摘要 2.2 不同研究類型的目的）。

探索

當我們所要研究的主題非常新穎、我們對其所知甚少或者是一無所知、以及尚未有任何人探討過這個主題時，我們就必須使用**探索性研究**（exploratory research）（請參閱範例 2.4 探索性研究）。進行探索性研究的目的，在於形成較爲精確的問題，以利於我們在未來的研究中， 進一步加以處理或探討。作爲研究問題的第一個階段，我們希望在進行探索性研究之後，能夠得到足夠的資訊，好讓我們能夠設計並執行接下來更具系統性與更加廣泛的研究。探索性研究很少產生明確的答案。探索性研究所處理的是「什麼」的問題，例如像是「社會活動的真正本質是什麼？」之類問題。此外，探索性研究相當難以進行，這是因爲探索性研究幾乎沒有什麼指導方針可以依循、研究中的每項事物似乎都具有潛在的重要性、研究步驟沒有明確的規範與定義、而且研究的方向經常在改變。

進行探索性研究的研究者必須充滿創意、心胸開放、而且有彈性；他們也必須採取一種追根究柢的立場，探索所有的資訊來源。提出具有創意的問題，及善加利用一些意外得到的發現（例如，那些可能具有重大涵義的意外因素，或是機會因素）。舉個例子來說，某位研究者可能原本預期，當人們移

探索性研究　此種研究的主要目的，在於檢視一個人們不甚瞭解的議題或是現象，發展出對此議題或現象的初步概念，以及發展出更為精確聚焦的研究問題。

重點摘要 2.2

不同研究類型的目的

探索性

- 逐漸熟悉基本的事實、環境背景、以及關心的事物
- 在心中產生關於研究情境的大致圖像
- 為未來的研究形成並聚焦研究問題
- 產生新的想法、推測、以及假設
- 確定進行研究的可行性
- 發展出測量以及找出未來資料的技術

描述性

- 提供一個詳盡的、高度精確的圖像
- 找出與過往資料有所衝突矛盾的新資料
- 創造一組類別，或是針對各種不同類別進行分類
- 澄清一系列的步驟或是階段
- 詳細記錄前因後果或是機制
- 描述某個研究情境的背景或是脈絡

解釋性

- 檢驗一項理論的預測或是原理
- 闡述並且擴展一項理論的解釋力
- 將一項理論延伸至新的議題或是主題
- 支持或是拒斥一項解釋或是預測
- 將議題或是主題連結至某個一般性的原則
- 在數個解釋中，確定何者為最佳的解釋

民到一個新的國家時，年紀較小的兒童所受到的負面影響或衝擊，會大於年紀較大的兒童所受到的影響或衝擊。然而，研究可能意外地發現，真正會因爲移民而使其生活產生混亂情況的，其實是某個特定年齡族群（介於六歲到十一歲）的兒童，而高於或是低於這個年齡族群的兒童，所受到的負面影響反而比較小。[16]

描述

關於某個社會現象，你可能擁有一個完整而成熟的想法，並且想要把此種想法描述出來。**描述性研究**（descriptive research）呈現出一個情境、社會環境，或是關係的具體圖像細節。大多數在學術期刊裡所刊出的社會研究、或者是那些被用來制訂政策的社會研究，都是描述性研究（請參閱範例 2.5 描述性研究）。

描述性研究 此種研究的主要目的，是利用文字或數字來「描繪圖像」，且呈現出輪廓、類別的分類方式、或是大致的研究步驟，以回答一些問題，好比像是「誰」、「何時」、「何處」、與「如何」等等。

在實務上，描述性研究與探索性研究有許多重疊與相似之處。描述性研究是由一個定義明確的議

範例　2.4

探索性研究

大多數的探索性研究所使用的都是質化的資料。一般而言，質化研究比較易於接受使用範圍廣泛的證據，以及發現新的議題。特羅辛斯基與布蘭克（Troshynski and Blank, 2008）曾經進行過一項探索性研究，其研究對象是從事非法販賣人口的人口販子。這項研究相當獨特，因為研究中的參與者都曾經從事過非法活動。該研究的作者曾經偶然地遇到某位特殊人士，而這位人士認識「人口販賣圈子內」的一些人。在長達三個月的期間內，作者們透過這位特殊人士而得以遇到五名人口販子，並且進行了多次的開放式訪談。作者們的目標在於探索人口販子看待這門生意的方式，並瞭解這些人口販子的背景。

其他的探索性研究則較為複雜。葛夫利（Gavlee, 2005）曾經針對波多黎各地區的種族分類進行過一項探索性民族誌研究。該項研究的動機是來自於先前的一些研究，因為在先前的研究中發現到，在巴西以及大部分拉丁美洲國家的居民，其處理種族問題的方式，與美國本土居民看待種族問題的想法大相逕庭。巴西人重視外在特徵勝於血統，因此也產生了許多浮動而且不確定的種族類別。該項研究的研究問題如下：波多黎各的居民所採用的是什麼樣的類別？而這些類別的分類原則又是什麼？葛夫利把焦點集中於波多黎各的一個小城市，他花了許多時間待在這個小城市裡，並且與二十四位居民進行開放式的訪談，以瞭解他們在談論到別人時所使用的詞彙和類別。接著，他請四十二位居民針對一組人的臉部照片進行分類，而他則使用電腦軟體來分析這些臉部的照片。他發現到，當地居民在分類時的主要方式，是根據外觀而非種族，而且他們是使用五種不同濃淡的顏色當作分類的類別。至於其他的身體外觀特徵（頭髮的質地、鼻子的形狀等等），也僅僅扮演了次要的角色。

有些探索性研究則使用了量化技術。克莉桑（Krysan, 2008）在一項探索性研究裡對調查資料進行了分析，而該項研究的主題是美國境內不同種族的人們如何尋找房子。該項研究詢問了底特律地區數百位民眾，關於他們最近尋找房子的過程，其中包括花了多少時間找房子、在找房子的過程中曾經看過多少間房子、及他們出價和申請的次數等等。克莉桑比較了租屋者和購屋者、以及白人和黑人在找房子策略（例如，和朋友、家人、或者是鄰居討論；注意在庭院裡的招租或是出售的立牌；尋找報紙或是網路上的消息；詢問不動產經紀人或是使用不動產搜尋服務等等）方面的不同。她檢視了百分比例並且發現了許多的相似性，但是，在求助於不動產經紀人、使用網路、以及找房子的困難度與時間長短方面，不同種族的人們之間有一些差異存在。人們通常傾向於向同一種族的不動產經紀人尋求協助。白人比較有可能會使用網路尋找相關資訊，而且在找房子時也比較傾向於選擇那些以白人為主的社區。相較之下，黑人在找房子時，對於地點的要求較為寬鬆，尋找的時間較長，而且也必須提出較多次的申請才能成功地找到房子。

題或是問題開始，並且試著正確而精準地描述此議題或者是問題。研究所得到的結果，將是對於某項議題的清晰圖像，或是對於研究問題的答案。例如，假設某項描述性研究所要聚焦的議題，是父母酗酒過度與虐待兒童之間的關係。如此一來，研究結果可能會顯示出，約有百分之二十五酗酒過度的父母，會在肢體與性方面虐待兒童，而相較之下，沒有喝酒習慣、或是僅喝一點酒的父母，則只有約百分之五的比例會虐待兒童。描述性研究試圖呈現一幅圖像，以描述人們或是社會活動的類別，並且將焦點集中於「如何」以及「誰」的問題上（例如，這件事發生的頻率有多高？有什麼人參與其中？）。探索新的議題、或者是解釋某事件發生的原因（例如，酗酒過度的父母爲何較容易虐待兒童的原因），並不是描述性研究關心的重點，因爲此類

範例 2.5

描述性研究

我們在第一章裡面討論過的羅威瑞（Lowery）與其同事（2007）針對無意識促發與學業表現所進行的實驗研究、艾德傑爾與崔比（Edgell and Tranby, 2007）針對宗教立場在解釋種族不平等現象方面，所造成影響的調查研究、以及范卡德胥（Venkatesh, 2008）針對幫派份子所進行的民族誌研究，都是描述性研究。上述每一項研究的主要目的，都是在描述模式，而不是在處理「為什麼」的問題、或是考驗一個現存的理論。

描述性研究的另外一個例子，是由安涅弗與卡倫（Unnever and Cullen, 2008）所進行的一項研究，其主題是對於死刑的支持。該研究的作者觀察到，許多民意調查都顯示出美國民眾對於死刑的支持與否，在不同的種族之間有著極大的差異。對於此種差異，許多人經常引用白人的種族歧視當作其背後的原因，然而，「目前沒有任何一個系統性的理論可以解釋，為什麼白人的種族歧視觀念會促使白人支持死刑」（page 1283）。該項研究的作者們，針對一份以全美國境內一千五百人為樣本所進行的調查資料，進行了次級資料分析（請參閱本章後面部分）。在經過統計分析之後，他們發現雖然有許多因素（例如傾向獨裁主義的個性、保守的意識型態、宗教信仰、以及反對平等主義的觀點等等）會促使人們支持死刑，但是最能夠準確預測白人支持死刑的因素，則是白人在種族歧視方面所得到的較高分數。在不具有種族歧視想法的白人中，支持死刑的程度和非裔美國人大約是相等的。作者們曾經簡短地討論過理論，但是他們只有在一般性的概念使用理論，而且主要是在描述死刑支持者所具備的特質。換句話說，他們並未直接地考驗任何理論、或是利用理論來產生任何關於此種情況的解釋（請參閱下一個部分）。

研究真正的焦點在於描述事物的樣貌。有許多的社會研究都是描述性研究，而進行描述性研究的研究者，則會使用大多數的資料蒐集技術，像是調查、田野研究、內容分析、以及歷史比較研究方法等等。

解釋

當面對一項眾人熟知且有適當描述的議題時，我們或許會想要知道事物*爲什麼*（*why*）會以現在的方式存在或發展。處理此種「爲什麼」的問題，正是**解釋性研究**（explanatory research）的目的。解釋性研究乃是奠基於探索性研究以及描述性研究之上，並且繼續進一步確認某事發生的原因（請參閱範例 2.6 解釋性研究）。除了針對某項議題提供一幅描述的圖像之外，解釋性研究更要進一步地去尋找原因以及理由。例如，描述性研究可能僅會詳細記錄酗酒過度且會虐待兒童的父母的人數，但是解釋性研究則會把焦點放在這些父母親*爲什麼*會虐待兒童的原因上。換句話說，我們的研究焦點在於，爲什麼酗酒過度會造成這些父母虐待兒童的原因。

在進行解釋性研究時，我們會採取多重的策略。在某些解釋性研究裡，我們會發展出新的解釋，然後提出實徵證據，以支持或拒斥此種新的解釋。在其他的解釋性研究裡，我們會大概地敘述兩個或是更多個彼此互相競爭的解釋，然後爲每種解釋提出證據，並且進行一種「正面對

解釋性研究 此種研究的主要目的，在於解釋事件發生的原因，並且去建立、詳細闡述、延伸、或是考驗理論。

範例　2.6

解釋性研究

在第一章裡提到過麥可卡蒙（McCammon）與同事（2008）曾經針對陪審團權利運動進行過一項歷史比較研究，而該項研究正是一項解釋性研究，因為該項研究把焦點集中於解釋為什麼陪審團權利運動在某些州比在其他州更為成功。而在第一章裡提到的麥可貝與索柏列斯基（McVeigh and Sobolewski, 2007）所進行的研究，也是一項解釋性研究，因為該研究的作者們，利用種族競爭理論與分割勞動市場理論，解釋為何各郡之間會有不同的投票行為。

解釋性研究通常會先概略描述一個現有的理論，然後去檢驗這個理論、將這個理論延伸至一個新的領域或群體。在過去的五十年裡，接觸假說（contact hypothesis）一直是一個相當廣為人知的社會心理學理論，而該假說主要是被用來研究不同種族之間的關係。該假說解釋了偏見的程度，以及負面態度的成因，也就是人們會由於無知以及負面的刻板印象，而傾向於對「外團體」（"out-group"）採取負面的看法。一旦人們接觸到、並且逐漸認識了外團體裡的成員之後，他們就會以較為正面的看法，取代原先的無知狀態和負面的刻板印象。此種情況回答了為什麼人們會對外團體抱持負面感受：這是因為人們缺乏與外團體的接觸。許多研究檢視了這個假說，而檢視的方式則為深入探究各種接觸的具體情況、以及外團體被視為具有威脅性的程度有多高。

李、法瑞爾、以及林克（Lee, Farrel, and Link, 2004）把接觸假說加以延伸，並利用該假設來解釋一個新的主題：那些在美國城市裡無家可歸的遊民。三位研究者檢視了十四項與這群遊民有關的衡量指標，其範圍包括，關於這些人的資訊（例如像是文章、電視報導等等）、研究者個人觀察結果、研究者個人與這群遊民的互動、曾有無家可歸的經驗、及家中有成員曾經無家可歸等等。三位研究者也發展出全面性的衡量指標，以測量某個人對於無家可歸的人們所抱持的看法。而這些看法則包括了對於人們為何無家可歸的想法、認為無家可歸的人具有危險性、對這些人感到同情並抱持正面情緒、與支持這些遊民的權利等等。在 1990 年，研究者以電話調查的方式，在美國的兩百個都會區，隨機抽樣訪問了 1,388 名成人。利用這份電話調查所得到的資料，研究者們發現了支持接觸假說的明確證據。相較於那些幾乎沒有接觸、或是根本不曾接觸過無家可歸的遊民的民眾，那些與遊民有越多接觸、或是有越密切接觸的人們，對這些遊民所抱持的看法是最為正面的，而且也最有可能去支持能夠幫助遊民的社會福利計畫。除此之外，研究者們也發現到，依據個人的種族、年齡、教育程度、以及政治意識型態的不同，人們對於遊民的看法也會有些微的變化與差異。

決」（"head-to-head"）的比較，以確定哪一種解釋較爲合理或可靠。而在其他一些解釋性研究裡，我們以一個來自於社會理論或是過往研究的現存解釋，當作研究起點，然後將此解釋加以延伸，解釋一個新的議題、背景環境、或是一群人，以決定這個解釋是否能夠站得住腳、是否需要修正、或是這個解釋是否只能適用於某些特定的情境。

個案內或是跨個案研究

根據我們探討的個案數目、以及在探討個案特色時的深度與程度，研究的類別也會因而有所變化。有時候我們會從一群數目相當龐大的個案或是母群中，仔細地挑選出或抽樣，得到數目較少的個案。然而，此類研究中所牽涉到的個案數目，仍

然可能高達數百或是數千個之多。而在其他的研究裡（尤其是實驗研究），我們則會去分析大約十來個人的情況，並會操縱他們的一些情況或是條件。而在另外一種研究類別中，我們則是密集而集中地檢視一個、或相當少數的個案（個案數目或許會少於十個）。雖然在研究中所探討的個案數目確實相當重要，但是更爲要緊的議題，則是該項研究的焦點究竟是在於探討個案內的、或者是跨個案的特徵。如同拉金（Ragin, 1994: 93）所觀察到的，「一般而言，在個案的數目、以及研究者能夠研究的個案特徵的數目之間，研究者必須作出一些取捨」。

「個案」的概念極爲重要，但是卻可能相當複雜。戈林（Gerring, 2007: 17）認爲個案「在定義方面是一團混亂」。之所以會有此種混亂複雜的情況，是因爲有太多事物都可以當作個案。個案可以由一項研究的觀點以及研究問題來加以確定。就形式上而言，一個個案在時間與空間方面，要有明確的界限，而且也經常被稱作一個「單位」或者是「觀察對象」。一個單獨的個體可以被當成是一個個案，而一個家庭、一間公司、或一整個國家也可以被當作是一個個案。在某一項研究中，被當作個案來加以探討的人事物，在另一個不同的研究中，可能無法被當成個案來探討。舉例而言，國家可以被當成個案，以探討一個國家的各個不同層面，或者在某個國家整體人口中許多個體的不同面向，也可以被當成個案來加以探討。

一個個案並非單純只是任何個人、家庭、公司、或國家而已，我們選擇某個個案的原因，是將它當成一個「事件類別」（“class of events”）的一部分，或因爲此個案屬於某個特定的類別（請參閱 George and Bennett, 2005: 17）。我們研究某個個案的原因，是因爲該個案是某種類別的一部分，而我們打算研究該類別的事物，以發展出相關知識，去瞭解在某種類別的個案中，造成相似性與差異性的原因。例如，我不會因爲我的鄰居艾利克斯住在我家隔壁，就把他當成一個研究的個案。然而，我卻有可能把艾利克斯當作一個相似個案類別裡的一個案例來加以研究。該相似個案的類別是指，那些因爲身體殘障而無法上班、只能依賴另一半上班賺錢、並且因而成爲全職「家庭主夫」的中年男子。同樣地，我也有可能把 1962 年發生的古巴飛彈危機，當成一個個案來研究。但卻可能作爲「國際危機管理與嚇阻情境」類別裡的一個案例加以研究。

不論是在任何一項研究裡，研究者都應該要先瞭解以下兩件事情：研究中要探討的個案數目，以及研究重點究竟是在少數個案之內、還是橫跨許多不同個案間，進行詳細的檢視與調查。

個案研究

個案研究 此種研究的研究對象是非常少數的單位或是個案，並在一段時間、或是橫跨不同的期間之內，針對這些個案的大量資訊進行相當深入的檢視。

個案研究（case-study research）主要是在詳細檢視少數幾個個案的許多特徵。這些個案可以是個人、團體、組織、運動、事件、或者是地理單位，而由個案所得到的資料是詳盡的、多元的、以及廣泛的。個案研究可以把焦點放在某一個特定的時間

點、或者是一段時間上。雖然大多數的個案研究都是質化研究，但是個案研究其實不一定非質化研究不可。相較之下，幾乎所有的跨個案研究（或者是非個案研究）都是量化研究。質化研究與個案研究之間並不完全等同，但是「幾乎所有的質化研究都會試圖以對個案深入的、詳盡的瞭解爲基礎，來建立其代表性」（Ragin, 1994a: 92）。[17] 在第一章裡所提過的范卡德胥（Venkatesh, 2008）針對都市幫派份子所進行的民族誌研究，就是一項個案研究，因爲該研究詳細描述了在南芝加哥地區一個有限的地理區域內，某一個幫派和周遭的一些具體事件，以及在八年內的關係發展過程。

在個案研究中，研究者會集中地探討一個或是少數幾個個案，並把焦點集中於每個個案裡，及其情境脈絡的許多細節之上。簡而言之，個案研究的研究者會詳細地檢視每個個案的內部特徵、以及其外在情境的許多細節。個案研究讓我們能夠把微觀層面、個人的行動與鉅觀層面、或大規模的結構與過程連結起來（Vaughan, 1992）。正如同華爾頓（Walton, 1992b: 122）所說的：「個案研究的邏輯在於證明一個因果關係的論點，而該論點則是關於在特定的社會情境中，一些社會力量將會如何去形塑，產生各種不同的結果。」

個案研究有許多的優點與長處。例如，個案研究能夠使我們的思想變得更加清楚，並且允許我們以特定的方式，將抽象的想法與我們仔細觀察的個案的具體細節連結在一起。除此之外，個案研究也使我們能夠根據實際的經驗，與被廣爲接受的證據標準，校正或是調整那些用來測量我們抽象概念的指標。個案研究的其他優點則與理論有關。正如同華爾頓（Walton, 1992b: 129）曾經提過的：「個案研究可能產生最好的理論」，而會發生此種情況的原因，有下列三項：首先，由於我們對於特定個案的深度細節已經極爲熟悉，因此我們就可以創造或是建立新的理論，並且能夠依據較複雜的個案或是新的情境，而把現有的理論加以改造。其次，當我們檢視特定的個案時，社會過程以及因果關係裡那些錯綜複雜的細節，就會變得比較清晰可見。此種情況讓我們能夠發展出較爲豐富的、更爲全面的解釋，更貼切地記錄下社會生活的複雜程度。除此之外，個案研究所提出的證據能夠更爲有效地描述複雜的、具有多重因素的事件／情境以及過程，而這些事件或過程，則是在一段時間與空間中所發生的。最後，個案研究也能夠在其中包含一個完整的情境以及多重的觀點。

總結而言，個案研究具有下列六項優點：[18]

1. *概念效度*（*Conceptual validity*）：個案研究可以幫助我們「找出」（"flush out"），確認我們最感興趣的概念或是變項，並且幫助我們在抽象的理論中，更加瞭解這些概念或變項的核心，抑或其基本意義。
2. *啓發式影響*（*Heuristic impact*）：個案研究具有高度的啓發性（也就是能夠提供我們更進一步的學習、發現、或解決問題的機會），能夠幫助我們建立新的理論、發展或是擴展概念、以及探索相關概念之間的界限。
3. *確認因果機制*（*Causal mechanisms identification*）：個案研究能夠讓我們清楚地看出，在社會過程及各種因素交互影響的作用機制中的細節。

4. *捕捉複雜性與追蹤過程的能力*（*Ability to capture complexity and trace processes*）：個案研究能夠有效地描述高度複雜的、具有多重因素的事件或是情境，並且追蹤在一段時間和空間之中的過程。
5. *校準*（*Calibration*）：個案研究讓研究者能夠依據可靠的、真實的經驗以及具體的標準，來調整抽象概念的測量指標。
6. *整體性闡述*（*Holistic elaboration*）：個案研究能夠針對一整個情境或是過程作整體性的完整闡述，並且允許研究者在研究中包含多重的角度或是觀點。

個案研究的焦點相當詳盡而瑣碎，但是卻能訴說一個較大的故事（請參閱範例 2.7 個案研究）。華爾頓在他針對加州地區歐文谷社區所進行的個案研究中曾經提到（Walton, 1992a）：「我試著……透過一個小個案來訴說一個大故事」（p. xviii）。該社區致力於進行一項社會抗議活動，因爲該社區想要控制相當重要的水資源，與社區的命運，而這項抗議活動則以不同形式，持續超過一百年之久。在華爾頓的研究中，他檢視了各種不同形式的資料，其中包括了直接觀察、正式與非正式的訪談、人口普查的統計資料、地圖、舊照片與報紙、各種不同的歷史文件、以及官方的正式記錄等等。

跨個案研究

大多數的量化研究會從大量的個案（30 個到 3,000 個）裡蒐集資訊，並且專注於這些個案裡的少數特徵。跨個案研究不會對每個個案進行鉅細靡遺的調查，而是橫跨許多個案，去比較研究者所挑選出來的少數特徵。換句話說，跨個案研究認爲每一個個案都具有一些研究者所感興趣的特徵。

雖然某些議題會比較適合某一種研究方法，但有些時候研究者可能會使用個案研究，及跨個案研究的研究設計，來研究相同的議題。舉例而言，假設我們想知道某個家庭如何決定是否要搬到另一個不同的城鎮居住。那麼，我們可以採取的策略之一，就是針對五個家庭進行個案研究。我們可以進行極爲詳細的觀察，並對每個家庭的決策制訂過程，進行深度訪談。我們可以採取的另一個策略，則是針對男主人的工作、家庭年收入、以及對於搬到另一個不同城鎮的決定這三者之間的關係，進行一項跨個案研究。我們或許會檢視 1,000 個不同的家庭，並確認在過去的五年裡，曾經搬過家的 250 個家庭、以及未搬過家的 750 個家庭的男主人的工作與家庭年收入。在這項跨個案研究中，每個家庭單位都有我們想要研究的三項特徵，也就是男主人的工作、家庭年收入水準、和搬家或不搬家的決定。跨個案研究的焦點在於各項特徵（工作、收入、以及搬家的決定）之間的關係，而不是每個特定家庭內所發生的事情。

範例　2.7

個案研究

或許你已經看過一部於 2002 年上映、並且曾經得獎的電影《戰地琴人》（*The Pianist*），該部影片的內容是關於一位波蘭籍的鋼琴家史匹曼（Wladyslaw Szpilman）的故事、以及 1943 年發生於波蘭華沙地區的猶太人民起義事件（譯註：《戰地琴人》是一部根據第二次世界大戰時，一位波蘭音樂家真人真事改編的電影，影片主角史匹曼 1911 年出生於波蘭。他是一位才華洋溢的猶太裔鋼琴家，而且自幼就追隨名師學習鋼琴。1931 年時，史匹曼到德國柏林的音樂學院深造，並寫下多首著名的鋼琴樂曲。他於 1935 年返回波蘭，並且受聘於波蘭國家廣播電台。1939 年 9 月，當他正在電台現場彈奏蕭邦的夜曲時，第二次世界大戰爆發，德國迅速進攻並佔領波蘭，而波蘭首都華沙遭受德軍無情的轟炸。從此之後，史匹曼一家人與鄰居都被德軍趕進在首都的「猶太特區」居住。居住在「猶太特區」裡的居民不僅出入受到嚴格的限制，連他們的一舉一動也都遭到嚴密的監視，尤有甚者，納粹軍官不時會對這些居民進行百般羞辱。在忍無可忍的情況之下，有一小群猶太人進行了報復行動計劃，但卻使得整個社區遭受到德軍幾近毀滅的轟炸與殘殺。為了逃避此種悲慘的遭遇，史匹曼冒著生命危險，成功地逃出「猶太特區」，並且躲進滿目瘡痍的首都公寓廢墟。為了不被納粹軍人發現，他只好每天晝伏夜出，躲在天花板的夾層中，而且一待就是好幾個月。然而，有一天，一位納粹軍官何思費（Wilm Hosenfeld）進入史匹曼躲藏的公寓，無意間發現這位逃脫的猶太人。此時史匹曼由於過度驚嚇而變得不知所措，此外，由於史匹曼長期躲藏、不見天日，整個人因而變得病奄奄的，臉色也極度蒼白。在詢問之下，何思費得知史匹曼以前曾經是一名鋼琴師，而在聽完他演奏一首美妙的曲子後，何思費不忍心殺害他，反而決定幫助他。直到六年後，第二次世界大戰結束，史匹曼仍在電台演出同一首曲目，然而，那名曾經幫助過他的納粹軍官何思費，卻已死於俄國的戰俘營。在德國納粹戰敗瓦解後，史匹曼一生表演過 2500 場以上的演奏，並獲得多次的世界大獎。最後，他於 2000 年 7 月 6 日以 88 歲高齡逝世。在他過世之後，他的兒子發現了他留下的日記手稿，並將其整理出版，書名為《The Pianist》，成為一本暢銷書。而本部影片即是根據這本書改編拍攝，其中的情節令人感動，也讓人們反思戰爭的殘酷，並且期待世界和平的來臨）。艾恩渥納（Einwohner, 2003）曾對一個單一事件——1943 年的猶太人民起義事件——進行了一項歷史個案研究，以檢視一個廣被接受的社會運動理論。該理論以下列三個概念為基礎：政治機會結構（political opportunity structure, POS）、威脅、以及動機框架（motivational frame）。政治機會結構是指，在機構裡及資源控制中的一整套選擇與限制。當新的機會出現時（例如，反對勢力出現了分裂、停滯、分散、或是缺乏補給的情況），那麼政治機會結構就會「出現缺口」，同時也會增加一項社會運動成長或是成功的可能性。政治機會結構的理論也承認威脅的存在。所謂的「威脅」（threat），其定義是指某項運動採取某些行動（例如，新的法令限制了抗議活動，而且有許多人因而被逮捕）、或是不採取某些行動時，所會增加的成本。第三個概念則是「動機框架」。框架指的是人們思考及看待某項事件的方式。而動機框架所指的就是，參與者在採取行動時，認為可以接受的理由，或者是具有道德正當性的事物。社會運動理論認為，當下述三項情境同時發生時，社會運動就會有所進展：當政治機會結構出現了缺口、威脅的程度很低、以及人們擁有一個能夠促使他們採取行動的框架時。

艾恩渥納（Einwohner, 2003）研究了 1943 年華沙猶太特區這個特定個案裡的日記資料以及歷史記錄。她發現了一個緊緊封閉的政治機會結構，與一個充滿巨大威脅的情境。居住於猶太特區裡的猶太人面對具有高度作戰能力、並佔有絕對優勢的軍事力量，而納粹也開始進行一項大規模屠殺猶太人的政策。因此，在進行一項成功社會運動的三個必要條件中，已經有兩項無法達到。但是猶太特區裡的猶太人卻形成了一個新的、而且相當激進的動機框架。猶太人把在鬥爭中壯烈犧牲的行為重新定義為他們唯一可以接受的、光榮的選擇。換句話說，猶太人不再把死亡視為一件可怕和應該避免的事，反而將在起義活動中壯烈犧牲的行

為視為是極度英勇的、有尊嚴的、與相當光榮的行動。對於在一場不可能勝利的戰鬥中被殺害這件事，猶太人將其重新定義為光榮行為，而且是為了每個猶太人跟整個猶太民族所必須作出的犧牲。因此，該項個案研究發現，雖然缺乏了由社會運動理論所預測的兩個極為重要的成功因素（也就是開放的機會以及較低的威脅），一項大規模的運動仍然應勢而生。事實上，在猶太起義運動中，猶太人可以說是完全沒有任何機會、而且面對著極高的威脅。在這個個案裡，猶太起義這項大規模社會運動所依賴的基礎，是所有的猶太人在一個完全絕望的情境裡，重新定義他們所必須從事的行動。因此，艾恩渥納所進行的這項詳盡的個案研究，修正了一個廣被接受、而且有大量文件支持的現存理論。

單一時間點或是多重時間點

在每一項研究裡，時間都是一個相當重要的面向。我們以橫斷性以及縱貫性兩種方式將時間面向整合到研究中。**橫斷性研究**（cross-sectional research）是在某一個特定的時間點蒐集資料，並且形成一種社會生活的「快照」（"snapshot"）。**縱貫性研究**（longitudinal research）則是在多重的時間點上蒐集資料，並且提供在一段時間內的事件、人物、或是社會關係等等「動態圖像」（"moving pictures"）。一般而言，縱貫性研究比較難以進行，而且也需要比較多的資源。研究者也許會在許多不同的時間點蒐集許多單位的資料，然後去尋找一些橫跨這些單位或是個案的模式。[19]

橫斷性研究

橫斷性研究可以是探索性的、描述性的、或是解釋性的，但是通常最適合於描述取向的研究。橫斷性研究通常是最爲簡單、而且也是花費最少的，但是卻很少能夠真正地捕捉到社會過程或是變化。在第一章所提及過的研究裡，由艾德傑爾與崔比（Edgell and Tranby, 2007）針對宗教立場在解釋種族不平等現象方面所造成的影響所進行的研究，以及由麥可貝以及索柏列斯基（McVeigh and Sobolewski, 2007）針對支持共和黨與民主黨州的現存統計資料所進行的研究，都屬於橫斷性研究。而在本章所提到的研究中，針對波多黎各地區的種族分類方式所進行的研究（Gavlee, 2005），以及針對底特律地區不同種族的人們如何尋找房子所進行的研究（Krysan, 2008），也都屬於橫斷性研究。而由安涅弗與卡倫（Unnever and Cullen, 2008）所進行的針對死刑看法的研究，也同樣屬於橫斷性研究。

要決定一項研究到底是屬於橫斷性研究或是縱貫性研究，並非總是件容易的事，因爲作出決定的判斷依據不僅僅是時間的長度而已。在第一章裡曾經提過羅威瑞以及同事（Lowery and associates, 2007）針對無意識促發所進行的實驗研究，而在該項研究的標題裡包含了「長期影響」（"long-term effects"）這四個字，雖然事實上該研究的時間長度僅有 4 天左右，但仍算是一項縱貫性研究。相對地，在由艾德傑爾與崔比（2007）

橫斷性研究 在進行此種研究時，研究者會檢視許多不同個案在同一時間點上的資訊。

縱貫性研究 在進行此種研究時，研究者會檢視許多單位或是個案，在不同時間點上的資料。

所進行的調查研究、及由麥可貝和索柏列斯基（2007）所進行的現存統計資料的研究中，雖然資料是在數天、甚至是數個月的期間之內所蒐集的，卻仍然只能算是橫斷性研究。有關促發效果的實驗研究之所以是一項縱貫性研究，並不是由於進行研究所花費的特定時間長度，而是因爲該研究的研究設計包含了時間面向在內。換句話說，該項研究的研究者在二個不同的時間點上蒐集資料，並在分析資料的階段，針對這些不同時間點上所蒐集的資料，進行比較。而在調查研究與現存統計資料的研究裡，研究者通常無法一次就蒐集到所有的資料。研究者認爲在蒐集資料時的時間差異並不會影響研究的結果，並且在其研究設計裡忽略了時間方面的差異。

縱貫性研究

我們可以將縱貫性研究應用於探索性的、描述性的、以及解釋性的目的上。雖然進行縱貫性研究通常比橫斷性研究更爲複雜、而且花費也較爲高昂，但是縱貫性研究的解釋力也更爲強大。由麥可卡蒙與同事（McCammon and colleagues, 2008）針對陪審團權利運動所進行的研究，就是一項縱貫性研究，而該項研究的焦點則在於數十年以來社會變遷的步調及模式。該項研究的作者們在多重的時間點上蒐集資料，而且在其研究設計裡，也把這些在不同時間蒐集而來的資料進行比較。

接下來，讓我們一起來探討三種不同類型的縱貫性研究：時間序列研究、同樣本多次研究、以及世代研究。

1. **時間序列研究**（time-series research）是一種縱貫性研究，在此種研究中，研究者針對屬於同一類別的人群或是單位，在不同的時間點上蒐集其相關資料。此種研究讓研究者能夠觀察到研究對象或是單位，在其特徵方面的穩定性或是變化，或是能夠追蹤各種條件在時間方面的變化（請參閱範例 2.8 時間序列研究）。

即使是在時間序列資料裡某個項目的簡單描述性資訊，也可能會揭露出相當多的訊息。舉例而言，關於美國從 1910 年迄今的出生率之時間序列資料（圖 2.1）顯示，在 1920 年代時，每位婦女的生育數開始出現了下降的現象，而且此種下降現象一直持續到 1930 年代以及 1940 年代早期，但是，在第二次世界大戰結束之後（1945 年），此種下降的趨勢出現了急遽反轉的情況。之後，在 1950 年代以及 1960 年代，出現了後來被稱爲「嬰兒潮」的生育率高漲的情況。一直到 1970 年代，生育率才又持續下降並達到穩定的狀況。時間序列能夠顯露出一些以其他方式不容易看到的變化。例如，自從 1967 年開始，「高等教育研究機構」（Higher Education Research Institute, 2004）每年都會針對進入美國大學就讀的學生蒐集大量的調查樣本資料，以利於大學進行應用研究。其中一項調查資料是有關於學生心目中認爲最重要事物的百分比例，而關於此項調查資料的時間序列結果（圖 2.2）顯示出，關於重要事物的優先順序，在 1960 年代和 1970 年代之間出現了明顯的反轉現象。當時有許多的大學生不再把發展出有意義的生命哲學當作是最重要的事物，反而轉向追求物質與財務方面的成功。

時間序列研究　一種縱貫性研究，在此種研究中，所蒐集的資訊是關於不同的個案或是人群在每個不同時段裡的資訊。

範例 2.8

時間序列研究

佩堤特與威斯騰(Pettit and Western, 2004)曾經進行過一項時間序列研究，主題是從 1964 年到 1997 年之間，美國黑人與白人的監禁率。該項研究發現，美國在 1980 年代時，監禁率曾出現大幅度增加的情況（增加了約 300%），而在這段期間裡，黑人被關進監獄的機率約比白人多了六到八倍左右。沒有上過大學就讀的年輕黑人比白人更有可能會被關進監獄，而且每三個年輕黑人就有將近一人曾經吃過牢飯。若是連高中都沒有讀完的年輕黑人，上述的比率則是增加為兩倍之多。藉由時間序列的結果比較，該項研究的作者讓人們看到，入監服刑的人數在不同種族間增加的比例是不對稱的，而且入監服刑人數的增加，主要是來自於美國整體人口中的某些特定部分。

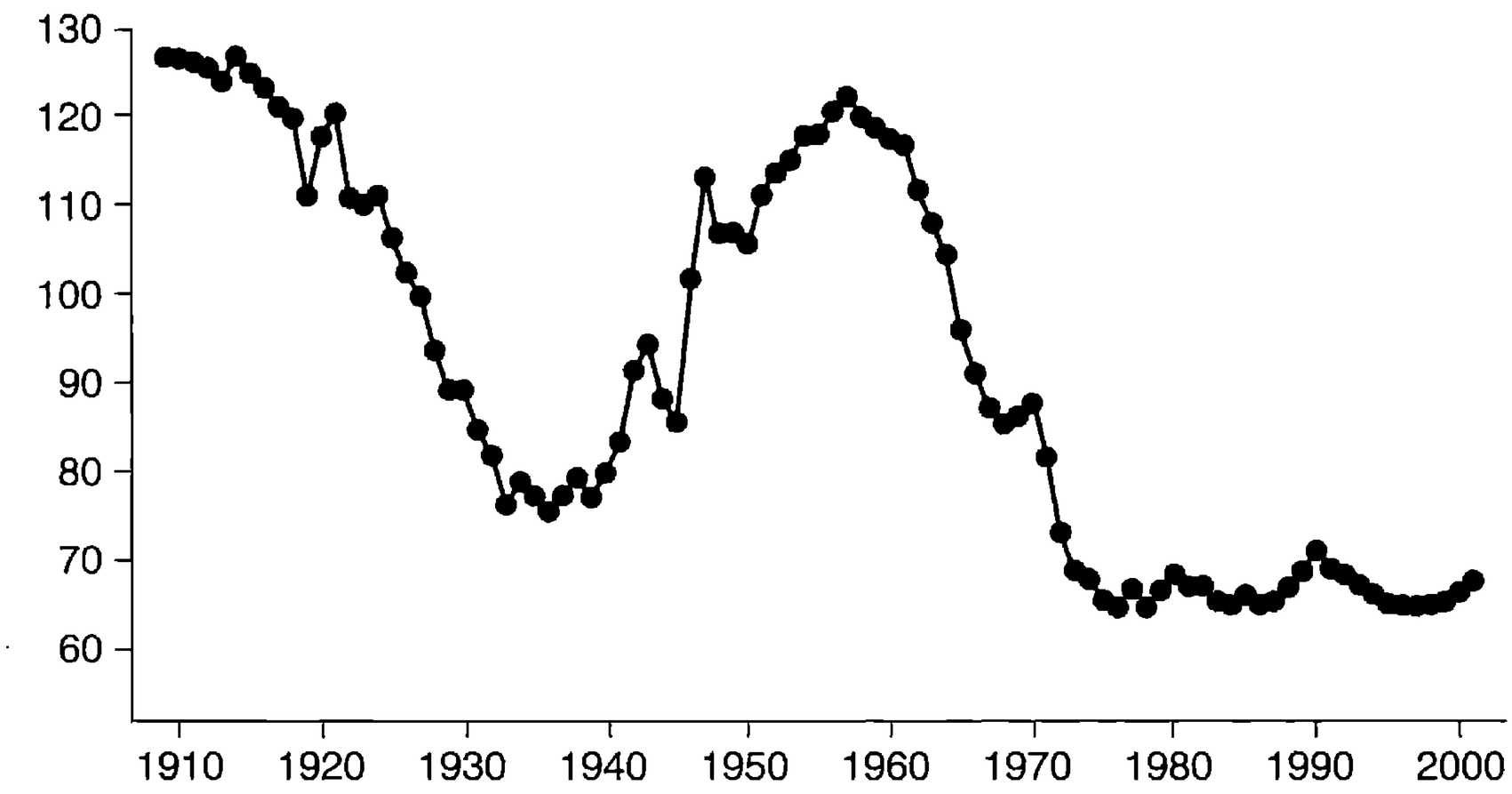


圖 2.1 美國從 1910 年到 2000 的出生率（每一千名 15-44 歲女性的生育數）

資料來源：作者由美國人口普查資料所計算得出

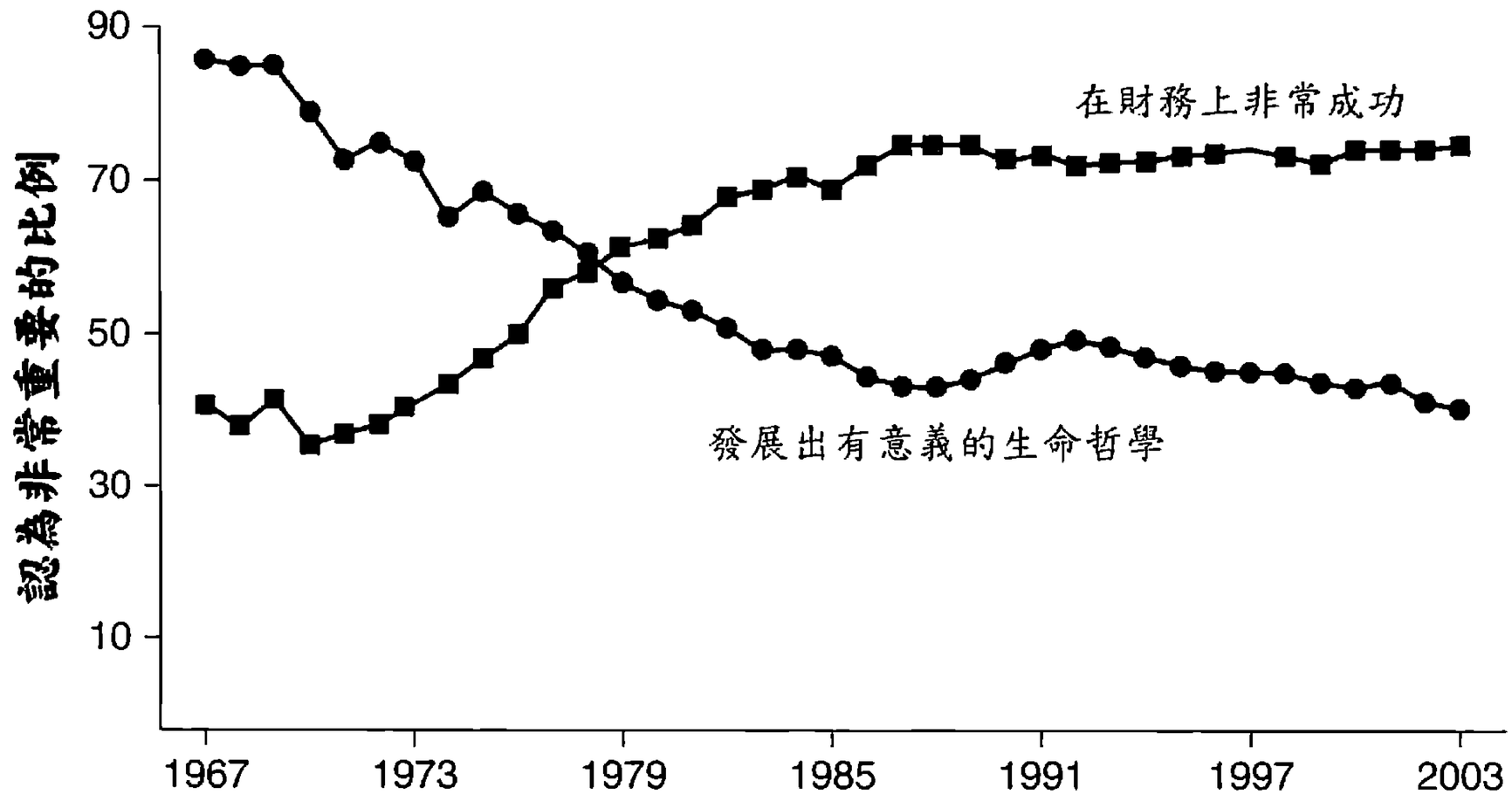


圖 2.2 1967 年到 2003 年之間，美國大學新鮮人的價值優先順序

資料來源： From Higher Education Research Institute. (2004). Recent findings, Figure 4. Retrieved September 25, 2004, from www.gseis.ucla.edu/heri/findings.html.

2. **同樣本多次研究**（panel study）是一種解釋力非常強大的縱貫性研究（請參閱範例 2.9 同樣本多次研究），但是也比時間序列研究更加難以進行。在進行同樣本多次研究時，研究者會對完全相同的一群受試者、團體、或組織，在不同的時間點進行觀察，抑或是蒐集資料。進行同樣本多次研究的困難度很高，花費也相當驚人。長時間追蹤同一群人是相當困難的，因爲有些人可能死亡、或是無法被尋獲。然而，無論如何，一項設計完善的同樣本多次研究結果，具有高度價值。即使是一項短期的同樣本多次研究，也能夠清楚地顯示出某個特定生活事件所造成的衝擊或是影響。然由於在同樣本多次研究中，我們所研究的是完全相同的一群人，所以我們可以從中得知許多明確的事物。例如，布魯爾等人（Brewer et al., 2005）研究了 2001 年 9 月 11 日發生的恐怖攻擊，對於人們的態度所造成的衝擊。研究者詢問人們對於其他國家的信任程度，並在三個不同的時間點上（2001 年 10 月、2002 年 3 月、以及 2002 年 9 月）針對同一群美國民眾進行了同樣本多次研究。他們發現在 911 恐怖攻擊事件過後，美國民眾對於其他國家所產生的感覺並非只是暫時性的。隨著時間流逝，美國民眾對於其他國家的不信任感逐漸增加，而且在一年之後，不信任感反而變得更高。此種情況顯示，911 恐怖攻擊事件終結了美國民眾對於其他國家的正面感受，並且引發了許多美國民眾對於外國人更深層的恐懼以及敵意。

3. **世代研究**（cohort study）與同樣本多次研究相當類似，但世代研究並非是觀察完全相同的一群人，而是研究在一段特定時期內共享類似生命經驗的同一類型的人（請參閱範例 2.10 世代研究）。世代分析是一種「明確的鉅觀分析」，換句話說，研究者把屬於某個類型的人當成一個整體，並深入檢視其重要的特徵（Ryder, 1992: 230）。在進行世代研究時，我們把焦點放在「世代」、或是某一個定義明確的類型上面。研究者所經常使用的世代類型，包括了所有在同一年出生的人〔稱爲*出生世代*（*birth cohorts*）〕、所有在同一時間被僱用的人、所有在同一年或是前後兩年之間退休的人、或是所有在某個特定年份畢業的人。與同樣本多次研究不同的是，在世代研究中我們不必找到完全相同的一群人，相反地，我們只須找出那些曾經經歷過相同生活事件的人即可。舉例而言，一項世代研究可能會比較三個婚姻世代，也就是分別在 1970 年、1990 年、以及 2010 年這三年當中結婚的所有人。然後檢視這三個世代在結婚典禮方面是否有所差異、在結婚時新娘是否已經懷孕、以及其他的一些特徵。

同樣本多次研究　一種縱貫性研究，在此種研究中，研究者會針對完全相同的個案或者是人群，在每個不同的時段裡重覆蒐集資料。

世代研究　一種縱貫性研究，此種研究追蹤關於某個特定類型的個案或是群眾的資訊，而這些個案或是群眾都是在某個時間點上、或是後續的一些時間點上有過共同的經驗。

資料蒐集技術

這個部分將針對主要的資料蒐集技術，進行一些簡短的概述，而在後續的章節裡，你將會讀到關於這些技術的詳細細節，也會學到應該如何使用這些技術。我們可以依據你所蒐集到的資料類型，把

範例 2.9

同樣本多次研究

在許多大型的美國城市裡，進入中學就讀的學生裡面，有高達百分之五十的人最後根本無法從中學畢業。奈利德、史東爾－埃比、以及佛爾斯頓柏格（Nelid, Stoner-Eby, and Furstenberg, 2008）研究了輟學的議題，而其方法則是把焦點集中於九年級學生身上。他們所使用的是「費城教育縱貫性研究」（Philadelphia Education Longitudinal Study, PELS）的同樣本資料，而在費城教育縱貫性研究當中，研究者們持續追蹤了某個中學學區裡約百分之十的青少年。在這個學區的學校裡，研究者們隨機挑選出一些學生與家長，並且在這些學生完成了八年級學業之後的那個暑假裡，對這些學生和家長進行了大約半個小時的電話訪談。之後，研究者再度持續地對同一批學生和家長進行訪談（以英文或是西班牙文進行），時間點分別是在這些學生就讀九年級的那個秋／冬學期（第二次調查）、九年級結束後的那個暑假（第三次調查）、以及之後每個學年結束時，一直到 2000 到 2001 年的秋／冬學期為止（也就是這一批學生原本應該已在中學就讀四年之後的六個月）。到了第四年度結束時，由九年級開始就讀的學生，有約百分之 48.9 的人已經畢業。這項研究試圖要去確認的是，在九年級時的課業落後情況，以及出缺席的情況，是否能夠大幅度地增加我們預測學生輟學情形的正確率。研究者們以統計方式分析了所蒐集到的資料，並且也發現，九年級的就學情況確實會對學生輟學的可能性有極大的影響，換句話說，在整個過程中，九年級是一個關鍵的「轉捩點」。許多最後輟學的學生，在處理這個社會與學業的過渡階段時，都遇上了困難。這些在社會調適方面有困難的學生，會表現出在行為與出席率方面的問題，而且他們之間大多數的人，在關鍵的九年級課程裡（數學和英語）都會不及格，這是因為他們根本還沒準備好去應付中學程度的課程。由於在此研究中會年復一年地去訪談同樣一批家長與學生，所以此研究是一項同樣本多次研究。

詹寧斯以及傑特納（Jennings and Zeitner, 2003）研究了公民參與（civic engagement）的議題，不過，他們把焦點集中於網際網路對於美國民眾的影響上面。這兩位研究者注意到，橫斷性資料顯示出網際網路使用者在公民參與方面的程度較高，然而，教育程度較高的人通常比較常使用網際網路，而且在公民組織的參與程度方面比較高。在過去所進行的研究中，無法判斷當人們使用網際網路的程度逐漸增加時，是否會影響到人們投入公民參與的程度。於是，他們進行了一項同樣本多次研究，該項研究在 1965 年針對一群高三學生進行了調查，並且分別在 1973、1982、以及 1997 年（此時這群參與者已經五十歲左右了）對同一群參與者進行後續的研究。利用這些資料，兩位研究者就能衡量出在網際網路普及之前以及之後的公民參與情況。網際網路是在 1982 年出現的，而在 1997 年時，使用網際網路已經是非常普遍的現象。在該項研究中，同時調查了原本接受訪談的參與者，以及這些參與者的後代子孫。至於衡量公民參與程度的指標，則包括了範圍相當廣泛的行為和態度。一般而言，兩位研究者發現，在網際網路出現之前就已經相當投入公民組織的人，比較有可能會使用網際網路，而對於那些經常使用網際網路的人而言，在他們開始使用網際網路之後，公民參與的程度也會隨之增加。對於從 1965 年就參與本項研究的人們（現在大約是五十歲左右）而言，一旦開始使用網際網路，那麼這些網路使用者在所有形式的公民參與的投入程度都會增加。然而，他們的後代即使會使用網際網路，卻比較不願意去當志工、或是參與當地社區的活動。在使用網際網路之後，較年長的人會比較年輕的人（尤其是那些使用網路的目的並非出自於關心公共事務的年輕人）更投入公民參與。

範例　2.10

世代研究

安德森以及菲特納（Anderson and Fetner, 2008）兩人曾經進行過一項跨國的研究，分別在 1981 到 1982 年、1990 年、以及 2000 年時對不同國家裡的民眾進行調查，並且檢視了美國及加拿大的民眾，對於同性戀的接受程度。兩位作者利用上述研究所得到的資料進行分析，並且發現對於同性戀的接受程度會隨著出生的世代與時間而增加。因此，出生於二十世紀晚期的人，會比出生於二十世紀早期的人更能接受同性戀，而且在較晚近的年代裡，所有人對於同性戀的接受程度都普遍提高了。例如，以出生於 1920 到 1929 年的人為例，他們在 1981 到 1982 年之間對於同性戀的接受程度，會比在二十年之後（也就是 2000 年時）的接受程度來得低。而出生於 1960 到 1963 年之間的人，當他們在 1980 年以及 2000 年被問及對於同性戀的接受程度時，他們的接受程度都比 1920 到 1929 年出生的人還要高，而且他們的接受程度也會隨著時間經過而逐漸地增加。在這項研究中引人關注的面向之一，是加拿大與美國之間的比較。在 1980 到 1982 年之間，不論是哪一個世代的民眾，加拿大人對於同性戀的接受程度都比美國人還要低。因此，當研究者在 1981 到 1982 年之間進行調查時，出生於 1920 年代、1940 年代、以及 1960 年代的加拿大人（當時這些人分別是六十幾歲、四十幾歲、以及二十幾歲），他們對於同性戀的接受程度都比美國人來得低。然而，當研究者在 1990 年以及 2000 年進行調查時，每一個世代的加拿大人對於同性戀的接受程度，竟然都比美國人還要高得多。事實上，與加拿大人相較之下，美國人在 1990 到 2000 年之間，對於同性戀的接受程度所增加的幅度，可以算是小巫見大巫。除此之外，加拿大人之中最年輕的那個世代（也就是出生於 1960 年代的人），他們對於同性戀接受程度的增加幅度遠遠大於其他世代的加拿大人、和他們出生於同樣世代的美國人。在進行了更為詳盡的分析之後，研究者發現，來自於鄉村地區、小型城鎮、以及大型城市裡的加拿大人，對於同性戀的接受程度都明顯增加了。但是，來自於鄉村地區與極小型城鎮裡的美國人，其接受程度並未增加；只有那些來自於較大型城鎮或是都市地區的美國人，其接受程度才有增加的趨勢。若是有某一位研究者，只研究了 1981 到 1982 年之間的橫斷性資料的話，他將只會看到美國人對於同性戀的接受程度整體略微上升，而且在不同世代之間，僅有些微的差異而已。若是另一位研究者只考慮了 2000 年時的橫斷性資料的話，那麼他將會發現，在不同世代之間有非常大的差異，亦會發現加拿大人對於同性戀的接受程度遠遠高於美國人。只有藉由縱貫性的方式來檢視這些資料，才有可能發現在這兩個國家當中，在不同的世代間、以及在不同時間點，民眾對於同性戀接受程度的變化是非常不同的。

在另一項世代研究裡，布烈特以及金恩（Bratter and King, 2008）檢視了一份資料，該份資料是於 2002 年時在美國境內進行一項全國性抽樣調查所得到的資料，調查對象是曾經結過婚、並對於他們第一任配偶的種族有可靠資訊的 15 歲到 44 歲的民眾（其中包括 1,606 名男性，以及 4,070 名女性）。作者們研究了婚姻世代（也是那些在某個特定年份、或是在連續的幾個年份裡結婚的所有人），比較了不同種族之間、以及相同種族團體的結婚對象。作者們調查了在幾年之後，這些參與者的婚姻是否仍然持續下去、或是已經結束了。在這項研究裡，作者們檢視了六個婚姻世代（1980 年之前、1980 年到 1984 年、1985 年到 1989 年、1990 年到 1994 年、1995 年到 1999 年、以及 2000 年之後）。在進行了跨世代的比較之後，作者們發現，不同種族之間的婚姻，其離婚率通常比較高。然而，此種現象並非發生於所有的婚姻世代中，只有在 1980 年代晚期結婚的人才會明顯地出現此種傾向。作者們也發現到，白人女性與黑人男性、以及白人女性與亞洲男性的結合，其離婚率會高於雙方都是白人的夫妻。然而，若是非白人女性與白人男性結婚、或者是西班牙人與非西班牙人結婚的話，則其離婚率約略等同於、或者是低於雙方都是白人的夫妻。

資料蒐集技術區分爲兩大類：*量化方式*，也就是所蒐集的資料是以數字的形式來呈現的；以及*質化方式*，也就是所蒐集的資料是以文字或是圖像的形式來呈現的。在處理特定種類的研究問題或是主題時，某些技術會比其他技術更能發揮效果。而要使研究問題能夠搭配適當的資料蒐集技術，則需要技巧、練習、以及創造力。

量化資料

實驗 **實驗研究**（experimental research）使用的是在自然科學研究中所發現的邏輯與原則（請參閱第九章實驗研究）。實驗可以在實驗室裡或是在真實生活中進行。實驗研究通常涉及數目相對較少的人（三十到一百），而且是在處理那些聚焦非常明確的問題。對於解釋性研究而言，實驗非常具有效力。

在大多數的實驗裡，例如像是你在第一章裡曾經讀過的促發與學業表現的實驗研究（Lowery et al., 2007），研究者會把實驗對象（在促發與學業表現這項研究裡，總共有七十名實驗對象）分成兩組或是更多組別。接著，研究者會以完全相同的方式來對待這兩組實驗對象，但是研究者會對其中一組施加特定的研究條件，也就是所謂的實驗「處遇」（treatment），而在羅威瑞等人所進行的實驗中，就是用一些和聰明有關的字眼來「促發」學生。接著，研究者詳細地測量兩組的反應。藉由控制實驗情境、並且只對其中一組給予實驗處遇的方式，研究者就能因此而得到結論，認定不同組別之間反應的差異是單獨來自於實驗處遇所致。

調查 身爲研究者，在許多的研究情境中（例如，實驗或田野研究等等），我們會利用問卷或是訪談的方式，來得知人們的信念或是意見。**調查研究**（survey research）利用書面的問卷或是正式的訪談，以蒐集大量民眾的背景、行爲、信念、或是態度等資訊（請參閱第十章調查研究）。一般而言，我們會在一定的時間長度裡，向大量的受訪者（100 人到 5,000 人）詢問數十道問題。在第一章裡，我們曾經討論過艾德傑爾與崔比（Edgell and Tranby, 2007）針對宗教立場在解釋種族不平等現象方面所造成的影響的調查研究，在該項研究中，研究者在 2003 年秋天，對 2,081 名受訪者進行約三十分鐘的電話訪談，以蒐集研究所需的資料。與實驗研究不同的是，在進行調查研究時，我們不會去操縱一個情境或條件以便觀察人們如何反應。我們所做的是，以相同的問題去詢問許多受訪者，並仔細地記錄他們的答案。通常，我們在選擇調查對象時，會使用隨機抽樣的方式（在第八章裡將會討論隨機抽樣的相關議題），如此一來，我們就能夠把我們從少數人（例如，大約 1,000 人左右）所得到的資訊，合理地推論到較大的群體（例如，數百萬人）上面。我們通常會用各種圖形或是表格來呈現調查資料，並且用統計方法來分析這些資

實驗研究 在實驗研究中，研究者通常會把研究參與者分成兩組，並去操縱其中一組的某些條件，然後再比較兩組的反應，以確定操縱這些條件是否會造成差異。

調查研究 一種量化研究，在此種研究中，研究者會有計畫地向大量的受試者詢問相同的問題，並且記錄下他們的回答。

料。一般而言，描述性研究是最常使用到調查技術的研究，解釋性研究偶爾會用到調查技術，至於探索性研究則很少會用到調查技術。

非反應式研究　在實驗研究和調查研究裡，我們會利用創造出實驗情境或是直接問問題的方式，積極主動地讓研究對象參與我們的研究。而此種作法被稱作*反應式*（*reactive*）方式，因爲研究的參與者會由於意識到自己正在參與一項研究，而以某些方式來作出反應。其他的量化研究則被稱爲**非反應式研究**（nonreactive research），因爲研究參與者並未意識到那些關於他們的資訊，正是研究中的一部分。非反應式研究共有以下四種類別：非介入性研究（unobtrusive research）、現存統計資料、內容分析、以及次級資料分析（secondary data analysis）（你將會在第十一章裡學到更多關於非介入性研究的內容）。次級資料分析是以統計的方式針對量化資料來進行分析，而這些量化資料則是之前就已經蒐集，並留存下來的（通常原本是來自於某項調查的結果）。在此，我們僅簡單地討論二種非反應式研究：內容分析以及現存統計資料。

內容分析　**內容分析**（content analysis）是一種用於探討包含於書面文件或者是其他傳播媒體（例如，像是相片、電影、歌詞、廣告等等）裡的內容、資訊、或是符號的技術。在進行內容分析時，我們會先確認所要分析資料的主體（例如，學校的教科書、電視上的節目、報紙上的文章等等），然後建立一套系統來記錄其內容裡的特定觀點，而這套系統可能包括去計算某些特定文字，或主題出現的頻率。在我們系統性地記錄下我們所發現的資訊之後，我們接著會進行分析，通常是使用圖形或是表格，來呈現分析後續所得到的結果。內容分析之所以被認爲是一種非反應式研究，是因爲這些內容的創作者並不知道是否會有人去分析這些內容。內容分析讓我們能夠去發現並記錄在大量的資料內容中，有可能被忽略的某些特徵。內容分析最常被應用於描述性研究中，但是也有可能被應用於探索性和解釋性研究之中（請參閱範例 2.11 內容分析）。

現存統計資料　利用**現存統計資料研究**（existing statistics research），我們可以找出一份先前已經蒐集好的資訊來源，而這些資訊通常是以正式的政府報告形式而存在（在第十一章裡將會詳細討論現存統計資料之研究）。我們會用新的方式來重新組織這些資訊，以探討一個研究問題。找出資訊的來源並確認這些資訊的品質，可能會是一件相當費時的工作。通常，當我們開始進行一項研究時，我們無法得知是否能夠取得研究所需的相關資訊。我們可以在探索性、描述性、或是解釋性研究中使用現存統計資料的研究，但是比較常用於描述性研究中。

非反應式研究　一種研究方法，在使用此種方法時，人們並未意識到他們正在被當成研究對象。

內容分析　在此種研究中，某種傳播媒體的內容會被系統性地記錄下來，並進行分析。

現存統計資料研究　在此種研究中，研究者會重新檢視，並以統計方式去分析先前已經由政府機構或其他組織所蒐集的量化資料。

範例 2.11

內容分析

在1999年發生了科倫拜高中（Columbine High School）的校園槍擊事件之後（譯註：此起事件發生於1999年4月20日，當時有兩名青少年學生攜帶槍械和爆炸物進入校園，在開槍射殺了12名學生和1名教師、並且傷害了其他24名師生之後，這兩名學生隨即開槍自殺身亡。這起事件被認為是美國歷史上最為血腥的校園槍擊事件之一，之後也引起了許多關於槍械合法化以及校園霸凌問題的爭論），羅倫斯與柏克蘭（Lawrence and Birkland, 2004）針對校園槍擊事件進行了內容分析。這兩位研究者所想要知道的是，媒體報導如何影響了和校園槍擊事件有關的立法問題。他們檢視下列四種資料來源的內容，也進行了編碼：在1999年4月及9月間，在兩份發行量最大的報紙上，曾經提到這起事件的文章；1999年的電視新聞報導；1999到2000年之間，在國會裡針對此項議題所進行的辯論；以及在1999到2000年之間，美國國會所通過的相關法案。兩位研究者發現，在媒體報導裡、與國會議員在辯論時所強調的一些引起校園槍擊事件的原因（例如像是流行文化的影響、以及同儕壓力等等），並未出現在法案當中，但是其他的議題卻出現在法案中（像是校園安全以及取得槍械的管道等等）。除此之外，某一項在媒體報導中並未被強調的議題（法律執行方面的措施），後來卻變成了國會議員辯論及立法時的焦點。

質化資料

質化資料的形式非常多元，包括照片、地圖、開放式訪談、觀察、以及文件紀錄等等。我們可以把這些資料簡化爲兩個主要的類別：田野研究（包括民族誌、參與式觀察、深度訪談），與歷史比較研究。

田野／實地研究 **田野研究**（field research）是指對一小群人進行長時間的民族誌個案研究（在第十三章裡將會進一步討論田野研究）。通常在一開始進行田野研究時，研究者只有一個鬆散的研究想法或是研究問題。接著，研究者會選擇一個社會群體或是研究地點、取得接觸或是進入的管道、然後在社會情境裡採取一個社會角色並且開始進行觀察。在進行田野研究時，研究者會在田野環境裡仔細地進行觀察與互動，時間可能長達數個月或是數年。研究者會逐漸熟悉那些被研究的對象，並且可能會進行一些非正式的訪談。研究者所蒐集到的資料形式，是每天所記下的詳細筆記。在進行觀察時，研究者會不斷地思考他們所觀察到的事物，並且修正關於這些事物的重要性的想法。最後，研究者會離開進行田野研究的地點，重新審視他們所記下的筆記，並準備撰寫書面的研究報告。田野研究通常被應用於探索性以及描述性研究，偶爾會被應用於解釋性研究（你在第一章裡所讀到的范卡德胥針對幫派份子所進行的研究，就是一項田野研究，另外也請參閱範例2.12 田野／實地研究）。

田野研究 一種質化研究，在進行此種研究時，研究者會長期地在一個自然情境裡直接觀察研究對象，並且記下相關的筆記。

歷史比較研究 一種質化研究，在進行此種研究時，研究者會去檢視事件的資料，以及在歷史上過往的以及／或是在不同社會裡的各種情況。

歷史比較研究 **歷史比較研究**（historical-comparative research）是數種相關研究類型的集合體。有些研究

範例　2.12

田野／實地研究

鄧尼爾（Mitchell Duneier, 1999）曾經進行過一項田野研究，而其研究對象是美國紐約市格林威治村的街頭小販。他進入現場的方式，是先在一個他已經認識的小販那兒瀏覽書籍，然後那名小販再將他介紹給其他的小販、乞丐、無家可歸的遊民、以及其他人認識。鄧尼爾在大約四年的時間內斷斷續續地觀察這些人，而且偶爾會充當販賣雜誌的小販或者是拾荒者。由於他身為白人、而且又是大學教授，而研究對象則是低收入的非洲裔美國人，僅能靠著在人行道上販賣二手書籍和雜誌維生，因此，他必須經過許多調適，才能瞭解這些低收入者的日常生活、並且得到他們的接納。除了觀察並且記錄下這些小販在人行道上的生活之外，鄧尼爾也進行了多次的非正式訪談、閱讀相關的文件、並且還請一位攝影記者拍下許多田野現場和當地民眾的照片。

在研究結束之後，鄧尼爾所作出的結論，對在社會控制和降低犯罪中的「破窗」理論（"broken window" theory）提出了批評。當其他人只看到混亂無序的街頭環境會造成偏差行為以及犯罪時，鄧尼爾卻在這些掙扎求生的窮苦人們身上，發現了一個充滿著榮譽、尊嚴、與創業活力的非正式社會生活。他注意到，那些屬於中上階級的政府官員和企業領袖所提倡一些法令及規定，通常有可能會破壞他所發現的流動的、健康的非正式社會結構。這是因為他們既不認識這些在社會底層求生存的人、也不瞭解這些人在人行道上的生活樣貌。對於政府官員以及企業領袖而言，他們處於一個中上階級的世界，而這個世界建立在擁有正式規定、正式程序、固定不變的科層體制、以及標準化職位的大型複雜的組織之上。由於他們只看到這些在人行道上討生活的小販無法密切地融入中上階級的世界，因此，他們就認為這些人造成了社會混亂。

會去探討過去某段歷史年代中，某個社會或是數個社會裡的社會生活的各種面向。其他的研究則會去檢視一個不同的文化、或是針對兩個或是更多個不同的文化類型進行比較（我們將會在第十四章中詳細討論歷史比較研究）。我們可能會把焦點集中在一段或者是數段歷史時期上、比較一個或是多個不同的文化類型、或是將歷史時期和文化類型混合來加以比較。如同田野研究一般，在研究剛開始時，我們通常只有一個鬆散的研究問題，然後在研究過程中逐漸修正並且聚焦。我們通常會混合使用許多不同類型的證據，其中包括現存的統計資料、文件記錄（例如，書籍、報紙、日記、照片、以及地圖等等）、觀察記錄、以及訪談。歷史比較研究可以是探索性的、描述性的、或者是解釋性的，但是通常以描述性研究居多。然而，並非所有的歷史比較研究都會依循著質化的取向，有些歷史比較研究可能會檢視，在一個不同的時間點或是不同文化裡的量化資料（例如，調查資料）。

你在第一章裡曾經讀到過一項關於美國的陪審團權利運動的研究、而在本章前述部分也曾經讀到一項關於華沙地區猶太人起義的研究（範例 2.2），這二項研究都屬於歷史比較研究。在這二項研究裡，研究者都去檢視了在某個國家或是文化裡的過往事件。除此之外，歷史比較研究也有可能會橫跨多個國家或是時間點（請參閱範例 2.13 歷史比較研究）。

圖 2.3　社會研究中的時間面向

範例　2.13

歷史比較研究

馬哈尼（Mahoney, 2003）曾經提出過一個與中南美洲國家（具體而言，是指那些曾經被西班牙帝國殖民過的十五個國家）有關的謎題。他觀察到，這些國家在 1900 年時從高度開發國家到未開發國家的排名相對順序，到了 2000 年時竟然維持不變。換句話說，在 1900 年時開發程度最低的國家（當時是玻利維亞），到了 2000 年時，其開發程度依然是最低的。此種排名相對順序極度穩定的狀況，和二十世紀時發生於中南美洲的戲劇性變化以及進步，形成了強烈的對比。當馬哈尼回溯西班牙帝國在十七世紀達到全盛時期的狀況，他發現在那段時期內處於核心地位的幾個殖民地，後來都變成了最為貧窮的國家，而那些當時處於邊緣、落後、而且貧窮的殖民地國家，到了十九世紀後期時，反而變成了開發程度最高、而且最為富裕的國家。

為了解決這個謎題，馬哈尼使用了兩種質化資料分析工具：路徑相依（path dependency）以及質化比較分析（qualitative comparative analysis, QCA）（之後在第十五章裡都會討論到這兩種工具）。他所擁有的資料包括了地圖、各個國家的經濟與人口統計資料、以及數百份針對個別國家所進行的歷史研究。他得到的結論是，那些地位處於最核心、而且最為繁榮的西班牙殖民地，都是位於天然資源極為豐富的地區（以便於帝國榨取資源並運往歐洲）、而且當地也存在著大量的本土人口（帝國可以強迫他們進行勞動工作）。在這些殖民地裡，當地的菁英階層逐漸成形，且建立了僵化的種族階層體制。菁英階層將經濟力量與政治力量集中於自身之上，並把社會上其他階層的廣大人口排除在外。此種體制一直持續到十九世紀，在十九世紀時，出現了新的政治事件、貿易模式、與經濟形勢。在 1700 年到 1850 年這段期間內，心胸開放又能夠接受新思想的菁英人士，未能在那些處於核心地位並相當繁榮的殖民地內成功地引起改變。對比之下，那些位於南美洲西班牙帝國邊緣地帶的殖民地，則比較不會受到僵化的體制所阻礙。在國家地位「大逆轉」（"great reversal"）的時期裡，那些能夠改革創新、並且與時俱進的菁英人士就因此趁勢崛起。在這個歷史的「轉捩點」（"turning point"）過後，在 19 世紀晚期，一些國家在社會經濟發展方面，取得了大幅領先的優勢。這些國家建立了能夠幫助推動它們自己向前發展的政治經濟體制以及機構，換句話說，這些國家「緊緊抓住」（"locked into"）一個特定的方向或是路徑，而這個方向或路徑，則為這些國家帶來不斷增加的回報和利益。

馬哈尼（Mahoney, 2003: 53）認為，「若是僅僅利用各國目前現有的特質去解釋各國之間差異，是不恰當的作法」。換句話說，若是利用那些只有在某一個時間點所得到的資料，就去進行橫斷性研究並試圖解釋各國之間的差異，那麼根本無法真正地捕捉到重要的長期動態過程。在解釋上述此種差異之時，較好的作法應該是要去思考久遠的歷史事件所造成的衝擊，並採取一個長期的時間觀點。

結論

本章提供了社會研究面向的概觀，你也已經看到，一項研究可用數種不同的方式來加以分類（例如，依照研究的目的、研究的技術等等來加以分類），而且這些研究面向彼此之間，也有相互重疊的部分（請參閱表 2.1）。我們可以把研究面向視爲是引導我們在認識社會研究這個知識領域時的「道路圖」。

在下一章裡，我們將把焦點轉向介紹社會理論。其實，在第一章裡早已稍微討論過社會理論，在本章中也再度提到相關的議題。而於第三章裡，你將學習到理論與研究方法之間要如何結合運用，也將會學習到數種不同類型的理論。

問題回顧

1. 什麼時候適合使用探索性研究？而探索性研究又能完成什麼樣的目標？
2. 描述性研究會產生何種類型的研究結果？
3. 何謂解釋性研究？解釋性研究的主要目的爲何？
4. 基礎研究和應用研究之間最主要的差異爲何？
5. 什麼人比較有可能去從事基礎研究？基礎研究所得到的結果比較有可能出現在什麼地方？
6. 請說明三種應用研究之間的差異。
7. 時間序列研究、同樣本多次研究、以及世代研究之間，有什麼差異？
8. 成本效益分析可能會有哪些潛在的問題？
9. 何謂需求評估？在進行需求評估時，可能會遇上什麼困難？
10. 請說明質化研究與量化研究間的差異。

注解

1. Abbott (2004:40–79) offers a more comprehensive and complex organization of methods.
2. See Finsterbusch and Motz (1980), Freeman (1983), Lazarsfeld and Reitz (1975), Olsen and Micklin (1981), and Rubin (1983) on applied research. Whyte (1986) critiques social research that is not applied. McGrath and colleagues (1982) discuss judgment calls relevant in applied research.
3. See Crespi (1987) and Dutka (1982) on the use of survey research in legal proceedings.
4. See Turner and Turner (1991:181).
5. For a brief introduction to evaluation research, see Adams and Schvaneveldt (1985:315–328), Finsterbusch and Motz (1980:119–158), and Smith and Glass (1987). A more complete discussion can be found in Burnstein and associates (1985), Freeman (1992), Rossi (1982), Rossi and Freeman (1985), Saxe and Fine (1981), and Weiss (1972).
6. See Oliker (1994).
7. Smith and Glass (1987:41–49) discuss PPBS and related evaluation research.
8. See Reinharz (1992:252).
9. See Cancian and Armstead (1992), Reason (1994), and Whyte (1989).
10. On participatory action research, see Cassell and Johnson (2006), Kemmis and McTaggart (2003), and Stoecker (1999).
11. Social impact research is discussed in Chadwick and associates (1984:313–342), Finsterbusch and Motz (1980:75–118), and Finsterbusch and Wolf (1981). Also see Rossi and colleagues (1982) and Wright and Rossi (1981) on "natural hazards" and social science.
12. See Becker and Vanclay (2003) and *Guidelines and Principles For Social Impact Assessment* by The Interorganizational Committee on Guidelines and Principles for Social Impact Assessment (1994). http://www.nmfs.noaa.gov/sfa/social_impact_guide.htm
13. See Burawoy and colleagues (2004).
14. Hammersley (2000) makes this generalizalist versus narrow practitioner distinction.
15. Babbie (1998), Bailey (1987:38–39), and Churchill (1983:56–77) also discuss explanatory, exploratory, and descriptive research.
16. See Guy and colleagues (1987:54–55) for discussion.
17. For discussions of case-study research, see George and Bennett (2005), Gerring (2007), Miller (1992), Mitchell (1984), Ragin (1992a, 1992b), Stake (1994), Vaughan (1992), Walton (1992b), and Yin (1988).
18. (see George and Bennett 2005:19–22; Gerring 2007; McKeown 2004; Ragin 2008:71–84; Snow and Trom 2002).
19. See Mitchell (1984) and Stake (1994).

第3章

理論與研究

什麼是理論？
社會理論與意識型態
社會理論的組成部分
理論的面向
動態的組合
結論

> 理論的主要功能之一，就是藉由概念的輔助，將各種經驗排序定位。而社會現象研究者所面對的數量龐大的各種「事實」中，理論也能挑選出相關的面向及資料。
> ——Lewis Coser, "*The Uses of Classical Sociological Theory*," p. 170

在美國整體民眾當中，經常吸煙的人口所佔的比例已經下降了。我們猜想，造成此種吸煙人口比例下降情況的可能原因，是因為有許多的公共運動一直在警告人們有關吸煙對於健康的危害。我們也發現，與教育程度較低、收入較低的人們相比之下，教育程度較高、收入較高的人們吸煙的可能性較低。根據社會資源理論的解釋，這是因為與教育程度較低、而且收入較低的人們相比之下，受過教育、而且收入較高的人們，比較常閱讀、眼光看得較遠、而且也擁有比較多的資源以針對生活方式作出調整。然而，吸煙不只是一個和健康有關的議題而已。它也可能成為一個象徵性的時尚宣言，以及和文化品味有關的生活方式議題。同樣地，教育程度與收入水準所暗示的，也不只是知識與資源而已，同時也意味著歸屬於不同的階級文化（也就是來自於不同社會階級的人們，在文化方面分辨彼此的方式）。文化品味理論認為，採取中上階級生活方式的人們不會去吸煙，因為在中上階級裡，吸煙在文化方面是比較不受歡迎的。對比之下，採取勞動階級生活方式的人們，就比較有可能會染上吸煙的習慣，部分是因為這是其階級文化的特色之一。階級文化的其他層面還包括了音樂品味。受過高等教育、收入水準較高的人們，傾向於偏好古典音樂，而教育程度較低、收入也較低的人們則偏好流行音樂以及重金屬音樂。就邏輯上而言，文化品味的理論也暗示了，由於不同階級的生活方式有所差異，因此，音樂方面的品味也和吸煙與否有所關聯。這正是潘培爾（Pampel, 2006）所發現的真實狀況。但是，其研究結果卻更加耐人尋味。不論是教育程度較高、收入較高的人們，或是教育程度較低、收入較低的人們，都同樣地喜歡爵士樂。而爵士樂的次文化，長久以來都包括了吸煙行為。與文化品味理論一致的是，潘培爾發現到，在相同的社會階級裡，那些愛好爵士樂的人比起那些不喜歡爵士樂的人，更有可能吸煙。

在前面這個部分所提出的、關於一個人的音樂品味，以及他或她的吸煙行爲之間的關聯性，或許會令人感到相當地意外。但是，此種情況卻能清楚地顯示出理論的力量，以及理論對於研究所具有的影響力。理論能夠幫助我們瞭解社會生活的複雜性，因爲理論不僅僅解釋了人們做某些事情的真正原因，同時也提供我們更加深入的洞見、並且爲我們指出進一步探討的方向。在前面提到的例子裡，文化品味的理論引導潘培爾去提出新的問題、並且重新檢視吸煙行爲，而從這個例子裡，我們可以清楚地看到，理論可以提供我們一些相關的概念，讓我們能夠以新的方式去進一步探討並思考社會世界。同時，這個例子也顯示出，不同的理論在解釋事件時，往往會有不同的、甚至是完全相反的方式。

很多初次接觸社會研究的學生通常會對理論感到相當恐懼，因爲他們覺得理論就像是一座迷宮，裡面充滿了晦澀難解的術語，以及許多與日常生活毫無關聯的抽象概念。我希望你能慢慢地改變此種看法，並也能夠漸漸地瞭解到，在理解你周遭的社會世界時，理論不僅是相當有用的，而且也是極爲重要的。理論具備了許多的功能：理論能夠澄清你的思緒、擴展你的理解範圍、使相關的討論變得更爲深入、並且能夠使分析結果變得更爲豐富。在促進知識發展、與組織我們進行研究的方法等等方面，理論扮演了一個極度重要的角色。在本章中將會簡單地介紹社會理論的基本觀念，而在後面的章節裡，也將會包含更多與理論相關的內容。

我的學生們經常會與我分享他們對於社會理論的焦慮以及困惑。造成他們困惑的其中一個來源，是幾乎沒有人瞭解到底社會理論的真正內容與意涵是什麼。而使得情況更加雪上加霜的是，*理論*（*theory*）一詞不但具有多重的意義，還具有數種不同的形式。即使是專業人士也會對理論的意義爭論不休，並爲理論賦予下列數種不同的意義。

1. *理論*就是一整套在邏輯上互相關聯的一般性命題，而其功能則在兩個或者是更多個變項之間建立關聯性。
2. *理論*就是針對一個特定社會現象所提出的解釋，並且會明確指出一組彼此之間具有因果關係的因素或是情境。
3. *理論*提供一個對社會現象真正有意義的洞見，其方法是提出一個具有啓發性的詮釋，並告訴我們「社會現象的本質」。
4. *理論*就是某一位知名的社會思想家真正所指的事物。
5. *理論*就是一個整體的世界觀，或是一種觀看、詮釋、以及瞭解世界上各種事件的方式。
6. *理論*就是一種以政治—道德觀點爲基礎所提出的批評，換句話說，理論代表了一整組的信念與價值觀，並且依據這些信念與價值觀，批評對手的立場和論點。
7. *理論*就是一種與關鍵問題或議題有關的哲學評論，而這些關鍵問題或是議題所指的是，我們如何發展關於社會世界知識的一些核心議題。

資料來源：Gabriel Abend, The Meaning of "Theory," *Sociological Theory*, Volume 26 Issue 2, May 28, 2008, pages 173-199.

理論之所以會造成困惑的另一個可能原因，就是我們大多數人在日常生活中經常會遇到並使用一些外表相似、但卻根本不是理論的解釋。理論確實是一種解釋，但卻不是解釋的唯一來源。各種解釋可以提供我們概念，讓我們瞭解各種不同的事物，並告訴我們何者才是重要的、爲什麼人們會做出某些事、與世界上各種事件是如何連結在一起。我們可以在和朋友的對話中、電視節目裡、政治人物和商業領袖的口中、報紙的報導中、甚至在電影裡面，聽到各式各樣的解釋。以上這些都是解釋，卻都不能夠算是由社會理論所提出的解釋。

當人們遇到他們不熟悉的抽象概念時，許多人會變得焦慮不安。我們都知道，這個世界上存在著一些我們可以觸摸或是看見（例如像是拿著這本書）的具體事件與實體物體，同時也有一些存在於我們腦海之中的抽象概念（例如像是自由跟正義的意義等等）。當我們遇到許多不熟悉的抽象概念、以及那些沒有明確定義的概念時，不論是否刻意，我們很快會感受到焦慮和挫折感。社會理論是由相互關聯的抽象概念所組成的，而其中有些概念與可觀察的世界，或大家所熟悉的概念之間，僅有鬆散的聯結關係。在我們真正地瞭解一個理論的概念，及看到這些概念間的聯結關係之前，討論這些抽象概念會使我們覺得不自在，這件事也就不足爲奇了。

造成困惑的最後一個來源，則與進行研究有直接的關聯。身爲研究者，我們之間有些人無法使理論變得清晰易懂且容易釐清。雖然這需要多花一點時間和努力，但是當一項研究中的理論既清楚又明白易懂時，我們就可以更加容易地去評估該項研究的優缺點。

什麼是理論？

在第一章中曾經討論過，所謂的*社會理論*就是一個由互相關聯的概念所形成的系統。社會理論濃縮、並且組織了關於社會世界的知識。我們也可以把社會理論當作是以系統性的方式來「說故事」，以解釋或是說明社會世界某些面向的運作方式，及其背後的原因。

許多關於社會理論的課程，會強調社會思想發展的歷史、或者是教導我們那些偉大的思想家所提出的學說或理論。古典社會理論家（例如像是涂爾幹、馬克斯、米爾斯、屠尼斯、韋伯等等）提出了許多創新的想法。這些理論家們發展出具有高度獨創性與廣泛的理論，爲之後許多世代的社會學家奠定了基礎，並徹徹底底地改變了我們看待及理解周遭社會世界的方式。至今我們仍然持續研讀這些大師的著作，因爲他們的著作提供了許多具創意的、且互相關聯的想法。這些大師提出許多深具洞見的想法、從根本上改變了我們看待社會世界的方式，而此種真正的天才，其實相當罕見。雖然他們的貢獻具有重大的價值，但理論並不僅侷限於這些古典社會理論家的著作。除此之外，理論也不僅是我們從一些近代居於領導地位的社會理論家〔例如像是亞歷山大（Jeffrey Alexander）、布爾迪厄（Pierre Bourdieu）、柯爾曼（James Coleman）、傅柯（Michel Foucault）、紀登斯（Anthony Giddens）、以及

高夫曼（Erving Goffman）等等〕所學習的知識。雖然許多新的想法和理論，是由理論家所提出或構思的，但是我們都可以使用理論。

理論並不是處於靜止不動的狀態。我們經常會修正舊理論，並發展出新理論。理論也會有各種不同的形式和內容。其中有些理論是範圍廣泛的思想系統，而有些則只是針對某個特定議題所提出的，較爲狹隘且具體的解釋。就理論的核心價值而言，我們會使用社會理論來使我們的思緒更有條理、更有系統，並使我們對於事物的瞭解程度更爲深入與廣泛。除此之外，由於理論能夠幫助我們組織知識，因此理論也變成一種方法，能幫助我們以更爲有效的方式，彼此進行溝通。

我們很可能都曾在日常生活中遇到理論，只是這些理論很少是明白易懂的、清楚地讓我們知道它們其實就是理論。舉例而言，談論社會議題的報紙文章或是電視報導，通常會依賴一些隱晦不明的社會理論，或是部分的理論。雖然這些理論可能膚淺且不完整，但是這些理論仍然可能包含各種假定、互相關聯的概念、以及各種解釋。例如，某篇新聞報導所討論的內容，可能是對於像同性婚姻合法化這類的議題，一般民眾的態度是支持抑或是反對。該篇報導或許會提供某種類型的社會理論，以解釋爲什麼讓同性婚姻合法化可能會引起爭議。該篇報導可能會提到，反對的意見是來自於宗教組織、與那些害怕破壞傳統社會價值觀的民眾。這樣的理論包含以下數種假定：宗教組織能夠影響新的法律、有些民眾會極力保存過往或是現存的社會規範、某些宗教組織或某些民眾對於和婚姻有關的法律抱持強烈的意見與觀點。於是，該篇報導可能會提出如下所述的解釋：來自於某些組織或是擁有強烈宗教信仰民眾的政治反對意見，能夠阻止經由選舉產生的政府官員去通過某些法案。然而，在日常生活當中，大眾媒體並不是唯一的理論來源。當政治領袖在討論公共議題時，他們也經常會引用社會理論。例如，當某一位政治人物說教育程度不足將會導致貧窮時，他就是在發表某種類型的理論。與我們在社會科學研究中所使用的理論相較之下，這些隱晦不明的、部分的理論比較不具有系統性、結構也較爲鬆散、而且也比較難以用實徵資料來加以評估驗證。

社會科學理論通常比一般外行者所提出的理論還要更爲複雜且抽象。然而，良好理論的原則之一乃是**簡要**（parsimony），而此項原則對我們是很有幫助的。所謂的簡要原則，就是指理論應該越簡單越好，而且一項良好的理論，應該具有最低限度的複雜性。良好的理論沒有多餘或是額外的要素。在兩個具有相同說服力的理論中，較爲簡單的那個理論，就是比較好的理論。

幾乎所有的研究都會在研究過程中使用到一些理論。真正的問題不在於我們*是否*在一項研究中使用了理論，而是我們*如何*使用理論、或我們使用的是何種*類型*的理論。與應用研究或是描述性研究相較之下，在基礎研究或解釋性研究中，理論具有較爲重要，並顯著的地位。當我們意識到應該如何將理論與研究緊密地結合在一起時，我們所進行的研究將會具有較良好的研究設計，及較爲強大的說服力或解釋力。當我們思考於一項研究中所做的事情時，理論也能

簡要 簡要所指的概念就是越簡單越好。當所有其他的條件都完全相等時，一個能夠用較低的複雜性去解釋事物的理論，就是較佳的理論。

夠幫助思緒變得更加敏銳。如果我們能夠清楚且明白地說明，在我們的研究中所使用的理論，其他人就能夠更加容易地閱讀，並且理解我們的研究。在判斷某項研究過於薄弱時的一個指標，即爲該項研究中所使用的理論是不清楚的、不完整的、或者是鬆散而無結構的。

社會理論與意識型態

許多人都會把社會科學理論和某種社會政治意識型態、或者是某種道德宗教的教條混爲一談，而這種情況其實是可以理解的。在日常生活中，我們會遇到許多的教條以及意識型態，而且這些教條或是意識型態和社會理論之間存在著一些相同的特質。在美國所發生的演化論與「神創論」（“creationism”）之間的爭辯，正好可以用來說明許多外行者對於科學理論的誤解。民意調查結果顯示出，有超過半數以上的美國民眾認爲在學校裡應該同時教導演化論及神創論，因爲這些民眾認爲這兩者都是「理論」（Pew Forum, 2005）。然而，演化論之所以有資格被稱作是一個科學理論，是因爲演化論具有邏輯連貫性、開放性、與其他科學知識的整合性、及其經過了實徵證據的檢驗。至於神創論（或者是將此種概念重新包裝而成爲所謂的「智慧設計」）則不具備成爲理論的資格。相反地，神創論僅僅是一種以道德宗教教條爲基礎的意識型態的一部分。

道德宗教教條是以信仰爲基礎的信念系統。此種教條所仰賴的是神聖的宗教教義或是經典，信徒們會接受這些教義或經典、將之視爲絕對的真理，且大多數都不會加以質疑。然而，此種教條只是一種意識型態的類型、或是一種非科學的信念系統。在許多公共議題的爭辯中，都會牽涉到意識型態，有可能是某種道德宗教的意識型態、或者是某種社會政治的意識型態，也可能兩者皆是。這些教條經常出現於大眾媒體之中（來自於各種不同的政治道德觀點的擁護者）、在企業或是利益團體的宣傳活動之中、由政治人物爲公共政策或是新法令所作出的辯護理由之中。

由於社會理論和意識型態之間具有許多共同的特質，因此許多人容易把意識型態誤認爲是社會科學理論。這兩者都告訴我們爲什麼事物會是現在這個樣子：例如，爲什麼會發生犯罪案件？爲什麼有些人比其他人更加貧窮？爲什麼某些地方的離婚率比較高？等等問題。這兩者也都包含了關於人類和社會世界的基本特質的假定，像是告訴我們什麼事物是重要的、而什麼事物又是不重要的。除此之外，這兩者也提出各種想法或概念的系統，並將這些概念連結起來。

科學社群承認，理論對於科學研究極爲重要。在澄清思緒、擴展並深化我們對於事物的瞭解程度、以及逐漸建立知識體系方面，良好的理論是不可或缺的。然而，科學社群認爲意識型態完全不同於理論，是一種非科學的世界觀。意識型態或許適合處理非科學的問題，但在評估有關事實的各種宣稱、或是在建立許多議題，關於社會科學問題的知識等等方面，意識型態不是一種具有正當性的方式。對於許多科學社群裡的人而言，意識型態會造成許多混亂和困惑的情況，而此種情況正好違反

了科學的基本原則。當社會科學駁斥了意識型態信念系統的某些層面時，意識型態的捍衛者有時會因此而對社會科學產生敵意。

作爲一種「近似於」（"almost"）理論的準理論，**意識型態**（ideology）欠缺了真正的科學理論所應該具備的批判性特徵。我們可以用以下的七種方式來區別意識型態及理論（請同時參閱重點摘要 3.1）：

1. *答案的確定性*（*certainty of answers*） 意識型態會讓許多人感到比較自在，因爲意識型態提供了絕對的真理以與確定的答案。意識型態能夠帶給人們一種充滿信心的感受及安全感。對比之下，社會科學理論僅能提供暫時性的答案，並承認不確定性的存在。對於科學理論所具有的持續不確定性、猶豫不決的狀態、以及暫時性，許多人會覺得很不自在。社會科學理論要求我們要能夠忍受模稜兩可的情況、不斷地提出問題、並能夠與持續的懷疑狀態和平共處。

2. *知識系統的類型不同*（*type of knowledge system differs*） 意識型態所提供的是一個幾乎沒有任何變化的封閉知識系統。這是因爲意識型態宣稱擁有所有問題的答案，因此不必作出任何改變。對比之下，科學是一種開放式的知識系統，總是在成長及變化當中。科學所提供的答案是不完整的，而且隨著我們獲得新的證據與知識，這些答案也會不斷地進行修正。我們會不斷地修改並重新思考理論。理論則是持續地演化、成長、朝向更高的層次發展，此種過程有時很緩慢、有時則很迅速，有時候會以直接的方式進行、而有時可能暫時走上回頭路，或轉換方向後，才會出現較大的進展。

3. *假定的類型不同*（*type of assumptions differs*） 意識型態和社會科學理論都包含了各種不同的假定。意識型態所包含的假定通常是固定不變的、沒有彈性的、而且不能加以質疑的。大多數意識型態的假定都是來自於下列三個來源之一：宗教信仰或是信念（例如，基督教或回教的某種特殊形式）、一種以價值觀爲基礎的立場（例如，自由主義者、社會主義者、或是法西斯主義者）、抑或某種特定社會階層的觀點（例如，既富裕又握有權力的菁英人士、無家可歸又身無分文的人們）。當意識型態來自於特定的社會階層時，意識型態就會去保障並促進屬於該社會階層的利益（例如，擁有萬貫家財的投資者、身無長物的貧苦人們）。對比之下，社會科學理論的假定則來自於科學社群內的開放辯論及討論，而且這些假定也會隨著時間而不斷演變。我們將在稍後的章節裡（在第四章及第十六章），詳細地審視關於價值中立與客觀性的議題。到目前爲止，我們已經可以看得出來，社會科學理論與意識型態最大的差異之一，在於社會科學理論試圖在假定方面保持中立客觀，或者，就算不能完全中立客觀，也會明白陳述其假定爲何，並且抱持著開放態度。

如同在此說明的，意識型態通常會反映出某個社會階層所抱持的世界觀。那麼，研究者及科學家

意識型態 一種非科學的準理論（quasi-theory），通常植基於政治價值觀，或者是對於各種假定、概念、以及不同概念與解釋之間的關係所具有的信念。意識型態是一個封閉的系統，該系統抗拒改變、無法以實徵資料直接加以駁斥、且會提出規範性的宣稱。

重點摘要 3.1　社會理論與意識型態的比較

相異性	意識形態	理論
答案的確定性	絕對、確定的答案，幾乎不容質疑	暫時性的、有條件的答案，而且答案是不完整的、開放式的
知識的類型	封閉的、固定的信念系統	開放的、不斷擴展的信念系統
假定的類型	晦澀不明的假定，以信仰、道德信念、或是社會立場為其基礎	清楚明白的、不斷變化的假定，以開放的、有資訊根據的辯論，以及理性的討論為基礎
規範性陳述的使用方式	混合了描述性宣稱、解釋、以及規範性的陳述	將描述性宣稱、解釋、及規範性陳述區隔開來
實徵證據	選擇性地使用證據，避免對其宣稱進行直接驗證，對於對立的證據採取抗拒、拒絕、或是忽略的態度	將所有證據皆列入考慮範圍，尋求對其宣稱的重覆驗證，根據新的證據而作出改變
邏輯連貫性	充滿矛盾及邏輯上的繆誤	最高層次的連貫性及一致性，避免邏輯上的繆誤
透明程度	避免透明性	鼓勵透明性

們所處的社會階層是否也可能對社會理論造成影響呢？有些人認爲，研究者必須保持超然公正的立場，與其理論中所有的特定社會利益保持距離；有些人則認爲，只要研究者願意開誠佈公、坦誠以告，那麼研究者可以在研究過程中的某些領域裡，擁有自己的社會政治觀點；也有人認爲，研究者在社會中佔據了一個獨特的「關係」（"relational"）位置（Mannheim, 1936）。所謂的*關係位置*（*relational position*），意思是指社會研究者來自於社會中各種不同的領域、高度意識到所有社會領域的完整範圍、並自覺地去反思他們所處的獨特社會地位。

4. *使用規範性陳述的方式不同*（*use of normative statements differs*）　意識型態包含許多規範性的假定、陳述、以及概念，而這些規範性的假定、陳述、或是概念則會進一步地促進規範性的立場或位置。*規範性陳述*（*normative statement*）所指的是一種包含了「應然面」（"what ought to be"）的陳述，此種陳述告訴我們什麼是應該的、恰當的、道德的、以及正確的，並告訴我們什麼是不應該的、不恰當的、不道德的、或是錯誤的。與社會理論相同的是，意識型態會告訴我們「什麼」及「爲什麼」，還會更進一步地告訴我們「應該是什麼」（請參閱資訊箱 3.1 關於離婚的解釋）。意識型態刻意模糊了下列事物之間的界線：描述性的、以事實爲基礎的聲明（例如，*這就是真正發生過的事件、或者是人們的生活方式*）；解釋說明（例如，*這就是發生這件事的原因、人們以此種方式生活的原因*）；以及規範性立場（例如，*這件事應該要發生、或者這就是人們應該有的生活方式*）。

對比之下，社會科學理論幾乎不會去提倡某種特定的規範性宣稱。社會科學理

資訊箱 3.1 關於離婚的解釋

意識型態對於離婚的可能解釋

在過去的三十年之間，美國社會已經經歷了道德與社會方面的分裂崩解。在過去，家庭的力量相當強大，媽媽們並未離家工作，而是花非常多的時間照顧她們的孩子與丈夫。由於宗教與道德方面的教誨，家庭的力量相當強大，離婚也相當罕見。然而，在最近幾年以來，道德敗壞的情況越來越普遍。民眾越來越不尊重宗教及道德方面的權威。負面行為、政府的政策、與大眾傳播媒體都削弱了家庭的力量，也導致離婚率逐漸上升。

評估 上述的解釋使用了許多概念，像是道德與社會方面的分裂崩解、家庭的力量、離婚、媽媽陪伴她們的孩子與丈夫的時間、道德敗壞、喪失敬意、以及媒體訊息等等。這些概念缺乏精確的意義和測量方式，且這些概念發生的時間點也充滿了不確定性。除此之外，這些概念相當模糊，並且是高度評價性的（例如，像是敗壞、分裂崩解、壞的等等）。要去驗證此種解釋將是件不容易的事，在上述解釋中的時間框架相當長，因此在這段期間，或許也會出現一些其他造成影響的因素。

社會理論的例子

一個家庭是否維持完整（換句話說，兩個結婚的成人並未離婚）、或是否強大（也就是表達對家庭成員的關愛，花許多時間相聚、花較多的時間養育兒童、表現出正向的社會互動模式等等），乃是取決於資源的多寡、以及社會和情緒方面的壓力程度。這裡所指的資源包括了物質因素（收入、教育、住宅）、社會因素（朋友與其他的家人、對於社區組織的參與）、認知因素（例如，所接受的學校教育、知識、對於時事的關注等等）、以及心理因素（正面的自我意象、成熟度、對於他人的尊重）等等。至於壓力則包括了對於未來的不確定感，及生活情境的不穩定（例如，家庭中有人就業狀況不穩定、健康情況不佳或逐漸惡化、曾經是犯罪案件的受害者、或在情緒方面有不穩定狀況）。擁有充足資源和壓力程度較低的家庭，會比那些資源不足且壓力過高的家庭，還要更為強大，而且力量較為強大的家庭，比力量薄弱的家庭更可能維持完整。

評估 上述這個解釋使用了四個概念：資源（三種類型）、壓力、家庭力量、以及維持完整。這個解釋也提出了這些概念的定義，及我們如何測量每個概念的方式。各個概念之間的關係相當明白易懂，也能夠用實徵方式來加以驗證。

論會提出描述性的陳述（「這就是世界運作的方式」）及解釋說明。在社會理論當中，確實有一些個別、單獨的規範性立場。我們可以把一個理論的描述（例如，有些人們正在挨餓）或是解釋（例如，有些人刻意囤積糧食以哄抬價格，而此種情況導致其他人因此挨餓），和一個或多個規範性立場（例如，沒有人應該挨餓，讓弱者挨餓可使人類變得更加壯大等等）連結起來。雖然在一個理論中不一定要同時具備描述、解釋、以及規範性立場，但就算某個理論同時具備了三者，三者之間的關係也不是嚴格或固定不變的。

簡而言之，在社會理論裡，規範性道德立場是與描述性陳述或解釋截然劃分開

來的，相對地，在意識型態裡，規範性立場是描述性陳述和解釋不可或缺的部分。換句話說，我們不可能將規範性立場從意識型態中移除。

5. *使用實徵證據的方式不同*（*use of empirical evidence differs*）　科學理論與意識型態之間的一個關鍵性區別，則與實徵證據有關。某種意識型態的支持者會選擇性地呈現以及詮釋證據，以保護某個意識型態的信念。通常這些支持者們會強調個人經驗、嚴格服從某種核心價值信念、或把宗教信仰當作是一種凌駕於詳細的實徵觀察結果的最終證據。作爲一種已經擁有「所有答案」的封閉信仰系統，意識型態會抗拒或否認任何與其矛盾的證據。即使是面對著具有壓倒性優勢的負面或對立的證據，意識型態也不會因而屈服，或者作出改變。從意識型態的世界觀來看，某種意識型態的信徒會選擇性地重新詮釋負面證據、把負面證據當作例外情況、乾脆斷言這些負面證據與意識型態所作出的宣稱之間，根本毫無關聯可言。某種意識型態的信徒總是能夠找到一些方法來駁斥對立的證據，他們的立場就是「不要用一些事實來混淆我的想法，我知道我自己是正確的」。事實上，當面對著負面證據時，某種意識型態的信徒會以恐懼與充滿敵意的方式，來回應那些不認同他們的人。

社會理論是信念和解釋的開放系統，並歡迎所有各種類型的證據。由於社會科學理論樂於接受持續的辯論、修正、或者改變，因此社會科學理論總是在持續不斷地演變發展當中。來自於不同研究的證據可能會支持、擴展、反駁、或者是修正一個理論。我們經常會利用實徵證據——*所有*相關的證據——來檢驗理論，不論是支持性的證據、對立的證據皆然。我們用證據來評估理論，而不是用證據來爲理論進行辯護。我們永遠無法事先知道，某項證據是否會支持某個理論。任何一項研究都可能發現一些相關證據，而且這些證據可能會顯示某個理論有其缺陷、需要進行修正。

在社會科學中，我們假定隨著時間經過，社會研究會產生累積性的知識與證據，也就是說，這些知識及證據會隨著時間而逐漸增加。由於研究及理論是逐漸累積的，因此，如果我們遇到任何負面證據，我們不會不加思索地就將一個理論棄置不顧。如果在多年的研究及許多研究成果之後，我們已經累積廣泛的實徵證據來支持一個理論，那麼，當新的負面證據出現時，我們或許會緩慢地校正該理論。話雖如此，任何負面證據仍會提出關於某個理論的一些問題。倘若新的證據一再與某個理論有所違背，我們就不得不去修正該理論，或甚至用新的理論加以取代。

6. *對於邏輯連貫性的要求不同*（*demand for logical consistency differs*）　意識型態在邏輯方面通常會有自相矛盾的情況，而且許多意識型態所仰賴的是循環式的推論。循環式推論有許多種不同的形式，其中有些是邏輯上的繆誤、或在真正的邏輯推論中所犯的錯誤。循環式推論只不過是一直以稍微不同的方式重覆同樣的陳述，或以較爲強烈的詞彙來當作「證據」、支持自己論點的推論方式。當某個意識型態被發現有邏輯上的矛盾或繆誤時，典型的回應方式之一，就是拒絕接受該項發現、或者是想盡辦法加以掩飾。對比下，身爲社會科學家，我們堅持理論在邏輯上必須具備連貫性。我們會時常試著去找出並排除所有邏輯上的繆誤。假如我們發現某項繆誤或是矛盾，我們會去修正該理論，或用一個不同的、沒有繆誤或者矛盾的理論加以取代。

7. *透明程度不同（transparency differs）* 意識型態和理論之間的差異，也會影響到我們進行研究的方式。在社會科學研究中，我們會察覺到一個理論的假定、概念、及關係，並且會在研究中明白地顯示。換句話說，我們所用的理論、該理論在研究中的地位，都是非常公開的。身爲科學家，我們不會隱藏任何資訊。社會科學研究不但會公開所使用的理論，也歡迎他人挑戰並進行開放的辯論。比較之下，意識型態經常會包含一些晦澀難解的、或是難以明確界定的特徵。意識型態經常會包含一些充滿神秘或秘密的領域，此外，意識型態所追求的是服從與順從，而不是嚴肅的挑戰或辯論。

社會理論的組成部分

假定

所有的理論都有其內在的**假定**（assumptions），也就是關於事物本質的陳述，而且這些事物本質是我們無法直接觀察、或無法以實徵方式加以評估的。假定是必要的起始點。在社會科學中，我們會作出關於人類本性（例如，人們在本質上傾向於相互競爭的、或善良且樂於合作的）、社會事實（例如，這件事是顯而易見的，或者包含了一些隱藏的要素）、或是特定現象與議題的假定。

假定的類型之一，就是*背景假定*（*background assumption*），而且如果我們想要繼續進行探討或研究，一定要有背景假定的存在。關於一些像是種族偏見（racial prejudice）之類的複雜社會議題理論，必須依賴數種隱而不顯的背景假定。其中有些與種族偏見有關的背景假定如下：一個社會裡的人群可以辨認出種族類別與種族差異；人們會以某個人在一個種族團體裡的成員身份來看待個體之間的差異；人們會將某些特徵、動機、及特質連結到某個種族團體的成員身上；而且人們也會去評估某個種族團體成員的特徵、動機、及特質的好壞。上述這些都是背景假定，因爲如果人們不會去區別不同的「種族」（也就是說，把某些身體外表的特徵與血統連結在一起）、或是從未把一些特徵連結到某個種族團體的成員身上等等，那麼*種族偏見*這個概念就會變得毫無用處可言。因此，這個概念以及解釋這個概念的理論，就必須建立在數種背景假定之上。

除了背景假定之外，我們或許也會有「*引導*」（"*tractable*"）假定（換句話說，此種假定具有引導的力量，並且允許我們更進一步地去發展某個論點）。引導假定可能成立、也可能不成立。如果我們想要研究種族偏見，我們或許會假定人們擁有不同程度的種族偏見，而有些人也許根本就沒有種族偏見。我們或許還會假定，某個人的種族偏見只會應用於其他種族團體上，而不會用於他們自己的種族團體。除此之外，我們也可能假定某個人所抱持的種族偏見是持續存在的，也就是種族偏見不會突然間立刻出現或是消失。

假定 在一個理論中尚未經過驗證的起始點或是信念，而在建立一個理論性的解釋時，假定是不可或缺的。

概念

*概念*是構築理論的基石。[1] 而所謂的**理論概念**（theoretical concept）就是一個我們可以用符號或文字表達的想法。在自然科學及數學裡，我們經常會用符號的形式來表達理論概念，像是希臘字母（例如，π 或是 Σ）或者是公式（例如，$s = d/t$；其中 s = speed 速度，d = distance 距離，t = time 時間）等等。對比之下，大多數的社會科學家都是用文字來表達他們的概念。雖然數學和自然科學中所使用的奇特符號使得許多人感到緊張，但是將日常生活用語以專門化的方式應用於社會科學的概念中，同樣也會產生混淆的情況。用文字來表達概念以及用符號來表達概念，這兩者之間的區別不應被過度誇大。畢竟，文字本身也是符號，只不過我們是用語言的方式來學習這些符號。

讓我們來看看一個簡單的概念範例，而且你早已相當熟悉這個概念：*高度*（*height*）。你可以說出高度這個詞，或是用 h 這個符號來代表高度。在這個文字裡的字母組合或者是這個字的發音，代表了在你頭腦中的一個想法。像是在中文裡的「*高度*」、在法語裡的「*hauteur*」、在德語裡的「*höhe*」、以及在西班牙語裡的「*altura*」等字都代表了完全相同的概念。就某種意義而言，語言就像是一種以聲音或是書面文字來代表想法的協定，而人們則會在其生活中的某個時間點學習這些聲音或是文字。因此，學習概念及理論的過程，其實就像是學習語言。[2]

概念存在於社會科學理論之外，換言之，概念在日常生活中無所不在，而且我們時時刻刻都會用到概念。前面所提到的「*高度*」就是來自於日常生活的一個簡單概念，但是這個概念所指的究竟是什麼呢？我們可能會發現到，要去使用「高度」這個概念是相當容易的，但是若要加以定義或是描述這個概念，卻十分困難。這種情況其實相當常見，也就是說，我們可能會使用許多概念，但是卻會發現到，若是要徹底想清楚這些概念的完整意義、並賦予這些概念明確而良好的定義，是一件相當困難的事。「高度」是一個關於物理關係的抽象想法，而且，作爲實質物體的特徵之一，高度所指的是某個物體從頂點到底部的距離。一般而言，當我們在定義概念時，通常會使用到其他的概念，並利用一些例子來加以說明。例如，我們可以利用頂部、底部、及距離這三個概念來定義高度，也可以利用物質世界裡的許多例子，來說明高度這個概念。

高度是一個所有人都非常熟悉的概念，所有的人、建築物、樹木、山丘、以及書本等等，都有高度。我們可以去測量任何物體或者是生物的高度，或是去比較不同物體或生物的高度。高度爲零的情況雖然很罕見，但仍可能存在。而且高度也可能隨著時間增加或減少。就像許多其他詞彙一樣，我們也會用數種不同的方式來使用「高度」這個詞與這個概念。我們會在許多不同的表達方式中，使用「高度」這個詞，像是戰爭最激烈的時候（也就是戰況達到頂點的時候）、盛夏（也就是夏天達到最高點的時候）、某事物最爲流行的時候（也就是流行潮達到頂點）

理論概念　一個已經被徹底思考過、仔細地定義過、並且在理論中充分說明的觀念或想法。

（譯註：在英文當中，**height** 一詞可以指「高度」或指「頂點」，因此才會有上述這些表達法）。

「高度」這個詞所指的是一個抽象的觀念，我們將這個觀念與其發音及書面形式的文字聯想在一起。事實上，這個詞語的發音與觀念之間，並沒有任何內在的關聯性。雖然此種關聯性是我們所武斷指定的，但是仍然非常有用。由於有了符號的關係，我們在對他人表達一個抽象的觀念時，就可以單獨使用符號來進行溝通。而這正是重點所在，換句話說，我們可以利用可見的符號，把我們腦海中那些抽象的、無形的概念傳達給其他人。

概念由兩個部分所組成：*符號*（單字、詞彙、或是書面的文字）與*定義*。我們以許多不同的方法來學習概念的定義。我們很可能會從父母親那兒學到高度這個詞，以及它所代表的觀念或定義。當我們學習如何說話及學習如何融入一個文化時，我們也會在這個過程中學會許多概念。我們的父母親很可能不會教我們各種概念在字典裡的正式定義。相反地，他們可能會透過一種長期的、非語言的、以及非正式的過程來教導我們。他們可能會使用許多例子，我們則會觀察並傾聽別人如何使用這個詞彙。當我們以錯誤的方式使用這個詞彙時，別人的臉上可能會出現困惑的表情，或者可能會有人糾正我們的用法。當我們正確地使用這個詞彙時，別人就能理解我們所要表達的意思。因此，到了最後，我們就能真正精通這個概念。上述的這個過程，正是我們在日常生活的語言中，學會大部分概念的過程。假如我們的父母親不讓我們接觸電視、把我們和其他人隔絕開來，並且教導我們物體從頂點到底部的距離叫作 *zodige* 的話，那麼當我們在與其他人進行溝通時，就會遇上問題。爲了要讓概念具有價值及用處，人們必須和其他人分享各種概念的符號、用語，以及各種概念的定義。

我們在日常生活中使用的大部分概念都只有模糊的、不清楚的定義。相同地，某個特定文化裡的人們所擁有的價值觀和經驗，也會影響、或者是限制日常生活中的概念。舉例而言，那些住在一個沒有電力的偏僻地區、而且從未使用過電話的人，就很難去理解電腦或者是網際網路這兩個概念。除此之外，有一些日常生活的概念（例如，像是邪靈或惡魔等），其根源則是來自於誤解、古代的迷信、或民間傳說。

日常生活中所使用的概念與在社會科學中所使用的概念之間，存在著一些差異，但是此種差異並不相當嚴謹或明顯。有一些社會科學所使用的概念，一開始是在研究中所發展出來的，並且有著精確的技術性定義。後來，這些概念逐漸滲透，進入了更大的文化與語言範圍內，隨著時間經過，這些概念的定義變得不那麼地精確、發展出其他的意義。舉例而言，像是性別歧視、生活型態、同儕團體、都市擴張、及社會階層等等概念，一開始其實都是由社會理論所發展出來的技術性概念。

社會科學的概念到底是來自於何處？許多社會科學中的概念在一開始，只是日常生活中的一些觀念或想法、個人的經驗、創造性的思考、每天的觀察結果而已。之後，可能有某人試著去詳細闡述某個觀念，並提出較爲明確的定義，而其他人開始討論這個觀念，試圖使這個觀念變得更加清楚、更加精確。有一些社會科學概念則來自於古典理論。在檢視了一些研究中的發現、或把許多不同情境中的發現與觀

念整合後，經過深度的沉思以及反省思考，人們有時會發展出一些新概念。整體而言，爲數眾多的社會科學概念形成一種特殊或專業的語言，而我們使用這種語言進行討論、分析、並檢視我們周遭的社會世界。許多人把這種語言稱爲*專業術語*（*jargon*），而這個名稱其實是帶有負面意義的。

在許多領域裡的專家們都會使用專業術語，因爲彼此進行溝通時，這是較爲簡略且方便的方式。物理學家、律師、藝術家、會計師、水電工人、動漫迷、種植蘭花的人士、及汽車修理工人，都會使用專門化語言，也就是專業術語。他們使用專業術語來指稱一些他們經常處理，抑或是遇到的觀念和物件，其中有些專業術語是不爲外界所熟知的。舉例而言，出版商和印刷業者也會使用一些專業術語，像是*無齊行指令的穿孔紙帶*（*idiot tape*）、*字體*（*fonts*）、*切割*（*cropping*）、*網版*（*halftone*）、*長條校樣*（*galley proof*）、*適壓印刷*（*kiss impression*）、*寡行*（*widows*）、以及*將鉛字面上的突出部分作平*（*kerning*）等等。對於同一個領域內的人而言，專業術語是一種快速、有效、而且有效率的溝通方式。然而，當人們誤用了專門化的語言，因而使他人產生混淆、排擠別人、或者是詆毀他人時，就會使專門化的語言產生負面的意涵，我們把此種帶有負面意義的專門化語言稱作*專業術語*。若我們對那些不瞭解專門化語言的人使用了專業術語，就無法達到溝通的目的，而且通常會招致他人的怨恨或不滿。

一旦我們學會了社會科學的概念，並用這些概念和那些知道其意義的人進行溝通，我們就會發現，在討論觀念與議題時，這些概念是有效率的、簡潔的、精確的溝通方式。然而，對於新手或是還未瞭解這些概念的外行人而言，當一場討論中充斥著各種不同的社會科學概念的詞彙，聽起來就像是一堆根本無法理解的專業術語。

抽象層次　概念會依據其**抽象層次**（level of abstraction）的不同而有所差異。有一些概念是非常具體的，而且所指涉的是一些我們能夠看得到、摸得著的物體，例如像是披薩、樹木、貓、行動電話、或大學裡的測驗等等。另外一些概念則是抽象的心智產物，與我們日常生活的經驗沒有直接關聯性。抽象概念所指涉的，是我們無法輕易經驗到、無法輕易表達的世界面向。然而，抽象概念仍然具有極大的價值。因爲這些概念可以組織我們的想法，並擴展我們的理解範圍。雖然我們無法直接看到像是愛國主義、社會資本、自尊心、情緒上的痛苦、恐慌、恐懼、認知失調、政治權力、組織權威等等抽象概念，但是我們仍然可能會「感受」到這些概念，抑或承認這些概念確實在日常生活中運作。

在定義一些簡單的、具體的概念時，我們會使用相當多的例子，並會指出許多可見的物理特徵。對較之下，複雜的、抽象的概念則需要正式的、像是字典一樣的定義，這些定義會結合數個其他的、較不抽象的、或是抽象層次較低的概念。雖然「高度」並不是一個非常抽象的概念，但是我們仍然使用了一些抽象層次較低的概念，像是*頂部*、*底部*、與*距離*等等概念，來定義高度這個概念。同樣地，「侵略行爲」

抽象層次　概念的特質之一，其範圍是從實徵而具體的（通常可以在日常生活經驗當中輕易觀察到），到非常抽象的（無法看見的心智產物）。

（*aggression*）這個概念也比我們可能用來定義它的其他概念（像是*揍人*、*打耳光*、*大聲尖叫*、*推擠*、*大吼大叫*、*用拳頭猛打*、*身體受傷*、或是*威脅要將某人打成重傷*等等）更爲抽象。除此之外，我們或許也可能利用像是*態度*或者是*刻板印象*等其他抽象概念，來定義*種族偏見*這個概念。

身爲社會科學家，我們通常會傾向於爲各種概念賦予比在日常生活中更爲精確的定義。我們會把某個理論中的概念，和研究以及實徵資料聯結。之所以會發生此種情況，是因爲只有當我們對於我們的觀念擁有清楚的、邏輯上連貫的定義時，知識才能向前發展進步。

對於促進知識發展與進行研究而言，擁有清楚的、明白的、及精確定義的概念是非常重要的。雖然有少數的研究或是理論性文章，會發展出全新的概念，但是我們通常都會依賴現有的概念。然而，由於許多的概念都有著多重定義，因此我們必須決定使用哪一種定義。甚至在我們決定好要使用哪一種定義之後，我們可能也會希望修正或是澄清現有的定義。

溫默（Wimmer, 2008: 973）曾經探討並重新定義*族群界線*（*ethnic boundary*）這個概念（族群界線是指能夠區隔不同族群團體的界線）。他把族群性（ethnicity）這個概念定義爲「一種主觀地感受到的歸屬感，此種歸屬感來自於一種共享文化和共同血統的信念。」這是許多關於族群性的定義的其中一種，其他人也已經使用這個定義。社會研究者一直在爭辯應該如何定義*族群性*和*種族*（*race*）這兩個概念。溫默認爲族群性是一個非常廣泛的觀念。他把*種族*和*民族性*（*nationhood*）這兩個概念定義爲附屬於族群性下的子類型。換句話說，種族是族群性的一種，是根據遺傳上的表現型（phenotype）特徵來加以分類，民族則是根據某個社群所具有的民族主義的渴望所分類。族群性的其他子類型還包括，以共享的宗教性、地區性、或是語言的傳統來加以分類的其他類型。

溫默（Wimmer, 2008）明白地拒絕使用一般人在日常生活中對於族群性或是種族的理解方式。美國民眾對於這些概念的理解不但是重疊的、模糊的，而且還是自相矛盾的（若想進一步參考最近的一些證據，請參閱 Hitlin, Brown, and Elder, 2007; Morning, 2009）。溫默避免在單一文化裡以概念被使用的方式來下定義，因爲這麼做會限制跨文化的比較，與理論的建立。他提到：

> 在有些社會裡，民眾之間有著表現型特徵方面的差異，但卻沒有激進的團體。在另外一些社會裡，雖然民眾之間沒有表現型特徵的差異，但卻有著彼此全然對立的激進團體。還有非激進化的族群辨別體系，此種體系就和美國境內的種族問題一樣具有排他性。（p. 975）

這個例子顯示出我們定義概念的方式，同時也突顯出了下列兩者的緊張對立關係：其中之一是民眾在日常生活中使用概念的方式；另外一個則是在社會理論和社會研究中所使用的概念。一般民眾在爲許多的概念下定義時，所用的方法是重疊的、模糊的、或者是自相矛盾的方式。而爲了要加深對於社會世界的瞭解，並且創造出

清楚的理論，我們需要精確的、沒有重疊的、不會自相矛盾的理論性定義。話雖如此，我們仍須研究一般大眾看待及思考這個世界的方式。如果我們借用一般民眾所使用的定義，那麼我們的定義或許會比較接近民眾在日常生活中使用這些概念的方式。但這些定義卻可能會是模糊的、重疊的、自相矛盾的。如果我們使用學術界的定義，這些定義可能無法密切地符合一般大眾對於這些概念的瞭解，但是我們的定義卻可以較爲精確、無重疊、也不會自相矛盾，能使我們的思緒更加清晰，讓知識獲得真正的進展。另一個會造成混淆的來源，則是一般大眾所使用的詞彙（例如，種族），與我們在社會理論中所使用的詞彙是完全相同的。最後，上述這些議題，意謂著我們身爲社會研究者，希望在腦海中，能對各種概念都非常清楚，而且都能小心謹慎地爲各種概念下定義。

總而言之，當我們在進行研究時，相當重要的研究任務之一，就是仔細地、精確地把觀念或是概念徹底釐清，然後賦予這些觀念或者概念，一個明確的、清楚的定義（我們將在第七章中更加詳細地討論這個議題）。此種理論化的思考過程，將爲進行研究提供極爲重要的基礎，並能夠促進我們對於周遭世界的瞭解。

單一概念與概念叢　使用概念時，我們很少把各種概念彼此孤立、隔絕開來。概念會形成相互聯結的群集，或者稱爲**概念叢**（concept clusters），不論是在日常生活中所使用的概念，或者是在社會理論中所使用的概念，同樣都有此種情況。許多理論會包含相關概念的集合，這些概念彼此是連貫一致的，並互相增強、補充。總而言之，這些概念的集合形成一個廣闊的意義網絡。舉例而言，在討論到*都市衰敗*（*urban decay*）這個概念時，我們可能會讀到一些相關的概念，像是都市擴張、經濟成長、都市化、郊區、主要城市、都市更新、貧民區、大眾運輸、犯罪率、失業、白人外移（White flight）（譯註：這個詞所指的是白人爲了避免與其他人種混居，並爲了逃避都市裡治安日益惡化的情形，因而大量向郊區遷移的情況）、以及少數民族等等。當我們同時使用這些概念，概念便會形成相互增強、補充的概念集合，而我們則能夠利用此種概念集合，來建立理論以及進行研究。

我們可以把在日常生活中與在社會理論中的概念簡化爲兩種類型。其中一種類型有著不同範圍的數值、數量、或總數，例如收入總數、溫度、人口密度、接受教育的年數、以及暴力程度等等，上述這些概念被稱作變項（variables），或是變項概念（variable concept），而我們將在第六章中進行較爲深入的探討。另外一種類型則表達了類別，或者是非變項形式的現象（例如，科層體制、家庭、大學學位、無家可歸、寒冷等等）。

簡單概念與複雜概念　概念除了可以分成具體與抽象概念、變項與非變項兩種類型之外，也可以被歸類爲簡單概念或是複雜概念。*簡單概念*（*simple concepts*）只有一個面向，而且只會在單一的連續體中變動。*複雜概念*（*complex concepts*）則有多重

概念叢　由一群相互關聯的概念所形成的集合，這些概念擁有共同的假定、彼此互相指涉、而且在社會理論中共同運作。

面向或是許多從屬的部分。我們可以把複雜概念拆解成數個簡單的、單一面向的概念。一般而言，越複雜的概念通常就會越抽象，而越簡單的概念則通常會越具體。不過，這只是一般性的通則，並非一成不變的定律。

以下是一個複雜概念的例子。魯施邁耶和同事（Rueschemeyer and associates, 1992: 43-44）曾指出民主這個概念具有三個面向：（1）定期、自由的選舉，且全民均具有投票權；（2）一個經由選舉產生的立法機關，其功能與目的在於掌控政府；以及（3）言論與結社的自由。他們也承認每個面向都會有不同程度的變化（每個人都能參與的、定期的和極爲開放，或是自由的選舉，相對於只有少數人被允許投票的、非定期的、限制重重的選舉）。在結合了這三個較爲簡單的概念或面向之後，魯施邁耶等人創造出一組包含不同政體類型的概念。三個面向都非常低的，就是極權政體，三個面向都很高的，就是民主政體。至於在三個面向中高低程度不同的組合，則是獨裁政體，或是自由的寡頭政體。由此可見，此種政體類型所指涉的概念，比三個面向所單獨指涉的概念，還要複雜得多。

另外一種概念複雜的類型是**理想型**（ideal type）。這是一種更廣泛、更抽象的概念，並組織了一群較爲具體的概念。理想型是一種純粹、抽象的模型，試圖定義我們正在討論的現象之核心本質。理想型也是一種心理的圖像，概略地描繪出我們所感興趣的事物的中心面向。雖然理想型比理論還要簡短，但卻能夠幫助我們建立一個完整的理論。理想型並不是解釋，因爲理想型不會告訴我們事物發生的原因和方式。質化研究者經常使用理想型，以瞭解那些可觀察到的現象與理想模型之間，契合程度如何。理想型的著名例子之一，是由韋伯（Max Weber）所提出來的，他發展出了科層體制這個概念的理想型（請參閱範例 3.1 韋伯的科層體制理想型）。他所提出的理想型把科層體制和其他的組織區隔開來。雖然沒有任何真實生活中的組織可以完美地吻合這個理想型，但是這個模型卻能夠幫助我們去思考以及研究科層體制。

理想型 一種概念分類的類型，理想型呈現了關於某個事件、過程、或觀念的純粹而抽象的模型。理想型通常會被用於建立社會理論以及分析資料。

概念分類 一種複雜的、多重面向的概念，此種概念擁有許多種次類型，而且這些次類型介於單一概念與完整的理論解釋之間。

類型學 一種理論性的分類方式或是準理論，其產生方式是利用跨類型的分類、結合兩個或是多個簡單概念，以形成一組相互關聯的次類型。

所謂的**概念分類**（concept classification），乃是介於一個簡單的概念以及一個完整的理論之間。[3] 概念分類可以幫助我們組織許多抽象、複雜的概念。藉由以合乎邏輯的方式，結合一些較爲簡單的概念的方式，我們可以創造出一個複雜概念的類型，而這就稱作分類。你可以看看一些例子，以更加瞭解這個觀念。分類的主要類型之一是**類型學**（typology），也稱作分類學（taxonomy），[4] 在這種分類方式中，研究者會以合乎邏輯的方式結合兩個或是更多個單一面向的、簡單的概念，如此一來，原本的兩個簡單概念就可以交互結合，形成一個新的概念。這個新的概念可以表達出原本幾個簡單概念之間的相互關係，或者重疊的部分。

墨頓（Merton, 1938）所提出的偏差行爲脫序理論（anomie theory of deviance）是一種被廣泛使用

範例　3.1

韋伯的科層體制理想型

- 科層體制是一種連續性的組織，並受到一套由規則所形成的系統管理。
- 行為是由超然的、客觀的規則所管理。
- 體制內講求勞力分工，不同的職位會被分派給能力不同的個人。
- 層級權威關係勝過一切，換句話說，處於較低層級職位者，必須受到較高層級職位者所控制。
- 管理行為和規則等等，都是以書面方式呈現，並以檔案方式留存。
- 個人並未擁有其職位，也不能夠購買或出售其職位。
- 為了確保個人對於組織的忠誠，所有工作人員一律由組織支付薪資，而不是直接向接受服務的當事人收取費用。
- 組織的財產與個人財產完全分開。

來源：Adapted from Chafetz. *A primer on the construction and testing of theories in sociology* (1978: 72). F. E. Peacock Publishers.

的類型學，既簡潔又優雅。該理論利用下列兩個較爲簡單的概念，讓我們可以同時瞭解非偏差行爲及偏差行爲：（a）由社會所定義的值得追求的目標；以及（b）人們用來達成這些目標的手段或方式。此種類型學所倚賴的是下列兩種關係：（1）人們是否接受或是拒絕這些目標；以及（2）人們是否用社會上能夠接受的手段或方式（也就是用合法的方式）來達到這些目標。根據這些概念，墨頓的脫序理論確認了順從行爲與數種偏差行爲的類型（請參閱表 3.1）。當人們接受了社會的目標（例如，像是獲得高薪），並使用合法的方式或手段，來達到這些目標時（例如，取得一份好工作並努力工作），就會產生順從行爲或是非偏差行爲。當情況並非如此，就會產生各種不同形式的偏差行爲。墨頓關於個人如何適應目標，和所使用達到目標方法之分類，總結了他的複雜概念，也賦予每一種次類型一個名稱。舉例而言，退縮這個次類型所描述的，是一個拒絕社會目標及達到目標的合法方式之人，像是長期酗酒者、或宗教的遁世者。此種類型的行爲偏差者，不但拒絕那些表面上受人尊敬及獲得物質資產（例如房子和車子等等）的社會目標，也拒絕達到此目標的合法方式或是手段（例如，做一個誠實的人、在工作方面努力表現等等）。

表 3.1　墨頓的個人適應類型

適應類型	社會的目標	制度性方法
I　順從	接受	接受
II　創新	接受	拒絕
III　儀式主義	拒絕	接受
IV　退縮	拒絕	拒絕
V　革命	以新的目標加以取代	以新的方法加以取代

另一種不同的概念分類則是建立在古典社會理論之上。萊特（Wright, 1978）更新了馬克思關於資本主義裡的社會階級理論，並在之後以來自於當代美國社會的實徵資料，測試他所更新過的理論。萊特發現到，對於馬克思而言，不平等及剝削是來自於對下列三種類型的資源的控制：（1）投資（也就是能夠產生利潤的財產或資本），（2）生產組織，與（3）勞動力（是指其他人所付出的勞力與工作）。萊特認爲，階級社會的組織方式會創造出不同的職位，而這些職位則會賦予權力（也就是命令其他人去工作）給那些擁有職位的人們。他也認爲階級社會的組織方式會創造出其他的職位，而這些職位則會賦予人們控制三種類型資源的力量（請參閱表 3.2）。位於能夠控制所有三種類型資源的職位的人，就成爲最有權力的人，成爲社會中的最佔優勢的社會階級。在市場經濟中，這就是資本家階級，而這個階級裡的成員則包括了主要的投資者、財產的擁有者、銀行或是企業的總裁。資本家會作出投資的決定（例如，是否以及要在何處建立一座新工廠）、決定如何組織生產活動（例如，使用機器人或廉價勞工）、並對其他人下達命令。而接近社會底層的階級，則是由工人所組成的。他們所佔據的職位，對於投資或是如何組織生產活動等等，完全沒有發言的權利。他們也缺乏對其他人的權威，且爲了要保住他們的工作，他們必須遵從其他人所下達的命令。至於協助資本家的經理人和督導者，則介於兩個主要階級之間。當馬克思在 19 世紀中期發展出他的理論時，經理人以及督導者只是個尚未完全成形的準階級，這個階級控制了部分、而非全部的主要社會資源。萊特的分類方式也指出另一個馬克思曾提到過的階級，也就是小資產階級（petite bourgeoisie）。而這個階級主要由小型的、自營的經營者或是農民所組成。這個階級的成員擁有並經營他們自己的事業，除了家庭成員之外，不雇用任何人。馬克思認爲這個階級將會逐漸萎縮，並且消失，但是時至今日，這個階級依然存在。就像墨頓一樣，萊特也結合了幾個簡單的概念（就是擁有的、或是缺乏的資源類型），以創造出一個在理論上強而有力、並相當複雜的分類方式（即爲在資本主義社會裡的社會階級結構）。

表 3.2 萊特的社會階級系統

社會階級	對於社會資源的控制		
	投資	*生產*	*勞動力*
資本家	+	+	+
經理人	−	+	+
督導者	−	−	+
工人	−	−	−
小資產階級	+	+	−

+表示能夠控制該項資源，−表示幾乎無法或不能控制該項資源

最後一個與概念分類有關的例子，則是來自於沃爾德（Walder, 2003），他想要瞭解從採行計畫經濟的共產政體轉變到採行市場經濟的後共產政體的過程。爲了達成目標，他利用了下列二個因素，創造出共四種類型的後共產政體的分類法：（1）對於獲得私人資產的限制，以及（2）發生政治變化的劇烈程度。他利用這兩個因素交互分類，以創造出一個概念上的類型學。之後，他利用這個類型學與其他的觀念，來解釋共產主義轉型的平順過程，並確認在各種不同的後共產社會中，會由哪些社會—政治團體取得權力（請參閱表 3.3）。不過，你必須要注意的是，概念分類本身並不是完整的理論性解釋，我們仍須增加其他理論方面的思想或觀念，才能瞭解爲什麼會出現這些結果。

範圍 不同的概念範圍也各自不同。有些概念的範圍非常狹窄，只能應用於特定的社會背景或活動，可能受到時間與空間的限制。我們無法輕易地把這些概念應用於一個特定背景之外的地方。而有些概念的範圍則非常廣泛。這些概念可以應用於許多不同的背景或是活動，而且能夠橫跨大幅度的時間與空間。一般而言，範圍廣泛的概念通常會比範圍狹窄的概念更抽象。

範圍狹窄的概念案例之一，就是「足球暴力」，這個詞所指的是由英國的足球迷以及一些其他歐洲國家的球迷所展現出來的暴力行爲。此種暴力行爲從 1960 年代晚期以來，已有加速發展的趨勢。這個概念在時間與地點上，是有其限制的。雖然其他擁有許多觀眾的大型運動賽事的狂熱者，也會發生暴動或是展現出暴力與破壞物品的行爲，但是足球暴力這個概念，其實是特別用來指涉那些易於暴動的英國足球迷。另外一個例子則是日本的「過勞死」（*karoshi*）現象，也就是因爲過度工作而死亡的情況。其實在歷史上以及不同的文化中，一直都有因過度勞動而死亡的人，但是過勞死這個概念所指涉的範圍卻是相當狹窄的，也就是指那些男性的白領工作者，因爲承受了巨大的社會壓力，而必須在長達一年、或者是更長的時間裡，不眠不休地連續爲公司長時間工作（例如，每天連續工作長達 16 到 18 個小時）。這個概念與 1970 至 1990 年代在日本公司裡的工作文化有關。對比之下，與此相似、但是範圍較爲廣泛的概念，像是體力勞動（physical labor）或者是辦公室工作（clerical work）等等概念，則能夠廣泛地應用於不同的歷史時期、抑或多種不同的文化背景。

範圍狹窄的概念通常最爲貼近具體的日常生活，此種情形也使得這些概念能夠

表 3.3　從共產主義經濟到後共產主義經濟的四種轉型路徑

		政治變化的劇烈成度	
		高	低
對於取得資產的限制	高	1	2
	低	3	4

容易地被人們所接受與承認。我們可以在這些概念裡，融入特定的情境脈絡特徵和某個社會背景特質。與此同時，此種作法也使我們比較難以將這些概念向外推論，或輕易地利用這些概念建立對於社會生活的普遍理論性理解。範圍廣泛的概念（像是社會參與、情緒方面的溫暖等等），則與範圍狹窄的概念有著完全相反的優點與缺點。這些範圍廣泛的概念能夠連結不同的時空背景，亦能促進我們的一般性理解。然而，此類概念會忽略在特定社會背景以及歷史情境下的重要脈絡細節。

關係

社會理論不只是假定與概念的集合體，社會理論也會詳細說明概念之間的關係。社會理論告訴我們這些概念彼此間是否相互關聯，而且，如果真的有關聯性，又是以何種方式產生。由於社會理論能夠描繪出假定、概念、及關係所形成的複雜集合體，因此，社會理論也能夠提供一幅完整的圖像，以說明某些特定的關係究竟存在或不存在。

關係的種類 除了告訴我們某些概念是否相關之外，理論也能詳細說明概念之間的關係。例如，某個理論也許能夠告訴我們，某種關係是強或是弱、直接或是間接、正向或是負向。理論或許能告訴我們，某個概念會增加或減少另一個概念，某個概念所造成的影響是立即而直接的、或者是延遲出現的。好的理論會指出某個概念是否爲另一個概念的必要（也就是不可或缺的、而且是必須的）先決條件，或只是充分（也就是有關聯，但是不一定會出現）條件。有時理論會告訴我們，某個概念與另外一個概念有關聯性，但只有在某些條件下才成立〔此種情況稱爲*權變關係*（*contingent relationships*），我們將在本書稍後部分進行討論〕。理論通常也會詳細說明關係運作的解釋形式（例如，因果的、結構的等等）（請參閱本章稍後的說明部分）。

命題以及假設 社會理論包含了關於概念間的命題關係。所謂**命題**（proposition）就是指一項理論性的陳述，以說明兩個或是多個因素，或是概念彼此相關，並說明此種關係的類型爲何。命題是一種看法或是信念，而且可能已經過驗證、也可能尚未經過驗證。進行研究的重要目的之一，就是找出某個理論的命題是否能夠符合實徵的證據或是資料。有些理論性命題是以假定的形式出現，而其他的理論性命題則是能夠以實徵資料來加以驗證。至於**假設**（hypothesis）則是一個能夠用實徵證據或資料加以驗證的命題。假設是關於某種關係的試驗性或暫時性陳述，這是因當我們在開始進行一項研究時，我們尚且無法確定某個假設在實徵世界中，是否能夠真正成立。經過許多情境之中，對某個假設重覆進行實徵評估後，我們越來越能夠確定此項假設的真實性。藉由以實徵方式評估一個假設，我們可以得知某項理論性命題是否得到支持，我們也許可以決定要修正該命題、或是把該項命題從理論

命題 關於兩個或是多個概念間，關係的理論性陳述。

假設 一個能夠用實徵證據或是資料加以驗證的理論性命題，該命題尚未以實徵證據來加以測試或驗證。假設最常被使用於演繹式的理論推理過程中，也可以被重新陳述為預測。

中完全移除。雖然有許多研究的設計是要去驗證假設的，仍然有少數類型的研究在研究設計中，完全不包含任何假設。

分析單位

社會世界中包含了許多不同的單位，像是個人、團體、組織、運動、機構、以及國家等等。研究者會適度地修改理論性概念，以應用於一個或是多個**分析單位**（units of analysis）之上。舉例而言，*侵略行爲*這個概念能夠被應用於以下數個不同單位之上：個人、團體、組織、或是國家。此種情況可以由下列陳述加以說明：傑米是一個具有*侵略性*的孩子；這支籃球隊昨晚的*侵略性*很強；XYZ 企業以具有*侵略性*的方式進入了一個新市場；聯合國譴責了某個國家對其鄰國所進行的*侵略行動*。從表面上來看的話，兒童所展現出來的侵略行爲（像是打了另一個四歲兒童一個耳光，以及用腳去踢老師等等），其實與運動隊伍所展現出來的侵略行爲（肢體接觸與阻擋對手）、公司企業所展現出來的侵略行爲（降低售價並針對競爭產品發動大型的廣告宣傳戰）、或是國家所展現出來的侵略行爲（調動軍隊及坦克車穿越國際認同的界線）是有所不同的。

當我們進行一項研究時，我們必須使某個概念契合於我們想要進行分析的單位的特定類型，就像手套必須契合於手的形狀一樣。而這就意味著當我們在設計研究與測量概念時（在第六章和第七章將有更進一步的討論），我們必須使概念和分析單位之間彼此契合。如果我們所考量的是一個可以適用於各不同分析單位的抽象概念，例如侵略行爲等等，我們就必須決定要集中探討的單位，並在實際進行研究之前，根據這個分析單位，修改我們定義某個概念的方式。

理論的面向

現在你已經知道了社會理論的許多內容，你也可以去思考社會理論的其他形式。由於社會理論具有許多不同的面向，因此社會理論有時會令人感到相當困惑。爲了把事情簡化，我們可以把社會理論區分爲下列五個主要面向：

1. *建構理論的方向*（*direction of theorizing*）可能是演繹的、或是歸納的。
2. *分析的層次*（*level of analysis*）可能是微觀的、鉅觀的、或是中觀的（介於兩者之間）。
3. *理論的焦點*（*theoretical focus*）可能是實質理論或是形式理論。
4. *解釋的形式*（*form of explanation*）可能是因果的、結構的、或者是詮釋的
5. *理論的範圍架構*（*range of a theory*）可能是實徵推論、中距理論、或者是理論框架。

乍看之下，上述這些面向似乎有點令人望而生

分析單位　研究者所研究探討的單位、個案、或者是社會生活的部分。在發展概念、以實徵方式測量或觀察概念、以及使用資料分析時，分析單位都具有關鍵性的地位。

畏。不過，幸運的是，在這些面向之中，研究者通常只會用到少數幾種主要的結合方式。隨著你漸漸熟悉這些面向，你將會發現，在你進行研究時，這些面向可以幫助你澄清及簡化你應用理論的方式。

建構理論的方向

在理想的狀況中，你能夠從兩個方向來建立並且測試理論：(1)從抽象思考開始，然後以合乎邏輯的方式將理論中的觀念與具體證據連結起來，或者是（2）從對實徵證據的具體觀察結果開始，然後利用證據進行推論，以建立更爲抽象的觀念。但在實際狀況中，大多數的研究者都是具有彈性的，並且通常會同時使用兩種不同的方向，也就是他們或許會在研究中的不同時間點上採取不同的方向（請參閱圖 3.1）。

演繹 在以**演繹取向**（deductive direction）來建立理論時，我們是從一些抽象的概念或者是一個理論性的命題開始，而這些概念或命題則會概略地描繪出一些概念之間的邏輯關聯性。接著，我們會用具體證據來評估這些概念及命題。換句話說，我們由觀念、理論、或者是心靈的圖像開始，然後朝向可觀察的實徵證據移動。在第二章中，討論過關於接觸假說（contact hypothesis）的研究，所使用的就是演繹取向的理論推理方式。在該項研究中，研究者們由一個理論性命題開始，而該命題內容如下：當人們缺乏與其他社會「外團體」成員之間的人際接觸時，人們會由於無知及負面的刻板印象，對這些外團體產生負面看法。研究者們把這個命題轉變爲一個可以加以驗證的實徵假設：當人們對某個外團體中的成員有更多的接觸、更多認識、更加熟悉時，就能夠減少人們對於這個外團體的負面想法、態度、及陳述。此種理論性推理過程是由抽象層次朝向一個具體、實徵的層次，而此具體、實徵的層次，則包含了特定而明確的外團體、社會接觸形式、想法或是態度。

歸納 在以**歸納取向**（inductive direction）來建構理論時，我們是從觀察實徵世界開始，然後仔細去反思所發生的事情，並且逐漸以更爲抽象的方式來進行思考，也就是說，我們會朝向理論性概念與命題移動。我們可以從一個一般性的主題及少數幾個模糊的想法開始，並以歸納的方式，把這些主題或是想法加以修正，或詳細說明，以成爲更加精確的概念。我們從實徵的觀察結果開始，然後朝向較爲抽象的思考方式移動。在第二章裡，曾經提到過一項針對紐約市的街頭小販所進行的研究，而在該項研究中，鄧尼爾（Duneier, 1999）所使用的就是歸納的思考方式。他只有在蒐集實徵資料的過程中、以及已經完成蒐集資料的工作之後，才去發展理論性的理解。他自己說過：「我開始從我在街頭上看到和聽到的事物之中，去獲得想法」（p. 341）。鄧尼爾（p. 342）也描述了這個過程，就像一個看過許多不同症狀的醫生所使用的方

演繹取向 一種用來發展或證實某個理論的方法，此種方法從抽象概念及理論關係開始，然後朝向更為具體的實徵證據移動。

歸納取向 一種用來發展或是證實某個理論的方法，此種方法由具體的實徵證據開始，然後朝向更為抽象的概念以及理論關係移動。

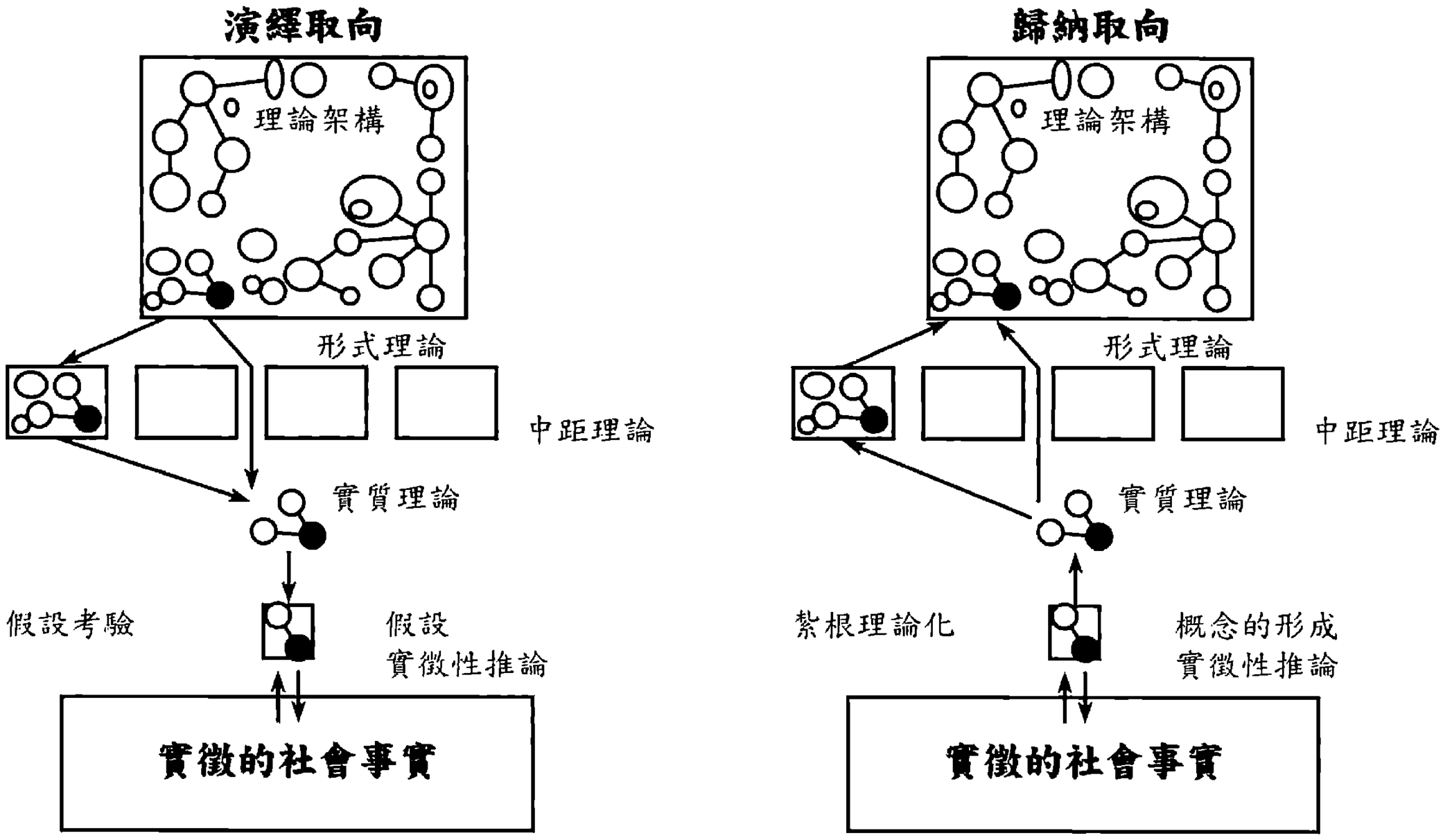


圖 3.1　演繹以及歸納取向的理論化過程

法。換句話說，只有當這名醫生分析過這些多樣性的症狀後，才能作出診斷、或作出合理而一致的判斷，以解釋在表面上許多明顯可見的症狀背後，真正潛在的病因。

許多研究者會使用一種特殊類型的歸納式理論化過程，此種類型被稱作**紮根理論**（grounded theory）。紮根理論牽涉到從根基開始，往上形成新的理論觀念，而非去測試現存的理論觀念。你將在第六章和第十四章中學到更多關於紮根理論的知識。

紮根理論是一種在質化研究中被廣泛使用的方法。當然，這並不代表紮根理論是唯一的方法、或是所有的質化研究者，都會使用紮根理論。紮根理論是「一種質化研究方法，此種方法使用一組系統化的程序，以發展出關於某個現象的、由歸納方式所推導而得到的理論」（Strauss and Corbin, 1990: 24）。紮根理論的目的在於建立一個忠於證據的理論，而且這是一種發現新理論的方法。利用紮根理論，研究者可以去比較一些彼此相異的現象，找出這些現象的相似之處。研究者也可以把微觀層次的事件當作是一個較爲鉅觀的解釋基礎。紮根理論與其他較爲實證主義導向的理論（我們將在第四章裡討論實證主義理論）之間，有著幾個共同的目標，這是因爲紮根理論所尋求的是一個可以符合證據的理論，而該理論是精確並且嚴謹的、能夠加以複製的、能加以推論的。紮根理論研究法所追求的，是利用橫跨不同社會情境所進行比較的方式，作出一些概括性的推論。

質化研究者也會使用一些紮根理論的替代方法。有些質化研究者會提供忠於資料提供者的世界觀的深度敘述，這些研究者挖掘出一個單一的社會

紮根理論　一種經常被使用於質化研究的歸納式社會理論類型。紮根理論通常會藉由對實徵觀察結果進行比較的方式，建立抽象的理論。

情境，以解釋維持穩定的社會互動的微觀過程。其他研究者的目標，則是提供關於某個事件或背景情境極度精確的描述。這些研究者會分析特定具體的事件或是背景環境，以獲得較大的社會動態變化過程的深入洞見。另外，有些研究者會應用一個現存的理論，分析特定的背景環境，他們會把這些特定的背景環境置入於一個鉅觀層次的歷史脈絡中。這些研究者會展現出鉅觀層次的事件間的關聯性，以及在微觀層次的情境和較大的社會力量之間的關聯性，而其目的在於，重新建構理論，與鼓舞社會行動（若想瞭解其他數種替代方法的摘要，請參閱 Burawoy, 1991: 271-287; Charmaz, 2003; Hammersley, 1992）。

分析的層次

社會事實存在於許多不同的層次之上，其範圍則從微觀層次到鉅觀層次。社會生活的微觀層次包括了少數個人之間短期的面對面互動，通常是在一個小型規模的背景中（例如，一名女性顧客在一家速食餐廳裡簡短地和一名員工以及排在她後面的男子聊天）。而在社會事實的微觀層次上，人們會從事直接的個人接觸，並且通常是在一個封閉的實體背景環境中進行互動。社會科學家發展出**微觀層次理論**（micro-level theory）與概念，這些理論及概念都已經被適度修正過，以用來分析微觀層次的社會事實。舉例而言，麥克法蘭（McFarland, 2004）就針對高中教室裡的破壞行為，發展出一個微觀層次理論。他仔細地觀察學生與老師間在教室裡的互動情形，而根據他的觀察結果，發現了支持者與反抗者各自有不同模式的行為，也會造成不同的結果（請同時參閱範例 3.2，以及歸納式微觀層次理論）。

微觀層次理論 一種社會理論，其焦點在於社會生活的微觀層次，而且通常是發生在相當短的時間之內（例如，個人或小型團體間的面對面互動與相遇）。

鉅觀層次理論 一種社會理論，其焦點在於社會生活的鉅觀層次（例如，社會機構與制度、社會的主要部分、整個社會、或者是世界上的不同區域），以及發生於長時間內的過程（數年、數十年、一個世紀或更久）。

中觀層次理論 一種社會理論，其焦點在於社會生活中階層次裡（例如，組織、運動、以及社群）的關係、過程、結構，與中等的時間長度內（數個月、好幾年、或者大約十年）所發生的事件。

至於與微觀層次處於兩個對立端點的鉅觀層次，則包含了大規模的社會事件（例如，像是十八世紀時，西方歐洲帝國主義強權與中國文明之間的相遇模式）以及整體社會機構與制度（例如，某個國家裡的整體犯罪司法系統）。**鉅觀層次理論**（macro-level theory）解釋了在大規模的社會單位裡運作的事件、過程、模式、與結構，通常涵蓋了數十年或是更長的時間，而且經常包含相當廣大的地理空間。在第二章的範例 2.13 裡，我們曾看過由馬哈尼（Mahoney, 2003）針對拉丁美洲國家在長達一個世紀的時間裡的變化，所進行的研究。該項研究正是鉅觀層次理論的例子。

介於微觀層次及鉅觀層次之間的，是中觀層次，也就是中間的層次。**中觀層次理論**（meso-level theory）的焦點集中於組織、社會運動、或是社群的層次之上。當我們檢視社會世界的不同層次時，我們會發展出與分析層次相對應的理論和概念。

範例　3.2

歸納式微觀層次理論

在她針對兩家差異極大的玩具店所進行的研究當中，威廉絲（Williams, 2006）曾在兩家玩具店各自工作過六個星期的時間，並進行了詳細的觀察。而她以歸納的方式，從她的觀察結果中，發展出一個微觀層次理論。她的目標是要從觀察日常生活中的互動，「以描述並且分析那些影響大型玩具店的規則」（pp. 19-20）。威廉絲觀察並詳細記錄了在日常工作場所的例行公事裡、非正式的玩具店規則裡、顧客與店員的互動裡，店員會以數百種以上的方式來更加善待男性、白人、以及高收入者。這些行動強化了盛行的社會階層體系：與女性、非白人、低收入者相比之下，男性、白人、及高收入的人處於優勢的地位，而且擁有特權。在兩家玩具店裡，所有的董事都是白人男性，而受雇去做「男性」工作（例如，像是安全警衛、搬運貨物的勞工、後勤組裝員等等）的人，全部都是男性（其中有一半是非白人），而受雇去做「女性」工作（例如，像是收銀員、顧客服務員等等）的人，則全部都是女性。唯一的例外情況，是在其中一家店裡的電子部門。該部門是一個獨立部門，而這在這個部門裡的員工都是亞洲男性。其中一家店屬於「高檔」的玩具店，而且販售要價不斐的玩具。另外一家店則比較像是大型批發倉庫，而且主要客群是勞動階級的民眾。在這兩家店裡，店員和經理都有完全相同的「顧客分類法」，也就是說，他們都會把白人男性顧客當作是潛在的「大客戶」，並把黑人男性客戶當作是潛在的竊賊。威廉絲的微觀層次理論顯示出，在差異甚大的背景環境裡，非正式的日常生活規則如何使得階級、種族、和性別之間的不平等現象更加根深蒂固。即使是平凡的買賣互動，也會持續地複製社會階層體系中的任何關係，而此種關係也幾乎從未改變過。

理論的焦點

我們會去建構、詳細闡明、測試或是驗證兩種類型的理論：實質理論以及形式理論。**實質理論**（substantive theory）的焦點在於社會事實中的特定內容或主題領域，像是家庭關係、偏差行爲、種族與族群關係等等。例如，在我們之前看過的例子裡，在馬哈尼（Mahoney, 2003）對於拉丁美洲的研究（第二章）裡的理論，其焦點在於經濟發展；而在威廉絲（Williams, 2006）對於玩具店的研究裡（請參閱範例 3.2），其理論焦點則在於社會不平等現象如何在每天的面對面互動中被不斷複製。

形式理論（formal theory）的焦點則是在於一般的過程或結構，而這些過程或是結構，能夠在多種不同的主題領域中運作，形成社會認同、參與衝突鬥爭、行使權力等等。一個關於取得資源的管道，掌握權力和權威的形式理論，或許能夠應用於數個不同的領域。這個理論或許能夠解釋，在先進的資本主義社會中，富有的企業擁有者，如何利用他們取得珍貴資源的權利，以維持所擁有的經濟和社會權力（請參閱表 3.2）；在從共產主義社會轉變到後共產主義社會的過程中，政府裡的菁英份子如何利用對資源的控制，把持他們所擁有的權力（請參閱表 3.3）；以及十九世紀時，處於資源控制僵化體制中的殖民地菁英們，如何緊抓權力不放，阻礙了之後的國家發展過程（請參閱範例 2.13）。在上述三個情況中，

實質理論　一種特別針對某個主題領域，量身打造的理論類型。

形式理論　一種具有普遍性、能夠應用於許多特定主題領域的理論類型。

都有相似的社會經濟動力在運作，換句話說，握有權力的菁英團體利用他們對於珍貴資源的擁有權與掌控權，以維持其掌握權力的地位，並抗拒任何對其權威的挑戰。

事實上，實質理論和形式理論之間，存在著相互重疊的部分。針對某一個主題而提出的實質理論，經常藉助於其他的形式理論，並與這些形式理論結合在一起，而某個形式理論也可能應用於數個不同的實質理論。如同雷德爾（Layder, 1993: 44）曾經說過的：「因爲有許多實質理論與形式理論的刺激，因而強化了理論的累積過程。」

每一種理論的焦點都會有其長處以及限制。實質理論能夠針對一個特定的主題領域，提供強大的解釋與說明，並包含了來自特定背景環境、過程、或事件裡的細節。然而，實質理論卻可能難以推論到不同的主題領域。與形式理論相較之下，實質理論中的概念抽象層次較低，而且範圍也比較狹窄。形式理論的長處，正是其橫跨多重主題領域及促進一般知識的能力。而形式理論的短處，則是由於形式理論通常並非來自於特定具體的議題，以及社會背景環境。因此我們必須將形式理論加以調整，才能看出它與一個特定的議題，或是主題之間的關聯性。形式理論可以幫助我們辨識與解釋，在許多不同主題間的類似特徵。由於形式理論比較抽象，因此也使得形式理論較爲複雜，且必須以一種純粹邏輯性的、分析性的形式，才能較爲容易地表達這一類的理論。

解釋的形式

預測以及解釋 理論最主要的目的就在於提出解釋與說明。然而，解釋一詞本身有兩種不同的形式：理論性、以及一般性。研究者通常把重心放在**理論性解釋**（theoretical explanation）上面，而所謂理論性解釋，是指一項邏輯性的論證，該論證告訴我們爲什麼某個事物會有某種特定形式，某件事爲什麼會發生。通常當我們提到理論性解釋時，我們所指的是一種一般性的規則或原則，而且我們會使用許多概念之間的聯結，把此種規則或是原則與一項理論性的論證聯結在一起。至於一般性解釋（ordinary explanation）則是在澄清某件事情、以一種舉例說明及一般人能夠理解的方式，來描述某件事情。舉例而言，一位好老師會以學生們能夠瞭解的方式來「進行解釋」。上述這兩種不同的解釋形式可以混合在一起，就好比我們在解釋（也就是使某件事物變得清楚明白）某項解釋（也就是一項牽涉到理論的邏輯性論證）時一樣。在我們進一步檢視理論性解釋的各種不同形式之前，讓我們先看看預測和解釋之間的差異，因爲許多人常常把這兩者混爲一談。

預測（*prediction*）是對某件事情將會發生的一種陳述，而*解釋*（*explanation*）則是以合乎邏輯的方式，將在一個特定情境中所發生的事情，與一項關於「事物如何運作」的較抽象或是基本的原則聯結起來，以回答「爲什麼」的問題。而上述的特殊情境，則被視爲是較有一般性原則的例子、或是一項具體案例。預測比解釋更爲容易，而解釋則比預測擁有更爲強大的邏輯力量，因

理論性解釋 一項邏輯性的論證或敘述，能夠藉由指涉較為一般性以及抽象原則的方式，說明為什麼某件事物會以特定的形式出現、或是某件事為什麼會發生。

爲好的解釋也能達成預測。一項具體的解釋，很少預測一個以上的結果，但是各種不同、互斥的解釋卻可能會預測出相同的結果。雖然預測的力量比解釋來得薄弱，仍有許多人著迷於預測所帶來的戲劇性預視能力。

我們可以用賭博的例子來具體說明解釋和預測之間的差異。假設我到一間賭場，並準確地預測了下一張將要出現的牌、或是在輪盤上所將要出現的下一個號碼，這將會造成轟動。我或許會贏得一大筆錢，至少一直到賭場的經理發現我一直在贏錢並把我趕出去爲止。不過，我用來作出正確預測的方法，會比我能夠作出準確預測這件事更加令人感興趣。此外，告訴你我如何能夠準確預測下一張牌，會比我具有準確預測的能力這件事，還更具有吸引力。以下是另外一個例子。你知道太陽每天早上都會「升起」，因此，你能夠預測，每天早上不論是不是有雲層遮掩，太陽總是會在某個時候從東方升起。但是，爲什麼太陽每天都會從東方升起呢？有一種解釋是，有一隻巨大的烏龜每天都會揹著太陽橫越整片天空。另一種解釋則是，某一位天神把祂手中的一支箭點燃，射過整片天空，而對我們凡人而言，那支點燃的箭就是太陽。當然，今日幾乎沒有任何人會相信這些古老的解釋了。而你很可能會接受的解釋，則是牽涉到地球的轉動及太陽（在我們太陽系裡的一顆星球）位置的一整套理論。在這個解釋裡，雖然表面上太陽看起來好像從東方升起，但事實上太陽本身並未移動。太陽所顯現出的移動狀態，其實是因爲地球的轉動所造成的。我們所在的這個星球，不但會繞著一條軸線進行自轉，也會繞著一顆在太空中相隔數百萬英里的星球進行公轉。以上所提到的三種解釋，都作出了相同的預測：太陽每天早上都會升起。正如同你可以看到的，即使是一個說服力薄弱的解釋，也能夠作出正確的預測。一個好的解釋仰賴一個充分發展的理論，並能夠經由實徵觀察結果而得到確認。

曾經獲得諾貝爾獎的物理學家溫伯格（Steven Weinberg, 2001: 47）爲解釋提供了一個「硬科學」（"hard science"）的觀點：

> 進行純粹而非應用研究的科學家，通常都會告訴大眾與提供資金的機構，他們的任務是找出某件事或其他事的解釋……在物理學有限的情境脈絡中，我認為一個人能夠……「辨別」解釋與單純描述這兩者之間的差異。當科學家說他們已經解釋了某種規律性時，他們所做的事其實只是單純的描述而已……當我們能夠證明某個物理原則能夠從另一個更為基本的物理原則演繹而來時，我們才算是**解釋**了這個物理原則（強調部分為本書作者另外添加）。

理論性解釋共有三種形式：因果關係的、結構性的、以及詮釋性的。這三種形式中的每一種，都在解釋或回答某件事爲什麼會發生的問題。而且每一種形式都會把一個具體個案和某種類型的一般性原則聯結在一起。

因果解釋　**因果解釋**（causal explanation）指出了在各種概念或是變項之間的因果關係。我們會在日

因果解釋　理論性解釋的一種類型，此種解釋是以原因與結果、或是某項因素產生某種結果的方式，來表達事件發生的原因與事物運作的方式。

常生活的語言中使用此種解釋，不過，日常生活中的語言通常比較草率而且含糊不清。以下是因果解釋的一個例子：你或許會說貧窮*導致*犯罪、或是社會道德風氣的敗壞，*導致*離婚率上升。這些都是很基本的因果解釋。不過，當社會科學家在討論到因果關係時，他們會盡力去以較爲正確及確切的方式，來加以說明。此外，社會科學家也試著確認此種因果過程的運作方式，與發生原因（例如，貧窮如何、以及爲什麼會導致犯罪）。

至少從十八世紀時的蘇格蘭哲學家休謨（David Hume, 1711-1776）開始，哲學家們就一直在針對「原因」（cause）這個概念進行辯論。有些人爭論說因果關係發生於經驗世界裡。雖然我們無法輕易地看見因果關係，但此種關係確實「存在於」客觀事實裡，而且我們能夠找到此種關係的間接證據。其他人則爭論說因果關係並不存在於客觀事實中，那只是一種存在於「我們腦海中」的心智建構物，而我們主觀地創造出因果關係的概念，目的在於幫助我們去思考，在客觀事實中所發生的各種事件。在不涉入哲學式論辨的情形之下，許多社會科學家已經針對因果關係提出不同的理論，進行了許多研究。

因果關係的必要條件 我們需要有三個要素才能建立因果關係：時間的順序、實徵方面的關聯性、以及排除其他看似可能的因素（請參閱範例 3.3 因果關係的三個要素）。第四個內含的條件，則是任何的因果關係都必須要合理或有意義，符合較廣泛的假定理論抑或架構。除了上述的三個必要條件之外，一個完整的解釋也必須詳細說明，造成該因果關係的機制，並概略描繪出因果鏈（a causal chain）。

1. *時間的順序*（*temporal order*）意思是指在時間上，原因必須發生在結果之前。此種屬於常識上的假定，建立了因果關係中的方向性：也就是從原因到結果。你或許會問，原因怎麼可能會出現在其所影響的結果之後呢？這當然是不可能出現的狀況。不過，時間順序只是因果關係的條件之一。在推論因果關係時，時間順序只是個必要條件、但卻非充分條件。有時候人們會犯下一個錯誤，就是在談論到「原因」時，單獨只以時間順序爲根據，而不考慮其他的因素。例如，在 1968 年時，就在太陽黑子出現一波頻繁活動之後一天，在十來個美國城市裡發生了好幾起種族暴動事件。此種時間上的先後順序，並無法在太陽黑子和種族暴動之間建立起因果關係。畢竟，所有先前人類的歷史都是發生在某個特定的事件之前。時間順序這個條件只是讓我們在考慮因果關係時，可以把一些在時間上較晚發生的潛在原因排除在外而已。

在橫斷性研究中，要建立時間的順序可能會相當不容易。例如，某一位研究者發現，接受過大量正式學校教育的人，會比其他人表現出較少的偏見。那麼，此種情況究竟是意味著接受較多的學校教育能夠降低偏見，還是因爲有高度偏見的人會避免接受學校教育呢？這裡還有另外一個例子。在我的課堂上得到高分的學生說我是一位優秀的老師。那麼，這是否代表我教學認真，學生收穫很多，因而使他們得到高分；還是因爲得到高分讓他們很高興，所以他們就投桃報李，說我是一位優秀的老師呢（也就是說，高分導致較高的評價）？這是一個雞生蛋、蛋生雞的問題。爲了解決這個問題，研究者必須帶入其他資訊，或是設計一項研究來驗證時間方面的順序。簡單的因

範例　3.3

因果關係的三個要素

我曾經讀過一則報導，內容是關於幾位政治人物參訪了位於芝加哥地區的一所天主教學校，而這所學校在教育兒童方面，比許多公立學校還要成功。隔天，這些政治人物召開了一場記者會，他們在會中主張制訂新的法令，並支持把部分徵稅收入的經費撥給天主教學校。由於我本身是個想要讓兒童獲得良好教育的人，所以我對於這件事抱持著高度的興趣。然而，身為一名社會科學家，我還是必須以批判的方式來評估這件事。這些政治人物的理論是，天主教學校比公立學校更能夠產生良好的學習效果。而且他們也提出了因果關係中的兩個要素：時間的順序（首先，兒童們進入一所天主教學校，然後他們的學習效果獲得改善），與關聯性（就讀於天主教學校的兒童，在課業方面的表現，比就讀於公立學校的兒童還好）。社會研究者們都知道這些資訊並不充分。他們首先會試著去排除一些可能的替代解釋，然後會試著去瞭解造成此種因果關係的機制（也就是說，到底天主教學校裡有什麼因素，能夠幫助學生學得更多、更好）。舉個例子，像是這些政治人物就無法排除下列這個可能的替代解釋：就讀於兩種不同類型學校裡的兒童，擁有不同的家庭環境。由於家庭環境會影響到學習的成效，所以此種情況則會導致學習成效方面的差異。如果就讀於兩種不同類型學校的學生，其家庭背景（例如，家長的教育程度及收入、家庭的宗教信仰以及信仰的虔誠程度、雙親家庭或是單親家庭、家長對於孩子教育的關心程度等等）完全相同的話，那麼這些政治人物的看法就是正確的。果真如此的話，那麼問題的焦點就在於，天主教學校到底做了什麼事，改善兒童的學習成效。然而，如果這些學生的家庭背景有著天壤之別，那麼這些政治人物便犯了大錯。不幸的是，政治人物很少接受過社會研究的訓練，而且大多數政治人物經常作出一些快速的、宣傳意味濃厚的決定，他們並未經過仔細的推理過程，也缺乏耐心去等候精確的實徵調查結果。幸運的是，社會學家柯爾曼（James S. Coleman）和其他人已經針對這個議題進行過深入的研究（請閱 Coleman and Hoffer, 1987）。

果關係是單向性的，意指其運作的唯一方向是從原因到結果。較爲複雜的理論會詳細說明*交互*作用的因果關係（reciprocal-effect causal relations），也就是一種*互*爲因果的關係、或者是同時發生的因果關係。例如，努力學習可以使得學生得到好成績，但是得到好成績也可以激勵學生繼續努力學習。理論通常會包含一些交互的或是反饋的關係，但是這些關係卻非常難以驗證。有些研究者會把單向性的關係稱爲*非循環*（*nonrecursive*）關係，而把交互作用關係稱爲*循環*（*recursive*）關係。[5]

2. 所謂的**關聯性**（association）指的是兩個現象以某一種模式一起發生、或是似乎一同發揮作用。人們經常把*關係*（*correlation*）這個詞與關聯性混爲一談。關係這個詞具有特定的技術性意義，必須符合一些特定的統計條件。而關聯性則是一個較爲一般性的觀念。相關係數（correlation coefficient）是一個統計方面的測量指標，主要是用來指出關聯性的強度，但是其實還有其他的方法，可以去測量關聯性的強度。有時候研究者會把關聯性稱爲*共同變異*（*concomitant variation*），因爲兩個變項會同時發生變化。圖 3.2 中所描繪的是 38 名來自低收入地區的人們，以及 35 名來自於較高收入地區的人們。你是否能在其中看出種族（以較淺和較深的顏色來代表不同的種族）和收入水準之間的關聯

關聯性　兩起事件、特質、或是因素一起出現的情況，也就是當其中之一發生或出現時，另外一個也會發生或出現。

性？有些人會把關聯性誤以爲是真正的因果關係。例如，當我還在讀大學時，只要是在星期五的考試，我都會得到高分，而若是考試在星期一，我則都會得到低分。因此，在考試的日期和考試分數的高低之間，就存在著關聯性。然而，此種關聯性並不代表考試日期導致了考試成績的高低。相反地，產生此種關聯性的真正原因，是因爲我每周末都會去打工大約二十個小時左右，所以在星期一時就會非常疲累。如果你無法找到事物之間的關聯性，那麼因果關係成立的可能性就很低，而這也正是爲什麼，你會想要找出事物之間的關係、與用其他方式來測量關聯性的原因。然而，僅僅因爲你發現了關聯性，並非就意味著你也找到了因果關係。關聯性是因果關係的必要條件、卻不是充分條件。換句話說，你需要找到關聯性來建立因果關係，但僅僅只有關聯性卻是不足的。

一項關聯性並不需要是完全的（也就是說，每次當某個變項出現時，另一個變項也會出現），才能顯示出因果關係。在剛剛那個牽涉到考試成績與考試日期的例子裡，如果我在十次星期五的考試中，得到七次 A、兩次 B、以及一次 C，而在十次星期一的考試中，我得到六次 D、兩次 C、以及兩次 B，那麼，這兩者之間仍然具有關聯性。也就是說，考試成績與考試日期之間確實存在著關聯性，但是兩者之間並不是完全相關。而在圖 3.2 中所顯示的種族與收入水準之間的關聯性，也同樣是一個非完全相關的例子。

3. *排除其他可能影響因素*（*eliminating alternatives*）的意思是，我們必須證明，某項結果確實是由原因變項、而非是由其他事物所導致的。此種情況也被稱爲*非虛假關係*（*no spuriousness*），這是因爲一個表面上成立的因果關係，若實際上是由其他尚未被辨識出來的因素所導致，就會被稱爲是一種虛假關係（a spurious relationship），而我們將在第六章裡進一步討論此種情況。雖然我們可以觀察到時間

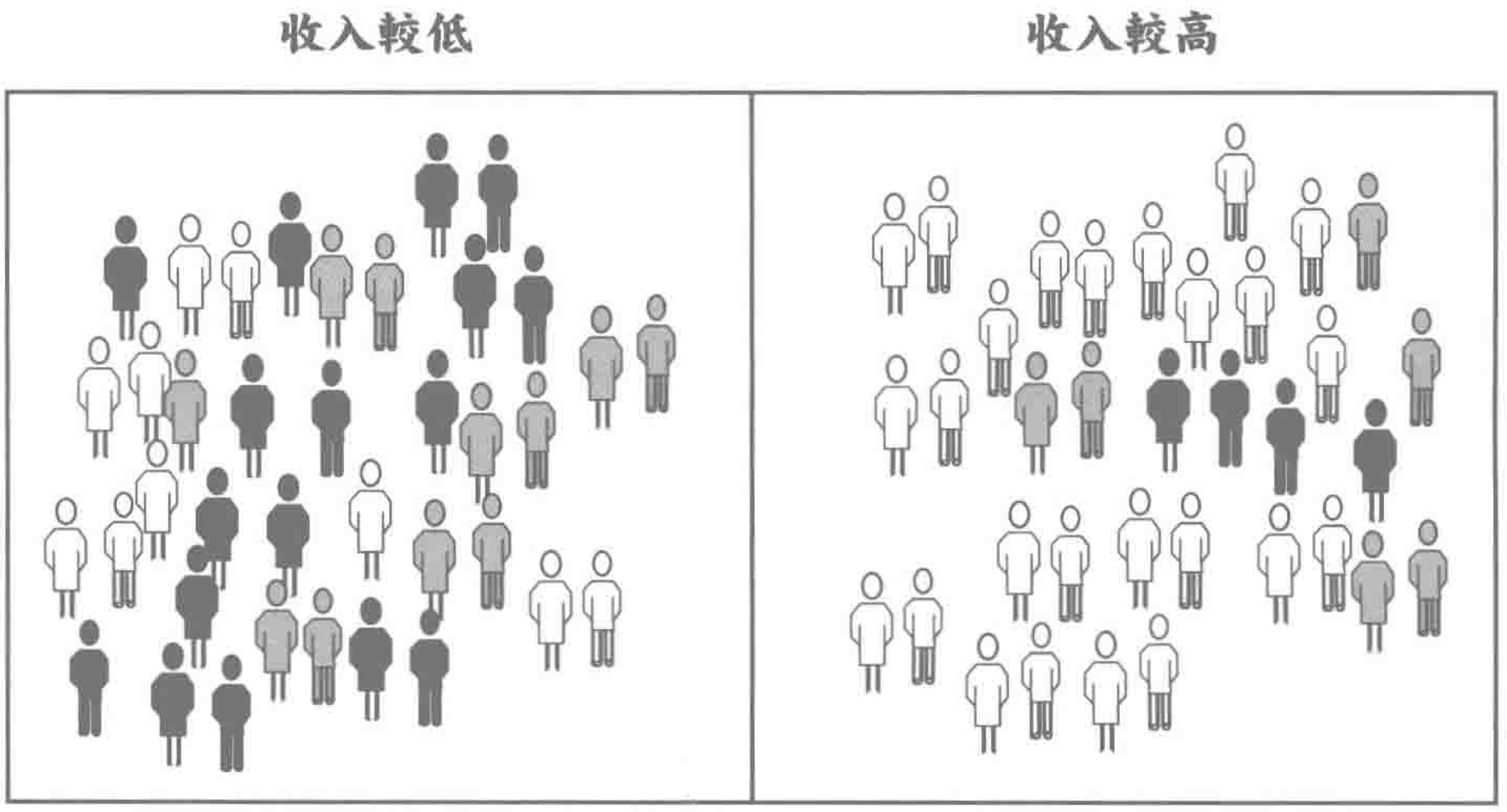


圖 3.2　收入與種族之間的關聯性

的順序以及關聯性，但是我們仍然無法在經驗上排除所有合乎邏輯的可能影響因素。排除所有可能的影響因素是一種理想狀態，而這意味著我們只能以間接方式來排除一些可能的影響因素，或是排除一些更爲明顯的替代解釋。在實驗當中，我們可以在研究設計本身加入控制組，以排除掉其他的替代原因。而且，除了主要的因果變項外，我們也會把實驗情境隔絕於所有其他變項的影響之外。至於在非實驗研究中排除替代原因的方式，則是試著去辨識可能的替代原因，並去測量這些原因。此種作法在調查研究中相當普遍。一旦我們測量了可能的替代原因後，我們就能夠利用統計技術去瞭解，原因變項或是其他變項是否對結果變項產生了作用。

4. *詳細說明在因果關係中的運作機制*（*specifying the mechanism in a causal relationship*）意思是指，當我們建立一項因果關係時，我們所必須說明的不只是兩個互相關聯的變項而已，也就是說，「一個令人滿意的解釋，要求我們也應該要詳細說明社會中的『齒輪與輪子』（"cogs and wheels"）」（Hedstrom and Swedberg, 1998: 7）。我們不能夠只有簡單地說某個自變項（independent variable）和依變項（dependent variable）之間具有關聯性，因爲這就好比說，此種關聯性是透過某種未知過程的「黑盒子」所形成的。因此，一個完整的因果解釋除了要確認因果關聯性的存在之外，也要詳加說明造成此種關係的**因果機制**（causal mechanism）。

例如，假設我們發現到，在一個人的社會階級和其健康情況之間，存在著很強的關聯性。我們可能會提出的「理論」，像是社會階級較高的人比階級較低的人活得較久，也較少生病。然而，僅僅只有說一個人的社會階級會影響其健康情況是不夠的。我們還必須解釋社會階級爲什麼、以及如何造成此種影響。簡而言之，我們應該正確地描述，社會階級到底如何造成健康情況的差異。我們或許會*猜想*，較高的社會階級提供人們較多的社會資源（知識、社會聯結、休閒時間、彈性的時程表等等），讓人們能夠享用健康食物、感受較少的壓力、從事運動等等，而這些事情都能使人們的身體變得比較健康。在這個例子裡，社會資源就是聯結社會階級和健康情況之間的機制（此處的資源指的是「擁有知識」、「認識對的人」、以及「擁有獲得各種機會的管道」等等）。

要看出在一個完整因果解釋中的機制，可能是一件相當困難的事，尤其是在自然科學中。爲了要解釋我們所能觀察到的現象，我們可能會假定，在次原子當中、或者在遙遠的銀河系當中存在我們看不見的機制。而隨著研究的進展，對於那些我們先前只能在理論中預測其存在的機制，我們逐漸能夠觀察到這些機制的大略樣貌。即使我們現在無法直接觀察到某種機制，我們仍然可以描述我們認爲此種機制運作的方式。

若是我們相信某一個過程聯結了輸入與結果，那麼我們就可以利用此種過程的模型來澄清因果機制。在經濟學中，市場是一個常見的機制，此種機制是一個在獨立的買方與賣方之間，進行交換的過程，而且每個買方與賣方都有各自的需求與資源。市場可以解釋供給—需求的關係是如何運作的。而在社會學裡，一個經常被使用的機制，就是墨頓的自我實現預言。當某個情境的定義激勵了一些行爲，而這些行爲使得錯誤的定義變爲真實情況

因果機制　因果解釋中的一個部分，用以詳細說明主要自變項如何影響主要依變項的過程。

時，就可以說出現了自我實現的預言。在某個預言中的「負向反饋」（"negative feedback"），在某個時間點上把人們的信念和行爲，聯結到之後所產生的結果。自我實現預言的經典例子之一，就是銀行發生擠兌的情況。例如，某間銀行的財務狀況可能非常穩定，但是，突然傳出一個錯誤的謠言，說這間銀行即將倒閉。雖然這個關於現狀的定義是不正確的，仍然使得許多人急急忙忙地去把自己的錢提領出來。由於人們在短時間之內提領了鉅額金錢，使得銀行的財務體質因而變得虛弱。而由於這間銀行的財務體質變弱，因此使得更多關於它即將倒閉的謠言甚囂塵上。這些新的謠言又造成更多人趕著把自己的錢提領出來。到了最後，不斷增加的恐懼感（對於現狀的錯誤定義）以及大量的提領行爲（基於上述錯誤定義而產生的行爲），就造成這間銀行真的倒閉了（錯誤的定義最後成真）。

有時候當我們在陳述理論時，是把理論當成像定律一樣的推論方式：當 X 發生時，Y 就會發生。然而，此種「理論」並不是完整的解釋（Elster, 1998），因爲此種理論仍然需要加上因果機制的說明。機制通常比一般性定律更爲具體，但是又比特定具體的例子更爲一般性。在一個完整的解釋裡，機制或許是關於機會或者是個人欲望的排列與整理，而這些機會與個人欲望，比特定的一個機會或一個欲望，更爲一般性。但是比起像是一句定律的陳述，則其一般性的程度又較低。除此之外，機制也會增加複雜性。我們不再僅僅使用簡單的定律（如果 B 的話，就會導致 R），而是在具體的情境中發現到，如果 B 的話，有時候會導致 R，但是其他時候則可能導致 P 或 D。機制能夠解釋爲什麼 B 不一定總是導致 R，有時候會導致其他的結果。舉例而言，我們可能會認爲，當經濟不景氣時（B），人們就會造反（R）。然而，當我們研究了許多具體情境之後，我們發現到情況似乎不一定總是如此。換句話說，當經濟不景氣時，人們有時候會造反，但是有時候人們會變得相當消極被動，並且認份地接受自己的命運（P），而在其他時候，人們可能會彼此爭鬥，並且走上自我毀滅的道路（D）。對於一個完整的解釋，我們必須在其中包含相關的機制，而該機制能夠告訴我們，經濟不景氣的情況在什麼時候會造成上述其他結果。

5. *概略描述因果鍊*（*outlining the causal chain*）是一種過程，其目的是用來評估組成因果鍊的每個部分。以下是一個因果理論中的一項關聯性：失業率上升，會導致兒童受虐案件增加。如果我們想要解釋爲何兒童受虐案件數量增加的原因，那麼我們可能會認爲，此種增加的狀況是由失業率上升所導致。爲了要「解釋」兒童受虐案件增加的現象，我們必須確認造成這種現象的主要原因。但是，如果要使解釋更加完整，我們還必須詳細說明造成此種現象的過程（也就是說，我們必須找出一個因果機制，並將此機制放入因果鍊之中）。在這個理論中的機制，就是那些失去工作的人們的處境。當他們失去工作，他們會覺得失去了自我價值，也會覺得壓力變大。接著，由於他們失去自我價值且感受到極大的壓力，他們比較易於感到挫折，也變得暴躁易怒。他們內在的社會控制力量逐漸變弱，生活模式也因而分崩離析。感到極度受挫、內在社會控制力量低落的人，可能會對那些與他們有親密個人接觸的人（例如，朋友、配偶、孩子等等）展現暴力行爲，藉此表達他們的不滿以及憤怒。尤其是當人們無法以行動去對抗那些造成其憤怒的真正來源（例如，雇主、政

府機構、或是「經濟力量」等等）時，此種情況會特別明顯。此種機制是一個較大的過程或因果鍊的一部分，而此種機制發生的時間點是在最初的原因（失業）之後、以及結果（虐待兒童）出現之前。

我們可以測試因果鍊中的每個部分。除了確認失業率與虐童案件是否一起發生之外，我們也可以同時考慮失業是否會增加挫折感、有挫折感的人是否會對家庭成員暴力相向等等。典型的研究策略之一，是把因果鍊依照不同的部分加以區分，然後再以資料評估因果鍊中的每個部分。

因果關係的圖形　我們可以利用文字、圖片、或是以圖文並用的方式，來表達因果關係及理論。我們經常會提出因果關係的圖形，以提供一個因果關係的簡潔圖像，因爲這麼一來，可以使得其他人更加容易地，很快一眼就能看出兩個變項之間的因果關係。此種符號式的表達方法，能夠補充言語描述上的不足，並能夠簡潔地傳達複雜的資訊。

最簡單的圖形，就是如圖 3.3（a）中所顯示的那種雙變項的模式。我們利用字母、圓圈、或是圖框來表示變項。傳統的作法是以英文字母 X 來代表原因，而以英文字母 Y 來代表結果，而箭號則顯示了因果關係的作用方向（例如，從原因到結果）。有時候，當原因不只一個時，我們會用下標的方式來表達不同的原因（例如：X_1，X_2）。我們會以帶有箭號的直線來表示變項之間的關係。因果關係是以直線來呈現的。傳統上，我們會用兩端都有箭號的曲線，來顯示兩個變項之間具有關聯性，但這並不代表兩個變項之間有因果關係的存在。

正向與負向因果關係　因果關係可能是正向的、也可能是負向的。若是研究者沒有特別加以說明，就意味著在原因變項及結果變向之間，存在著正向關係。所謂**正向關係**（positive relationship），是指如果原因變項的值越高，則結果變項的值也會越高。例如，一個人接受正式學校教育的時間越長，則此人的預期壽命也會越長。至於**負向關係**（negative relationship），意指當原因變項的值越高時，結果變項的值反而會越低。例如，隨著一個人接受正式學校教育的時間越長，則此人所抱持的偏執態度和偏見就會越少。在圖形裡，我們以＋號代表正向關係，而以－號代表負向關係。圖 3.3 所顯示的，是一些可以用圖形來表示的因果關係範例。研究者通常不會使用圖形來表示如圖 3.3（a）中那種非常簡單的關係，但是，研究者們已經發現，隨著因果關係中的變項數量，與複雜度逐漸增加，使用圖形來表示因果關係，是非常有幫助的作法。

有些時候，某個原因會先透過其他中介變項或是條件，才會對結果造成影響。這意味著原因只有在某些條件、而非其他條件之下，才會發生作用。例如，在允許個人自由以及合法離婚的現代社會中，早婚通常會導致離婚，但是在極爲傳統的社會，

正向關係　在兩個概念或測量變項之間的一種關聯性，當其中之一的值增加時，另外一個的值也會隨著增加；當其中之一出現時，另外一個也會隨著出現。

負向關係　在兩個概念或是測量變項之間的一種關聯性，當其中之一的值增加時，另外一個的值則會隨著減少；或是當其中之一出現時，另外一個則不會出現。

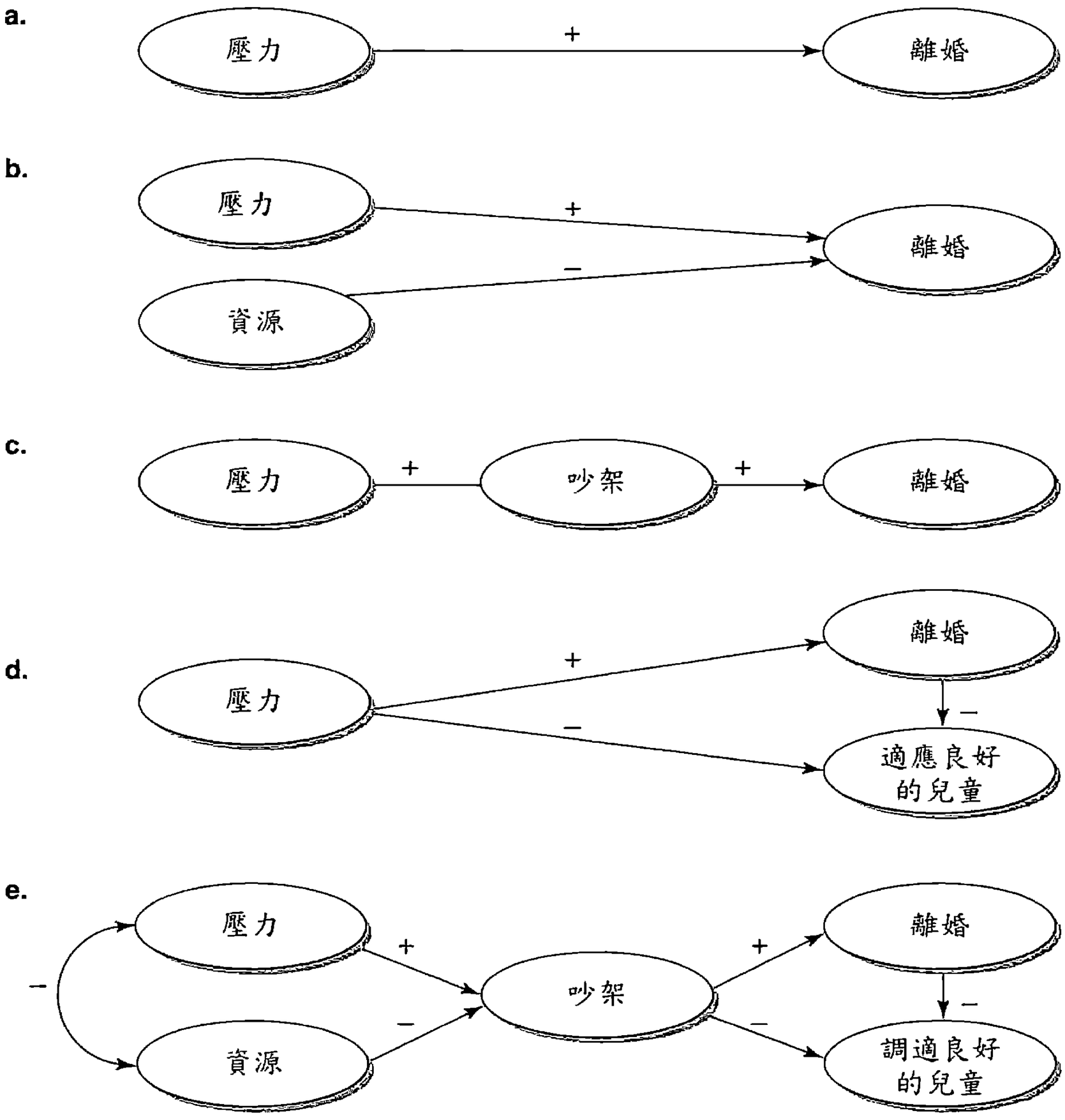


圖 3.3　因果關係圖形

每個圖形中因果關係的解釋

a. 壓力（財務壓力、社會壓力、情緒壓力等等）的程度與一對夫妻離婚的可能性呈現正向關聯。

b. 壓力的程度與一對夫妻離婚的可能性呈現正向關聯，但是資源（財務資源、社會資源、情緒資源等等）的數量則與離婚的可能性呈現負向關聯。

c. 壓力的程度與一對夫妻吵架的頻率呈現正向關聯，而吵架的頻率則與離婚的可能性呈現正向關聯。

d. 壓力的程度與一對夫妻離婚的可能性呈現正向關聯，但是與這對夫妻擁有情緒上調適良好的兒童的可能性呈現負向關聯。除此之外，離婚過程本身對於兒童在情緒方面的調適會造成負面影響。

e. 壓力程度與資源數量彼此之間呈現負向關聯（也就是說，擁有較多資源的人通常較不會經歷壓力的感受、或比較能處理壓力）。壓力程度與一對夫妻之間的吵架頻率呈現正向關聯，但是資源數量則與吵架頻率呈現負向關聯。吵架的頻率與離婚的可能性呈現正向關聯。夫妻吵架及離婚這兩者，與這對夫妻擁有情緒上調適良好的兒童的可能性呈現負向關聯。

則無此類現象。因此，在基本的因果關係中，發揮中介作用的第三個因素，在圖形中就會以第三個字母或是圓圈來表示，這個字母或圓圈會位於原因和結果中間，也就是那條帶有箭號的直線中間。第二章中，你曾經讀過一項關於接觸假說的研究，在該項研究中所使用的，正是因果解釋。因爲該項研究的研究者發現，越多的接觸，越能導致負面態度的減少（請同時參閱範例 3.4 解釋種族衝突）。

結構性解釋　在一項因果關係的解釋裡，一個或是多個因素，可能引起其他因素的反應，這就好比一顆滾動的球撞上了其他的球，然後使得這些球也開始滾動一樣。對比之下，**結構性解釋**（structural explanation）的邏輯，則是找出一個社會過程、事件、或在一個較大型的結構裡，因素所處的位置。這個結構就像是一張蜘蛛網、一個有輪輻的輪子、抑或是一個內部零件相互連結在一起的機器。結構性解釋在解釋社會生活時，所使用的方式是，說明某個部分如何與整體較大的結構結合在一起。一項因果解釋可能會告訴我們，B 之所以會發生，是因爲 A 所導致的。但是，一項結構性解釋卻可能會告訴我們，B 發生的原因，是因爲 B 位於一個較大型的結構之內，而這個結構可能會阻擋、或是提供 B 接觸結構中其他領域的通道。

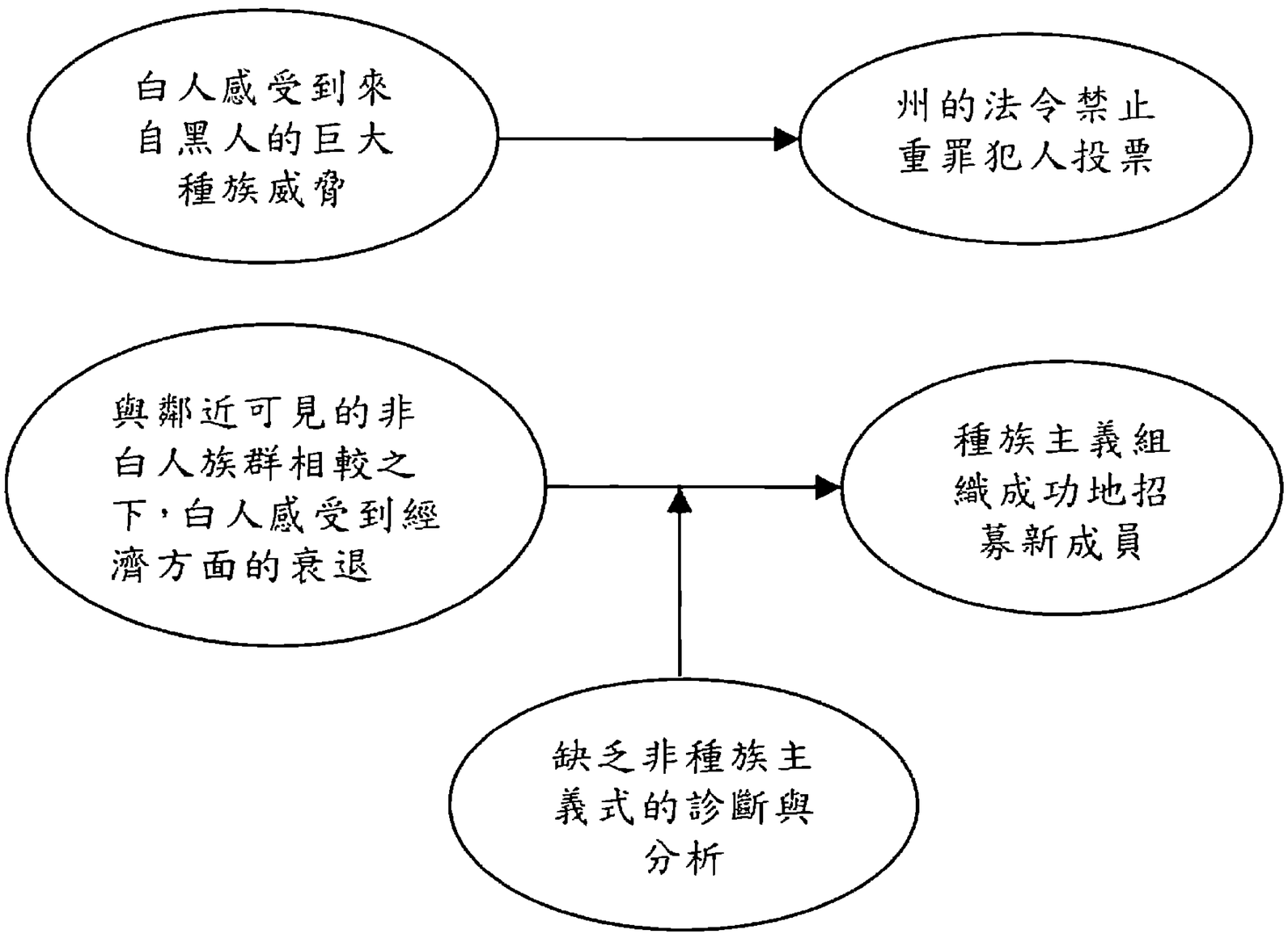
結構性解釋　一種關於事件發生原因、事物運作方式的理論性解釋的類型，其表達方式是概略描述出一個整體的結構，並強調在結構中不同地位之間的位置、相互依賴性、距離、或關係等等。

範例　3.4

解釋種族衝突

貝倫斯、烏干、以及孟札（Behrens, Uggen, and Manza, 2003）三人提出了一個關於美國境內重罪犯被褫奪公權的因果解釋。他們發現，在所有先進國家中，對於曾因為犯過罪而被判刑的人而言，美國的相關投票法規是限制最嚴格的。在美國各州層級上，投票法規的差異非常大：有些州完全沒有任何限制；其他州則禁止曾經入監服刑的重罪犯人投票；而在有些州裡，曾經入監服刑過的重罪犯，終身都被禁止投票。這些作者們延伸了一個現存的理論，也就是「種族威脅假說」（"racial threat hypothesis"），以解釋為什麼有些州的投票法規具有極高的限制性，而其他州的投票法規卻沒有此種現象。其他人則已經發展出這個理論來解釋種族之間的經濟競爭。在某些州裡，由於居民絕大多數都是白人，因而在民權時代之前阻止黑人擁有任何投票權利，但在民權法案通過之後，當地的白人就無法再阻止黑人去投票。在上述這些州裡，這些作者們發現了高度的種族威脅，而此種威脅的存在形式是那些隨時可能會發洩其憤怒、而且相當強大的黑人族群（例如，眾多的黑人人口、以及許多被囚禁在監獄裡的黑人）。作者們檢視了限制性的投票法規通過的年份、該法規所包括的限制類型、以及在該州人口和監獄人口中的黑人比例。在此項研究中，監禁的議題具有相當重要的角色，因為黑人成為重罪犯的可能性遠比白人還要高。該理論認為，「種族威脅」程度最高的州，將會擁有限制性最高的投票法規，這是因為具有限制性的投票法規，可以視為等同間接拒絕賦予黑人任何投票權。作者們詳細記錄了時間順序，並發現了在種族組成及限制性法規之間，存在著關聯性。而此種情形正好符合種族威脅假說的內容。在這項鉅觀層次的研究裡，主要原因是在監獄裡的大量黑人罪犯，主要的結果是具有限制性的投票法規，而分析單位則是州（請參閱本範例後面所附圖形）。

麥可貝（McVeigh, 2004）也利用因果解釋來研究為什麼在美國境內，白人的種族主義組織在某些地區比在其他地區還要容易成功。種族主義組織的訴求對象是曾經歷過向下社會流動的白人，而其提供的訊息則是把白人所遭遇到的種種困境怪罪於非白人身上。麥可貝所提出的假設是，種族主義組織最為成功的地方，應該是當地的情況與種族主義者所宣稱的狀況最為符合的地方。更確切地說，他預測在那些種族組成較為多元、經濟情形不穩定、以及收入不平等的狀況逐漸加劇的地區裡，白人種族主義者所傳遞的訊息應該是最能引起共鳴的。除此之外，他也預期，在分析整體情況時，那些不包含任何替代性解釋的、而且不包含任何非種族主義解釋的訊息，將會是最為成功的訊息。他認為，在白人教育程度最為混雜的地區裡，替代性解釋的數量應該會是最多的，因為教育程度較高的白人會把引起經濟衰退的其他替代性理由與訊息（例如，像是全球化的競爭、科技的改變、以及缺乏相關的技能等等）散播給其他的白人。他所提出的假定是，下列兩種原因的結合，能夠解釋種族主義組織在某些地區的成功：白人在經濟方面落後於鄰近可見的非白人族群，及缺乏對於此種情況的非種族主義式的診斷與分析。在這項研究中，麥可貝所測量的是各郡的經濟情況、種族主義組織的成功程度、及各郡之間白人不同的教育程度，而各個不同的郡則是本項研究中的分析單位（請參閱本範例後面所附的圖形）。



採用結構性解釋的三種主要理論類型爲次序理論、網絡理論、以及功能理論（請參閱圖 3.4）。

1. **次序理論**（sequential theory）強調的是事件發生的順序或是次序。此種理論確認了在一個隨時間而呈現的發展模式中，必要的早期步驟與可能的後續步驟。次序理論描繪出一整組有條理、有次序的階段或是步驟。幾乎所有的人、組織、或事件，都會依循一定的次序而逐步發展。對於某個特定過程而言，像是兒童的道德發展、親密關係的培養、家庭的組成、都市的擴張、組織的成長或凋零、衝突的升高或是

次序理論 一種使用結構性解釋的理論，此種理論會概略描繪一個次序性的模式，並詳細說明各種事件發生的順序、階段、步驟、或是時期。

解決、以及社會的發展等等，都可能會有單一的發展途徑、或者是範圍狹窄的其餘發展途徑。除了辨識出一個過程中的步驟或階段之外，次序理論也會解釋每個步驟之間的移動速度、在某個階段的停滯狀況、在一個過程中可能會引發不同方向，或步驟的關鍵轉捩點。次序理論可能會指出必要及選擇性的步驟、或是某一個特定的先前步驟如何限制了可能的後續步驟。次序理論並不是一個因果理論，因爲位於一個較早步驟的事件，不會*導致*事件沿著次序移動。相反地，一個有明確階段性的次序結構，限制了可能發生的事件。因此，一個次序理論可能會說，除非階段 A 已經完成了，否則不可能會移動到階段 B，而從階段 A 進行到階段 C 的唯一方式，就是先通過階段 B。

在第二章裡，你曾經讀過馬哈尼（Mahoney, 2003）關於拉丁美洲國家的研究，而該項研究所使用的就是次序理論。馬哈尼發現到，在某個拉丁美洲國家發展過程中的一個早期階段，也就是在殖民主義的那個時代裡，形塑了該國在之後階段的發展路徑方向。歐斯特爾、強森、以及摩爾提默（Oesterle, Johnson, and Mortimer, 2004）在他們針對年輕人之間的志工活動所進行的同樣本多次研究，提出了一個次序理論。作者們採取了一種「生命進程」（"life course"）的觀點，根據此種觀點，「在不同的生命階段，各種角色與活動的意義也會有所變化」（p. 1124）。因此，在一個人生命中某一個特定階段所發生的事件，其所造成的影響，會和同一事件發生在其他階段時完全不同，而且此一事件的影響也會形塑了在後來階段中所發生的事件。作者們注意到，轉變爲成人的過程，是一個極爲重要的階段，因爲在此階段中，人們必須去學習新的社會角色及身爲成人的期望。作者們檢視了一群九年級學生（約爲 15 到 16 歲）的同樣本資料，資料的蒐集時間從 1988 年開始，一直持續長達九年的時間，研究中的受試者年紀從 18 到 19 歲，成長到 26 到 27 歲左右爲止。作者們發現，在受試者先前的活動階段，會強烈地影響到後續階段發生的事件。那些在早期階段裡（18 到 19 歲左右）一直在工作、或是專職在照顧孩子的人，比起那些主要活動是全心忙於學校課業的人，在後來階段裡（26 到 27 歲左右）去從事志工活動的可能性明顯低落許多。除此之外，在早期階段有參與志工活動的經驗，也能用來預測一個人在後來的階段裡是否會去參與志工活動。

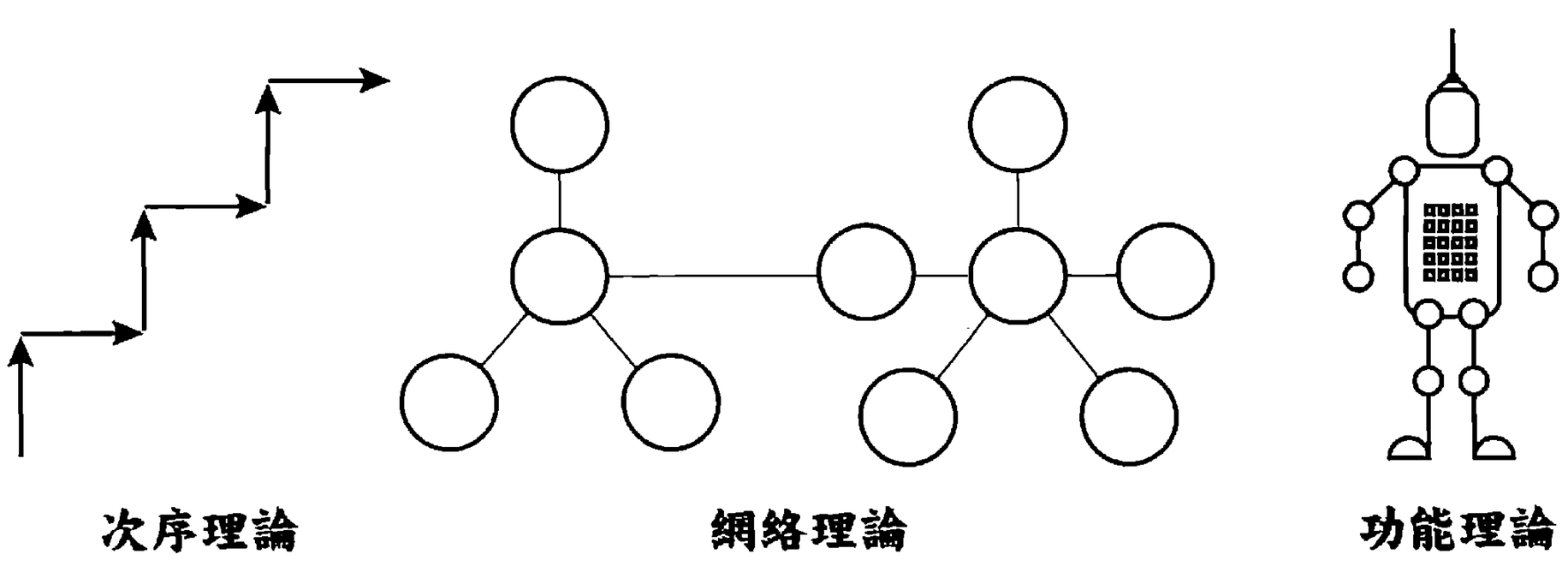


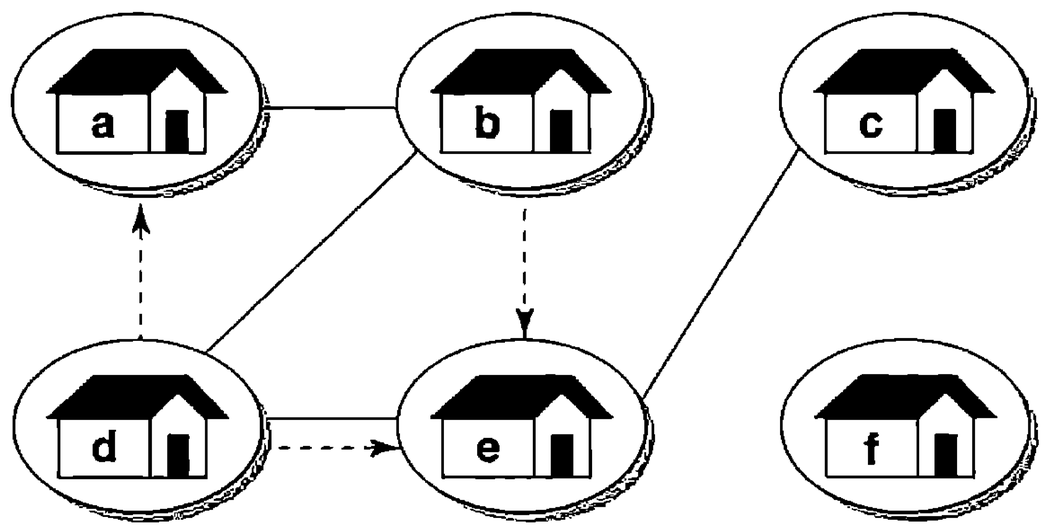
圖 3.4　結構性解釋的各種形式

2. **網絡理論**（network theory）以在網絡中的位置來解釋社會關係。網絡理論提出解釋的方式，是去指涉一個網絡內的關係位置，或在位置之間聯結的大小與形狀、類型以及存在，各種聯結的重疊部分或密度，網絡裡的中心點，網絡中的位置或節點之間的流動。[6]所謂的位置，可能是在由人們、組織、城市、國家組成的關係網絡中，各個不同的點或節點。位置以及網絡的結構，可以幫助我們解釋在一個網絡中的溝通便利性、權力關係、階層關係、以及流動的速度等等。

網絡理論家在提出解釋時，通常會指出一個模式、一組語法規則、或是結構等等。而他們所提出的解釋，顯示出事件如何契合於一個較大的模式、或是一個更爲大型的聯結系統。網絡理論是一種推論的形式，此種推論形式類似於用來解釋人們爲何以某種特定方式使用語言。例如，某種語言可能有一組語法規則，說明了 X 會與 Y 一起出現，或者是一個句子裡必須要有一個名詞及一個動詞。在解釋此種情況時，也等同於指出那條涵蓋此種情境的語法規則。

許多研究檢視了社會網絡、並描繪出網絡的結構，以當作解釋社會的一種方式。恩特韋索、佛斯特、林德法斯、與金田（Entwisle, Faust, Rindfuss, and Kaneda, 2007）研究了泰國某些地區中，在部分村莊裡的網絡。作者們發現，在不同的村莊裡，透過親屬關係或是其他社會聯繫，而把人們聯結在一起的網絡都各自不同：「某些村莊裡的網絡相當稀疏，但在其他村莊裡則相當密集；在某些村莊裡網絡滲透性很高，在其他村莊裡則較低。除此之外，此種網絡之間的變異性，具有相當高的重要性」（p. 1524）。對於和鄰近村莊之間的關係、在某個村莊裡的經濟活動、以及人們是否會遷出某個村莊等等方面，這些網絡具有相當大的影響力。網絡結構會形塑各種活動的流動，及村莊之間的合作程度。爲了具體說明這些研究發現，作者們提供了一個關於六個家族的圖形。在圖形中，實線代表了彼此之間有血緣關係的兄弟姐妹，而虛線則代表那些來幫忙收割稻米的人們。圖形中顯示出 a、b、d 及 e 這四個家族會一起工作。在 a 和 d、或是 b 和 e 之間，沒有直接的家族關係，但因爲這些家族透過在網絡中的 d，而有了間接的連結，因此這些家族之間才會共同合作。網絡所產生的關鍵影響之一，是對於社會凝聚力的影響。如同作者們所說明的：「平均而言，當某些網絡中的活動者擁有更多與他人的聯繫時，其凝聚力會高於那些擁有較少聯繫網絡的活動者……當網絡的凝聚力越強時，資訊透過社會聯繫而傳達到所有成員的可能性、及網絡成員之間合作進行活動的可能性也會隨著變高。」（p. 1508）換句話說，網絡會影響某個村莊裡活動發生的方式。更重要的是，內部許多密集的網絡聯結，其社會凝聚力會高於鬆散的網絡。所謂的*凝聚力*（*cohesion*），指的是相較於住在網絡稀疏的村莊裡的人們，那些住在網絡密集的村莊裡的人們，會分享資訊、共同合作、更加快速而輕鬆地完成各種工作（*American Journal of Sociology*, 2007: 1515）。

網絡理論　一種使用結構性解釋的理論，在此種理論中，強調的重點是在一個互相聯結網絡中的位置與聯結，以及該網絡的形狀或是整體模式。



—— 兄弟姐妹
--➤ 幫忙收割稻米的人

*資料來源：*The Construction of a Global Profession: The Transnationalization of Economics, by Marion Fourcade. *American Journal of Sociology*, Volume 112 Number 1 (July 2006): 145- 94 (page 151).

3. **功能理論**（functional theory）所採用的概念，是一個由一組互相依賴的關係所組成的系統。系統中各個不同的部分，會與其他的部分互相依賴，而且在結合起來之後，所有的部分會成爲一個整體而一起發揮作用。某個部分的成功或是失敗，會對其他部分及整個系統造成影響。系統可能是指一個家庭、一個社會團體、一個正式的組織、或一整個社會。功能理論認爲，系統若要長時間存續下去，必須要達到平衡或是均衡的狀態，以持續進行順暢的運作。如果有一個重要的部分失效了，系統就無法完成一項重要的功能，除非能夠找到該項重要功能的替代品。系統中的各個部分通常專精於、能夠更有效率、或是有效地完成不同的系統需求抑或功能，也因此符合一種特定的勞力分工模式。功能理論解釋各部分的方式，是指出這些部分以何種方式，融入所有功能結構之中。就像是在人體裡或是機器人內部的不同部分，每一個部分（例如，頭、手、腳等等）都會去執行特殊化的功能。

關於社會變遷的功能理論認爲，社會經過各種不同的發展階段而向前移動，從傳統移向現代。隨時間經過，社會分工變得越來越精細、越來越複雜，並且也會伴隨著個人主義的出現，而逐漸演變出一個更加專業化的勞力分工模式。這些發展會使整體系統，變得更有效率。專業化及個人主義可能會造成一些分裂的情況，因而需要作出系統方面的調節。此種分裂的情況或許會弱化執行系統功能的傳統方法。然而，新的社會關係類型將會浮現，以取代傳統方法，且這些新的關係類型也將執行相同的功能，以滿足系統爲了延續下去而產生的各種需求。

卡米金（Kalmijn, 1991）曾經利用一項功能性解釋，來說明美國民眾在選擇婚姻伴侶時，其選擇方式的變化。他所採用的是現代化理論，而該理論主張，現代化（工業化、都市化、以及世俗化）的歷史過程形塑了社會的發展。由於受到現代化所影響，人們在做事情時比較不會依賴傳統的方法。因此，就像家庭對於年輕人的控制力變弱一樣，傳統

功能理論　一種採取結構性解釋的理論類型，在此種理論中所強調的，是各種相互依賴的部分如何融入一個整體系統、並且執行各種功能以延續該系統。而在系統中則有各種不同的部分，這些部分在整體系統中扮演了互補的、及專門化的支援角色。

的宗教信仰及本土社區的聯繫，也逐漸變弱。人們不再終生住在一個小型的、同質性高的社區裡。年輕人脫離其父母和本土宗教組織而試著獨立生活。爲了要能繼續運作下去，每個社會都有一套作法來組織人們選擇其婚姻伴侶、以及找到與他們擁有共同基本價值觀的伴侶的方式。在過去，父母與宗教在選擇婚姻伴侶的過程中，扮演了重要角色，但是在現代社會，人們在大多數時間內都不會待在小型地區環境中，而是花較多的時間在學校環境。在學校環境中，尤其是在大學裡，人們會遇上一些未婚的人，而這些未婚的人都可能成爲他們的婚姻伴侶。在現代社會裡，教育是一個主要的社會化媒介，因爲教育會影響個人未來的收入、道德信念與價值觀、及休閒活動的興趣等等。隨時間經過，美國境內目前的潮流與趨勢，是人們與來自相同宗教背景的人結婚的情形越來越少，而與教育程度相仿的人結婚的情況則逐漸增加。在傳統社會中，家庭及宗教組織所具備的功能之一，是使人們融入社會的道德價值觀、連結人們與其結婚伴侶。然而在現代社會中，此項功能已經被高等教育所取代了。

詮釋性解釋 **詮釋性解釋**（interpretive explanation）的目的就在於促進我們對事件的理解，其作法則是把我們想要解釋的事物（例如，社會關係、事件、文化習俗等等）置於一個具體的社會脈絡與背景之中，而這些社會脈絡與背景擁有一個意義系統。詮釋性解釋的目標，是要讓其他人在心理上，能夠領會社會世界某些領域的運作方式、與把我們所要解釋的事物，置於該社會世界之中。而達到此目標的方法，則是幫助他人理解，在一個完整的世界觀及意義系統中，我們所要解釋的事物爲何。人們所擁有的主觀世界觀會形塑他們的行爲方式，因此，我們的目標就是要分辨出別人的推理方式和對於事物的看法。此種過程類似於去解碼一篇文字或是文學作品，因爲在這些文字或文學作品中，意義其實是來自文化符號系統的脈絡。

弗特瑞爾以及西米（Futrell and Simi, 2004）曾經利用一個詮釋性解釋來研究美國的白人權力運動。作者們的焦點在於運動、集體認同、或是一種共同的「我們」的感受。他們檢視了分散在許多組織裡（例如，三 K 黨、基督教認同團體、亞利安民族、或新納粹團體等等）的種族主義運動成員，而且這些組織裡的成員都是處於較大型主流社會的邊緣人。作者們也進行了調查，以瞭解當激進的信念可能會使這些成員失去工作並毀了他們大部分的人際關係時，這些成員會如何傳播他們的信念、並仍然參與激進主義。從 1996 年到 2003 年之間，作者們訪談了五十六名激進主義分子，也蒐集了許多資料。根據這些資料與記錄，作者們發現，這些成員會去參與小型的聚會（例如，研習團隊、儀式派對等等），而在聚會當中，他們會重新宣誓對於該團體的承諾，也不鼓勵服從於外界的主流價值觀。這些聚會都很小型、具有包容性、而且都是植基於持續性的個人關係之上。在這些聚會裡，成員們覺得他們可以安全地、公開地表達他們的種族意識型態。家庭成員和親密的朋友們也都支持這

詮釋性解釋 一種理論性解釋的類型，內容是關於事件發生的原因，及事物運作的方式，其表達方式則是利用社會建構的意義與主觀世界觀來進行說明。

些「文化的避風港」（“cultural havens”）。因此，這些成員就創造出並找出「自由的空間」，在這個空間裡，他們能夠在一群想法類似的人們之間，確認自己的種族信念。在這些表面上看來似乎是「正常」的活動裡（在家自學、研習團體、露營旅行、派對等等），成員們有機會表達他們的政治信念，因此，他們也可以藉由此種方式來減少他們自己與外面世界間的距離。他們建立了一個保護性的社會環境，讓他們可以維持並頌揚一種激進的意識型態與認同，亦能夠在表面上把此種意識型態與認同，僞裝成主流社會所認同的樣子。

理論的範圍架構

理論性的陳述在範圍架構方面也各自不同。其中的一個極端是**實徵推論**（empirical generalization），這是一種狹窄的陳述，所依賴的是具體的概念，並且符合實質理論。除此之外，實徵推論也是一種低層次的描述性陳述，其內容是關於一項據信在實徵方面能夠運作的關係。此種陳述的推論範圍可以超出某一個具體的個案或是一些觀察結果，但是超出的幅度不會太多。舉例而言，那些在年紀很輕時（低於 21 歲）就結婚的人，比那些在年紀較大時（超過 31 歲）才結婚的人更有可能離婚。我們也許會想詳細說明那些造成離婚現象的歷史、文化、或其他的條件，並藉由此種方式來證明推論的合理性。如果實徵推論包括了一項解釋，該項推論會變得簡單而具體，但仍然不是一個完整的社會理論。例如，早婚者之所以離婚的可能性較高，是因爲他們較不成熟的緣故。

中距理論的理論範圍較廣泛，並且在一個實質理論或形式理論中使用較多的抽象概念。關於一個離婚的**中距理論**（middle-range theory），可能包括與一些較爲抽象的概念連結的實徵推論。換句話說，離婚問題或許會成爲不穩定的婚姻關係，此較大觀念中的一部分，而結婚年齡則可能與某個人在生命循環中所處的階段、以及他或她所學習到的社會角色相連結。維持婚姻關係可能會被放入一個由其他的社會力量（例如，性別意識型態、社會接受或拒絕、影響離婚的法令、朋友或是家族團體、宗教壓力等等）所組成的情境脈絡裡。某項研究也許會詳細闡述並測試這個中距理論的特定部分，而長時間累積該理論中，部分大量的正面實徵證據，也有助於該理論發展成爲一個完整解釋。

理論架構（theoretical frameworks）〔也被稱爲*理論系統*（*theoretical systems*）或者是*典範*（*paradigms*）〕的範圍最大，而且處於和實徵推論相反的另一個極端。一個理論架構不只是一個形式理論，或是實質理論而已，還包含了許多特定的形式理論及實質理論。而這些理論可能擁有共同的基

實徵推論　一種狹猛的、準理論的陳述，該陳述是利用一些不非常抽象的概念來表達實徵的模式、或是描述實徵的規律性。

中距理論　一種介於一般性架構及實徵推論之間的社會理論，此種理論的抽象程度和範圍都相當有限，而其形式則為以實徵方式能驗證的一種陳述。此種陳述能夠與可觀察的現象連結在一起。

理論架構　一種非常廣泛的理論系統，其中包含假定、概念、及特定的社會理論。

範例 3.5

卡米金在「變動的邊界」（"Shifting Boundaries"）研究中的理論層次

理論架構 結構功能論認為，工業化以及都市化的過程，會把人類社會從傳統的形式改變為現代的形式。在這個現代化的過程當中，社會制度和習慣也會隨之演變，此種演變包括那些滿足社會系統基本需求、將人們社會化以接受文化價值體系、以及規範社會行為的制度與習慣。原本在傳統社會中，滿足需求和維繫社會系統的制度，將會被現代制度所取代。

形式理論 世俗化理論認為，在現代化的過程中，人們對於傳統宗教信仰及地方社區連結的依賴情形也會跟著改變。在傳統社會中，賦予社會地位的制度或機構（家庭、教會、以及社區）也會控制社會化的過程，規範社會生活。而在現代社會中，這些制度或機構，會被世俗的制度或是機構（例如，教育、政府、媒體）所取代，而這些新制度或機構會賦予*人們以成就為導向的地位*。

中距實質理論 一項關於通婚模式的理論認為，現代社會中的年輕人花較少的時間待在小型、地區性的環境中。而在此種環境，家庭、宗教、及社區都具有強大的影響力。相反地，年輕人花越來越多的時間待在學校環境裡。在這些環境中，尤其是在大學裡，年輕人有機會遇到其他未婚者。因此，在現代社會中，教育已經變成了一個促進社會化的重要媒介。教育會影響人們未來的收入、道德信念與價值觀、及休閒活動的興趣。因此，年輕人在選擇婚姻伴侶時，已較不會以共同的宗教或是地區連結，作為擇偶的考量條件，而是越來越以相同的教育程度作為考量重點。

實徵推論 美國人過去曾經傾向於和擁有類似的宗教信仰的人結婚。此種作法目前已被和教育程度相似的人結婚所取代。

本假定，及一般性概念。在社會學裡有幾個主要的架構，[8]我們在觀察與思考社會世界時，這些架構可以當作我們的指引方針或康莊大道。因爲這些理論架構提供了假定、概念、以及各種形式的解釋。舉例而言，每個架構可能會有它自己關於家庭、犯罪、或社會變遷的一套理論。有些架構（例如，象徵互動論）偏向於微觀層次的分析，而其他架構（例如，衝突理論）則以鉅觀層次見長。具體的研究很少會測試或比較整體架構。比較常見的情況是，某項研究只會去尋找特定的證據，以測試在某架構中的一個理論的部分（例如，在關於犯罪的衝突理論中的一項命題）。範例 3.5 是卡米金所提出的理論層次，而此範例清楚地說明了，在卡米金針對婚姻伴侶選擇方式的變化情形所進行的研究中，其理論的範圍與架構爲何。如同你將在第四章中看到的，每一個理論架構都和一個進行研究的整體取向有關。而在「資訊箱 3.2 社會學中主要的理論架構」裡，則簡短地描述了在社會學四個主要理論架構的假定關鍵概念。

動態的組合

你已經看到理論的許多面向了（請參閱重點摘要 3.2）。只有一些天真的、新進的研究者，才會錯誤地以爲理論與進行研究之間毫無關聯性，或是認爲我們只要單純地蒐集資料就可以了。如果我們試著在不使用理論的情況下進行研究，我們可能會發現，當我們要設計一個研究時，會覺得漫無目的、毫無頭緒。我們可能會浪費

資訊箱 3.2　社會學中主要的理論架構

結構功能論

主要概念　系統、均衡、功能失調、勞力分工。

關鍵假定　社會是由不同部分互相依賴所組成的系統，而且這個系統處於一種平衡或均衡的狀態。隨著時間流逝，社會將從簡單的型態，演化成複雜的型態。而在複雜的型態中，社會將有高度專業化部分。社會中不同的部分將會滿足社會系統中的不同需求或功能。而將整個社會結合在一起的，則是價值觀或價值系統方面的基本共識。

交換理論（也稱作理性選擇理論）

主要概念　機會、回報、認可、平衡、信用。

關鍵假定　人與人之間的互動，類似於經濟方面的交易活動。人們會給予並得到資源（象徵性的、社會認可的、或物質的），並且會試圖使其得到的回報最大化，同時避免痛苦、花費、以及難堪。交換關係會傾向於達到平衡。如果交換的關係不平衡，擁有信用的人就能支配其他人。

象徵互動論

主要概念　自我、參考團體、角色扮演、知覺。

關鍵假定　當人們進行社會互動時，會傳達並接受象徵性的溝通。人們會創造出對於彼此與社會環境的知覺。人們的行動主要依據他們的知覺。至於人們對於自己以及對於他人的看法，則依據與他人的互動而來。

衝突理論

主要概念　權力、剝削、鬥爭、不平等、異化。

關鍵假定　社會是由許多團體所組成的，而這些團體間的利益彼此互相衝突。對他人進行強迫、以及得到權力的企圖，在人類關係中一直都是存在的面向。握有權力者為了維持其權力，會刻意散播種種的迷思，或在必要時使用暴力。

時間去蒐集一些沒有用處的資料、缺乏明確觀念、並陷入模糊而茫然的思考陷阱中。除此之外，我們可能會發現，要組織論點、聚焦研究議題、對他人清楚地描述我們所進行的研究，也是一件相當困難的事情。

造成上述這些困境的原因其實很簡單。理論能夠提供我們如何進行調查及思考某個主題的架構，也能給予我們概念、提供基本的假定、引導我們朝向重要的問題、並建議一些瞭解資料意義的方法。理論幫助我們找出關聯性，並且看出研究結果所具備的更廣泛的重要性。打個比方，理論能夠幫助我們看見一整片樹林，而不是只看見一棵樹。

幾乎在所有研究中，理論都佔有一席之地，但其重要性則有所不同。一般而言，理論在應用與描述性研究中的重要性，會低於在基礎及解釋性研究中的重要性。理論在應用與描述性研究中所扮演的，或許是間接性角色。這是因爲在這些研究中，

重點摘要 3.2

社會理論的組成部分以及面向

社會理論的四個組成部分

1. *假定*
2. *概念* 在抽象層次（具體及抽象）、單一概念與概念叢、簡單與複雜（例如，分類、類型學）、以及範圍（狹窄與廣泛）等方面會有所不同
3. *關係* 關係的形式、命題、以及假設
4. *分析單位*

社會理論的五個面向

1. *建構理論的方向* 演繹的（從抽象到具體）或是歸納的（從具體到抽象）
2. *分析的層次* 微觀的、中觀的、鉅觀的
3. *理論的焦點* 實質理論或是形式理論
4. *解釋的形式* 因果的、結構的（次序的、網絡的、功能的）、或詮釋的
5. *理論的範圍架構* 實徵推論、中距理論、或是理論架構

概念通常比較具體，研究的目的也不在於產生一般性的知識。然而，我們仍會在描述性研究中，使用理論以使概念更精確、評估理論的假定、並間接地驗證假設。

理論不會維持不變，也就是說，理論是暫時性的、也會開放讓研究者加以修正。理論會以兩種方式成長爲更精確、以及全面性的解釋，藉以說明社會世界的組成和運作方式。隨著我們辛苦地以清楚而合乎邏輯的方式進行思考，理論也會不斷向前發展，然而，此努力畢竟有其極限。而某項理論獲得重大進展的真正方式，則是與研究發現交互參照與進行互動。

科學社群會依據實徵研究的結果，擴展並修改理論。如果我們採用的是演繹取向，理論將會引導我們的研究設計與研究結果的詮釋。我們可以根據研究結果來反駁、擴展、或修改理論。只有不斷地進行實徵研究來驗證理論，我們才真正能有信心地認爲理論的某些部分爲真。然而，某個理論的核心命題和中心原則就比較難以驗證，也比較不常被反駁。而在經過一個緩慢的過程之後，伴隨著反面證據逐漸累積，且無法在邏輯上相容於原本的理論時，我們可能會決定放棄或是徹底地改變這個理論。

如果我們採用的是歸納取向，我們所依循的過程會稍微不太一樣。在以歸納取向來建構理論時，我們會先從少數的假定及一些廣泛方向性的概念開始。隨著我們逐漸地蒐集和分析資料，理論會從頭開始慢慢地發展成形。在某個特定領域內，理論會以相當緩慢的方式，一個概念接著一個概念、一個命題接著一個命題地浮現出來。這整個過程類似於漫長的懷孕。隨著時間慢慢經過，各種概念及實徵推論會逐漸出現並且發展成熟。很快地，各種關係也變得清晰可見，然後我們就能把來自於各種不同研究的知識編織在一起，成爲一個較抽象的理論。

由於理論能夠提供解釋，因此理論與研究是密不可分的。不同的理論能夠提供不同的解釋，而各種不同類型的解釋則告訴我們，這個世界以不同的方式在運作。

有些研究會評估某一個理論，而其他的研究則會擴展某一個理論，或是發現某一個理論有一些不完整的地方，加以補足。例如，你已經在本章開頭的部分讀到，單獨用教育和收入這兩項因素，並不足以解釋吸煙行爲的差異。還有一些研究在一開始時就先提出兩個或多個彼此互斥的理論性解釋，並試著讓這些理論性解釋相互競爭，以觀察何者最能清楚地解釋發生的事件。

有些時候，在某一項研究裡會去比較由兩個或是多個理論性解釋所提出的不同預測。例如，克瑞傑（Kraeger, 2008）就曾經比較過兩種不同的解釋，這兩種解釋是關於高中男生所表現出的反社會行爲（例如，逞兇鬥狠與偏差行爲），以及參加學校運動校隊這兩者之間的關係。而他所比較的兩種理論，分別是社會控制理論及社會學習理論。社會控制理論認爲，參與學校的體育活動可以減少反社會的行爲，這是因爲學校的體育活動是一種由成人所管理的、在制度方面獲得認同的行爲。體育活動會在青少年之間建立社會聯結，並讓他們服從於傳統的行爲模式。如果這些男生有偏差行爲，就會失去運動員的資格，並伴隨著在同儕間社會地位下降。由於參與體育活動必須付出相當多時間，如此一來，這些男生比較沒有餘裕去從事反社會行爲。除此之外，有系統的學校體育活動，也能促進有利於社會的價值觀，像是團隊合作及公平競爭等等。社會控制理論認爲，關於高中男性運動員的暴力行爲報告，可歸咎於少數不合群者，而這些人不但缺乏足夠的自我控制能力，也無法融入於社會中。

對比之下，社會學習理論則認爲，我們從同儕與家人身上學習到利社會行爲或是反社會行爲。高中的體育活動會同時促進利社會及反社會的價值觀，像是忍受痛苦繼續比賽、不要接受限制、以及崇拜非學業方面的成就等等。有一些比賽或是體育活動，像小雞賽局（the game of chicken）（譯註：在此種比賽中，兩個人在同一車道上開車衝向對方，看看誰先轉動自己的方向盤使得車子離開車道，就是小雞，也就是懦夫、膽小鬼），會比其他的比賽更傾向於鼓勵侵略性的肢體行動、使用恐嚇威脅的手段、對於同伴的忠誠與支持、以及「性格競爭」（“character contests”）等等。至於學習理論的一個分支，男性理論（masculinity theory），則發現某些運動是「超級男性」的運動（像是足球、橄欖球、以及冰上曲棍球等等，至於像是游泳、徑賽、棒球、和網球等運動就不屬於這個類別）。超級男性的運動強調從事個人暴力行爲，比方像是把身體當成武器使用、粗暴的肢體接觸、及對他人身體的支配力量等等。這些運動把在同儕間的成功與聲望，和一種特定形式的「男子氣概」（“maleness”）連結在一起。此種形式的男子氣概隔絕於其他形式的男子氣概之外，並把其他形式的男子氣概都稱爲是「虛弱的」或是「女性化的」。社會學習理論及男性理論都預測，那些參加了超級男性運動、或是有朋友參加了超級男性運動的男生們，比起那些參加了其他運動，或者完全不參加高中體育活動的男生，更有可能從事一些諸如逞兇鬥狠的反社會行爲。

這兩個理論提出了完全不同的預測：（1）參加學校的體育活動或是有同儕參加這類的活動，可以減少反社會行爲；（2）參加某些類型的體育活動或是有同儕參加這類的活動，會增加反社會行爲。克瑞傑（Kraeger, 2008）檢視了一份關於美國 6,397 名高中學生的資料。他仔細調查了參與十二種特定體育活動、或有朋友參加這些體育活動的高中男學生，以確認參加這些體育活動與學生們從事反社會行爲（逞兇鬥狠或是其他的偏差行爲）之間是否有任何關聯性。研究發現顯示，有許多朋友參加超級男性運動（像是橄欖球）的高中男生，尤其是那些自己本身也很熱衷於這些運動的高中男生，確實比較有可能從事逞兇鬥狠的行爲。對比之下，參與其他體育活動（像是網球）的高中男生，就比較沒有逞兇鬥狠的傾向。因此，他的研究發現比較支持社會學習理論，而不是社會控制理論。

結論

在本章裡，你學到了有關社會理論的許多知識，包括社會理論的組成部分、目的、以及類型等等。把理論和研究截然二分，只是一種人爲的、不自然的分法。理論的價值、及理論對進行一項良好研究的必要性，現在應該很清楚了。在進行研究時完全不使用理論的研究者，很少能作出高品質的研究。而且他們經常發現自己陷於困境之中。同樣地，當理論家在建構理論時，如果沒有把理論和研究連結在一起、或是沒有將理論和實徵事實結合在一起，那麼就可能會淪爲只是胡亂思考或臆測的困境。你現在應該已經熟悉科學社群、研究的面向、以及社會理論了。在下一章，你將進一步瞭解研究者從事社會科學時，所採用的不同研究取向。

問題回顧

1. 具體概念與抽象概念之間有何差異？試舉例說明之。
2. 研究者如何利用理想型及分類等方法，來詳細說明各種概念？
3. 概念如何包含內在的假定？試舉例說明之。
4. 以歸納取向和演繹取向建立理論時，兩種方法之間的差異爲何？
5. 請描述微觀層次、中觀層次、以及鉅觀層次的社會事實之間的差異爲何。
6. 請討論預測與理論性解釋之間的差異爲何。
7. 因果關係的三個條件爲何？其中有哪一個條件從未被完整地證明過？爲什麼？
8. 爲什麼研究者要使用圖表來顯示因果關係？
9. 結構性解釋和詮釋性解釋之間的差異爲何？
10. 社會學中的主要理論架構在研究中所扮演的角色爲何？

注解

1. For more detailed discussions of concepts, see Chafetz (1978: 45–61), Hage (1972: 9–85), Kaplan (1964: 34–80), Mullins (1971: 7–18), Reynolds (1971), and Stinchcombe (1973).
2. Turner (1980) has provided an interesting discussion of how sociological explanation and theorizing can be conceptualized as translation.
3. Classifications are discussed in Chafetz (1978: 63–73) and Hage (1972).
4. For more on typologies and taxonomies, see Blalock (1969: 30–35), Chafetz (1978: 63–73), Reynolds (1971: 4–5), and Stinchcombe (1968: 41–47).
5. *Recursive* refers to a procedure that can repeat itself indefinitely or an iterative process that reoccurs with a feedback loop. Applied to a causal relationship, recursive suggests that a cause (X) operates on an effect (Y) to produce an effect (Y), but this process repeats with the effect (Y) at a later time, itself acting as a cause influencing the original cause (X).
6. Network theory is discussed in Collins (1988: 412–428), Fuhse (2009), Galaskiewicz and Wasserman (1993), and Schweizer (1997).
7. A basic introduction to functional explanation can be found in Chafetz (1978: 22–25).
8. See Craib (1984), Phillips (1985: 44–59), and Skidmore (1979). Chapter 1 of Bart and Frankel (1986) also offers an elementary introduction. Jasso (2004) offers a tripartite model of social science knowledge that consists of empirical analysis, theoretical analysis, and framework analysis, arguing that the advance of knowledge takes place on all three levels and their interrelationship.

第4章

方法論的意義

哲學基礎
三種取向
實證主義社會科學
詮釋取向社會科學
批判取向社會科學
女性主義以及後現代研究
結論

在社會科學中易於引起混亂之處——現在應該已經顯而易見了——就隱藏在長久以來關於科學本質的爭議之中。

——C. Wright Mills, *The Sociological Imagination*, p. 119

許多人常常會有這樣的疑惑：社會科學到底是不是*真正*的科學？這是因爲在他們眼中只有自然科學（例如，物理、化學、以及生物學）才是真正的科學。*科學*的意義對於我們進行社會科學研究的方式，造成極爲重大的影響。我們可以用下列兩種方法來定義科學：（1）目前仍在進行研究的科學家所做的事，和科學制度與其機構的運作方式；（2）在哲學家們細心剖析之後，所認定的二十一世紀科學的核心意義。有件事是相當清楚的，也就是在社會學及科學哲學的許多研究中，都告訴我們，社會科學的實踐與意義，比大多數人所想的要更爲細緻而且複雜。正如同柯林斯（Collins, 1989: 134）所說，「當代的科學哲學並未摧毀社會科學。當代的科學哲學並未說科學是不可能的，而是提供我們一幅更具有彈性的科學本質圖像。」

關於*是什麼因素使得社會科學成爲一門真正的科學*這個問題，人們已經爲此爭辯了相當長久的時間，而且這個問題與學習社會研究也有密切的關聯性。這個問題連結了各種不同的社會科學，並讓我們去思考在自然科學及人文科學之間，是否存在著分裂的，或者是一致的狀態。哲學家如孔德（Auguste Comte）、涂爾幹、休謨、馬克思、彌爾（John Stuart Mill）、以及韋伯之類的偉大社會理論家，都曾經仔細地思考過這個問題。然而，雖然人們已討論及辯論這個問題超過兩個世紀，這個問題至今仍困惑我們。很顯然地，這個問題並沒有一個簡單的答案。

此問題之所以沒有單一答案，是因爲進行科學的方式並非只有一種，相反地，科學有著多重不同形式，或數種*可供我們選擇的取向*（*alternative approaches*）。「取向是種一般性的詞彙，其範圍比理論或方法論還更廣泛。取向包括了認識論抑或是關於知識理論的問題、與研究的目的，不論該目的是去瞭解、解釋、或規範性評估……」（Della Porta and Keating, 2008: 1）。每一種社會科學的取向都是以哲學的假

定爲基礎，而且對於什麼事物能夠構成最佳的研究這個問題，都有自己的特定立場。在不同的國家，我們可以在社會科學領域中發現各種不同的取向。不過，就如同艾本德（Abend, 2006）所一直主張的，不同國家中，主流的社會研究取向可能會有極大差異。更明確地說，在美國蔚爲主流的取向，在其他地區可能不會被社會科學家們所廣泛地接受或是使用。

一開始的時候，你可能會對同時有這麼多種取向感到相當困擾，但是，一旦你學習這些不同的取向之後，你將會發現，研究以及理論的其他面向將會變得更清楚。如果你能知道特定的研究技術（例如，像是實驗以及參與觀察等等）所依賴的邏輯與假定，那麼這些研究技術就會變得更有意義。除此之外，當你在閱讀社會研究報告時，取向也能夠幫助你瞭解你可能會遇到的各種不同觀點。同樣重要的是，當你面對你可能想要進行的不同研究種類時，取向可以讓你在各種不同的選擇之間，作出明智的抉擇。因爲你或許會對某種取向覺得比較自在一些。

學習關於各種取向的知識，並不是一件簡單的事。當你讀到一份研究報告時，作者很少會明白地告訴你，研究中所使用的是什麼取向。許多專業的研究者，通常對於其他可供選擇的取向，也都只有模糊的概念而已。這是因爲當他們在接受研究方法的訓練時，是以間接的方式學習某個特定取向的原則及假定（Steinmetz, 2005a: 45）。各種不同的取向可以在不同的社會科學和應用領域裡發揮作用，並對於研究的進行方式產生很大的影響。[1]

我在此所呈現的主要取向都是理想型，而且我也特別突顯出這些取向之間的差異，好讓你可以更加清楚地看到每種取向的特色與內涵。雖然各種取向在發揮作用時，所依賴的是不同的核心假定、互斥的原則、及截然不同的優先次序。但是我們在進行研究時，仍然可以使用超過一種以上的取向，且從中獲益良多。每一種取向都會以它自己的方式，對知識的進展作出重大貢獻。正如同羅斯和梅塔（Roth and Mehta, 2002）所主張的，我們可以用不同的取向去研究同一個社會事件，並從我們所使用的每一種取向當中，學到許多知識。不只關於我們想要研究的社會事件，還關係到一些最重要的問題、相關資料的類型、以及在創造知識時的一般方法等等，每種取向都會提供我們一個不同的觀點。

哲學基礎

在這一章裡，我們將要把哲學中的抽象議題與具體的研究技術銜接在一起。這些抽象議題指出了良好的社會研究所該具備的條件、爲我們進行研究的原因提出證明、將道德－政治價值與研究連結起來、並且也指導我們一些合乎倫理道德的研究行爲。各種不同的取向都是一個相當廣泛的架構，讓所有研究者可以在這些架構當中去進行研究。考契（Couch, 1987: 106）將不同的取向總結如下：

> 這些研究傳統在本體論及認識論方面的立場，為當代社會學中一些更為激烈的爭論提供了基礎……每一邊都宣稱自己所提倡的思考架構，提供獲取

> 關於社會現象知識的工具，而且每一邊都認為其他人的努力充其量只能算是弄錯方向而已……在關於應該注意哪些現象、如何去接近並處理各種現象、及如何分析各種現象等方面，這些不同的立場都有各自不同的看法。

在上述這段引文中提到了兩個哲學方面的領域，分別是本體論與認識論。不論研究者是否承認，所有的科學研究都是以源自這兩個領域的假定及原則作爲基礎。我們不需要先深入討論各種不同的哲學假定，然後才去進行研究。然而，當我們在進行研究時，我們會在這些不同的假定之間作出抉擇。多數人會不加質疑地就去接受這些假定，然而，當你能夠知覺到這些假定的存在，你就能夠更加清楚地瞭解到，哪些事物構成了你在研究方面的各種選擇基礎。不同的哲學假定會突顯出社會研究中各種不同取向的差異，以及造成這些差異的原因。

本書並不是一本討論科學哲學的書籍，儘管如此，研究方法論仍然是以本體論與認識論的假定作爲基礎的。一旦你瞭解了這些假定，你就能夠開始看出在社會科學家之間造成爭議及差異的根源所在。藉由思考各種假定和坦白說明假定的方式，你將會成爲一個更好的研究者，這是因爲去反思與察覺到各種假定（而不是毫無知覺地全盤接受這些假定），將能夠幫助你更加清楚地進行思考。如同柯立爾（Collier, 2005: 327）所說的：

> 現存的各種科學，尤其是社會科學，無法置身於哲學的影響之外。許多的社會科學從一開始時就採取了關於科學本質與樣貌的某種哲學立場，並試著模仿此種立場。尤有甚者，各種社會科學的實踐者經常忘記這些哲學上的假定或是前提……因此，也就把這些假定或前提轉變成無可質疑的、不能挑戰的信條。

在進行研究的實務活動以及知覺到科學中的基礎哲學議題這兩者之間的勞力分工，已經帶來了不幸的後果。大多數從事實務活動的研究者，都集中心力於精通特定的研究技術，此種情況已經把「關於實徵研究的本質與進行方式這個問題，留給社會科學的哲學家去處理」，而此種分歧的狀況「模糊了一幅原本應該更加精確的、關於現有研究活動範圍的圖像：也就是各種相異的取向之間的實際情況……」（Mihic, Engelmann, and Wingrove, 2005: 483）。

現在，我們要先討論一下這兩個哲學的領域，以及在這兩個領域之間的一些基本區別，而這些區別則與社會科學研究中的主要取向有著直接的關聯性。

本體論（ontology）關注的議題是存在的事物，或是實在（reality）的基本本質。當我們在進行一項研究時，我們會作出一些假定，而這些假定是關於我們所要研究的事物爲何、以及該事物在世界中的位置。本體論中的兩個基本立場分別是實在論者（realist），與唯名論者（nominalist）。實在論者認爲這個世界「就在那兒」（“out there”）。換句話說，這個世界是依照原本就存在的類別組織

本體論　哲學的一個領域，該領域主要是在討論存在的本質，或者是真正存在的事物。換句話說，本體論所關心的問題就是真正存在的事物為何，以及實在的基本類別是什麼。

而成的，只是等著我們去發現而已。實在論者認爲，「真正的世界」獨立於人類以及人類對於世界的詮釋之外，如此一來，要去評估真正的世界所存在的事物，就變得比較不那麼困難。我們可以用一句老生常談來說明此種狀況：「你所看見的，就是你所得到的」。批判實在論者（critical realist）是實在論者裡的一個次團體，他們對上述的假定作了一些修正。批判實在論者認爲，要直接去捕捉、瞭解「就在那兒」的實在，並不是件容易的事。而我們對於客觀實在所進行的探索與研究，很容易變得扭曲並且混淆不清。這是因爲我們先入爲主的想法、主觀性、或者是文化詮釋，都會污染我們與實在的接觸。因此，批判實在論者增加了一些預防措施或調整校正，以控制上述的詮釋方式所造成的影響。

唯名論者則認爲人類從未真正直接地體驗過一個「就在那兒」的實在。我們與所謂的「真正的世界」的接觸經驗，總是必須透過一組詮釋的透鏡或策略、以及內在的主觀性才得以發生。主觀的與文化的信念，會影響我們所看到的事物及我們體驗實在的方式。我們個人的生活經驗和文化的世界觀，總是把我們的經驗組織成各種不同的類別與模式，且此種情況在我們毫不知情的狀況下發生。唯名論者承認某些詮釋的策略比其他策略更爲晦澀難懂。然而，他們仍然認爲我們永遠無法完全地移除詮釋的透鏡。除此之外，我們總是會受到重重的限制，因此無法大幅度地超越我們內在的想法、文化背景、以及主觀性。

讓我們用一些比較具體的例子，來說明實在論者和唯名論者之間的抽象區別。例如，當一個實在論者看到一張地毯時，他會說呈現在他面前的實在是一張地毯，也就是某個可以用來舖在地上、並讓人走在上面的東西。而當實在論者看到一個人的臉部特徵、頭髮、以及皮膚膚色時，他可以認出這個人屬於這個世界上某一個特定的種族團體。而在深入地檢視過一個人的身體之後，像是骨骼、生殖器官、胸部、荷爾蒙的化學檢驗結果、毛髮的覆蓋部分等等，實在論者就可以看出這個人在生物學上是屬於男性或是女性。相比之下，當唯名論者看到一張地毯時，他會問這可能是什麼東西。他會去問這個東西是由什麼所組成的、如何被製造出來、要以何種方法來使用、爲什麼這個東西會在這兒、及特定的歷史文化背景和人類使用這個東西的方式，如何形塑了我們所看到的這個東西。這張地毯只是一個可以讓他擦腳和走在上面的東西嗎？有沒有人會整天坐在上面、睡在上面、並且在上面吃東西呢？有沒有人會把這張地毯掛在牆上，以使房間維持溫暖呢？這張地毯有沒有可能是一件藝術作品，可以讓人們欣賞、並帶來審美方面的樂趣呢？人們會不會把這張地毯視爲一件宗教物品並膜拜它呢？當唯名論者看到一個人膚色與臉部特徵時，他會覺得相當困惑。爲什麼會有這些種族差異呢？當種族這個概念可能因爲文化及歷史時期的不同，因而產生極大的變化時，這些種族類別可能包含什麼事物呢？同樣地，當唯名論者看著人類的身體，他會因爲外表方面的差異、模稜兩可的情況而感到擔憂。每個人是否都很明確地屬於生物學上的二種性別之一呢？生物及外表方面的差異，與社會中的性別以及社會差異之間，契合程度有多高呢？就如同種族類別，性別方面的類別、把人們作出區別的各種事物，都會隨著文化和時代的不同，而有極大的差異。由唯名論者眼中看來，人們通常把主觀的觀點強加於可見的外表特徵上。因

此，不同的人們所看到的事物，會有很大的差異。

我們可以把實在論者及唯名論者在本體論方面的假定放在一條連續的直線上（請參閱圖 4.1）。一位純粹的實在論者會說，我們所看見的，就是真正存在的事物，而且我們可以輕易地捕捉這個事物，以產出客觀的知識。批判實在論者則會比較小心一些，並承認主觀的以及文化的詮釋，有可能影響我們與實在接觸的部分經驗。一位溫和的唯名論者會說，主觀的及文化的因素，會大幅度地影響我們與物質以及社會世界接觸的經驗，而我們永遠無法完全徹底地排除這些因素。一位極端的唯名論者會說，我們對於所有物質與社會經驗的理解，都非常依賴詮釋的與文化的因素，以致於如果沒有這些因素，所有的經驗都會失去意義。此外，我們不可能獲得任何形式的客觀知識。

認識論　哲學的領域之一，該領域所關心的議題是知識的創造，而且其焦點在於我們如何認識我們所認識的事物、或什麼才是達到真理的最有效方法。

認識論（epistemology）所討論的議題，是關於我們如何認識身邊的世界，或是什麼事物能使得一個關於這世界的宣稱爲真。我們瞭解或認識這個世界的方式，其實根源於我們在本體論方面的假定。認識論的內容包括我們需要做些什麼事以產生知識，以及一旦我們產生科學知識，此種知識看起來會像什麼樣子。

如果我們採取的是實在論者的立場，那麼我們就可以藉由仔細觀察實在的方式，生產知識並瞭解實在。實在論者認爲，有一個「就在那兒」的實徵世界，而這個世界存在於我們的內在想法與對它的知覺之外。當我們蒐集了實徵證據，我們會發現，有些關於實在的想法，可以得到證實、或是與證據一致。而其他有些想法則可能錯誤，因爲這些想法缺乏支持性的實徵證據。當我們調查、研究實徵世界時，我們可把事實與迷思，或是幻覺區隔開來，並產出客觀知識。在我們結合並且組織那些已被證實的想法後，我們將會發現一些廣泛的原則或是法則，以解釋實在所包含的內容及運作的方式。我們可以用演繹的方式來產出新知識，在過程中以實徵資料去驗證一些早就存在的、關於實在的想法和臆測。我們也可以用歸納的方式產出新知識，也就是先蒐集實徵證據，再把這些證據組織成較高層次的一般性推論。不論是用演繹的方式抑或歸納的方式，在經過一段時間之後，針對那些關於各種實在領域的想法，我們都能辨別出這些想法中到底何者爲真、何者爲假。

如果我們採取的是唯名論者的立場，那麼，仔細地進行觀察就無法產生關於實在的知識，因爲詮釋以及主觀的看法會嚴重影響所有的觀察結果。對於我們可能會觀察的人們而言，同樣的情況也會發生在他們身上。換句話說，他們的詮釋及主觀的看法也會影響他們的言語和行爲。我們與其他人所體驗到的實在，是由一個行動以及詮釋的連續過程的結果所建構而成的，這些行動與詮釋則發生於特定的地點和

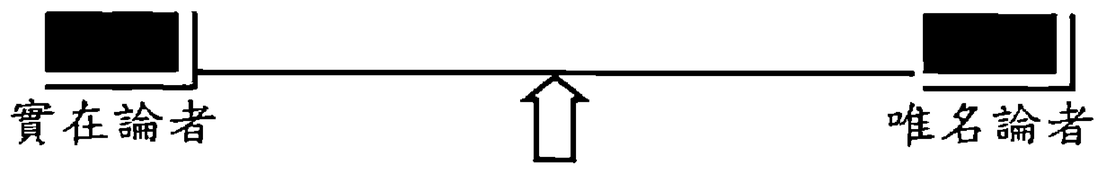


圖 4.1　本體論的假定

時間。我們不可能把一個客觀的、「就在那兒」的實在，與我們所作的詮釋、或是時間和空間所造成的影響完全分隔開。我們最好的作法，就是坦白地承認我們自己的觀點及詮釋。我們或許會試著發現其他人在過日常生活時，他們內在、主觀的看法以及詮釋。我們不可能創造出關於社會生活的普遍法則，也就是能夠應用於所有人及所有地方的法則。我們所能生產的、關於這個世界的最佳知識，便是針對特定背景中的特定人們，提出一些經過詳細思考的詮釋。我們可以提出一些詮釋，以說明在某一個特定的背景中，我們認爲其他人正在做的事、和我們認爲他們這麼做的理由。在創造社會科學知識時，我們必須以歸納的方式去觀察、詮釋、以及反思其他人在特定的社會情境脈絡中，所說的話及所做的事，而且我們同時也必須反思，我們本身的經驗與詮釋。

三種取向

科學是一種人類所創造的產物，而且不像寫在某個石頭上的神聖文字般、可以作爲代代相傳的事物。一直到十九世紀早期，只有那些從事理論性臆測的哲學家以及宗教方面的學者，才寫過關於人類行爲的文章。早期的社會思想家主張，我們可以利用源自於科學的原則來研究社會世界。這些思想家認爲，只要對社會世界進行嚴格的、系統性的觀察，並結合仔細的、合乎邏輯的思考過程，便能夠提供一種嶄新的、有價值的知識形式。

慢慢地，我們能利用科學原則來檢視社會世界的這個想法，逐漸被廣泛接受。於是，接下來的議題，就是如何進行科學研究，以探究社會實在。簡單的作法之一，就是直接從自然科學（例如，物理學、生物學、化學）中借用，並盡可能地去模仿、調整自然科學中的假定以及研究方法。

許多社會研究者欣然接受此種作法，但是，此種作法卻會造成數種困境。首先，即使是在自然科學家之間，對於科學一詞真正的意義爲何，目前依然爭辯不休。所謂的*科學方法*（*scientific method*），只不過是一組寬鬆的抽象、模糊的原則，而且這些原則僅能提供有限的指引。除此之外，從事研究的科學家們也會同時使用數種不同的方法。其次，有些人認爲，人類具有一些質化的特質，這些特質與自然科學中所研究的物體類型（像是星體、岩石、植物、化學的化合物、及魚類等等）之間存在著明顯的差異。人類擁有能夠進行思考與學習的能力，且人類能夠察覺到自身的存在與過去的事件，還擁有動機以及理性。因此，有些人不禁要問，在研究人類的特質時，到底是只對自然科學方法進行調整就好，還是需要一種完全不同的、特殊種類的科學？

在本章中所討論的三種取向，都是從許多具體論點中所濃縮而成的核心想法。[2] 因此，這三種取向都是理想型。在實務層面上，身爲社會研究者，我們或許會把來自於每種取向的不同元素加以混合使用。然而，這些取向仍然代表在關於進行社會研究方面的不同觀點及假定。[3] 這些取向都是一直發展、成長的立場，提供我們各種

觀察、測量、及瞭解社會實在的不同方法。

爲了不使相關的討論內容變得過於複雜，我們把這三種取向的假定和觀念整理爲針對十個不同問題所提出的答案（請參閱圖表 4.1）。

典範　一種理論及研究的一般性組織架構，其內容包括了基本的假定、關鍵的議題、良好研究的模範、與尋找答案的方法等等。

這三種取向分別是*實證主義社會科學*（*positivist social science*）、*詮釋取向社會科學*（*interpretative social science*）、及*批判取向社會科學*（*critical social science*）。大多數正在進行中的社會研究都是以前兩者爲基礎，而實證主義（positivism）則是最早且最被廣泛使用的取向。至於其他兩種非實證主義的取向，則代表關於社會科學的不同觀點及假定，而且都已發展超過一個世紀之久。

每一種取向都和不同的社會理論，及互異的研究技術有關。然而，這些取向和科學、社會理論、以及研究技術之間的連結，其實並不十分嚴謹。這些不同的取向其實類似於一種研究方案、研究傳統、或是科學典範。**典範**（paradigm）一詞是由孔恩（Thomas Kuhn, 1970）所提出並加以推廣的，而典範所指的是一種基本的理論和研究方向。目前，關於*典範*一詞的定義，仍然眾說紛紜。一般而言，一個科學典範就是一個與思考方式有關的完整體系，其中包括了基本的假定、尚待回答的重要問題或尚待解決的謎題、研究者所使用的研究技術、以及良好的科學研究範例。在社會科學中，實證主義一直是佔據優勢地位的典範，尤其自從 1945 年實證主義在美國被採用之後。人類學及歷史學是最不具實證主義色彩的學術領域，而經濟學及實驗心理學，則是最具有實證主義色彩的學術領域。至於政治學與社會學則介於上述兩個極端的中間。在社會學中有好幾個彼此相互競爭的典範，[4]但是，「除了曾經有段關於認識論方面的簡短混亂時期外……自從 1945 年之後，幾乎都是實證主義的天下」（Steinmetz, 2005a: 25）。

圖表 4.1　十個問題

1. 進行社會科學研究的最終目的為何？
2. 社會實在的基礎本質為何？
3. 人類的本質為何？
4. 關於人性能動作用（human agency）（自由意志、意志力、以及理性）的觀點為何？
5. 科學與常識之間的關係為何？
6. 社會實在的解釋或理論是由什麼所構成？
7. 如何確定某個解釋的真假？
8. 良好的證據或事實資訊的特徵為何？
9. 社會科學知識的關聯性及作用為何？
10. 社會政治價值由何處滲透到科學中？

實證主義社會科學

實證主義社會科學（positivist social science, PSS）的應用最廣，而且廣義的實證主義正是自然科學所採用的取向。事實上，大多數人都會不假思索地認爲實證主義取向*就是*科學的同義詞。實證主義有許多不同的版本，而且在科學哲學裡及研究者之間，也已有一段長久的歷史。[5] 然而，對於許多研究者而言，實證主義已經演變成一種他們極力避免的負面標籤。根據透納（Turner, 1992: 1511）的觀察，「*實證主義*不再具有清楚的指稱對象，但是，顯而易見，對於許多人而言，被稱爲是一個實證主義者並非是一件好事。」實證主義的社會科學具有許多種不同的變化種類，其名稱包括了邏輯實徵主義（logical empiricism）、既有的或傳統的觀點、後實證主義（postpositivism）、自然主義（naturalism）、涵蓋律則模式（covering law model）、以及行爲主義（behaviorism）等等。史坦梅茲（Steinmetz, 2005b: 227）則把「過去半個世紀以來，在美國社會學界裡風行一時的本體論、認識論、方法論等各方面的假定，所形成的特殊集合體」稱爲*方法論的實證主義*（*methodological positivism*）。

西歐的哲學家在十九世紀晚期與二十世紀初期，發展出實證主義，當時先後有兩位英國哲學家在其著作裡，概略地描繪實證主義科學的基本原則。而這兩位哲學家及其相關著作，分別爲休謨（David Hume, 1711-1776）的《人性論》（*A Treatise of Human Nature*, 1739-1740），以及彌爾（John Stuart Mill, 1806-1873）的《邏輯系統》（*A system of Logic*, 1843）。而社會學的法國創始者孔德（Auguste Comte, 1798-1857）則寫下了《實證主義哲學》（*The Course of Positive Philosophy*, 1830-1842）一書，在書中詳細地闡述社會科學實證主義的原則。另一位法國的社會學家涂爾幹（Emile Durkheim, 1858-1917）則在他所寫的《社會學方法論》（*Rules of the Sociological Method*, 1895）一書中採用了實證主義的假定，而這本書也成爲早期社會研究者最重要的教科書。

> 實證主義樹立了一種特定的科學模型，也就是價值中立、追求基本原則、以及去發現因果法則等等……這些都被認為是使得自然科學如此成功的特徵，因此，實證主義的假定是，如果社會科學能夠模仿自然科學的這些特徵，或許也能得到類似的成功。（Collier, 2005: 328）

實證主義牽涉到數種社會理論和結構功能、理性選擇、及交換理論的架構。實證主義的社會科學研究者偏好量化資料，且通常會使用實驗、調查、統計之類的研究技術。這些研究者所尋求的是嚴謹的、精確的測量結果，以及「客觀的」研究。因此，他們會仔細地分析由測量結果所得到的數據，以考驗因果關係的假設。在許多領域裡（例如，公共衛生管理、犯罪司法、市場調查、政策分析、方案評估等等）的研究者，都相當依賴實證主義的社會科學。

實證主義社會科學（PSS） 社會研究中的三種主要取向之一，此種取向強調的是去發現因果法則、仔細進行實徵觀察、以及進行價值中立的研究。

1960 年代與 1970 年代之間，在英國、加拿大、斯堪地那維亞、以及美國等地區，主要的社會學期刊幾乎都被實證主義社會科學的相關文章所佔據。而到了 1980 及 1990 年代時，相關文章的數量在歐洲的社會學期刊中，已經急遽地下降，但在北美的期刊中仍然居於主流地位。[6]

在實證主義裡，「科學只有唯一一*個*邏輯，而任何想要獲得『科學』這個名號的知性活動，都必須遵循此種邏輯」（Keat and Urry, 1975: 25）。因此，社會科學與自然科學都必須使用同樣的方法。基於這樣的觀點，社會科學及自然科學間的任何差異，都必然是由於社會科學的不成熟和其探討的主題所致。另外還有一種假定認爲，到了最後，所有的科學，包括社會科學在內，都會變得越來越接近最先進的科學——物理。由於探討的主題不同（例如，由於研究的物體不同，所以在地質學中所需要使用的技術，就和天體物理學及微生物學的技術不同），因此在不同的科學之間，仍然會有一些差異存在。然而，所有的科學仍然共享一套原則與邏輯。

實證主義的社會科學是一種*有組織的方法，用以結合演繹的邏輯與針對個體行爲所進行的精確實徵觀察結果。其目的在於發現並確認一套能夠用來預測人類行爲模式的可能因果法則*。

問題

1. *進行社會科學研究的最終目的爲何？*

進行研究的最終目的，是要獲得科學解釋，也就是發現、並詳細記錄關於人類行爲的普遍**因果法則**（causal laws）。正如同透納（Turner, 1985: 39）所說的，「社會世界依循著一些抽象法則發展，我們可以透過仔細蒐集資料的方式，來驗證這些抽象的法則」，而研究者必須「發展出一些抽象原則，以描述關於社會世界不變及永恆的特性」。科學家們投身於知識，永無止盡的追尋。當我們獲得更多知識與發現新的複雜事物，我們會發現還有更多值得學習的知識和事物。有些實證主義社會科學的版本認爲，人類永遠無法窮究所有的事物，只有上帝才擁有一切知識。然而，上帝賦予人類追求知識的能力，因此我們有義務盡我們所能，發掘知識。

2. *社會實在的基礎本質爲何？*

現在的實證主義者採取實在論的本體論。他們主張實在或是事實「就在那兒」，並等待人類去發現。人類的知覺能力及智力或許有些缺陷，且實在或許也相當難以確切掌握。但是，實在真正地存在著、有其運作的模式、也有其自然的規則。若不以這樣的假定爲前提（也就是說，如果這個世界雜亂無章、而且沒有任何規律性），就不可能會有任何邏輯以及預測可言。科學能夠讓人類發現自然界的秩序與法則。「科學裡基本的與觀察的法則，被認爲是真實的、主要的、及確定的，因爲這些法則內嵌於自然世界的構造之內。因此，發現一條新法則，

因果法則　在社會理論的因果解釋中所使用的一般因果規則，而在實證主義社會科學，發現因果法則是最為主要的目標之一。

就像發現美洲新大陸一樣，這兩者早就已經存在，只不過等我們去發現而已。」（Mulkay, 1979: 21）

實在論的本體論〔又稱爲*本質主義*（*essentialist*）、*客觀主義*（*objectivist*）、或是*實徵實在論*（*empirical realist*）〕關於實在的假定，充斥在一般人的常識思維裡，尤其是在英國及歐洲社會裡。此種假定的內容是，我們能夠看到並碰觸到的事物（也就是實徵的實在）其實不會過度複雜。我們所觀察到的事物，反映出世界上各種事物、人們、和關係更深層的本質。換句話說，這是一種「所見即所得」（"what-you-see-is-what-you-get"）、或者是「讓我看到」（"show-me"）之類的立場。事物的本質就和它們所呈現出的外表一致，由世界的自然法則所創造出來。因此，種族、性別、時間與空間的長短大小，就是「如實存在」的。此種觀點有許多言外之意。例如，男性之所以比女性更容易犯罪，是因爲某種牽涉到「男性特質」的事物在作祟。另一個與時間有關的假定，則認爲時間是線性的，也就是時間是在一條直線上流動的。發生在過去的事件，總是與現在所發生的事件有所不同，因爲時間只往單獨一個方向流動，也就是往前朝向未來流動。

其他的實證主義社會科學的假定，包括社會實在是穩定不變的，而關於我們對社會實在的知識，則逐漸累積增加。隨著時間的流逝，社會實在的核心規律性並不會隨著改變，我們今日所發現的法則，到了未來仍可適用。而知識所具有的累積性則意味著，我們可以同時研究社會實在的許多個別部分，然後把這些知識的片段累積組合起來，以獲得對於整體社會世界的圖像。而隨著時間經過，我們會累積越來越多的知識，並不斷地擴展我們對於這個世界的瞭解。

3. *人類的本質爲何？*

實證主義社會科學假定，人類是一種自私自利、追求歡樂並避免痛苦的理性動物。單一的原因會在每個人身上造成相同的效果。藉由去觀察那些我們從外在事實中看到的人類行爲，我們就可以瞭解人們。外在事實中可以看見的事物，比存在於內在主觀事實的事物更爲重要。有時候，此種觀點被稱作「**人類的機械模式**」（mechanical model of man）、或行爲主義取向，而此種模式意味著人類回應外在力量的方式，就像物體回應外加的物理力量一樣。涂爾幹（Durkeim, 1938: 27）曾經說過：「社會現象就是事物，也應該被當成事物加以研究」。此種強調可觀察的外在事實的觀點，暗示研究者不必檢視那些不可見的、內在的動機。

人類的機械模式 在實證主義社會科學裡所使用的人性模式，此種模式認為，只要觀察人們的外在行為、記錄作用於人們身上的外界力量，就足以提供對於人類思想與行為的適當解釋。

決定論 一種關於人性能動作用，以及因果關係的取向，該取向假定人類的行動主要是由外在力量所引起的，而且這些力量能夠被我們所發現。

4. *關於人性能動作用（自由意志、意志力、理性）的觀點爲何？*

實證主義社會科學強調各種關係的**決定論**（determinism），並去尋找能夠產生效果的決定性原因或是機制。實證主義社會科學會仔細地調查研

究，那些作用於個人、團體、組織、或社會之上的外在力量、壓力、以及結構等等，是以何種方式產生結果的（例如，行爲、態度等等）。實證主義社會科學不重視個體主觀的或是內在的理性，也不重視任何自由選擇與意志力的觀念。心智過程的重要性，遠遠低於那些超出個人控制範圍的結構性力量或情境。而這些外在的力量或情境，會對個人的選擇及自由造成極大影響。雖然個別的人們或許相信他們能夠自由地行動、並作出任何決定，但是實證主義者所強調的，卻是強大的社會壓力和情境，這些壓力及情境會對人們造成影響，並形塑人們大多數的（就算不是全部的）行動。即使是使用理論選擇解釋的實證主義者，其焦點也不在於個體進行思考與作出決定的過程，而在於辨識出各種不同的情境，以幫助他們預測人們將會以何種方式作出選擇。實證主義者認爲，一旦他們知曉外在的因素，那麼，個體大部分的思考過程，就會依循像是機械一樣的理性決策邏輯來進行。

幾乎沒有任何實證主義者會信奉嚴格的、或絕對的決定論。在此種決定論裡，人們僅須永遠以相同的方式，成爲僅能作出回應的機器人或玩偶。相反地，因果法則是具有機率性的。也就是說，因果法則適用於大型團體、或是發生於許多情境中，因此研究者可以估計某種預測行爲的發生機率。換句話說，因果法則使我們精準地預測在一個大型團體中，某種社會行爲的發生機率。因果法則無法預測在某個特定情境中，某個特定的個人會不會表現出某種特定的行爲。然而，研究者可以說，在 X、Y、Z 三種條件下，有百分之九十五的機率，半數以上的人，會表現出某種特定的行爲。舉例而言，研究者無法預測史密斯先生在下一次的選舉中會把票投給哪一位候選人。然而，在研究了關於史密斯先生的一些事情，在運用政治行爲法則後，研究者可以精確地宣稱，史密斯先生（以及與他類似的人們）有百分之八十五的機率會投票給候選人 C。這並非意味著史密斯先生不能夠把票投給任何一個他喜歡的候選人。更確切地說，這只是代表了他的投票行爲具有一定的模式，且受到外在社會力量的影響。

5. *科學與常識之間的關係爲何？*

實證主義社會科學認爲，在科學與非科學之間的分隔界線相當清楚明白。在追尋真理的許多方法中，科學是最特別的一種，也是「最佳的」方法。科學知識優於其他獲得知識的低劣方法（例如巫術、宗教、占星術、個人經驗、傳統等等），而且終究會取代這些方法。科學會從常識當中借用一些觀念或想法，但是科學會取代常識中那些草率的、邏輯上不連貫的、沒有系統的、或充滿偏見的部分。由於科學社群具有特殊的規範、科學態度、及科學的技術，因此能夠持續地生產「真理」（“Truth”），相對地，常識則只能夠偶爾並不一致地，對知識作出一些貢獻。

許多實證主義的研究者創造出完全新穎的詞彙，這些詞彙比在日常生活常識中所使用的詞彙，在邏輯上更爲一致、更經過仔細考慮、也更爲精練優雅。實證主義研究者「在一開始就應該構想出一些新概念，而不依賴過去原有的想法……科學偏好精確性，而精確性只能在一種以紀律爲基礎的語言中，才有可能存在。在日常生活裡那種模糊、不精確的語言中是無法達到精確性的」（Blaikie, 1993: 206）。而涂爾

幹在他所寫的《社會學方法論》一書中，也警告研究者要「堅決地拒絕使用那些在科學之外所形成的概念」，並且「使自己遠離那些會對一般人心智造成重大影響的謬誤想法」（引自 Gilbert, 1992: 4）。

6. *社會實在的解釋或理論是由什麼所構成的？*

實證主義的社會科學所提供的是**律則的**（nomothetic）（*nomos* 在希臘文當中是指*法律*）解釋，換句話說，此種解釋是以一個一般法則所形成的系統爲基礎。科學在解釋爲什麼社會生活會是現在這種形式時，其方法就是去發現背後的因果法則。律則的解釋一般會採取如下形式：Y 是由 X 所引起的，因爲 Y 和 X 都是在因果法則中的特定例子。也就是說，一項實證主義社會科學的解釋，其所陳述能夠適用於、或是涵蓋關於社會生活具體觀察結果的一般因果法則。而這也就是爲什麼實證主義社會科學所使用的解釋模式，會被稱爲「**涵蓋律則模式**」（covering law model）的原因。

實證主義的社會科學假定，法則的運作方式乃是依據嚴謹而且合乎邏輯的推理過程。研究者連結因果法則，並且能夠以演繹的方式連結他們所觀察到的許多事實。許多實證主義者相信，他們或許總有一天能夠用原則、推論、假定、以及定理來構成一些正式的符號系統，並且用這些符號系統來表達社會科學中所有的法則以及理論。有朝一日，社會科學理論看起來也會如同數學以及自然科學中的理論一般。

人類行爲法則應該是普遍有效的，並能夠適用於所有的歷史時期及所有文化。如同之前已經提過的，法則是針對大型群體以機率的形式呈現。舉例而言，實證主義的社會科學，在解釋多倫多地區 2010 年犯罪率上升時所提到的因素（例如，離婚率上升、對傳統價值觀的認同感逐漸下降等等），也應該能夠適用於任何地點與任何時間，像是 1890 年代的布宜諾斯艾利斯、1940 年代的芝加哥、或是 2020 年代的新加坡等等。除此之外，這些因素也會以合乎邏輯的方式遵循某種一般性法則（例如，傳統道德秩序的崩潰導致了犯罪率增加）。

7. *如何確定某個解釋的真假？*

實證主義發展於西方思想史上的啓蒙時代（也就是在大約中世紀之後）[7]，因此也包含了一項重要的啓蒙思想：人類可以利用理性去辨識真理，並把真理和虛假區隔開來。此外，透過理性的應用與對真理的追求，人類的處境最終可以逐步獲得改善。隨著知識不斷增加、無知不斷減少，人類的處境將會不斷地提升。此種認爲知識可以隨著時間而不斷累積的樂觀信念，在實證主義者區別出正確與錯誤解釋的過程中，扮演了重要的角色。

律則的 一種在實證主義社會科學中所使用的解釋類型，此種解釋極度依賴因果法則，以及像是律則一樣的陳述與相互關係。

涵蓋律則模式 一種實證主義社會科學的原則，該原則認為一些高層次的、非常抽象的理論，能夠涵蓋並允許演繹許多低層次的、較為具體的情境。

實證主義社會科學的解釋必須合乎兩個條件，也就是說，這些解釋必須（1）沒有邏輯上的矛盾，

以及（2）與所觀察到的事實一致。不過，這樣還不夠充分，除了上述的兩個條件之外，這些解釋還必須能夠被複製[8]，也就是任何一位研究者都能夠複製或重製他人的結果。這個條件對於整個創造知識的系統，設下了一個檢驗的標準。由於這個條件重覆地以堅實的、客觀的事實來驗證解釋，因此也確保了研究者的誠實。在各種相互對立的解釋中，存在著一種公開的競爭。在此種公開競爭裡，我們必須運用公正的規則、正確地觀察中立的事實、並嚴謹地應用邏輯。隨著時間經過，當不同的研究者進行獨立的理論檢驗，並累積發現時，科學知識就能逐漸累積增加。例如，某位研究者發現，在美國加州的聖地牙哥地區，失業率上升和兒童受虐案件增加之間，存在著關聯性。不過，我們不能僅僅憑著一項研究的結果，就武斷地認定在失業和虐待兒童之間，具有因果關係。在證實一項因果法則時，還必須由其他研究者在其他地方進行獨立的研究、及仔細謹慎地測量失業情況與虐待兒童案件間的關係，若這些研究者也發現相同的關係，此項因果法則才能夠得到證實。

8. *良好的證據或是事實資訊的特徵為何？*

實證主義的社會科學所採取的是二元論觀點，也就是假定冰冷的、可以觀察的事實，基本上與觀念、價值、或理論截然不同。實徵事實存在於個人的觀念或想法之外。我們可以用我們的感官（視覺、嗅覺、聽覺、以及觸覺）、或是能夠延伸這些感覺的特殊工具（例如，望遠鏡、顯微鏡、以及測量放射線的蓋氏測量器等等），來體驗這些實徵事實。有一些研究者把這個概念表達為兩種不同的語言：實徵事實的語言與抽象理論的語言。如果人們對於事實的看法有所分歧，此種意見不合的情況，必然是因為使用測量工具的方式不夠正確、草率或不恰當的觀察結果。「科學的解釋牽涉到對於現象進行精確且正確的測量」（Derksen and Gartell, 1992: 1714）。利用我們的感官所獲得的、對於可觀察的實在的知識，比其他的知識（例如，直覺、情緒感受）更為優異。此種知識讓我們可以把關於社會生活的正確與錯誤的觀念，明確地區隔開來。

實證主義者賦予實徵觀察一種優越的地位，且認為我們都共享關於實徵世界的相同基本經驗。此種想法意味著事實知識並非以個人的觀察與主觀推理為基礎，事實知識必須能夠與他人溝通和共享，只要是獨立地觀察事實的理性人們，在主觀上都將認同這些事實。此種情形稱作**互為主體性**（intersubjectivity），也就是對於可觀察的事實，人們在主觀上都會承認這些事實的存在。

許多實證主義者也都贊同奧地利裔哲學家巴柏（Sir Karl Popper, 1902-1991）在《科學發現的邏輯》（*The Logic of Scientific Discovery*, 1934）一書中所提出的證偽原理（falsification doctrine）。巴柏主張，任何對於知識的宣稱「永遠無法被證明或完全檢證，這些宣稱只能被反駁」（Phillips, 1987: 3）。證明某個因果法則的證據，不只要有成堆的支持性事實，還要去尋找那些違反該因果法則的證據。舉一個經典的例子說明：如果我想要檢驗「所有的天鵝都是白的」這

互為主體性　在實證主義社會科學中評估實徵證據的一種原則，也就是不同的人都能認同那些利用感官所獲得的、關於實徵世界的知識。

個宣稱，而我也找到了一千隻白天鵝，那麼，我仍然未能完全地證實這項因果法則或模式。只要找到一隻黑天鵝，也就是一項反面的證據，就能反駁我的宣稱。這意味著研究者也要尋找反面的證據，在尚未找到任何反面證據之前，他們最多也只能說：「到目前爲止，我還沒有找到任何反面證據，所以這個宣稱很有可能是正確的。」

9. *社會科學知識的關聯性及作用爲何？*

實證主義者試圖瞭解社會世界運作的方式，以使人們能夠控制社會世界，並作出正確的相關預測。簡而言之，當我們發現了人類行爲的法則時，我們就能夠利用這項知識來改善社會情境。此種工具性形式的知識（我們在第二章裡曾經討論過）把研究結果視爲工具，而人們可以利用這些工具，滿足自己的欲望與控制社會環境。因此，實證主義社會科學採取了一種**工具性導向**（instrumental orientation）。在此種導向中，知識的相關性，正是取決於該項知識賦予人們能夠主導或是控制周遭世界中事件的能力。

對於知識的應用，實證主義社會科學採取的是**技術官僚觀點**（technocratic perspective）。技術官僚這個字是結合了技術（technology）以及官僚（bureaucracy）這兩個字所形成的。實證主義社會科學認爲，在經過多年的專業訓練之後，研究者會發展出深度的技術與專門知識。作爲一名專家，研究者試圖滿足大型官僚組織（例如，醫院、商業組織、政府機構等等）在資訊方面的需求。這些類型的組織所提出的問題，通常傾向於應該如何提升組織運作的效率、應該如何有效地達成組織的目標之類的問題。扮演著技術專家的角色，研究者的主要工作是去回答由別人所提出的問題，而*不是*去提出不同的問題、挑戰這些問題的基本假定、或者去反抗由那些掌控官僚組織的領導者所設立的目標。

工具性導向 一種對於社會知識的手段與目的之導向，在此種導向中，知識就像一個工具，人們可以利用這個工具來控制他們的環境、或達成某些目標。知識的價值就在於達成目標的過程中，所發揮的用處。

技術官僚觀點 一種應用的導向，在此種導向中，研究者毫無異議地接受由政府、企業、或者是官僚所提出的任何研究問題以及對於研究範圍的限制，以不加批判的方式為這些人進行應用研究，並且順從地為官員提供他們作出決策時所需要的資訊。

價值中立的科學 一種實證主義社會科學的原則，也就是社會研究應該以一種單獨植基於實徵證據的客觀方式來進行，而且不應受到道德與政治價值觀的干擾。

10. *社會政治價值是由何處滲透到科學中？*

實證主義社會科學所堅持的目標是**價值中立的科學**（value-free science）。*客觀*一詞具有兩種意義：（1）觀察者對他們所看見的事物擁有一致的看法，以及（2）科學知識並非是以價值觀、意見、態度、或信仰爲基礎的。[9] 實證主義者把科學視爲社會的一個特殊的、與眾不同的部分，而這個部分不受到個人的、政治的、或宗教的價值影響。科學能夠在那些影響人類活動的社會，和文化力量之外獨立運作，因爲科學牽涉到使用嚴謹的理性思考及系統性的觀察，並能夠超越個人的偏見、偏差、價值觀。因此，科學社群的規範與運作方式，使得科學能夠

確保其客觀性。研究者接受並內化這些規範，把這些規範當作是他們成爲科學社群成員的部分必要條件。科學社群擁有一個複雜的制衡（check and balance）系統，以防範價值方面所產生的偏差。研究者所該扮演的適當角色，就是成爲一名「公正無私的科學家」（"disinterested scientist"）[10]。實證主義的社會科學對於人們看待倫理議題以及知識的方式，產生了極大的衝擊：

> 科學知識的實證主義理論的影響程度之大，使得該理論已經成為所有知識的標準，以致於道德的洞見和政治的承諾都已被貶為不理性的、或被矮化成只是主觀的偏好。倫理的判斷現在也被認為只是個人的意見而已。（Brown, 1989: 37）

摘要

實證主義社會科學被廣泛地宣揚是完全等同於科學的，因此，極少有人去意識到實證主義社會科學裡各種假定的起源。在十八世紀及十九世紀時，發展出這些假定的西歐哲學家，都曾經接受過宗教方面的訓練，他們當時所處的文化歷史背景也帶有特定的宗教信仰。當你在之後的章節裡閱讀到與量化研究技術以及測量方式有關的內容時，許多實證主義社會科學的假定，也會再次出現。實證主義的取向意味著，研究者在進行研究時，會從一個因果關係開始，而此項因果關係則是由該名研究者以合乎邏輯的方式，從一般理論中可能的因果法則所演繹而得來。該名研究者以合乎邏輯的方式，把抽象的概念和社會世界的精確測量結果連結起來。當研究者在測量社會世界的各種面向、檢視證據、複製他人所進行過的研究時，必須維持超然的、中立的、以及客觀的立場。這些過程將能針對在理論中所概略描繪的社會生活法則，進行實徵檢驗與確認。圖表 4.2 則提供了實證主義社會科學的簡略摘要。

實證主義社會科學是在何時、又是因爲什麼緣故而成爲主流取向呢？這可是個既冗長又複雜的故事。許多人把此種情況視爲一種自然的進展，或純粹是知識無法避免的演進過程。實證主義社會科學之所以能夠擴展，主要是因爲發生在較大的政

圖表 4.2　實證主義社會科學摘要

1. 社會科學的目的是發現法則。
2. *本質主義*的觀點之一，是就實徵方面而言，實在是顯而易見的。
3. 人類是能夠進行理性思考、並具有個人主義傾向的動物。
4. 在考量人性的能動作用時，所採取的是決定論的立場。
5. 科學知識和所有其他的知識都不同，也比所有其他的知識更為優越。
6. 解釋是*律則*的，並透過演繹的推理方式向前進展。
7. 藉由能夠讓其他研究者以*複製*的方式，而使解釋獲得證實。
8. 社會科學的證據必須具備*互為主體性*。
9. 對於知識採取*工具性導向*的立場，而在使用知識時，則以*技術官僚觀點*為主。
10. 社會科學應該是*價值中立*而客觀的。

治社會情境脈絡裡的許多變化所致。實證主義先是在美國成爲主流，接著，在二次世界大戰之後，由於美國成爲居於領導地位的世界強權國家，因此實證主義也陸陸續續地在許多國家裡，成爲社會研究的主流模式。在 1920 年代時，美國的社會學界出現了一股朝向客觀主義（objectivism）的推力，而客觀主義是更加嚴格的實證主義派別。隨著研究者從社會改革導向的研究（此種研究所使用的是較不正式或較不精確的研究技術），轉向以自然科學爲模範、以一種「價值中立」的方式使用嚴謹的研究技術時，客觀主義也就隨之日漸蓬勃發展。研究者創造了測量個體外在行爲的嚴謹技術，以產生能夠進行統計分析的量化資料。客觀主義取代了原本行動導向與偏向質化的地域性研究。客觀主義之所以能夠迅速成長，主要因爲研究者之間爲了爭取名望和地位而互相競爭。而其他種種的壓力，其中包括了來自私人基金會（像是福特基金會、洛克斐勒基金會等等）的資金贊助、那些想要避免非傳統政治力干擾的大學行政人員、研究者想要獲得嚴謹專業形象的渴望、擴張中的政府以及企業官僚體系的資訊需求等等。上述這些壓力結合在一起之後，重新定義了社會研究。由社會改革者（通常是女性）所進行的那些較不具技術性的、應用的地域性研究，其聲勢與光采就被大學各系所裡的男性教授所進行的非政治的、精確的量化研究所掩蓋了。[11] 二次世界大戰之後，在聯邦政府進行大規模擴張時期所作的關於研究資金的決策，也把社會科學推向實證主義的方向。

詮釋取向社會科學

詮釋取向社會科學（interpretive social science, ISS）的起源可以追溯到德國的社會學家韋伯（Max Weber, 1864-1920）及德國哲學家狄爾泰（Wilhelm Dilthey, 1833-1911）。在他的鉅著《人文科學導論》（*Introduction to the Human Sciences*）一書中，狄爾泰主張有兩種在本質上完全不同的科學：*自然科學*（*Naturwissenschaft*）以及*人文科學*（*Geisteswissenschaft*）。前者以抽象解釋（Erklärung 或 abstract explanation）爲基礎，後者則是以對於在特定歷史情境中，人們日常生活經驗的同理心理解（emphathetic understanding）爲基礎。而同理心理解又可以稱作***體悟***（verstehen）。韋伯主張社會科學應該研究有意義的、或是有目的之社會行動。他相當認同體悟的觀念，並覺得我們必須研究那些會形塑個體內在感受，與引導個體決定以某種方式行動的推理判斷以及動機。

詮釋取向社會科學（ISS） 社會研究的三種主要取向之一，該取向強調有意義的社會行動、社會建構的意義、以及價值相對主義。

體悟 一個來自於德文的詞彙，意思是指同理心的理解（也就是對於共享意義的深刻理解）。而在詮釋取向的社會科學裡，體悟也是社會研究的主要追求目標。

當我們在討論所謂的「社會行動」時，是指人們的行動在主觀意義方面與其他人的行為具有關聯性。例如，兩位自行車騎士無預期地相撞時，他們的行為就不應該稱為社會行動。但是，他們在相撞之前，可能曾企圖閃躲對方，而此

> 種企圖就可以被定義為社會行動……社會行動並非社會學因果解釋中唯一具有意義的重要行動，但卻是「詮釋社會學」最主要的研究對象。（Weber, 1981: 159）

詮釋取向的社會科學與**詮釋學**（hermeneutics）有關，這是一種起源於十九世紀的意義理論。這個詞彙來自於希臘神話中一位叫做赫密士（Hermes）的神，而他的工作就是把神的旨意傳達給凡人。因此，詮釋學這個詞「就字面上而言，是指把晦澀難解的事物變得明白易懂」（Blaikie, 1993: 28）。詮釋學被大量應用於人文科學之中（像是哲學、藝術史、宗教研究、語言學、文學批評等等）。詮釋學強調對文本（*text*）進行非常仔細的、詳盡的閱讀，以獲得廣泛而深刻的理解。文本可指一段對話、書面文字、或是圖片。我們會進行「一次閱讀」以發現蘊藏在文本中更深刻的、更豐富的意義。每一位讀者都會把他或她的主觀經驗帶進文本中。在研究文本時，研究者或是讀者會試圖去吸收、進入文本所呈現的整體觀點，然後再發展出對於每個部分與整體關係間的理解。換句話說，我們幾乎無法從表面上輕易地看出真正的意義，只有透過詳盡檢視與研究文本、仔細思考文本的許多訊息、以及尋找文本各個部分的連結，我們才能瞭解文本的真正意義。[12]

詮釋取向的社會科學有數種不同的變化型態，其中包括了詮釋學、建構主義、俗民方法論、認知取向、唯心論、現象學、主觀主義、以及質化社會學等等。[13] 有一種詮釋取向與 1920 到 1930 年代社會學當中的芝加哥學派有關，而該學派主要屬於象徵互動論。人們通常會把詮釋取向的社會學直接稱作質化研究，因爲大多數採用詮釋取向的研究者，都會使用參與觀察與田野研究這兩項研究技術。而這兩種技術都需要研究者投入許多時間，去與他們的研究對象進行直接的個人接觸。其他採取詮釋取向之研究者，則會以非常詳細的方式去分析對話的文字記錄稿、或是研究與行爲有關的錄影帶，其目的在於，尋找微妙的非言語溝通，以瞭解人們在這些情境脈絡中進行互動的細節。實證主義的研究者可能會針對數千人精確地測量所挑選出來的量化細節，並使用統計方法來分析所蒐集到的資料，而詮釋取向的研究者，則可能與一群人共同生活一年的時間，並蒐集大量且非常詳盡的質化資料。如此一來，這名研究者就能夠深入地瞭解人們如何在日常生活中創造意義。

詮釋取向的社會科學所關心的是，人們如何與彼此進行互動和相處。一般而言，詮釋取向意指*透過在自然情境中對人們進行直接而詳細的觀察方式，對於具有社會意義的行動進行系統性的分析，其目的則爲瞭解並詮釋人們創造與其維持社會世界的方式。*

詮釋學　一種與詮釋取向的社會科學有關的方法，此種方法起源於對文本資料進行宗教性以及文學性的研究，而在此種研究中，對於文本以及其中各部分與整體之間關係的深入探究，能夠揭露出更為深層的意義。

問題

1. *進行社會科學研究的最終目的爲何？*

對於詮釋取向的研究者而言，進行社會研究的目的就在於發展出對社會生活的理解，以及發現人

們如何在自然環境中建構意義。詮釋取向社會學的研究者想要得知何種事物對於其研究的人們而言，是有意義的或是有關的，這些人們又是如何體驗日常生活的。爲達成這個目標，研究者會非常深入地，逐漸認識處於某個特定社會情境中的人們，並努力以置身此一情境之人的觀點，看待此一社會情境。研究者試圖以最爲深入的方式瞭解研究對象的感受與詮釋，並試著透過研究對象的眼睛來看各種事件。哈潑（Harper）是一名詮釋取向的研究者，他曾經對一名住在鄉下的修車廠主人威利（Willie）進行過爲期十年的研究。而他在總結這項研究的目標時，曾經說過：「本項研究的目的在於分享威利的觀點。」

詮釋取向社會學的研究者所研究的是**有意義的社會行動**（meaningful social action），而不僅僅是人們可見的、外在的行爲。社會行動是指那些人們賦予主觀意義的行動，以及具有目的或意圖的活動。由於非人類的物種並不具文化和規畫事物的推理能力，因此，社會科學家應該研究人類所獨有的社會行爲。研究者必須考量到社會行動者所具有的理性能力，與行動的社會情境脈絡。例如，像是眨眼之類的身體反射動作，雖然是一種人類行爲，但卻鮮少是一種具有意圖的社會行動（也就是爲了某種特定理由或人類動機而做出的動作）。但是，在某些情境中，這個動作卻有可能是一種社會行動（也就是故意眨眼對某人示意）。除了具有目的之外，行動還必須具有社會性，而「某種行動若要被視爲具有社會性，並成爲社會科學家感興趣的研究主題，則行動者必須賦予該動作主觀意義，該動作也必須與其他人的活動有所關聯」（Blaikie, 1993: 37）。

大多數的人類行動都不具有任何內在意義，只有在一個由共享某個意義系統的人們所形成的社會情境脈絡中，這些行動才能獲得意義。共同的意義系統使人們可以把某種行動詮釋爲一個具有社會關聯性的動作或是行動。舉例而言，在某種情境中，與其他人一起舉起某隻手指，就能夠表達社會意義。至於此種動作所表達的具體意義（例如，指示方向、表達友誼、或是一種粗鄙的手勢），則取決於社會行動者所共享的文化意義系統。

2. *社會實在的基礎本質爲何？*

詮釋取向的社會科學把人類社會生活視爲是一種由人類所完成與建構的事物。身爲社會性生物，人們會採取有意義的行動來進行互動，並刻意地以此種有意義的行動，來創造出社會實在。實證主義者認爲社會生活「就在那兒」、而且等待人們去發現，與此種觀點對比之下，詮釋取向的社會科學採取了比較偏向唯名論的本體論。換句話說，社會實在主要是人們所察覺與認知的事物，而社會實在之所以存在，是因爲人們體驗到社會實在，並爲社會實在賦予意義。社會實在是流動的、脆弱的，而人們則是在一個溝通與協商的持續過程中，隨著與其他人進行互動而建構出社會實在。在面對周遭人們及事件時，人們會依賴許多未經證實的假定，並會使用那些他們視爲理所當然的知識。當人們與其他人進行互動

有意義的社會行動　指在社會環境中的社會行動，人們會賦予此種行動主觀方面的意義與重要性，而詮釋取向的社會科學則把此種行動視為是社會實在中最重要的面向。

且建構意義時，社會生活就會在人們的主觀經驗當中成形。捕捉人們對於現實的主觀感受，藉此真正地瞭解社會生活，是一件極爲重要的事情。在詮釋取向的社會科學裡，「我們有可能去接近、瞭解其他人，不過，只能透過間接的方式：我們一開始所體驗到的是手勢、聲音、及行動，而且只有在逐漸瞭解的過程中，我們才能一步步地從外在的符號進入深層的內在生活」（Bleicher, 1980: 9）。

在詮釋取向社會科學裡的**建構主義導向**（constructionist orientation），則假定人們是出於他們彼此之間的互動與信念來建構實在。沒有任何內在的本質會導致人們所看見的實在。例如，當你看到一把椅子時，在那個物體之內並沒有任何的「椅子特質」這種東西。相反地，在你眼中所看到的椅子，其實是來自於在某個特定的社會及時間中的人們所定義、接受、並理解爲椅子的東西。當然，在你面前的確實是一個由木頭、金屬、或是布料所製成的實體物體，組合成一種特別的形狀。然而，你所看見的、並認爲是實徵實在的那把椅子，事實上只是來自於一些告訴你要把這個物體定義爲椅子的文化社會過程。

一般而言，人們在社會世界所看見以及體驗到的事物，都是社會建構的。然而，僅僅因爲人們的經驗是社會建構的，並非意味著這些經驗就是虛幻的、無形的、或是無足輕重的。一旦人們把社會創造當成事實或真實的，這些創造就會產生非常真實的影響或是後果。例如，如果社會建構的實在告訴我，那個搬到我隔壁公寓的人曾經犯過暴力罪行，並且隨身攜帶槍械，那麼不論我這個建構出來的想法是否符合真實的情況，我都會根據我的想法而採取對應的行動。對於建構主義者而言，人們生活於、相信、並接受建構的實在，此種建構的實在與物理實在有所關聯，但是兩者之間卻或多或少有著清楚的差異。

建構主義者注意到，人們會把周遭的社會世界「視爲理所當然」，並在行動時把社會世界當成彷彿是自然的、客觀的、及固定實在的一個部分。舉例而言，人們已接受一周共有七天的觀念，因此，很少有人瞭解到，其實一周可以是非常不同的。在某些文化裡，一周可能有三天、五天、或甚至十天。以七天爲一周，並不是個物理實在，但是人們卻將之視爲理所當然，並認爲這是實在的一個自然的、固定不變的部分。我們現在所接受的一周七天的觀念，只是個社會建構而已，人們是在特定地區及特定歷史情境下，創造出這個觀念的。

實證主義社會科學所使用的語言，與實在之間有直接的連結，而且實證主義者也試圖使得這個語言盡可能地變得純粹、合乎邏輯、和精確，好讓這個語言能夠正確無誤地反映實在。對比之下，建構主義者把語言視爲組成社會建構的主要部分。當我們在學習語言時，我們就是在學習以特定的方式思考及觀看這個世界。語言與真正的實在之間，幾乎沒有什麼直接的關聯，且語言之中包含了特定的世界觀，此種世界觀會影響我們觀看和體驗世界的方式。此種差異性會繼續影響其他人的社會概念，像是性別及種族等等。例如，歐洲的英語系社會把性別分爲兩種類別，並且依據膚色的深淺而把種族分成

建構主義導向　一種對於社會實在的導向，該導向假定人們所創造的、使用的信念與意義，會從根本上形塑人們所認定的現實。

六種類別。實證主義社會科學的實在本體論認爲，性別以及種族都是真實的（也就是說，男性與女性、或是各個種族，都是在現實中基本的差異）。相較之下，建構主義者則認爲，語言與慣性思考方法會支配人們所看見的事物。在歐洲人眼中，或許會看到這個世界只有兩種性別及六個種族，但是在其他的文化裡，可能會看到超過兩種以上的性別以及不同數目的種族，且是除了用膚色以外的特徵來當作區隔不同種族的基礎。相對於實證主義社會科學對於「冰冷的硬事實」的要求，建構主義者所強調的，是人們創造社會建構、把社會建構當成彷彿是真正「事物」來使用的過程。[14]

實證主義社會科學假定每個人都是以完全相同的方式在體驗這個世界。詮釋取向則是質疑人們是否能夠以完全相同的方式來體驗社會或物理實在。對於詮釋取向社會科學的研究者而言，以下這兩個問題都是相當關鍵的問題：人們如何體驗這個世界？他們會創造並分享意義嗎？詮釋取向的社會科學會指出許多的例子，在這些例子裡，有一些人曾經看過、聽過、甚至碰觸過完全相同的實體物體。然而，他們對這個物體卻各自擁有不同的意義或是詮釋。詮釋取向的研究者主張，實證主義者把體驗世界的單一方式強加於其他人身上。對比下，詮釋取向的社會科學假定，對於人類經驗或實在的多重詮釋是可能存在的。簡而言之，詮釋取向的社會科學對於社會實在的定義如下：社會實在是由透過日常社會互動而建構意義以及創造詮釋的人們所組成的。

3. *人類的本質爲何？*

一般人都投入於一種透過社會互動來創造意義系統的持續過程。他們接著會利用這些意義詮釋他們的社會世界，並且理解他們的生活。人們的行爲或許會有其模式及規律，但是此情況並非一些等著我們去發現的、事先已存在的法則所導致。人類行爲的模式是來自於不斷演變的意義系統或社會傳統，而這些系統或是傳統，則是人們在進行社會互動時所創造出來的。對於詮釋取向的研究者而言，重要的問題是：人們真正信以爲真的事物是什麼？人們認爲與自身有所關聯的事物又是什麼？人們又是如何定義自己正在做的事情？

詮釋取向的研究者想要發現的是，對於從事各種行動的人們而言，行動所代表的意義爲何。這些研究者認爲，試圖從一些抽象的、邏輯的理論去演繹社會生活，其實沒有太大意義，因爲這些理論往往與一般人的日常感受及經驗沒有太多關聯。人們對於行動都有自己的理由，而我們則須瞭解人們所使用的理由。即使個體的動機是不理性的、帶有強烈的感情、並包含了錯誤的信念及偏見，這些動機仍然相當重要、應該列入考慮。有些詮釋取向社會科學的研究者認爲，只有在科學社群瞭解人們如何創造及使用意義系統、常識如何發展、以及人們如何把他們的常識應用於各種情境時，實證主義者才能發現他們所要尋找的法則。其他的詮釋取向社會科學的研究者則認爲，此種人類社會生活的法則根本就不存在。因此尋求這些法則，根本等同於是緣木求魚。例如，對於詮釋取向社會科學的研究者而言，當實證主義者試圖發現一些人類行爲的法則，以解釋失業會造成虐待兒

童的情況時，此種企圖充其量只能算是一種不成熟的舉動，如果從壞處來看，此種企圖或許還具有危險性。相反地，詮釋取向的研究者想要瞭解人們在主觀方面如何體驗失業、失去工作對他們的日常生活具有什麼樣的意義。同樣地，詮釋取向的研究者也想要瞭解，那些對兒童施虐的人會如何解釋自己的行爲、他們虐待兒童的理由是什麼、及他們對於虐待兒童這件事的感受爲何。詮釋取向的研究者會探究失業所具有的意義，還有虐待兒童的理由，以瞭解在那些直接受到失業影響的人們身上，到底發生了什麼事。

4. *關於人性能動作用（自由意志、意志力、以及理性）的觀點爲何？*

實證主義社會科學所強調的是決定論的關係，以及外在力量。而詮釋取向社會科學則是強調自願性個體的自由選擇，有時候也稱作*人性能動作用*（*human agency*）。詮釋取向的社會科學採取**唯意志論**（voluntarism）觀點，因而把人們視爲具有意志力（也就是能有意識地作出選擇）。社會環境及主觀的觀點有助於形塑個人所作的選擇，但是人們能夠創造並改變這些環境，擁有發展或形成新觀點的能力。詮釋取向的研究者強調考量個人決策過程、主觀感受、以及瞭解事件的方法這三者的重要性。在詮釋取向的社會科學裡，相較於實證主義者所強調的外在客觀情境以及結構力量，這個內在世界及一個人觀看和思考的方式，如果不是更爲重要，至少也是同等地重要。

5. *科學與常識之間的關係爲何？*

實證主義者認爲常識不如科學，相較之下，詮釋取向的社會科學則認爲，在日常生活中，一般人通常是運用常識來引導他們自己。常識是許多日常生活理論的累積，而人們則是運用這些理論來組織或解釋世界上的事件。我們有必要瞭解常識，因爲常識中包含了人們在從事日常例行的社會互動時所使用的意義。

詮釋取向的社會科學認爲，常識以及實證主義的法則，是用來解釋這個世界的可能方式。也就是說，這兩者明顯是不同的意義系統。不論是常識或科學法則，都無法提供所有問題的答案。相反地，詮釋取向的研究者認爲，科學法則以及常識兩者在其各自所屬的領域中，都非常重要。換言之，我們是爲了不同的目的、以不同的方法創造出科學法則與常識的。如果一般人在日常生活中，單獨只以科學作爲其行動的基礎，那麼他們就無法正常地過生活。舉例而言，在煮蛋時，人們必須使用非系統性的經驗、習慣、及猜測。如果要嚴格地運用自然科學的法則，那麼人們就必須先知道那些與加熱開水有關的物理定律、以及與雞蛋內部物質的變化有關的化學定律才行。即使是自然科學家，當他們不是在自己的專業領域中「研究科學」時，也一樣要運用到常識。

在瞭解人們時，常識是一個相當重要的資訊來源。一個人所具有的常識，乃來自於一種實用的導向、一組關於這個世界的假定。由於人們必須使用常識完成事務，因此他們必須假定常識是真的。詮釋哲學家舒茲（Alfred Schutz, 1899-1959）把此種現

唯意志論　關於人性能動作用以及因果關係的一種取向，該取向假定人類行動是基於主觀的選擇以及個人的理由。

象稱爲**自然態度**（natural attitude），也就是假定這個世界在你出生之前就已經存在了，而且在你離開人世之後，這個世界仍然會繼續存在下去。人們會根據他們在與他人進行社會互動的過程中所創造出來的意義系統，發展出許多方法，來維持或是再製他們對於實在的感知。

6. *社會實在的解釋或理論是由什麼所構成的？*

實證主義社會科學的理論試著去模仿自然科學的理論，因此或許會包含演繹原則、定理、相互關聯的因果法則。相對地，詮釋取向社會科學的理論則揚棄了相互關聯的因果法則以及命題，轉而敘述一個故事。詮釋取向的社會科學描述並詮釋人們過日常生活的方式。雖然詮釋取向社會科學的理論可能包含一些社會科學的概念以及有限的推論，但卻不會過度地遠離被研究者的生活經驗與內在現實。

詮釋取向的社會科學是個殊的以及歸納的。所謂**個殊的**（idiographic），是指這種取向會針對某個事物，提供一種象徵性的再現或是「豐厚」（"think"）的描述。因此，一份詮釋取向的研究報告，讀起來可能像是在讀一本小說或傳記一樣，其中充滿了各種細節的詳細描述，但卻少有抽象的部分。就像是對於文學作品的詮釋一樣，詮釋取向的研究報告也有其內在的連貫性，並以文本爲基礎。而在此處的文本，所指的是被研究的人們有意義的日常生活經驗。

詮釋取向社會科學理論的目的，在於提供一種詮釋性的解釋（請參閱第二章）。藉由揭露意義、價值觀、詮釋策略、以及生活規則的方式，詮釋取向的社會科學試圖讓讀者深刻地感受到另一個人的社會實在。例如，詮釋取向社會科學的理論或許會描述人們在一個環境中，辨識並詮釋其經驗時所使用的主要典型。所謂**典型**（typifications）就是人們在分類與組織其經驗的成串生活事件時，所使用的一種非正式模型、策略、或是一組信念。

詮釋取向社會科學的理論就像是一幅地圖，概略地描繪出一個社會世界，並描述地域性的習俗及規範。舉例而言，一份關於職業賭徒的詮釋性報告，可能會告訴讀者關於職業賭徒的生涯和他們日常關心的事物。這份報告會描述所研究的特定人物、進行研究的地點、觀察的活動、以及賭博時所應用的策略等等。因此，讀者就可以得知職業賭徒的交談方式、他們看待別人的方式、及他們的恐懼和抱負是什麼。研究者可能會提供一些推論與組織化的概念，但是整份報告的主體則是對於賭博世界的詳細敘述。理論與證據會交織在一起，以創造出一個具有一致性的整體，而各種概念及推論，則與其情境脈絡緊密地結合。

自然態度　詮釋取向社會科學所使用的一個觀念，該觀念假定常識理解的世界是穩定的、真實的，而且在沒有劇烈變化的情況下從過去一直持續到未來。而我們之所以會這麼做的原因，乃是出自於完成日常工作的實際需要。

個殊的　一種應用於詮釋取向社會科學的解釋類型，在此種類型中，一個解釋就是關於一個社會情境或環境的深入描述或者是圖片。其中具有豐富的具體細節，但是卻沒有太多抽象概念。

典型　一種非正式的模型或計畫，人們在日常生活中用以進行分類，以及組織他們所經驗的一連串事件與情境。典型通常是普通知識或常識的一部分，能夠簡化並且幫助人們組織一連串的生活事件與其複雜性。

7. *如何確定某個解釋的真假？*

實證主義社會科學會以合乎邏輯的方式，從理論進行演繹、蒐集資料、並分析事實， 整個過程都能夠讓他人加以複製。對於詮釋取向的社會科學而言，一個理論是否爲真，取決於該理論對被研究的人們是否具有意義、該理論是否能夠讓其他人進入被研究者的現實裡。如果研究者傳達了對於其他人的推理過程、感受、以及看待事物之方式的深入理解，該理論或描述就是正確的。或許我們可能作出預測，但是此種預測僅能發生於兩個非常親密的人之間，例如一對已經結婚很久的夫妻。一項詮釋取向的解釋，詳細記錄了行動者的觀點，並將之轉譯成一種讀者能理解的形式。史馬特（Smart, 1976: 100）將這此種情形稱爲**適當性假定**（postulate of adequacy）：

> 適當性假定宣稱，如果一項關於人類行動的科學解釋，就像呈現給個別行動者（演員）的劇本，那麼這項解釋必須能夠讓該名行動者（演員）瞭解、能夠由該名行動者（演員）轉化為行動、並能讓其同伴以日常生活中的常識性詮釋，就能夠完全理解。

就像一位訴說異地見聞的旅行者般，研究者也並非當地人士。而這種外界的觀點，永遠也比不上被研究的人們所可能提供的內部人士解釋。話雖如此，越能夠接近當地人的解釋，當然會越理想。例如，檢驗一份關於職業賭徒的詮釋性研究真實性的方法之一，就是找一些職業賭徒來閱讀這份研究報告，並證明其正確性。一份好的報告，能夠提供讀者關於職業賭博世界的豐富資訊，如此一來，如果讀者充分吸收這份報告內容的話，那麼當這名讀者日後遇上一名職業賭徒時，他對於賭博行話、生涯展望、及生活型態的深入瞭解，或許會使得這名賭徒忍不住好奇，這名讀者是否也是一名職業賭徒。

8. *良好證據或是事實資訊的特徵爲何？*

在實證主義裡，良好的證據是指那些可以觀察的、精確的、且獨立於理論以及價值之外的證據。對比之下，詮釋取向的社會科學則認爲，在瞭解社會意義時，特定脈絡及意義特徵是不可或缺的。關於社會行動的證據，不能從該行動所發生的情境脈絡中抽離出來，也不能無視於相關的社會行動者賦予該行動的意義。如同韋伯（Weber, 1978: 5）所說的：「當我們透過同情的參與而能夠適當地捕捉行動發生的情感脈絡時，就是正確地對他人感同身受、或是理解他人的想法。」

對於詮釋取向的社會科學而言，事實是變動不居、而且內嵌於一個意義系統之中的。此外，事實也不是毫無偏見的、客觀的、或者中立的。事實是偶然發生的、有其特定情境脈絡的，而且事實取決於特定情境中的特定事件，與特定人群的結合。實證主義社會科學假定，中立的局外人能夠去觀察人們的行爲、並看到清楚客觀的事實。對於這樣的假定，詮釋取向的社會科學則認爲，這是一個必須加以處理的問題：人們如何能夠觀察在社會生活中

適當性假定　一項詮釋取向社會科學的原則，亦即各種解釋都應讓被研究者只利用常識就能夠理解。

的曖昧狀況、並且賦予意義呢？詮釋取向的研究者認爲，社會情境裡充斥著模糊與曖昧不清的情況。大多數的行爲或陳述都可能有數種不同的意義，而且能夠以多重不同的方式加以詮釋。在社會生活的流動過程中，人們不斷地重新評估在情境中的各種線索，並且賦予意義，以使自己能夠瞭解各種事件，直到他們「知道發生什麼事情」爲止。例如，假定我看到一名女士伸出她的手，而且掌心朝前。即使是這麼一個簡單的動作，也帶有多種可能的意義。如果我不明白整個社會情境，我就無法確切地知道這個動作的意義。這個動作可能意味著她正在擋開一名潛在的扒手、吹乾她的指甲油、招呼一輛計程車、欣賞一枚新的戒指、要來往的車輛停下來、或是在一家熟食店的櫃檯前面點五個貝果。[15] 只有當人們將某個動作或某句話語所發生的社會情境脈絡納入考量時，才能夠爲這個動作或是話語賦予適當的意義。

詮釋取向社會科學的研究者很少詢問調查的問題、累積許多人的回答、或者宣稱要獲得對於某些問題具意義的事物。對於詮釋取向社會科學的研究者而言，每個人對於調查問題的詮釋，都必須被置入於一個情境脈絡之中（例如，個人之前的經驗，或是進行調查訪談時的情境），而且個人答案的真正意義，也會依據訪談或詢問的情境不同，有所變化。此外，由於每個人賦予問題與答案的意義多少會有些不同之處，因此把許多人的答案結合起來時，所得到的只是無意義的資訊而已。

在研究一個環境或是資料時，俗民方法論學派的詮釋取向研究者，通常會使用**懸置法**（bracketing）（譯註：又可譯爲括弧法、放入括弧、存而不論）。這是一種心智的運作方式，作法是研究者先辨識出一些在社會科學中被認爲是理所當然的假定，並將其擺在一旁。詮釋取向社會科學的研究者會去質疑、並重新檢視那些對於牽涉在其中的人們具有「明顯」意義的日常事件。例如，在一個辦公室的工作環境中，一位年近三十歲的男性員工對一名男性研究者說：「今晚下班後，我們一群人要聚在一起打壘球。你要不要一起來？」在上述的情境中，*沒有明白說出來的*，是這名研究者應該要瞭解壘球的規則、擁有壘球手套、並在比賽前先把西裝換成運動服。懸置法揭露出什麼是「每個人都知道」的事：也就是大家都心照不宣、很少說出口的事。懸置法能夠使社會場景中那些促成其他事件的重要特徵變得顯而易見，也能揭露出潛藏在各種理解之下的架構，而社會行動正是以這些架構爲基礎來進行的。

懸置法　詮釋取向社會科學的研究者所採用的一種策略，作法是先辨識出一些在社會科學中被視為理所當然的假定，然後將這些假定擺在一旁、或是暫時擱置。藉由辨識並分隔出人們在日常生活中使用的一般的、「明顯」的意義，研究者就能更加深入地瞭解這些意義所扮演的角色。

9. *社會科學知識的關聯性以及作用爲何？*

詮釋取向的社會科學家想要瞭解這個世界運作的方式，如此一來，他們才能對其他人有深刻的瞭解、欣賞人們實際生活經驗的廣泛多樣性、並更深入地認識共通的人性。詮釋取向社會科學的研究者並未把知識當作是一種工具或手段，相反地，他們試圖去捕捉一般人的內在生活以及主觀經驗。此種人文主義的取向，其焦點在於人們如何在日常生活中處理各種實際的事務、而人們如何把社會知識當作是種實用的成就。

根據詮釋取向社會科學的**實用性導向**（practical orientation），社會科學知識的相關性，在於此種知識能夠以真實的、全面性的方式，反映出一般人在日常生活的情境中做事的方式。詮釋取向的社會科學也強調要融合知識創造的社會脈絡，並創造出一種反身形式的知識（我們曾經在第二章裡討論過此種知識形式）。

對於新知識的使用與應用，詮釋取向社會科學的研究者通常會採取一種**超驗觀點**（transcendent perspective）。所謂的*超驗*（*transcend*），是指要超越一般物質的經驗和感覺。而在社會研究中，這意味著不要停留在表面或可以觀察的層次，而要深入到人類經驗的內在，以及主觀層次。超驗觀點不把人們當成研究者要研究的外在物體，而是鼓勵研究者去檢視人們複雜的內在生活。此外，超驗觀點也不鼓勵研究社會情境的當前狀態，而是要求研究者應該檢視人們主動建構並轉變現存情境的過程。詮釋取向社會科學的研究者，試圖去熟悉並參與那些被研究的人們，以獲得對他們的親密熟悉感。超驗觀點強調，研究者與被研究的人們應該共同合作，以創造相互的理解，及其對環境發揮影響力。

10. *社會政治價值是由何處滲透到科學中？*

實證主義社會科學的研究者呼籲要摒棄價值立場，並在無政治色彩的環境中進行研究。對比之下，詮釋取向社會科學的研究者則主張，研究者應該反思、重新檢視、分析個人的觀點與感受，將之當作研究他人過程中的一部分。詮釋取向社會科學的研究者必須（至少是暫時性地）去感受或分享其研究對象所擁有的社會與政治承諾、或價值觀。這就是爲什麼在價值方面，詮釋取向社會科學的研究者採取了**相對主義**（relativism）立場的原因。

詮釋取向的社會科學質疑價值中立的可能性，因爲詮釋取向的研究認爲價值與意義是滲透於所有地方，以及所有事物之內的。實證主義社會科學所宣稱的價值中立，不過就是另一種意義系統和價值，意即實證主義科學的價值。詮釋取向的研究者採取相對主義的立場，不認爲任何一組價值是較爲優越或是劣等的。研究者應該要清楚地承認自己的價值立場，並坦白地將其價值立場公諸於世。

摘要

許多年以來，詮釋取向的社會科學一直忠實地扮演著實證主義反對者的角色。雖然有些實證主義的社會研究者，願意接受詮釋取向，並認爲詮釋取向在探索性研究（請參閱第二章）中是相當有幫助

實用性導向　一種對於社會知識的實用主義導向，在此種導向中，人們會將知識應用於日常生活之中。至於知識的價值，則是在於能夠與人們實際的日常生活理解與選擇結合在一起的能力。

超驗觀點　研究者與研究對象共同發展研究、檢視研究對象的內在生活，以獲得對於研究對象的親密熟悉感，此外，研究者也會和研究對象密切合作，以創造出共同的理解。

相對主義　詮釋取向社會科學所使用的一種原則，意指沒有任何一個觀點或是價值立場，優於其他觀點或是價值立場。而且對於抱持這些觀點或是價值立場的人們而言，這些觀點與價值立場同樣都是合理而有效的。

批判取向社會科學（CSS） 社會研究的三大主要取向之一，此取向強調的是對抗表層的扭曲現象、多重層次的社會實在、以及為人們賦權，並以價值為基礎的行動主義。

的。但是卻很少有實證主義者認爲詮釋取向合乎科學性。在往後的章節裡，當你讀到田野研究時，會再次讀到關於詮釋取向的觀點，而學習歷史比較研究時，也會提到一些相關內容。詮釋取向是一些社會研究技術的基礎，這些技術包括能夠敏銳感知情境脈絡的技術，瞭解他人觀看世界方式的技術，與一些較關心達到同情理解、而非像是測試人類行爲理論法則之類的技術。圖表 4.3 則提供了關於詮釋取向的摘要。

批判取向社會科學

批判取向社會科學（critical social science, CSS）的各種版本被稱作*辯證唯物論*（*dialectical materialism*）、*階級分析*（*class analysis*）、及*批判結構主義*（*critical structuralism*）等等。[16] 批判取向的社會科學結合了律則的與個殊的兩種取向，它認同詮釋取向對於實證主義社會科學的許多批評，但也加入了一些自己的批評，而且在某些觀點上，與詮釋取向的社會科學並不一致。我們可以追溯此一取向的源頭到馬克思（Karl Marx, 1818-1883）及佛洛依德（Sigmund Freud, 1856-1939）。之後，阿多諾（Theodor Adorno, 1903-1969）、弗洛姆（Erich Fromm, 1900-1980）、以及馬庫色（Herbert Marcuse, 1898-1979）等人更進一步地闡述並豐富批判取向。批判取向的社會科學通常會和衝突理論、女性主義分析、激進的精神心理療法聯想在一起，也被認爲和 1930 年代德國的法蘭克福學派所發展出的批判理論有著密切關係。[17] 批判取向的社會科學批評實證主義科學對於理性的運用是狹隘的、違反民主的、而且也是失去人性的。阿多諾曾在他所寫的兩篇文章〈社會學與實徵研究〉（*Sociology and Empirical Research*, 1976a）與〈社會科學的邏輯〉（*The Logic of the Social Sciences*, 1976b）中概略地提及此種觀點。另外一位屬於這個學派、且仍然健在的知名代表人物哈伯瑪斯（Jurgen Habermas, 1929-），則在他的著作《知識以及人類旨趣》（*Knowledge and Human Interests*, 1971）一書中，更進一步地擴展了批判取向的社會科學。而在教育的領域裡，弗雷勒（Paulo Freire, 1921-1997）所著的《受壓迫者的教育學》（*Pedagogy of the Oppressed*, 1970）也被認爲歸屬於批判取向的社會科學之列。

另外一個例子則是法國社會學家布爾迪厄（Pierre Bourdieu, 1930-2002），以及他的著作，如《實作理論綱要》（*Outline of A Theory of Practice*, 1977）等等。[18] 布爾迪厄既拒絕實證主義者所倡導的那種客觀的、彷彿法則般的量化實徵研究，也不接受詮釋取向社會科學那種主觀的、唯意志論的取向。他主張社會研究必須是具有反身性的（換句話說，就是除了研究的主題之外，也需要研究並批判研究本身），亦必定是帶有政治色彩的。他同時也認爲，研究的目標之一，就是去揭露和除去日常事件中的神秘部分。

圖表 4.3　詮釋取向社會科學摘要

1. 社會科學的目的在於瞭解情境脈絡下的社會意義。
2. *建構主義*的觀點認為實在乃是由社會建構的。
3. 人類是會進行互動的社會性生物，並創造與強化共享的意義。
4. 在人性能動作用方面採取*唯意志論*的立場。
5. 科學知識異於其他形式的知識，但彼此之間沒有孰優孰劣的問題。
6. 解釋是*個殊*的，而且會透過歸納的推理過程，不斷向前進展。
7. 利用關於研究對象的*適當性假定*來驗證解釋。
8. 社會科學的證據是偶然的、有其特定情境脈絡的、且通常必須使用*戀置法*來加以驗證。
9. 對於知識採取*實用性導向*，並以*先驗觀點*來應用知識。
10. 在價值立場方面，社會科學應該採取*相對主義*的觀點。

詮釋取向的社會科學批評實證主義社會科學未能處理人們在真實生活中的意義以及其思考與感知的能力、忽視了社會情境脈絡、也違反人性。批判取向的社會科學認同大多數此種對於實證主義社會科學的批評，也相信實證主義社會科學只是在維護現狀而已。然而，批判取向的社會科學也對詮釋取向的社會科學提出批評，認爲後者過於偏向主觀性與相對性，把人們的想法當作比實際狀況（例如，真正的貧窮、壓迫、及暴力等等）更爲重要。批判取向的社會科學亦認爲，詮釋取向的社會科學過於強調地域性、微觀層次、短期的處境，反而忽視了更廣大的、和長期的結構性條件。對於批判取向的社會科學而言，詮釋取向的社會科學不關心道德、且又是被動的。詮釋取向的社會科學沒有採取強烈的價值立場，也未能積極主動地幫助人們看清周遭的謬誤假象。而批判取向的社會科學則主動地涉入人們的生活，幫助一般人改善生活。總結而言，批判取向的社會科學把社會科學定義爲*一種探究的批判過程，此過程穿透表層的假象以揭露物質世界的真實結構，其目的在於幫助人們改變處境，並爲自己建構一個更美好的世界。*

問題

1. *進行社會科學研究的最終目的爲何？*

依據批判取向社會科學的觀點，進行研究最主要的目的，並非只是單純研究社會世界而已，而是要去改變社會世界。批判取向社會科學的研究者進行研究的目的在於批判並改變社會關係，而其方法則是去揭露社會控制、權力關係、與不平等現象的底層源頭。藉由揭露各種情境的方式，批判取向的社會科學賦予人們力量，尤其是那些在社會中不具有權力及被邊緣化的人。更明確地說，批判取向的社會科學想要打破迷思、揭露隱藏的真相、並協助人們改善他們的生活。對於批判取向的社會科學而言，進行研究的目的在於「以一種特定的方式來解釋社會秩序，使自己成爲一種引導社會秩序大幅度轉型的催化劑」（Fay, 1987: 27）。

批判取向社會科學的研究者會提出一些許多人不敢正視的問題、揭露各種虛僞的情況、並深入調查各種狀況以刺激基層民眾採取行動。「所有科學的重點，也就是

所有學習的重點，在於改變且發展我們的理解、減少幻象……學習就是去減少幻象與無知。學習能夠幫助我們從至今尚未被察覺到的限制、教條、以及謬誤之中解放出來」（Sayer, 1992: 252）。

舉例而言，批判取向社會科學的研究者，可能會進行一項與租屋行爲中的種族歧視有關的研究，而其研究問題可能是：白人房東會不會拒絕把房子租給少數民族的房客呢？批判取向的研究者不會只發表一篇研究報告，然後就等著市政府裡的公平住屋部門去採取行動。這名研究者會把研究報告投稿到報紙上，並且拜會基層組織以共同討論研究的結果與發現。此外，這名研究者會與社會運動人士合作，並以社會正義之名發起政治行動。當基層的人們到房東的辦公室裡進行監督、向房東提出讓少數民族承租公寓的大量申請、或者是組織一場遊行到市政府去要求採取行動時，批判取向的研究者預測這將能迫使房東把房子租給少數民族。因此，研究的目標就在於彰權賦能。金奇洛與麥拉倫（Kincheloe and McLaren, 1994: 140）曾經這麼說過：

> 要能最清楚地瞭解批判研究，就必須從幫助個人彰權賦能的脈絡裡著手。一項研究若想高舉「批判」之名，就必須結合一種對抗某個特定社會裡、或在該社會裡，某個特定範圍內的不公正企圖之情形。如此一來，研究就變成一種促進改革的力量，不會因為被貼上「政治的」標籤而進退維谷、也無懼於促進解放的意識。

2. *社會實在的基礎本質爲何？*

批判取向社會科學認同實證主義社會科學前提裡的一些面向，實徵主義的實在，的確獨立存在於我們的感知外。除此之外，批判取向社會科學也同意詮釋社會科學取向的強調重點，也就是經由我們的主觀經驗、文化信仰、社會互動，去建構我們所認爲的實在。批判取向的社會科學採取了一種批判的實在本體論，此種本體論把實在視爲是由多重層次所組成的，其中包含了實徵層次、真實層次（the real）、以及事實層次（the actual）。[19] 我們可以利用感官去觀察到實徵層次的實在。然而，我們所體驗到的表面實徵層次，其實是由更深層的結構及因果機制所產生的，這些結構與機制是在我們無法觀察的層次裡運作的。隨著時間經過，理論與研究能夠幫助我們瞭解，那些在真實層次裡運作的結構，及其在事實層次運作的因果機制。而在事實層次運作的因果機制，會產生結構上的修改。

我們可以直接觀察到真實層次的結構，此種結構並非固定不變、而是會隨著時間而演變的。同時我們可以修正這些結構。例如，在真實層次的性別結構，會形塑表層人們的特定具體行爲，而這些行爲都是我們可直接觀察到的。只要以理論性洞見與詳細調查，研究者就能逐漸揭露這些深層的結構。但是由於這些結構會產生變化，因此這項工作也變得相當複雜。位於較深層次裡的結構，不會在實徵層次上產生直接及立即的表象。這些深層結構可能一開始並不會發揮作用、或處於靜止不動的狀態，之後可能開始產生作用並且浮現在表面上。除此之外，各種不同的結構之間，並非彼此孤立隔絕。當我們反制或抵消某些結構的作用時，可能會同時壓抑另

一個結構的表層作用、或使其變得更爲複雜棘手。

在事實層次運作的因果機制可能具有內在的矛盾，並以一種自相矛盾的方式運作，因而導致結構方面的衝突。這些機制可能包含了一些力量或是過程，這些力量或過程在表面上看起來可能是相反的、或相互衝突，但是事實上卻都是一個更大過程的一部分。利用一個生物學方面的類比，可以幫助我們更加清楚瞭解這個觀念。我們把出生和生命視爲與死亡對立，然而，從我們出生的那一天起，就開始了死亡的過程。而且隨著我們的身體逐漸老化及衰敗，我們活著的每一天都不斷地朝著死亡前進。在生命與死亡之間，存在著內在的矛盾。只要我們活著，就會不斷地朝向生命的對立面——死亡——前進。活著及死去在表面上看來似乎是互相對立的，然而，這兩者事實上只是同一個過程的兩個不同部分。發現且瞭解此種被稱作**辯證**（dialectic）的弔詭過程，正是批判取向社會科學裡最重要的任務。

批判取向的社會科學認爲，我們對於實徵實在的觀察與經驗都不是純粹的、中立的、直接的。相反地，各種觀念、信仰、以及詮釋都會扭曲或影響我們所觀察到的事物和觀察的方式。我們對於實徵實在的知識能夠捕捉到事實真正存在的方式，然而這卻是一種不完整的捕捉方法，因爲我們的經驗會受限於各種觀念以及信仰的影響。批判取向社會科學主張，我們對於實徵實在的經驗總是取決於理論或是概念。換句話說，我們所擁有的理論或概念，不論常識的或科學的，都使我們對於實徵實在的某些特定層面較爲敏銳、左右了我們認爲在社會實在當中與我們有所關聯的事物、並影響我們分類與劃分特徵的方式。隨著時間經過，新的理論洞見及概念，將會使我們辨識出實徵實在表層中的更多層面，並增進我們對於更爲深層實在的瞭解。

總結而言，實證主義社會科學所強調的，是外部實在運作於人們身上的方式，而詮釋取向社會科學則較爲重視社會實在的內在主觀建構。至於批判取向的社會科學則認爲，有一個預先建構的、並非由我們所創造的深層實在。這個深層實在其實在我們經驗它或思考它之前，就已經存在了，且會對人們產生真實的影響。與此同時，我們則建構了許多觀察與思考的方法，這些方法形塑了關於我們實徵實在的經驗。我們的思想可以引導我們去採取行動，這些行動將可以改變位於較深層次的實在裡的結構。批判取向社會科學把我們瞭解實在的能力視爲一種互動過程，在這個過程中，想法、經驗、及行動會隨著時間經過，彼此進行互動。

批判取向社會科學注意到，社會變遷與衝突並非總是顯而易見、或是能夠被輕易地觀察到的。社會世界充滿了假象、迷思、以及扭曲。由於人類的感官具有許多限制，因此我們對於世界的初步觀察結果只是局部的，而且通常具有誤導作用。表面實在所呈現的外貌，並非必然來自有意識的欺騙。各種物體、事件、或是社會關係直接可以感知的特徵，幾乎不會揭露一切真實情況。這些假象讓社會上的某些團體，得以把持權力並剝削他人。德國的社會學家及政治思想家馬克斯，便以強而有力的方式，說出下列這段話（Marx and Engels, 1947: 39）：

辯證　批判取向社會科學所強調的一種改變過程，在此種過程中，社會關係包含無法解決的內在矛盾，而隨著時間經過，這些矛盾將會引發劇烈的混亂，並徹底地重新建構社會關係。

> 在每一個時代裡，統治階級的思想，就是主流的思想……擁有可以供其使用的物質生產工具的階級，同時也掌控了心智生產的工具。於是……那些缺乏心智生產工具的人們，其思想就只能被他人所支配。

批判取向的社會科學認爲，雖然主觀的意義是重要而真實的，客觀的關係也會形塑社會關係。因此，批判取向的研究者會深入去探究社會情境，將其置入於一個較大的歷史脈絡之中。

舉例而言，假設有一位詮釋取向社會科學的研究者，研究一位男性老闆與女性秘書之間的互動過程，那麼這名研究者將會提出一份既豐富又精彩的研究報告，以說明這兩者間的行爲規則、詮釋機制、及意義系統等等。對比之下，批判取向社會科學的研究者則會先從一個觀點（例如，女性主義）切入，並注意到一些在詮釋取向的描述中，所忽視的議題。例如，爲什麼老闆是男性，而秘書則是女性？爲什麼老闆和秘書這兩種角色所擁有的權力是不平等的？爲什麼大型組織所創造出的這些角色遍布整個社會？此種不平等的權力關係在歷史上如何形成，秘書爲什麼總是由女性來擔任？爲什麼老闆可以講一些讓秘書覺得難堪的粗俗笑話？根據老闆面對的日常生活情境（老闆享有優渥的薪資、鄉村俱樂部的會員資格、新車子、氣派的豪宅、退休金規劃、股票投資等等），秘書面對的日常生活情境（低時薪、嗷嗷待哺的孩子、支付帳單的壓力、唯一的娛樂只有看電視等等），老闆與秘書這兩個角色之間如何產生衝突？秘書是否能夠聯合其他人來挑戰她的老闆，以及其他老闆的權力？

3. *人類的本質爲何？*

實證主義社會科學把人類視爲哺乳動物，把焦點集中於人類作爲一理性行動的個體及其行爲上。詮釋取向社會主義則認爲，人類在本質上是社會性的生物，具有創造並維持社會意義的能力。批判取向社會主義則承認，人類是受到社會結構影響的理性決策制訂者，也是能夠建構意義與社會結構的、具有創造力的生物。社會先於人們存在、並獨立於人們而存在。然而，只有在人們主動積極的參與之下，社會才能夠存在。人們創造了社會，而社會也創造了人們，之後，人們則在一個持續的過程中創造社會。[20] 因此，人類存在於一個持續的關係過程（relational process）之中。

批判取向社會科學注意到，人類可能會被誤導並且擁有尚未完全發揮的潛力。造成此種情況的原因之一是**物化**（reification）。所謂物化，是指在我們創造了某些事物之後，我們變得與其疏離、看不見我們與這些事物的關聯或關係時，所發生的情況。當我們切斷了我們與創造物之間的關聯時，我們就無法在這些創造之中看到我們自己，而只是把它們視爲無關的、外在的力量，這些力量反而掌控了我們。由於「忘記」並看不見彼此之間的關聯性，我們對於自己所創造出來的事物就失去了控制的力量。人類擁有極爲龐大的潛力，但是這些潛力卻經常無法完全發揮。這是因爲我們感到從我們自己創造的信仰、條件、情境中掙

物化 批判取向社會科學中所使用的一個概念，所指的是，人們疏離了自己所創造的事物、看不見自己與這些事物之間的關聯性，並把這些事物當成是無關的、外在的力量。

脫出來，是一件極爲困難的事。爲了要發揮所有的潛力，人們必須超越直接的表面現象，並突破他們所物化的事物。如此一來，人們才能看到他們其實擁有改變各種現狀的巨大能力。

4. *關於人性能動作用（自由意志、意志力、以及理性）的觀點爲何？*

批判取向社會科學結合了決定及唯意志論，並強調**有限的自主**（bounded autonomy），也就是能動作用與結構之間應該如何相互配合的方式。所謂有限的自主，意指自由意志、選擇、決策制訂等等，都不是毫無限制或開放的。相反地，這些事物若不是必須維持在有嚴格界限的選擇範圍之內，就是被限制於一些界限之內，而這些界限可能是文化或是物質的界限。批判取向社會科學的研究者會辨識出一個選擇的範圍、或至少是人們認爲實際上可供他們選擇的事物。而允許人們在這個範圍內作選擇時，可以運用一些自由意志。人們可以作出選擇，但是這些選擇會被限制於人們認爲可能的事物範圍內。物質因素（例如，天然資源、身體能力等等）和文化主觀的結構（例如，信仰、核心價值觀、意或是根深蒂固的規範等等）決定了人們認爲可能或不可能的事物，人們亦會根據他們認爲可能的事物採取行動。

塞威爾（Sewell, 1992）觀察到社會結構同時是文化的及物質的。一個人所看到的、想到的、或感覺到的事物（也就是文化），會形塑這個人在物質世界裡的行動。物質的物體、情境、以及資源等等，都取決於文化的基模。研究者承認「所謂堅實的資料，本身就是需要加以詮釋的文化產物」（Sewell, 2005: 190）。如果一個人的世界觀把某項行動定義爲不可能的、把某項物質資源定義爲無法取得的、或是把某個選擇視爲受到阻礙的，那麼他或是她的「自由意志」就會受到限制。如果由於文化上的理由，某個人不認爲昆蟲是一種食物的來源、或不認爲同時擁有三個太太在道德上是可以接受的，那麼，文化信念就會限制了他在物質資源方面的使用方式，並使他不可能去採取某些行動。物質因素與主觀文化因素會彼此交互作用。文化主觀信念會定義哪些物質資源是可取得或可利用的，因此限制了選擇的可能性，而物質環境則能夠形塑人們的文化主觀經驗和信念。在某些情境下，人類集體行動能夠改變物質環境及文化信念的深層結構，當此種情況發生時，就能擴展意志與選擇的範圍。

5. *科學與常識間的關係爲何？*

批判取向社會科學認爲常識包含了**錯誤意識**（false consciousness），認爲人們經常會被誤導，並作出一些與他們在客觀實在中的最佳利益背道而馳的行動。客觀實在隱藏於迷思及假象的背後。詮釋取向社會科學認爲錯誤意識是毫無意義的，因爲錯誤意識這個概念暗示著，一個社會行動者使用了一個錯誤的、或與客觀實在脫節的意義系統。詮釋取向社會科學主張，人們創造並使用各種意義系統，而研究者只能描述這些意義系統，卻不能判斷這些

有限的自主　批判取向社會科學裡對於人性能動作用以及因果關係的一種取向，該取向假定人類行動是基於主觀選擇與理智，卻只能在可辨識的有限範圍內採取行動。

錯誤意識　批判取向社會科學所使用的一種觀念，即是人們對於實徵環境及他們真正的利益，經常抱持錯誤的、或被誤導的想法。

系統的價值高低優劣。批判取向社會科學則主張，社會研究者應該要去研究主觀想法以及常識，因爲這些事物會形塑人類的行爲，但是這些主觀想法與常識卻包含了許多迷思與假象。這些迷思與假象會蒙蔽一個客觀的世界，在這個客觀的世界裡，對於資源和權力的控制其實是相當不平等的。

批判取向的研究者所談論的結構，其實並不容易看得清楚。研究者首先必須說明、解釋這些結構，並揭開覆蓋於表象之上的神秘面紗。光是仔細的觀察還不夠，因爲這樣無法告訴我們應該要觀察什麼事物，且光是觀察一個假象，並不足以驅除這個假象。研究者必須善加利用理論，挖掘隱藏在表面之下的真正關係、去觀察各種危機與劇烈衝突的時期、徹底調查各種互相聯結的關係、回溯過去以及思考未來的可能性。揭露出社會實在的更深層次，是一件困難重重、但卻極爲重要的事，因爲表面的實在充斥著各種意識型態、迷思、扭曲、以及錯誤的表象。「常識傾向於將各種社會現象當作習以爲常的現象，並且假定現存的一切都是必須存在的。建立在對常識不加批判而全盤接收的基礎之上的社會科學……只會不斷地複製這些錯誤」（Sayer, 1992: 43）。

6. *社會實在的解釋或理論是由什麼所構成的？*

在演繹法與歸納法之外，批判取向的社會科學使用逆推法來創造解釋性的評論。美國哲學家皮爾斯（Charles S. Peirce, 1839-1914）利用延伸其他推理形式的方式而發展出**逆推法**（abduction）。一般的作法是由許多的觀察結果或一個理論前提作爲開端，但是逆推法捨棄了此種作法，反而先去「試探」（"tries on"）可能潛在的一條規則，並詢問從這條規則可以推導出什麼。各種觀念和觀察結果都會被放置到一些可供選擇的架構內，然後，研究者會逐一加以檢視，提出「如果……會怎麼樣」（"what-if"）的問題。利用逆推法，研究者就能依序應用並評估各種不同架構的功效，在這個過程中能夠以具有創造力的方式，針對各種資料及觀念，重新賦予脈絡或是加以重新描述。

逆推法 一種建立理論的方法，在使用此種方法時，研究者會把數個可供選擇的架構應用於資料及理論，而在每個架構裡，這些資料與理論都會被重新描述並加以評估。

解釋性批評 一種使用於批判取向社會科學的解釋類型，在此種批評中，解釋同時說明了實際情況（或告訴我們為什麼事件會發生），並批評了這些情況（或指出不一致之處、揭露各種迷思、或指明矛盾之處）。

逆推法幾乎不會產生一個單一的、明確的真理，相反地，隨著逆推法進展到更深的理解時，能夠排除一些可能的選擇。就某些方面而言，逆推法是所有人類感知的一個面向。逆推法類似於一位明察秋毫且具有創意的偵探，在解決一樁犯罪案件時所用的方法。也就是先蒐集各種資料（線索），並把這些資料放進各種可能的狀況（可能是什麼事物導致了這起犯罪案件）當中。考量各種可能的狀況，會爲相同的觀察結果賦予新的意義。因此，逆推法的意義，就在於對觀念及資料不斷地進行重覆評估，其方式則是應用各種可能的規則或計畫，並從每一種狀況中瞭解其可能的結果。

解釋性批評（explanatory critique）是由下列這

個前提開始的：當我們研究社會生活時，我們既要研究該事物的「本身」（"itself"），也要研究人們如何去思考、或是理解我們正在研究的「事物」。實際的情況及人們對於實際情況的看法都具有關聯性，而且這兩者不一定互相吻合。解釋性批評具有實際的、道德的、以及政治的意涵。因此種批評可能與流行的看法大異其趣。此種解釋同時解析了實際情況（或是告訴我們爲什麼事件會發生），並且批評這些情況（或指出不一致之處、揭露各種迷思、或指明矛盾之處）。

當我們以解釋性批評來描繪社會情況時，我們通常可以啓發新的想法並幫助解放人們。由於此種解釋可以揭露出超越表面層次的實在層面，因此人們有可能會意識到社會的底層結構。解釋性批評揭露了深層的因果機制，一旦這些機制被揭露出來，人們就能瞭解如何影響這些機制，以改變更大的社會結構。如此一來，解釋性批評就爲人們指出了一條採取行動與達成社會改變的途徑。

7. *如何確定某個解釋的真假？*

實證主義社會科學會演繹假設、複製觀察結果以考驗假設、然後將結果結合起來，以決定要支持或是拒斥因果法則。詮釋取向社會科學則會去詢問，對於被研究的人們而言，意義系統以及行爲法則是否合理或具有意義。批判取向社會科學驗證理論的方式，則是精確地描述那些由底層結構所產生的情況，並應用這項知識來改變社會關係。批判取向社會科學的理論教導人們有關他們本身的經驗，幫助人們瞭解他們自己的歷史角色，而人們也可以利用這些理論，改善社會情況。

批判取向社會科學的理論能夠爲實際行動提供豐富資訊，而批判取向社會科學的理論，也能依據使用之後的結果進行修正。批判取向社會科學的理論會不斷成長，與它試圖去解釋的世界產生互動。由於批判取向的社會科學試圖以透視隱藏結構的方式，解釋並改變這個世界，因此對於某個解釋的檢驗方式就不會是靜態的。檢驗理論是一個動態的、持續進行的過程，在這個過程中不斷地應用並修正理論。在透過行動以消除無知，擴大洞見的持續過程之中，知識也隨之不斷地豐富增長。

批判取向社會科學區別理論好壞的方式，是將某個理論付諸實踐，然後利用其應用之後的結果來重塑理論。**實踐**（praxis）這個概念意謂著，只有當某個解釋能夠幫助人們瞭解這個世界，並且以採取行動的方式改變世界時，這個解釋才會受到重視。如同塞爾（Sayer, 1992: 13）所主張的，「知識主要是透過行動獲得的，其中包括試圖改變我們的環境（透過勞動或工作）的活動，以及與其他人彼此互動的活動。」

批判的實踐試圖消除在研究者與研究對象之間的分立，以及科學和日常生活的區隔。舉例而言，一位批判取向社會科學的研究者發展出一個關於租屋歧視的解釋。這名研究者會利用這個解釋試圖改變現狀，再根據應用的結果來檢驗這個解釋。如果這個解釋認爲，深層的經濟關係導致歧視，而且房東拒絕把房子租給少數民族的原因，是因爲租給非少數民族的房客可以得到較多利潤的話，那麼，若是採取政治行動可以使得把房子租給少數民族也同樣有利可圖，則

實踐　一種在批判取向社會科學中用以評估解釋的方式，其作法是將理論性解釋置入真實生活的實踐當中，並利用後續的結果修正解釋。

此種政治行動應該就能改變房東的行爲。對照之下，如果這個解釋認爲，事實上是深層的種族仇視導致房東對少數民族產生歧視，那麼以利益爲考量的行動就無法奏效。於是，這名批判取向的研究者將會透過新的研究、結合新的政治行動，將種族仇視當作造成房東歧視行爲的真正原因，並且對此加以檢視。

8. *良好的證據或是事實資訊的特徵爲何？*

實證主義社會科學假定有不容置疑的中性事實，而且所有具備理性的人們都會同意這些事實的存在。實證主義的二元論核心信念認爲，社會事實就像外在的客體，也就是說，社會事實的存在是獨立於價值或理論之外的。詮釋取向社會科學則認爲，社會世界是由人們所創造出的意義所組成，而在其中的人們則不斷地創造與協商各種意義。詮釋取向社會科學拒斥實證主義的二元論，改而強調主觀性，而證據就是那些存在於相關人們主觀理解之中的各種事物。批判取向則試著去弭平客體與主體間的鴻溝。批判取向社會科學認爲，物質狀況的事實是獨立存在於主觀知覺之外的，然而，事實卻非理論中立。相反地，事實需要詮釋，而詮釋則來自於一個由價值、理論、以及意義所組成的架構中。

例如，有一項「事實」是，美國花費在健康照護上的費用，佔國民生產總額（gross national product, GNP）的百分比例高於其他任何一個先進工業化國家。然而，在預防嬰兒死亡率方面，美國的排名卻在第二十九位（每一千位新生嬰兒當中，約有七名嬰兒死亡）。批判取向社會科學在詮釋這項事實時，指出美國有許多公民沒有健康照護的保障，也沒有任何體系可以照顧到每一個人。此一事實包括健康照護是透過一個複雜的體系分配到某些人身上，而這個複雜體系的組成成分則包括，以營利爲主的保險公司、藥廠、醫院、與其他從當前的制度設計中獲得龐大利益的人士。在這個體系中，某些強大的團體因從中牟利而越加富裕，但是，社會中有許多較爲弱勢或較爲貧困的階層，卻只能得到品質低劣的健康照護，有些人甚至完全沒有得到健康照護。批判取向社會科學的研究者會一邊關注這些事實，一邊詢問到底誰從中獲得利益、而誰又從中受到傷害或是被剝削。

理論可以幫助批判取向的研究者發現新的事實，並將這些事實去蕪存菁。理論就像是某種地圖般，告訴研究者該到何處尋找新的事實，且一旦發現了新的事實之後，又該如何詮釋這些事實。批判取向認爲，理論在自然科學中也能發揮同樣的功用。例如，一位生物學家透過顯微鏡觀察紅血球細胞，這個「事實」也是植基於有關血液與細胞的理論，以及這名生物學家在顯微鏡現象方面所受過的訓練。若是沒有相關的理論與教育訓練，這名生物學家所看到的就只是一些無意義的紅點。因此，顯而易見的是，事實與理論是彼此密切相關的。

批判取向的社會科學指出，只有部分理論能夠幫助我們發現並瞭解關鍵事實。理論所依賴的基礎，是關於世界樣貌的信念與假定、及一套道德與政治的價值觀。批判取向的社會科學主張，有些價值是優於其他價值的。[21] 因此，爲了要詮釋事實，我們必須瞭解歷史、採取一套特定的道德與政治價值、並明白該到何處尋找深層的結構。不同派別的批判取向社會科學，則會提供不同的價值立場（例

如，馬克思主義與女性主義）。

9. *社會科學知識的關聯性以及作用爲何？*

批判取向社會科學的研究者瞭解世界運作的方式之後，他們就會把主觀理解和數種分析客觀現狀的方法結合起來，以揭露隱藏的力量和未被辨識出來的不公正情形。對於批判取向社會科學而言，知識並不是一種可以讓人們任意操縱的工具，也不是要捕捉並描述人們的內在、主觀的經驗。相反的，知識意味著積極主動地參與世界的運作。知識可以把人們從過往思考方式的桎梏中解放出來，並幫助人們掌控周遭的事件。簡而言之，知識並不是一個讓人們擁有的事物，而是一個結合了逐漸增加的意識與採取行動的過程。

反身—辯證導向　批判取向社會科學所使用的一種關於社會知識的導向，在此種導向中，主觀層面及客觀層面會被結合在一起，以提供單一層面所無法獲得的深刻洞見。而知識作為一種過程，其價值就在於能夠把進行觀察、反思觀察結果、以及採取行動三者結合在一起。

轉型的觀點　此種觀點認為，研究者應該深入並超越實在的表面層次。進行探索，以轉變主觀的理解，並提供洞見幫助人們瞭解，從事社會與政治行動或許能夠大幅度地提升、改善人們的生活。

批判取向社會科學的研究者混合了工具性導向及實用性導向的不同層面，並且也調和了實證主義者強調的外在、實徵的實在，以及詮釋取向社會科學所重視的內在、主觀的實在。批判取向的社會科學利用反身性知識（請參閱第二章），以提供「第三條路」，也就是**反身—辯證導向**（reflexive-dialectic orientation）。這個新的導向「不是兩種不同觀點的結合或折衷，而是代表一種完全屬於它自己的立場與觀點」（Denermark et al., 2002: 202）。反身—辯證導向不把外在和內在的實在視爲彼此對立的兩方，而是把這兩者當成一個單一動態整體的兩方，這個動態整體則處於一種不斷成形的過程之中。單獨只看外在或是內在導向的作法，根本是不完整的。這兩方其實就像整體一樣地共同運作，並交織在一起、彼此互相影響。

對於知識的應用，批判取向社會科學所採取的是一種**轉型的觀點**（transformative perspective）。所謂的*轉型*（*transform*），是指*徹底地改變*、*重新組織基礎的結構*，並*突破當前的限制*。此種觀點超越了表面層次的實在，結合了主觀的理解及外在的實在，然後利用重新獲得的意識作爲投入行動的基礎，所採取的行動則具有修正外在現狀，與未來意識的潛力。知識的相關性，在於知識能夠讓人們與意識連結起來，使人們投入具體的行動、反思這些行動的後果與影響、然後在一個持續進展的循環當中，把意識提升到更新的層次。

10. *社會政治價值是由何處滲透到科學中的？*

批判取向的社會科學具有行動主義者的傾向。社會研究是一種道德政治活動，此種活動要求研究者必須表明其價值立場。批判取向的社會科學反對實證主義社會科學中所強調的價值中立，並認爲那只是一種迷思而已。此外，批判取向的社會科學也攻擊詮釋取向社會科學的相對主義。在詮釋取向的社會科學裡，天才眼中的實在與白癡眼中的實在是同樣有效而重要的。我們幾乎沒有任何（如果真的有的話）

基礎或是標準，可以用來評判對立的實在或是相互衝突的觀點。例如，詮釋取向的研究者不會指責種族主義者的觀點是錯誤的。因爲不論任何觀點，對於相信此種觀點的人而言，都是真實的。批判取向的社會科學主張，正確的觀點只有一種、或其數量非常稀少。至於其他的觀點，若不是完全大錯特錯，就是具有誤導性質。所有的社會研究都*必然*要以一個價值，或是道德立場爲其出發點。對於批判取向的社會科學而言，追求客觀並不等同於價值中立。客觀性要求以一種未經過扭曲的、真實的方式來描述社會實在。「批判取向的社會科學，挑戰了科學必須被嚴加保護以免受到政治干擾的信念。批判取向的社會科學主張，某些政治——追求解放性社會變革的政治——可以提升科學的客觀性」（Harding, 1986: 162）。

批判取向社會科學主張，否認研究者帶有某種觀點的想法，本身就是一種觀點。這是一種技術人員的觀點，也就是進行研究、不理會道德問題、滿足贊助者的要求、並且遵從命令。此種觀點認爲科學只不過是 一種任何人都可以使用的工具或器材。當德國納粹的科學家進行殘忍而不人道的實驗，宣稱他們「只是奉命行事」而且「只是科學家」、因此沒有任何過失時，上述的技術人員觀點便招致了嚴厲的批判。實證主義社會科學採取了此種研究取向，並生產出技術官僚的知識，而此種形式的知識，最適合讓握有權力的人用來支配或是控制其他人。[22]

批判取向的社會科學拒斥實證主義社會科學與詮釋取向社會科學，因爲這兩者要不是對於世界過度冷漠、就是僅止於關心這個世界的研究，而不對這個世界採取任何行動。批判取向的社會科學主張，知識就是力量。社會科學的知識可以被用來控制人們，此種知識也可以被隱藏在象牙塔之內，供知識份子把玩，抑或可說，此種知識亦能提供人們，以幫助他們掌控並改善自己的生活。一名研究者所研究的事物、進行研究的方式、以及應用研究結果的方式，都會牽涉到價值觀與道德，這是因爲知識對於人們的生活具有實質影響效果。那些研究瑣碎行爲的研究者、無法穿透表層進行深入探究的研究者、或是把研究結果深埋在大學圖書館裡的研究者，皆是作出與道德有關的選擇。而他們所作出的選擇，即是從研究對象身上取得資訊，但卻沒有讓研究對象參與其中、或解放他們（請參閱資訊箱 4.1 擴展個案方法以及批判取向的社會科學）。批判取向的社會科學質疑此種選擇的道德性，即使此種選擇並非有意識所作出的選擇。

摘要

雖然很少有全職的學術研究者會採用批判取向的社會科學，但是社區行動團體、政治組織、及社會運動經常依循批判取向社會科學的方法。除此之外，批判取向的社會科學甚少出現在學術期刊中。雖然批判取向的研究者可以使用任何一種研究技術，但是他們通常傾向於使用歷史比較研究法。他們之所以偏好歷史比較研究法的原因，是因爲此種研究法強調改變，而且能幫助研究者揭露深層的結構。批判取向的研究者與其他取向的研究者之間的差異，其實不在於他們所使用的研究技術，而在於他們處理研究問題的方式、他們所問的問題類型、與他們進行研究的目的。圖表 4.4 提供了批判取向社會科學的摘要。

資訊箱 4.1　擴展個案方法及批判取向的社會科學

布若威（Michael Burawoy, 1998）的擴展個案方法（extended case method），就是一個批判取向社會科學的例子。布若威說此種方法把反身性科學應用於民族誌或是田野研究之上。反身性科學是批判取向社會科學的一種類型，此種科學主張社會研究應該是研究者及被研究的人們之間所進行的一場對話。因此，互為主體性不應該只是像實證主義中一樣，只存在於科學家之間。更確切地說，互為主體性應該要發生在研究者與研究對象之間。布若威指出了反身性科學的四項特徵：

1. 研究者必須和研究對象或研究參與者進行互動。在彼此互動的過程中所產生的分裂與混亂，都有助於人們揭露、並更加瞭解社會生活。
2. 研究者採取研究對象或是研究參與者在特定情境下的世界觀，但並不僅止於此。研究者還會同時加入許多來自於個別研究對象，以及特定情境的觀點，並將這些觀點融入於更為廣闊的社會過程當中。
3. 研究者看待社會世界的方式，是同時採取由內而外的方式（亦即從被研究者的主觀觀點）、與由外而內的方式（亦即從作用於人們身上的外在力量觀點）。
4. 研究者持續地建立並重建理論。此種情況發生於研究者與研究對象之間的對話、及其與科學社群內其他研究者間的對話。

布若威利用擴展個案方法研究尚比亞的礦工。他主張最適合實證主義社會科學的情境，是那些人們在其中「無力抗拒更大的經濟與政治系統」的情境（p. 30）。換句話說，就是那些人們在其中受到宰制、而且對於自己的生活幾乎沒有掌控權的情境。批判取向社會科學所適用的情境脈絡，則是人們在其中試圖去抗拒或減少權力差異，以及宰制的情境脈絡。批判取向的社會科學特別強調解放的情境，在此種情境中，人們會產生質疑，或挑戰生活中那些造成權力差異與控制的外在力量。

圖表 4.4　批判取向社會主義摘要

1. 社會研究的目的是要揭露隱藏的結構，以解放人們並且賦予人們權能。
2. 社會實在具有多重的層次。
3. 人們擁有尚未發揮的潛力，且會被物化所誤導。社會生活是具有關係性的。
4. 對於人性能動作用採取*有限的自主*的立場。
5. 科學知識是不完美的，但是能對抗*錯誤意識*。
6. 使用*逆推法*來創造*解釋性批評*。
7. 透過*實踐*來驗證解釋。
8. 所有的證據都是以理論為基礎，而且有些理論能夠揭露出較深層類型的證據。
9. 對於知識採取*反身—辯證導向*，並使用*轉型的觀點*。
10. 社會實在以及對社會實在的研究，必然包含著道德與政治的面向，而在促進人類自由及彰權賦能方面，各種道德與政治立場是不平等的。

女性主義以及後現代研究

另外兩個其他的、比較不是那麼知名的研究取向，分別是女性主義及後現代社會研究。這兩種取向都對實證主義社會科學提出批判，也提出了以詮釋取向社會科學和批判取向社會科學爲基礎的替代取向。此外，這兩種取向都是在 1980 年代之後，才逐漸爲世人所知。

女性主義研究

進行*女性主義研究*（*feminist research*）的研究者大多數是女性，這些研究者抱持著女性主義的自我認同，有意識地採取女性主義的觀點。這些研究者使用多種不同的研究技術，試圖爲女性發聲，並努力地去改正那些佔據優勢地位的、以男性爲主的觀點。像是《對抗生命衝擊的女人》（*Women's Ways of Knowing*）（Belenky et al., 1986）（譯註：本書在國內由蔡美玲翻譯，並由遠流出版公司出版）之類的作品，則主張女性不論是在學習與表達自我方面，都跟男性有明顯的差異。

女性主義研究假定女性的主觀經驗與男性有所不同。[23] 許多女性主義研究者認爲實證主義社會科學只是呈現出男性的觀點，而這種觀點的特徵是客觀的、合乎邏輯的、任務導向的、並且是工具性的。此種觀點反映出男性所重視的是個人競爭、對於環境的支配與控制、以及「堅硬的事實」。此外，此種觀點也反映出一種家長制的傾向，該傾向強調的是發現那些作用於世界的力量，而不是去尋找與世界互動及合作的方法。

對比之下，女性所重視的是調適及逐漸地發展出人與人間的聯結。她們把世界視作一張由互相聯結的人際關係所交織而成的網，充滿由信賴的感覺以及彼此的義務所連結在一起的人們。女性重視社會生活中主觀的、同理心的、過程導向的、以及包容的面向。女性主義研究同時也是行動導向的，並試圖促進發展女性主義的價值（請參閱資訊箱 4.2 女性主義社會研究的特徵）。

資訊箱 4.2　女性主義社會研究的特徵

- 擁護、提倡女性主義價值立場和觀點。
- 拒絕在假定、概念、以及研究問題中的性別主義（性別偏見）。
- 在研究者以及被研究的人們之間建立同理心的聯結。
- 敏銳地察覺到性別與權力關係如何滲透於社會生活的各個層面。
- 在研究過程中融入了研究者的個人感受與經驗。
- 在選擇研究技術及跨越不同學術領域的界限方面具有彈性。
- 承認在人類經驗中存在著情感以及互相依賴的面向。
- 是一種行動取向的研究，企圖促進個人和社會的改變。

女性主義研究者主張，大多數的非女性主義研究都是歧視女性的。發生此種情況的原因，是因爲整體大環境的文化信念及男性研究者所擁有的優勢所導致。因此，大多數的研究都是把男性的經驗推論到所有人身上、忽視性別是一種基本的社會差異、只關注男性的問題、以男性作爲參照點、並完全接受傳統的性別角色。舉例而言，一位傳統的研究者會認爲，當一個家庭裡的成年男性找不到一份穩定的工作時，這個家庭就會面臨失業的問題。然而，當同一個家庭裡的女人在外面找不到一份穩定的工作時，此種情況卻不會被認爲是家庭問題。同樣地，傳統的研究者經常使用「未婚媽媽」這個概念，但是卻沒有人使用「未婚爸爸」這個類似的概念。

女性主義取向認爲，研究者基本上也是具有性別的。也就是說，研究者必然有其性別，而這個性別會形塑研究者體驗實在的方式，也會影響他們所進行的研究。除了性別對於個別研究者所造成的影響之外，基本的理論假定以及科學社群也都顯現出性別化的文化脈絡。性別所造成的影響在文化中隨處可見，也形塑了基本的信念和價值。而在進行科學探索的社會過程中，這些信念與價值都是不能被輕易地加以孤立與隔絕的。[24]

女性主義研究者不追求客觀或是中立，而會去與他們所研究的人們進行互動並且共同合作。他們也會將自己的個人生活及專業生活融合在一起。例如，女性主義研究者可能會試圖去理解受訪者的經驗，同時也與受訪者分享他們自己的感覺與經驗。此過程可能會在研究者及受訪者之間形成私人情誼，而這份情誼可能會隨著時間經過而越加緊密。雷恩哈茲（Reinharz, 1992: 263）主張：「模糊了正式關係與私人關係之間的分際，正如同移除了研究計畫與研究者生活之間……的區隔，此種情況就算不是全部、也是大多數女性主義研究的特徵。」

即使是在生物科學的領域內，女性的觀點、女性想與研究對象建立親密關係的渴望，也都產生了衝擊與影響。女性主義研究者傾向於避免使用量化分析與實驗。這些女性主義的研究者會使用多樣化的研究方法，通常是質化研究與個案研究。戈莉克（Gorelick, 1991）就曾經批判過許多女性主義研究者過於偏好詮釋取向社會科學的現象，由於她認爲詮釋取向社會科學只侷限於研究對象的主觀意識，因而無法揭露出隱藏的結構。戈莉克希望女性主義研究者能採取更具批判性的取向，並更加堅決地鼓吹支持社會變革。

女性主義研究者拒絕實證主義者所宣稱的價值中立。例如，黎絲蔓（Risman, 2001）就曾經批判過一項研究，因爲該項研究試圖以幾乎完全是生物因素的方式來解釋性別之間的差異。她認爲（p. 606）「實證主義的模式不僅僅在這個例子裡無法承認，亦排除特定的政治價值。除此之外，此種價值中立的科學不但是一個不可能達成的目標，也是一個不合宜的目標，因爲此種科學扭曲了研究及出版。」此外，她也提到（p. 609），「價值的中立性就像是一件斗篷，隱藏了（甚至連科學家自己也被蒙蔽）那些深植於我們文化與民俗智慧中的各種價值，以致於這些價值似乎皆消失於無形中。那些相信自己正在一個非關政治的、價值中立的科學領域中工作的研究者們，通常在無任何意識作出決定的情形下，完全忽視了主流假定塑造了他們研究問題的此一個事實。」

後現代研究

後現代研究（*postmodern research*）是範圍較大的後現代運動的一部分，而後現代運動則包含了藝術、音樂、文學、文化評論等等。後現代運動起始於人文學科，以存在主義、虛無主義、無政府主義等多種哲學，以及海德格（Martine Heidegger, 1889-1976）、傅柯（Michel Foucault, 1926-1984）、尼采（Friedrich Nietzsche, 1844-1900）、沙特（Jean-Paul Sartre, 1905-1980）、維根斯坦（Ludwig Wittgenstein, 1889-1951）等人的思想爲基礎。後現代主義是對現代主義的反動。所謂*現代主義*（*modernism*），指的是在啓蒙時代出現的基本假定、信念、與價值。現代主義所依賴的是邏輯的推理過程、對於未來抱持樂觀想法並相信進步、對於科技與科學擁有極大信心、信守人道主義的價值（亦即，在判斷想法或觀念的好壞時，是以這些想法對於人類福祉的影響爲依據）。現代主義認爲，大部分的人對於真、善、美，都有彼此能夠認同的一套標準。[25]

後現代研究者不認爲在藝術、人文學科、抑或是社會科學之間應該要有任何區別。他們與批判取向社會科學擁有共同的目標，便是揭穿社會世界的神秘面紗，並意欲解構或拆除表面的假象，揭露出隱藏的結構。彷彿是詮釋取向社會科學的一種極端形式，後現代主義不信任抽象的解釋，認爲研究所能做到的，僅僅只有描述而已，而且所有的描述皆同樣有效。某位研究者所作出的描述，既不優於、也不劣於任何其他人的描述，其所陳述的也僅僅是這位研究者的個人經驗。然而，與詮釋取向和批判取向社會科學不同的是，後現代主義企圖徹底顛覆社會科學。極端的後現代主義者不相信我們可以建立一門關於社會世界的科學、不信任所有系統性的實徵觀察結果、亦懷疑知識是否真可加以推論，或是隨著時間發展逐漸累積。他們認爲知識具有多種形式，屬於特定的人們或是特定地點所獨有。羅斯勞（Rosenau, 1992: 77）曾經說過：

> 幾乎所有的後現代主義者都拒絕把真理當作一個目標或理想，因為真理是現代主義的象徵……真理涉及秩序、規則、以及價值，並且依賴邏輯、理性、理智，而上述全都是後現代主義者所質疑的事物。

後現代主義者反對以一種超然的、以及中立的方式呈現研究結果。當有人在閱讀一份研究報告時，研究者或是該份報告的作者，不應該把自己隱藏起來，而是應該清楚明白地讓自己出現在報告當中。因此，一份後現代研究的報告，就類似於一件藝術作品，其目的就在於刺激他人、帶來樂趣、引發回應、挑起好奇心。後現代研究報告的呈現風格通常是劇場式的、富有表情的、以及戲劇化的，而其呈現的形式則可能是一本小說、一部電影、或一齣戲劇。後現代主義者認爲，在溝通傳達由研究者所創造出來、和社會生活的知識有關時，較佳的方式是透過短篇故事、小型喜劇、或是音樂，而不是透過學術期刊的文章。這些喜劇、故事、或音樂的價值，就在於所訴說的故事能夠激起閱讀者或是接觸者的內在經驗。後現代主義是反菁英

主義的，且也拒絕利用科學來進行預測與作出政策方面的決定。除此之外，後現代主義反對那些利用實證主義科學來強化權力關係，以及強化科層體制控制力量的人（請參閱資訊箱 4.3 後現代社會研究的特徵）。

結論

本章提出了兩個重要的概念。首先，依據關於科學目標與社會實在本質的不同哲學假定，本章提出了社會科學中數種相互競爭的研究取向。其次，不同研究取向的理想型，則以不同的方式，回答了關於研究的一些基本問題（請參閱表 4.1）。大多數的研究者主要採取其中的一種取向來進行研究，但是許多人也會結合一些其他取向中的元素。

要記得的是，你可以用這些取向中的任何一種，來研究相同的主題，但是每一種取向在進行研究的方式上，都各自不同。此種差異可以用下列四個國家中少數民族受到的歧視，與工作競爭的主題加以說明：在澳洲內地的原住民、在加拿大西部的中國人、在美國中西部的黑人、以及在倫敦的巴基斯坦人。

實證主義社會科學的研究者首先會從一個關於少數民族與多數民族關係的一般理論中，演繹出一些假設。而這個理論的形式，很可能是因果關係的陳述或預測。接下來，研究者可能會從現有的政府統計數據中去蒐集資料，或進行一項調查，以精確地測量該理論所指出的因素，像是初次接觸的形式、多數民族與少數民族數量的比例、或是族群差異的可見度等等。最後，實證主義社會科學的研究者會使用統計技術，以正式地檢驗該理論對於歧視程度和工作競爭激烈程度所作出的預測。

詮釋取向社會科學的研究者則會與來自於這四個國家的少數民族，以及多數民族的特定人士親自面對面談話，並且對他們進行觀察。這位研究者會利用對話與觀察的方式，瞭解每一個族群所認定最主要的問題。以及族群成員是否認爲歧視或工作競爭是他們日常生活中最關心的事。研究者會把人們表達的內容，置入於他們日

資訊箱 4.3　後現代社會研究的特徵

- 拒絕所有的意識型態及有組織的信念系統，包括所有正式的社會理論。
- 極度仰賴直覺、想像、個人經驗、以及情感。
- 無意義感以及悲觀主義，認為世界永遠不會進步。
- 極端的主觀性，認為心智與外在世界之間沒有任何區別。
- 狂熱的相對主義，認為詮釋的數量是無窮無盡的，沒有任何一個詮釋優於其他詮釋。
- 擁護多樣性、混亂、以及不斷變動的複雜性。
- 拒絕研究過去或是不同的地區，因為只有此時此地才具有關聯性。
- 相信因果關係是不能被研究的，因為生活太過複雜，而且變化太過快速。
- 堅持研究永遠無法真正地呈現社會世界裡所發生的事。

表 4.1 社會研究三大取向之間的差異摘要

	實證主義	詮釋取向社會科學	批判取向社會科學	女性主義	後現代
1. 進行研究的理由	發現自然法則，以讓人們可以預測並且控制事件	瞭解並且描述有意義的社會行動	打破迷思並且賦予人們權能以去改變社會	賦予人們權能，使人們能夠促進照顧他人及平等的價值觀	表達主觀的自我、像是開玩笑一般的、達到娛樂以及激勵他人的效果
2. 社會實在的本質	社會實在是穩定的、事先存在的模式或是秩序，此種模式或秩序可以被人們所發現	社會實在是關於某一個情境的流動性定義，此種定義是由人際互動所創造出來的	社會實在具有多重層次，而且是由隱藏的、底層的結構所決定的	社會實在是依據性別所建構的權力關係，此種權力關係使得人們遭受壓迫	社會實在是混亂而且變動不居的，沒有真正的模式或者是偉大的計畫
3. 人性的本質	自私而且理性的個體，而且會受到外在力量所影響	會創造出意義的社會性生物，而且會不斷去理解自己所處的世界	具有創造力以及適應力的人們，擁有尚未實現的潛力，而且被假象所困	具有性別的生物，擁有尚未實現的潛力，但是此種潛力經常被看不見的力量所困住	有創造力的、動態的生物，具有尚未實現的潛力
4. 人性能動作用	強大的外在社會壓力形塑了人們的行動，自由意志幾乎只是一種假象	人們具有強大的意志力，人們可以發展出意義，並且擁有選擇的自由	有限的自主及自由選擇，但是會受到結構性的限制。不過，此種限制是可以移除的	以性別為基礎的結構性限制會侷限了選擇，但是新的想法以及行動則可以突破這些限制	人們具有強大的意志力，而所有的結構都只是虛幻的
5. 常識所扮演的角色	常識與科學之間有著清楚的區別，而且常識在效力方面遜於科學	常識就是由一般人所使用的、強大的日常生活理論	常識是一種錯誤的信念，隱藏了權力以及客觀現狀	常識是一種錯誤的信念，隱藏了權力以及客觀現狀	常識就是社會實在的真正本質，且優於科學的或是官僚形式的推理方式
6. 理論的特徵	理論就是一個邏輯的、演繹的系統，該系統由互相聯結的定義、規律、以及法則所構成	理論就是對於一個群體的意義系統如何產生與維持的描述	理論就是一種批評，揭露出真實的情況並幫助人們採取行動	理論就是一種批評，揭露出真實的情況，並且幫助人們看到建立一個更好的世界的方法	理論就是一場表演或是一件藝術表現的作品，能夠娛樂、震憾、或是激勵其他人

7. 判定為真的解釋	在邏輯上與法則聯結在一起，並以事實為基礎	能夠引起被研究者的共鳴、或是被研究者覺得正確的事	提供人們改變世界所需要的工具	提供想法或是工具，以幫助人們從壓迫的關係中解放	沒有任何解釋比其他解釋更為真實，對於接受某項解釋的人而言，該項解釋就是真的
8. 良好的證據	以他人能夠複製的精確觀察結果為基礎	蘊藏於流動的社會互動脈絡中	由一個能夠識破表層假象的理論所提供	由能夠揭露性別結構的理論所提供	具有美學的特質，並且能與人們的內在感受產生共鳴
9. 知識的相關性	採取工具性導向，知識使人們能夠掌握並且控制事件	採取實用性導向，知識幫助我們以同理心的方式接納或分享他人的生活世界以及經驗	採取辯證導向，知識讓人們能夠看到並且改變深層的結構	知識提高人們的意識，並且賦予人們權能去做出改變	正式的知識沒有任何特別的價值，知識可以娛樂人們，或是帶來個人的樂趣
10. 價值的地位	科學是價值中立的。除了在選擇主題時，價值是沒有任何地位的	價值是構成社會生活整體的一部分：沒有任何群體的價值是錯誤的，最多只是彼此不同而已	所有的科學在一開始時都必然有其特定的價值立場，有些立場是對的，而有些立場則是錯的	價值是研究中不可或缺的部分，而女性主義的價值是明顯受到偏好的價值	價值是構成研究整體的一部分，但是所有的價值立場都是平等的

常事務（例如，繳房租、捲入家庭紛爭之中、惹上官司訴訟、生病等等）的脈絡之中。在研究者瞭解了少數民族或多數民族對於歧視的看法、他們如何找到工作、其他族群的成員又是如何找到工作、以及他們實際上做了些什麼事來獲得，或是保住工作之後，研究者就會以其他人也能理解的方式，來描述這些發現。

至於批判取向社會科學的研究者，則會從檢視較大的社會及歷史脈絡開始進行研究。這些脈絡包括數種因素，像是英國殖民主義者對澳洲的入侵和澳洲被當成監獄殖民地的歷史、導致中國民眾遷移到加拿大的經濟狀況、美國奴隸制度的遺緒及在民權方面的奮鬥、大英殖民帝國的興衰起落以及前殖民地的人民移入英國的現象等等。研究者會從一個道德—批判的立足點來提出問題：多數民族是否真的歧視少數民族，並在經濟上剝削少數民族？研究者會仔細查詢各種資料來源，以論證剝削行爲的潛在模式，並衡量每一個國家的歧視案件數量。研究者可能也會檢視不同族群間之收入差異統計資料、親自查訪生活處境，並且隨同人們去應徵工作、進行調查以瞭解人們現在的想法。一旦研究者發現歧視如何阻礙了少數民族找到工作，就會把研究結果提供給少數民族的組織、公開宣揚研究發現、把研究結果發表在少數

民族的成員所閱讀的報紙上，以揭發真實的狀況，並鼓勵他們採取政治社會行動。

就社會研究法這門課程而言，上述這些關於三種取向的資訊，對你究竟有什麼意義呢？首先，這意味著進行社會科學研究沒有單一的、正確的取向。當然，這並不是說你怎麼做都可以，也不是說各種取向之間沒有取得共識的基礎（請參閱資訊箱 4.4 社會科學三種主要取向的共同特徵）。更確切地說，這意味著進行社會研究的基礎尚未穩固。換句話說，目前有希望脫穎而出的取向並不只有一個，而這種三強鼎立的狀況可能會一直持續下去。對這三種取向的認識，將會有助於你閱讀研究報告。一般而言，研究者會以某種取向爲基礎，但他們卻很少明白地告訴你，他們究竟是採用哪一種取向。

其次，這三種取向意味著，當你在進行研究時，你試圖要達到的目標（例如，發現法則、指認出深層的結構、描述意義系統等等），將會因爲你所選擇的取向而有所不同。這三種取向與在第二章中所討論過的研究類型之間，僅僅存在著鬆散的對應關係。例如，實證主義社會科學較有可能會進行成本效益分析，詮釋取向社會科學的研究者則傾向進行探索性研究，至於批判取向社會科學的研究者，則偏好行動取向的研究。當你在進行社會研究時，若能對這些取向有所認識，對於應該進行何

資訊箱 4.4 社會科學三種主要取向的共同特徵

1. *都是實徵的*：每一種都是植基於可觀察的社會實在，包括人們所看見的景像、聲音、行為、情境、討論、以及行動等等。研究絕對不會建立在虛構與想像的基礎之上。
2. *都是系統性的*：每一種取向都強調一絲不苟及仔細謹慎的工作，也都拒絕雜亂的、低劣的、或是草率的思考與觀察。
3. *都是理論性的*：雖然理論的本質有所不同，但是每一種取向都強調要使用觀念及找出特定的模式。沒有任何一種取向認為社會生活混亂無序。每一種取向都認為解釋或是瞭解是可能達成的。
4. *都是公開的*：每一種取向都認為研究者的工作成果必須坦白誠實地傳遞給其他的研究者，也就是研究成果應該要明白清楚地公開、而且應該要與他人分享。每一種取向都反對以藏匿的、私密的、或秘密的方式進行研究。
5. *都是自我反思的*：每一種取向都主張研究者必須思考他們所做的事，並應該要有高度的自覺，絕對不能以盲目的或是不加思索的方式來進行研究。研究牽涉到嚴肅的反思，並且要求自我覺察。
6. *都是無止盡的過程*：每一種取向都認為研究是持續前進的、演進的、改變的、提出新問題的、以及追尋各種線索的過程。沒有任何一種取向認為研究是靜態的、固定的、或是封閉的。目前的知識或研究程序並非「僵硬死板」和確鑿無誤的。不論是知識或研究程序，都會持續地變化，也會對於新的思考方式和行為方式保持開放態度。

由此可知，雖然三者之間存在一些差異，但是這三種取向都主張社會科學應該透過自我反思，以及開放性的持續過程，致力創造以系統化方式蒐集的、以實徵為基礎的理論性知識。

種類型的研究，你將能作出較爲明智的決定。

第三，在社會研究中所使用的各式各樣研究技術（例如抽樣、訪談、參與觀察等等），最終都是以來自這些取向的假定和觀念爲基礎。通常，當你看到作者提及某一項研究技術的時候，作者都不會特別說明這項技術原本的背景推理過程。在認識了這些不同的取向後，你可以更清楚地瞭解特定的研究技術背後的原理或原則。例如，精確的測量技術及實驗研究的邏輯，就是源自於實證主義，至於田野研究，則是以詮釋取向爲基礎。

到目前爲止，我們已經看到研究過程的整體運作方式、不同類型的研究與理論、以及社會研究的三種基本取向。現在你應該已大致瞭解社會研究的基本樣貌了。在下一章裡，你將會看到如何找出特定研究計畫的報告。

問題回顧

1. 根據三種不同研究取向的看法，進行社會研究的目的爲何？
2. 每一種取向如何定義社會實在？
3. 根據每一種不同的取向觀點，人性的本質爲何？
4. 在每一種取向當中，科學以及常識之間的差異爲何？
5. 依據每一種取向的看法，社會理論的定義爲何？
6. 每一種取向如何檢驗一個社會理論？
7. 每一種取向對於事實的看法爲何？又該如何去蒐集事實？
8. 每一種取向分別是以何種方式評價中立科學的可能性？試解釋之。
9. 詮釋取向與批判取向的社會科學，在批評實證主義的方法上有何類似之處？
10. 在第一章中所提到的科學模型與科學社群，與本章中的三種取向各自有何關聯？

注解

1. This book is primarily concerned with sociology (Steinmetz, 2005a). For anthropology, see Kean (2005); educational research, see Bredo and Feinberg (1982) and Guba and Lincoln (1994); for psychology, see Harré and Secord (1979) and Rosnow (1981); for political science, see Hauptmann (2005) and Sabia and Wallulis (1983); and for economics, see Hollis (1977), Mitchell (2005), and Ward (1972). A general discussion of alternatives can be found in Nowotny and Rose (1979).
2. See especially Friedrichs (1970), Giddens (1976), Gouldner (1970), and Phillips (1971). General introductions are provided by Harré (1972), Suppe (1977), and Toulmin (1953).
3. Divisions of the philosophies of social science similar to the approaches discussed in this chapter can be found in Benton (1977), Blaikie (1993), Bredo and Feinberg (1982), Fay (1975), Fletcher (1974), Guba and Lincoln (1994), Keat and Urry (1975), Lloyd (1986), Miller (1987), Mulkay (1979), Sabia and Wallulis (1983), Smart (1976), and Wilson (1970).
4. For discussions of paradigms, see Eckberg and Hill (1979), Kuhn (1970, 1979), Masterman (1970), Ritzer (1975), and Rosnow (1981).
5. In addition to the works listed in note 3, Halfpenny (1982), Steinmetz (2005), and Turner (1984) have provided overviews of positivism in sociology. Also see Giddens (1978). Lenzer (1975) is an excellent introduction to

Auguste Comte.

6. See Gartell and Gartell (1996, 2002).

7. From Bernard (1988:12–21).

8. See Hegtvedt (1992).

9. For a discussion, see Derksen and Gartell (1992:1715).

10. See Couch (1987). Also see Longino (1990:62–82) for an excellent analysis of objectivity in positivism.

11. For a discussion, see Bannister (1987), Blumer (1991a, 1991b, 1992), Deegan (1988), Geiger (1986), Gillespie (1991), Lagemann (1989), Ross (1991), Schwendinger and Schwendinger (1974), Silva and Slaughter (1980), and Smith (1996).

12. For a further discussion of hermeneutics see Bleicher (1980) and Schwandt (1994; 1997). Sewell (1996; 2005) also discusses the significance of "reading" text.

13. In addition to the works in note 3, interpretive science approaches are discussed in Berger and Luckman (1967), Bleicher (1980), Cicourel (1973), Garfinkel (1967, 1974b), Geertz (1979), Glaser and Strauss (1967), Holstein and Gubrium (1994), Leiter (1980), Mehan and Wood (1975), Silverman (1972), and Weber (1974, 1981).

14. See Roy (2001:7–13) on the essentialist versus constructionist orientation.

15. See Brown (1989:34) for more examples and explanation.

16. In addition to the works in note 3, critical science approaches are discussed in Burawoy (1990), Dickson (1984), Fay (1987), Glucksmann (1974), Harding (1986), Harvey (1990), Keat (1981), Lane (1970), Lemert (1981), Mayhew (1980, 1981), Sohn-Rethel (1978), Veltmeyer (1978), Wardell (1979), Warner (1971), and Wilson (1982).

17. For a discussion of the Frankfurt School, see Bottomore (1984), Held (1980), Martin (1973), and Slater (1977). For more on the works of Habermas, see Holub (1991), McCarthy (1978), Pusey (1987), and Roderick (1986).

18. See Swartz (1997) on Bourdieu.

19. For discussions of realism, see Bhaskar (1975), Miller (1987), and Sayer (1992).

20. For discussion of critical realism, see Archer et al. (1998), Bhaskar (2003), Danermark et al. (2002), and Groff (2004).

21. See Sprague and Zimmerman (1989) on feminists' privileged perspectives of women and see Rule (1978a, 1978b) on constituencies that researchers favor.

22. See Habermas (1971, 1973, 1979) for a critical science critique of positivism as being technocratic and used for domination. He has suggested an emancipatory alternative.

23. See Olsen (1994).

24. See Evelyn Fox Keller's (1983) biography of Barbara McClintock and her other essays on gender and science (1985, 1990). Also see Longino (1990), Chapters 6 and 7.

25. From Brannigan (1992).

第5章

如何回顧文獻以及進行合乎倫理的研究

文獻回顧
社會研究中的倫理
結論

> 但由於我們始終無法活在一個免於世俗困擾與跳脫歷史的時代，因此，我們的問題不是要如何去處理應該是「真理」的知識，而是應該處理人們如何求知的問題，而這個問題牽涉到人們對於自己在時間與社會中處於何種位置的認識與瞭解。
>
> ——Karl Mannheim, *Ideology and Utopia*, p. 188

在他針對芝加哥都市住宅計畫中的販毒幫派所進行的田野研究中，范卡德胥（Venkatesh, 2008: 185-186）瞭解到：「在一頭栽進這項研究長四年之後，我終於注意到，如果我繼續做我一直在做的事，就有可能會惹上許多麻煩……我去和一位律師見了面，而且也瞭解到許多重要的事情。第一，如果我得知某項會傷害到任何人的計畫，我就有義務去通知警方……在這其中沒有所謂的『研究者與當事人之間的保密協定』之類的事情，此種保密協定通常是律師、醫生、或者是牧師才擁有的特權。這意味著如果我真的被傳喚到法庭作證指控這些幫派時，我在法律上就有提供證詞的義務……這個法律方面的忠告幫了我很大的忙，因為這個忠告讓我認真而嚴肅地去評估我所進行的研究……」

你已經準備好要針對幫派這個主題來設計一項研究。當你把這個廣泛的主題聚焦爲一個具體而明確的研究問題（例如，在都市住宅計畫中的販毒幫派，是否會爲其他住戶提供服務或保護，抑或只是在剝削這些住戶？）時，你遇上了兩個議題。在實務上，把一個主題聚焦成爲一個研究問題的過程，會恰好與回顧文獻的過程有重疊之處。接著，在你蒐集有關幫派的資料時，你必須要怎麼做才能符合倫理規範？具體的倫理議題，取決於研究問題和資料蒐集技術。以人群爲研究對象所產生的議題，在調查研究、實驗研究、以及田野研究中，是最爲顯著的。而在現存文件研究、次級資料分析、內容分析、或是歷史比較研究中，則是最不顯著的。在一些可能會侵犯到個人隱私、牽涉到違法行爲的爭議性議題，或是領域裡，倫理議題會比在「安全的主題」（“safe topics”）中更爲重要而且引人注目。爲了研究違法的幫派，你不只需要保護自己，免於遭受任何肢體方面的攻擊，同時也須注意到法律方面的問題。理想上，與在上面方框中所提到的范卡德胥所進行的研究不同的是，你不必在進行

研究整整四年之後，才瞭解到你的研究牽涉到複雜的法律與倫理議題，還因此而必須改變研究的方向。

前面幾章已經討論了研究的基礎：科學社群的運作方式、各種不同的研究類型、理論與研究之間的連結、以及進行社會研究的廣泛取向等等。在這一章與接下來的三章裡，我們將討論一些當你開始進行研究時所可能遇到的實際問題。其中包括如何回顧相關文獻、考量倫理議題、設計一項研究、測量社會世界的各種面向、以及決定要蒐集何種資料等等。

文獻回顧

在進行研究時，一個早期而且不可或缺的步驟，就是回顧與你的研究問題有關的累積知識，而這個文獻回顧的步驟，對於所有的研究問題與所有類型的研究都是必要的步驟。如同生活中的其他領域般，在你開始親自處理某項議題之前，先觀察別人對這個議題已經累積的知識，會是相當明智的作法。一些老生常談的道理，可以用來證明這個建議：不要浪費時間「做白工」（"reinvent the wheel"），而在你開始去做某件事情之前，記得要「先作功課」（"do your homework"）。不論你是某項研究報告的讀者、或是要自己開始進行一項研究，上述的重點都很有道理。

首先，我們會由文獻回顧所能達成的多種目的開始，此外，我們也會討論何謂*文獻*（*literature*）、要到何處尋找文獻、以及文獻中所包含的內容爲何。接下來，我們將會探索各種以系統性方式進行文獻回顧的技術。最後，我們會討論如何撰寫文獻回顧，以及文獻回顧在研究報告中的地位。

進行文獻回顧是基於以下觀念：知識會不斷地累積，因此，我們可以從他人所完成的事物中學習，並以此爲基礎而繼續累積知識。文獻回顧所依賴的原則，是把科學研究視爲一種集體的努力。其中有許多研究者不但作出貢獻、也和彼此分享研究成果。雖然有一些研究可能會變得特別重要，而有一些個別的研究者有機會變成家喻戶曉的名人。但是，一項研究其實只是整體創造知識過程中，極微小的一個部分而已。今日的研究乃是以昨日的研究爲基礎。而我們閱讀研究的目的，則爲從中學習、比較、複製、或批判這些研究。

文獻回顧在範圍與深度方面有許多不同的變化。不同種類的文獻回顧，在達成四種目標中的某一項時，可能會發揮較爲強大的作用（請參閱資訊箱 5.1 文獻回顧的目標）。即使是經驗豐富的研究者，也必須花上好幾年的時間，才能針對一個問題廣泛的相關研究文獻，進行大規模專業摘要與回顧。另一方面，同樣一位研究者卻可能在僅僅一周內，就完成一份特定領域內的、焦點較爲狹窄的文獻回顧。在開始進行文獻回顧時，你必須先挑選一個主題領域或是研究問題、決定你可以花多少時間和力量在這項研究上、確定適當的深度層次、並判斷最適合這項研究的文獻回顧類型（請參閱資訊箱 5.2 六種文獻回顧的類型）。在一份具體的文獻回顧中，你可以一一結合每種類型的特色。

資訊箱 5.1　文獻回顧的目標

1. *展現出對於某一知識體系的熟悉度，並建立可信度*。文獻回顧可以讓讀者知道，研究者不但熟知這個領域內的相關研究，也知道重大議題為何。好的文獻回顧可以增加讀者對於研究者的專業素養、能力、以及背景的信心。
2. *顯示出先前研究的途徑，以及目前的計畫與這些研究之間的關聯性*。文獻回顧可以概略描繪出針對某項問題的研究方向，同時顯示出知識的發展。好的文獻回顧能夠把研究計畫置於一個脈絡之中，利用與知識體系之間的連結，展現研究計畫的相關性。
3. *整合並且摘要某個領域內已知的知識*。文獻回顧把不同的研究結果放在一起，將這些結果整合起來。好的文獻回顧能夠指出哪些部分是先前的研究同意或是不同意的，而且還有哪些重大的問題尚待解決。此外，好的文獻回顧也能蒐集到某個時間點為止所有已知的事物，並指出未來的研究方向。
4. *從他人的努力學習並刺激新的想法*。文獻回顧能夠讓我們得知他人已經發現的事物，可以讓研究者從他人的努力當中獲益。好的文獻回顧會指出盲點，建議我們進行複製的假設。此外，好的文獻回顧也能揭露那些值得仿效的程序、技術、與研究設計。如此一來，研究者就能更加專注於假設上，並獲得新的洞見。

資訊箱 5.2　六種文獻回顧的類型

1. *脈絡回顧*　這是一種相當常見的文獻回顧類型，在此種回顧中，作者會將一個特定的研究連結到一個較大的知識體系。脈絡回顧通常出現在研究報告的開頭部分，以作為該項研究的簡介。並把該項研究置入一個較大的架構內、顯示出該項研究如何延續，建立在一個發展中的思想派系或是研究之上。
2. *歷史回顧*　這是一種特殊的回顧類型，在此種回顧中，作者會追溯一個議題的發展過程。此種回顧可以被併入理論回顧或是方法論的回顧中，以顯現出一個概念、理論、或是研究方法如何隨著時間經過，演變成形的發展過程。
3. *整合式的回顧*　這也是一種相當常見的文獻回顧類型，在此種回顧中，作者會針對某一個主題提出並摘要整理目前的知識狀態，突顯在此主題領域裡，過往的研究者所同意以及不同意的部分。此種類型的文獻回顧通常會與脈絡回顧結合，也有可能被當作單獨的文章出版，以供其他研究者使用。
4. *方法論的回顧*　這是一種整合式回顧的特殊類型，在此種回顧中，作者針對各種不同研究所使用的研究方法，比較其相對優點，並顯示出不同的方法論（例如，研究設計、測量方式、樣本等等）如何導致不同的研究結果。
5. *自我研究的回顧*　在此種類型的回顧中，作者會展現出自己對於某個主題領域的熟悉程度。自我研究的回顧通常是某個教育計畫或課程規定的一部分。
6. *理論回顧*　這也是一種特殊的回顧類型，在此種回顧中，作者會提出數個針對相同主題的不同理論或概念，並根據這些理論或是概念的假定、邏輯連貫性、以及解釋的範圍等等，進行比較。

文獻後設分析

文獻**後設分析**（meta-analysis）是一種特殊的技術，主要是用以創造整合式的回顧，或方法論的回顧。[1] 後設分析牽涉到蒐集大量先前已進行過的研究細節，並將結果整合。在進行後設分析時，通常會依循下列五個步驟：

1. 針對一個特定主題或是研究問題，找出所有可能的或潛在的研究
2. 發展出一致的標準，並以相關性／品質來篩選這些研究
3. 確認並記錄每個研究中的相關資訊
4. 綜合且分析這些資訊，以成爲概略的發現
5. 根據這些發現，寫出摘要式的結論

對於量化研究的後設分析而言，在第三個步驟的相關資訊通常包括了樣本大小、變項的測量方式、方法論的品質、以及變項的效果大小。而在第四個步驟裡，研究者會用統計技術來分析上述的這些資訊（請參閱範例 5.1 量化研究的後設分析）。至於質化研究的後設分析，則有一些稍微不同之處。換句話說，在第三個步驟裡的相關資訊，包含了質化的描述，而這些描述會被編碼成一整組的類別，而在第四個步驟裡，研究者會以質化的方式來綜合這些結果，以揭露一再重覆的主題（請參閱範例 5.2 質化研究的後設分析）。

除了利用後設分析來辨識出橫跨許多研究的主要發現之外，我們也可以利用後設分析來確認一個研究個案的作者如何定義並使用主要概念。例如，佛克森以及湯普森（Fulkerson and Thompson, 2008）這兩位研究者曾經檢視在 1988 年到 2006 年這十八年的期間內，所有與「社會資本」（"social capital"）相關的概念。他們在 450 份的學術期刊中找到 1,218 篇相關的文章，這些文章在其標題或摘要中都包含*社會資本*這個詞彙。接著，他們以七種方法將這些文章進行編碼，這麼做的目的在於爲這個概念下定義，也能確認文章中所引述的關於社會資本這個概念的「創始學者」。除此之外，他們也使用了統計技術來分析一些模式，這些模式顯現出，在不同時期以及不同專業領域內，對於這個概念定義的使用情況。

何處尋找研究文獻

研究者會發現到，研究報告可以用數種不同的格式來呈現：書籍、學術期刊的文章、論文、政府文件、以及政策報告。除此之外，研究者也會在專業協會的會議上以報告的形式來提出他們的研究發現。在這個部分，我們將會討論每一種不同的格式，並提供一個簡單的指引，好讓你知道應該如何取得這些文獻。

後設分析 一種文獻回顧的特殊類型，在此種類型中，研究者會組織整合來自許多研究的結果，並利用統計技術來確認這些研究中的共同發現。

期刊 研究結果有可能會出現在報紙、流行雜誌、電視或廣播節目、以及網路新聞摘要中。但是上述

範例　5.1

量化研究的後設分析

程以及陳（Cheng and Chan, 2008）兩位研究者曾經針對 133 個關於工作不安全感的研究進行了後設分析，而他們想要知道的，是工作的不安全感對於健康情況會造成什麼樣的影響。兩位研究者考量了下列三個因素：工作年資（也就是某個人做同一份工作的時間有多長）、年齡、以及性別。他們的目的在於瞭解工作年資、年齡、性別這三個因素，可能會以何種方式減弱或增強工作不安全感對健康所造成的影響。首先，他們在數個研究資料庫中，以工作安全感與工作不安全感這兩個關鍵詞彙，搜尋從 1980 年到 2006 年之間所出版的所有研究，以找出可能的相關研究。此外，他們也用人工的方式檢閱了十五種學術期刊，尋找未出版的論文，並詢問一些重要學者，是否有任何他們曾經進行過、但沒有出版的研究。接著，這兩位研究者利用一套選擇的標準篩選出可能的研究。若要被歸類為相關的研究，那麼研究報告必須以英文撰寫、在使用工作不安全感這個詞彙時要符合這兩位作者所下的定義、而且也要包含所有相關的變項。在他們找出 133 個可以接受的研究之後，兩位研究者對每一個研究結果進行了編碼。而進行編碼的資訊則包括樣本大小、關鍵變項的測量方式、變項間的相關性、以及統計效果的大小。接下來，兩位研究者以統計方式，分析了編碼後的資訊。最後，根據他們對於結果所進行的統計分析，這兩位研究者所得到的結論是，相較於年紀較輕、經驗較少的員工，年紀較大、與工作年資較長的員工，比較容易因為工作不安全感，遭受更多身體與心理健康層面的負面影響。

這些媒體中出現的研究結果，都不是你在撰寫文獻回顧時可以派上用場的完整研究報告。因爲這些研究結果的內容，都是記者針對一般閱聽大眾所準備的，所以皆是經過挑選並濃縮之後的摘要。因此這些研究結果的摘要欠缺許多我們在審慎評估一項研究時，所必須得知的重要細節。教科書與百科全書也只是列出濃縮過後的摘要，以提供簡要的介紹給那些新接觸到某項主題的讀者。上述這些研究報告的摘要都不適用於文獻回顧之中，因爲在這些摘要中，許多關於研究的重要細節都付諸闕如。

乍看之下，在已出版的學術文章之海中進行探索與搜尋，可能是件令人望而生怯的工作。當被要求進行「文獻回顧」時，許多剛接觸社會研究的學生會在網路上用搜尋引擎去尋找相關的文章，或是去翻閱一些他們熟悉的、但卻是非專業的，以及非學術的雜誌或報紙文章。社會科學的學生必須學會如何分辨詳盡報導各種研究的學術出版品、及爲一般大眾所撰寫的流行、外行的娛樂文章，或是新聞報導（請參閱表 5.1）。換句話說，他們必須捨棄外行的公共資料來源，轉而仰賴爲專業人士所撰寫的嚴肅學術出版品。

專業的研究者會以下列數種形式中的一種來呈現研究結果：學術研究書籍〔通常被稱作*專題著作*（*monographs*）〕、學術期刊上的文章、經過編輯的學術書籍中的部分章節、以及在專業會議上所提出的研究論文或報告等等。在爲初次學習某個科目的學生所撰寫的教科書裡、爲一般大眾所撰寫的新聞式摘要刊物裡出現的文章，都只是簡化過的、刪節過的、與「事先消化過」的研究報告版本。不幸的是，在這些經過簡化的摘要裡，原本相當完整的一項研究，可能會被以一種不完整的、甚至是扭曲的方式呈現出來。因此，研究者必須找出原本的學術期刊文章，才能真正看

到作者的完整描述、及其原始資料所呈現的圖像。

在撰寫正式的研究報告時，高年級的大學生與研究生應該要仰賴學術文獻，也就是刊登在學術期刊中的原創性文章。不幸的是，學生們可能覺得有些學術文章過於艱澀或是過於專門化，不容易讀懂。不過，往好處想的話，這些文章才是「貨真價實的本尊」(“real McCoy”)(譯註：這個詞在英文中是指「正牌的東西或人物」)，也就是原本的文章，而不是另一個人對於原本文章的解讀或誤讀。

研究者或許也會發現一種非研究的期刊，其中有對某些主題或研究問題所作出的評論。這些期刊其實只是討論輿論的雜誌（例如，《美國展望》、《Cato 期刊》、《評論》、《國家》、《國家評論》、《新共和國》、《紐約書籍評論》、《政策評論》以及《公共利益》等等）。而在這些期刊裡，許多專業人士會爲受過教育的與其他的專業人士撰寫簡短的文章，以表達他們的意見、信念、價值傾向，以及臆測。不過在這些期刊中，並不包含原本的實徵研究或是實際的科學研究。這些期刊有可能被歸類爲「學術期刊」（相對於一般期刊而言），而且可能也有「經過同儕審查」的程序，但是這些期刊卻未包含實徵研究的原始完整報告。例如，《政策評論》這份期刊所報導的主

範例　5.2

質化研究的後設分析

馬爾斯頓以及金恩（Marston and King, 2006）曾經針對 1990 年到 2004 年之後所出版的 268 項質化研究，進行了後設分析。而這些研究的主題則都與青少年的性行為有關。這兩位研究者意圖理解的，是青少年之間的性行為可能以何種方式影響愛滋病毒的感染和散播。因為幾乎一半以上的愛滋病毒感染案例，都是發生在這個年齡族群之內。此外，這兩位研究者想要檢視質化研究的原因，是因為他們所感興趣的是當青少年從事性行為時所發生的事、從事性行為的理由、以及發生性行為的情境脈絡。對比之下，大部分的量化研究只有檢視了一些簡單的、孤立的問題，像是青少年使用保險套的比例之類的問題。他們找到了在 1990 年到 2004 年之間以英文出版的所有研究，這些研究提供了關於 10 歲到 25 歲間的青少年性關係的質化實徵證據。作者們所挑選出來的研究，是那些焦點在其他議題上面（例如藥物濫用）、但同時也包含性行為相關議題的研究。作者們搜尋了數量龐大的文章與書籍的資料庫，也調查了英國境內 150 間學術單位圖書館的館藏目錄。他們先以標題進行搜尋，並且發現了 5,452 篇可能符合條件的研究報告，然後再依據標題的相關程度，把研究報告的篇數縮減到 2,202 篇。最後，他們再根據一套篩選標準，把數量進一步地縮減到 268 項研究（其中包括了 246 篇期刊文章以及 22 本書）。而他們的篩選標準，是排除那些關於兒童性虐待及從事性交易的研究、或是那些無法取得全文的研究。除此之外，他們也把各種文獻分成兩種類別：其中一種是主要的而且品質良好的文獻（例如，非常具體、清楚地描述性行為的過程和情境脈絡），而另一種則是次要的文獻（例如，只是關於態度的報告，缺乏支持陳述的證據等等）。在這 268 份文獻當中，有 121 份被歸類為主要的文獻。馬爾斯頓及金恩所使用的方法是比較主題分析法（comparative thematic analysis），在進行此種分析時，研究者會先檢閱一些文獻或研究，並進行編碼，而這些文獻或是研究，都代表了研究中所發現的主題（例如，對女性所施加的暴力、對於難堪場面的恐懼等等）。然後再把這些編碼合併成廣泛的整體主題。最後，這兩位作者總共辨識出七個廣泛的主題，像是在決定社會期待（例如，女性應該要保持貞潔，男性則沒有這方面的問題；一般人認為男性所追求的是身體的快感，而女性所追求的則是浪漫愛情等等）方面，具有重要地位的性別刻板印象等等。

表 5.1　各種刊物的類型

類型	範例	文章作者	目的	優點	缺點
經過同儕審查的學術期刊	《社會科學季刊》（*Social Science Quarterly*）、《社會力量》（*Social Forces*）、《當代民族誌期刊》（*Journal of Contemporary Ethnography*）	專業的研究者	為專業人士報導實徵研究，以及建立知識	由於細節相當完整，因此品質最好、正確度最高、而且客觀程度最高	過於技術性、閱讀難度高、必須有背景知識、不一定能夠跟上最新的趨勢或是潮流
半學術的專業刊物	《美國展望》（*American Prospect*）、《社會》（*Society*）、《美國人口統計資料》（*American Demographics*）	教授、專業的政策制訂者、政客	與受過教育的民眾一起分享以及討論新的發現與其影響	大致上是正確的，而且也較為容易閱讀	缺乏完整的細節以及解釋、經常會在討論中摻雜個人的意見
新聞雜誌以及報紙	《華爾街日報》（*Wall Street Journal*）、《基督教科學箴言報》（*Christian Science Monitor*）、《新聞週刊》（*Newsweek*）、《時代雜誌》（*Time*）	受到敬重的新聞記者	以一種易於閱讀、容易理解的方式，為一般大眾報導時事	容易閱讀、易於理解、而且非常能跟上時代的脈動	半正確的、不完整的、扭曲的、或者是片面的看法
嚴肅的輿論雜誌	《國家》（*Nation*）、《人間事》（*Human Events*）、《公共利益》（*Public Interest*）、《評論》（*Commentary*）	教授、專業的政策制訂者、政客	為受過教育的民眾提供帶有特定價值的想法或是意見	文章的書寫與推理過程都相當仔細謹慎	片面的看法、並且高度偏向特定的價值立場
一般大眾閱讀的流行雜誌	《君子》（*Esquire*）、《Ebony》、《Redbook》、《富比士》（*Forbes*）、《財星》（*Fortune*）	新聞記者、其他的作者	娛樂大眾、或為一般民眾呈現與討論時事	容易閱讀，而且也容易取得	通常較為膚淺、不正確、而且也不完整

題相當多元，其中包括了法律執行、犯罪司法、國防和軍事、政治、政府與國際關係、以及政治學等等。而主要的保守派「智庫」——美國傳統基金會（the Heritage Foundation）——則會刊載一些題材，以作爲保守派人士針對重大政治議題進行辯論的論壇。有些時候，大學教授及一些進行正式研究的專業研究者，也會把自己的意見與想法投稿到這些期刊上。然而，在使用這些期刊時，必須非常謹慎小心，因爲這些期刊所呈現出來的只是辯論、意見、與判斷，並不是嚴謹實徵研究的正式報告。如果你想根據實徵研究（例如，一項實驗、調查資料、或是田野研究等等）寫出一

份研究報告，你就必須依賴專業的資料來源。如果你使用的是一篇表達個人意見的隨筆文章，你就必須以這種方式來看待這篇文章，而不能把此種文章和實徵的社會科學研究混爲一談。

研究者會使用特殊的電腦搜尋工具，找出在學術文獻裡的文章，而他們也必須學會用以指明資料來源的特殊格式或引述方式。專業的社會科學家經常會利用搜尋工具來取用、並建立一個不斷成長的研究與科學知識體系。對於所有認真閱讀研究的人及研究者而言，都必須精通下列幾種非常重要的技巧：知道如何找出相關的研究；辨識、閱讀、以及評估研究；還有適當地引用學術來源。

學術期刊 學術期刊是在文獻回顧裡最主要的資料來源。在學術期刊裡，充滿了經過同儕審查的研究報告（請參閱本書最後一頁所列出的社會科學學術期刊一覽表）。這些期刊幾乎只能在大學的圖書館裡才能找到。你可以回想一下，我們在第一章裡曾經提過，大多數的研究者散播新發現的方式，就是透過學術期刊。因此，學術期刊可以說是科學社群溝通系統的核心。

有一些學術期刊是很專門化的，且內容只有一些書籍評論，提供了針對其他學術書籍的評論和評價〔例如，《當代社會學》（*Contemporary Sociology*）、《法律與政治書評》（*Law and Politics Book Review*）等等）、或只有文獻回顧的文章（例如，《社會學年度回顧》（*Annual Review of Sociology*）、《心理學年度回顧》（*Annual Review of Psychology*）、以及《人類學年度回顧》（*Annual Review of Anthropology*）等等〕，而在此種期刊中，研究者所寫的都是一些關於「領域現況」的文章。以文獻回顧爲主的刊物，能夠針對某一個主題提供有用的介紹或是概述。許多學術期刊都同時包括文獻回顧、書評、研究報告、以及理論文章等各種類型的內容。

沒有任何簡單的方法或是某種「合格」的標章，可以幫助我們輕易地區隔學術期刊與其他刊物，或立刻就辨別出研究報告與其他類型的文章之間有何不同。爲了能夠正確地辨認出研究報告，你必須培養判斷的技巧，詢問經驗豐富的研究者或是專業的圖書館員。儘管如此，學習如何分辨不同類型的刊物，仍是你必須精通的重要技巧。分辨各類型刊物的最佳方法之一，就是大量閱讀學術期刊中的文章。

在不同的學術領域，學術期刊的數量也有很大的差異。心理學擁有超過 400 種以上的學術期刊，社會學大約有 250 種學術期刊，政治學與傳播學的學術期刊數量則略少於社會學，人類學—考古學及社會工作各自大約有 100 種學術期刊，都市研究以及婦女研究約略只有 50 種學術期刊，至於犯罪學，則大約只有十多種學術期刊。在「純粹」學術領域裡的期刊數量，通常會比在像是行銷與社工等「應用」領域中的期刊數量還多。每種期刊每年所刊登的文章數量，則從數十篇到超過 100 篇不等。

你或許會很好奇，是否有任何人曾真的把所有文章讀過一遍。有一項研究發現，在一份有 379 篇社會學文章的樣本裡，有百分之四十三的文章在出版之後的第一年之內，曾經被其他研究所引用，而在出版之後六年內，則有百分之八十三的文章曾經被引用過。[2] 正如同在第一章中所提過的，各種學術文章的聲望及接受比例各自不同。聲望卓著的期刊在投稿的研究報告中，大約只接受百分之十的文章。在社

會科學領域內，整體的文章退稿比例高於其他的學術領域，而且這個比例仍然在上升當中。[3] 不過，這並非意味著研究者所進行的研究品質日趨低落。更精確地說，其實是審查過程變得更爲嚴謹、審查標準正在日漸提升當中、而且也有更多的研究正在進行中。而此種情況也意味著，在受到高度敬重的期刊上刊登文章的競爭，已經變得越來越激烈。

你可以在網路上找到許多學術期刊文章的完整全文。一般而言，你必須透過圖書館才能取得這些內容，因爲圖書館會支付特別的訂閱費用給線上文章搜尋服務或是來源工具（source tool）。有一些期刊或出版社則只提供有限的文章、或是要求使用者付費購買閱讀文章的權利。舉例而言，我可以免費在網路上閱讀《社會科學季刊》裡的最新文章，但是當我試圖要在線上閱讀《政治與社會》這本期刊裡的一篇文章時，我便會被要求必須支付每篇文章二十五元美金的費用。不過，如果我透過任教的大學圖書館的網路，來閱讀這篇文章，就不必支付任何費用。

文章搜尋服務可能會提供學術期刊文章完整而正確的副本。例如像「JSTOR 資料庫系統」及「Project MUSE 電子期刊」都會提供正確的文章副本。但是其內容僅限於少數的學術期刊，並只提供前一年的文章內容。其他的來源工具，像是「Anthrosource」（人類學期刊全文資料庫）、「Proquest 資料庫系統」、「EBSCO HOST」、或是「Wilson Web 資料庫系統」等等，都會提供近期文章的全文版本。大多數網路文章的電子版本都與紙本版本的格式完全相同。除了利用來源工具以搜尋文章資料庫外，你也可以選擇一份特定的期刊，並快速地瀏覽目錄，尋找特定的議題。此種作法在產生關於研究主題的新想法、以具有創意的方式來思考一個既定的主題、或者是把一個想法擴展到新的領域時，都是非常有助益的。每一種線上來源工具都有自己的一套搜尋程序，與學術期刊的列表。不過，目前沒有任何一種來源工具，能夠提供所有學術期刊在所有年份中的全部文章。

有一些最近才出現的、僅有網路版本的學術期刊，被稱爲「*電子期刊*」（*e-journals*）（例如，《線上社會學研究》（*Sociological Research Online*）、《現代社會心理學研究》（*Current Research in Social Psychology*）、以及《世界體系研究期刊》（*Journal of World Systems Research*）等等），這些電子期刊裡的文章也都經過同儕審查的研究。到了最後，學術期刊的網路版本可能會全面取代紙本的版本。不過，就目前而言，大約有百分之九十五的學術期刊仍都有紙本版，且其中大部分的期刊在網路上都有全文版可供讀者使用。然而，幾乎所有的線上期刊文章系統或是資料庫，皆要求你透過圖書館的網路取得線上服務，因爲圖書館會支付使用這些服務的年度費用。另外，某些期刊和某些年份的文章，目前仍無法在網路上取得或閱讀。

一旦你找到一份包含實徵研究的學術期刊，你接著還得找到特定的文章。你必須確定，某篇特定文章所呈現的是某項研究的結果，這是因爲在同一份期刊裡通常會刊登數種不同類別的文章。要辨識出量化研究的文章會較簡單些，因爲在這類的文章裡，通常會有說明研究方法及資料的部分，也會有表格、統計公式、以及數字表。質化研究的文章則比較不容易被分辨出來，而且許多學生會把質化研究的文章和理論性論文、文獻回顧的文章、討論想法的論文、政策建議、書籍評論、及法律

案件分析等等文章混爲一談。若想正確地分辨這些不同類別的文章，你必須大量閱讀各種文章，才能夠瞭解研究的不同變化型態及其經驗。

大部分的大學圖書館都會設置學術期刊以及雜誌專區，或者，在某些情況下，有些圖書館會把期刊和書籍擺放在一起。你可以看看圖書館館藏的配置圖、或是請圖書館員告訴你放置學術期刊的位置。最新的幾期（看起來就像雜誌一樣）通常會被放置在一個「當期期刊」的單獨區域，這麼做是暫時擺放在那兒以供讀者查詢資料，直到圖書館收到該卷的每一期爲止。

圖書館會把許多不同領域的學術期刊和一般流行的、非學術性的期刊擺放在一起。這些館藏都屬於期刊類，或屬於圖書館員所用術語中的「分期期刊」（"serials"）。因此，你將會發現流行的雜誌（例如，《時代雜誌》、《道路與車道》、《柯夢波丹》、以及《亞特蘭大月刊》等等）就和天文學、化學、數學、文學、社會學、心理學、社會工作、以及教育等等學術期刊擺放在一起。圖書館會在館藏書目系統中，以名稱排序列出所有期刊，也會提供該圖書館所訂閱的期刊一覽表。

學術期刊的出刊頻率鮮少低到一年只出版一期、或是高到每個星期出版一期。多數的期刊每年大約出版四到六期。以《社會科學季刊》這本期刊爲例，就像其他同樣在名稱中有*季刊*這個詞的期刊般，一年的出版次數就是四次。爲了協助讀者找到需要的文章，每一期期刊都會有出版日期、卷數、期數。有了這些資訊之後，在尋找某篇特定的文章時，就會比較容易一些。這些資訊與諸如作者、標題、和頁數等等細節，被稱爲一篇文章的**引用**（citation）資訊，並且會被運用在參考書目或引用的作品列表中。一份期刊在剛發行的當期，會從第 1 卷第 1 期開始編號，然後再逐漸增加。大多數的期刊都依循著類似的系統，但是仍有爲數不少的例外情況。因此，你依然必須仔細注意引用資訊。對於大部分的期刊而言，每一卷代表收錄一整年份的文章。如果你看到某一份期刊已經出版到第 52 卷，這很有可能表示這份期刊已經發行長達五十二年之久了。除此之外，大部分（但並非全部）的期刊都是從一月份展開新的出版週期。

大多數期刊在編頁碼時是以卷計算，而不是以期計算。每一卷的第一期通常由第 1 頁開始編頁碼，而頁碼則逐漸增加累積到整卷完畢。例如，第 52 卷第 4 期的第 1 頁，在期刊上的正式頁碼可能是第 547 頁。大部分期刊都會有每一卷的索引，以及每一期的內容目錄，在其中會列出標題、作者姓名、與該篇文章的起始頁數。每一期中所包含的文章數量，可能少到只有一或二篇文章，也可能多達五十篇文章。大多數期刊的每一期通常會有八到十八篇文章，而每篇文章的長度則大約由五頁到五十頁不等。每篇文章通常會附上**摘要**（abstracts），也就是在文章的第一頁簡短地總結該篇文章的內容，或是把所有的摘要集中來，放置在該期的最前面。

許多圖書館不會保留過期期刊的實體紙本版本，而只保留電子版本或是微縮膠片版本以節省空間和支出。由於每個領域可能有數百種學術期刊，

引用　關於某項學術刊物的細節，能夠幫助人們快速地找到該項學術刊物。

摘要　某篇學術期刊文章的簡短總結，通常出現在該篇文章的開頭部分。摘要也是一種找到期刊文章的參考工具。

而每種期刊一年的訂閱費用大約是 100 到 3,500 美金，所以只有大型研究圖書館才有經費訂閱其中大部分的期刊。你也可以透過*館際借閱服務*（interlibrary loan service）來向其他圖書館借閱某篇文章的副本。館際借閱服務所指的是圖書館之間互相借調書籍或資料的一種服務系統。不過，只有少數圖書館會允許讀者外借最新幾期的學術期刊。

如果你到圖書館找到了期刊區，於走道間漫遊、並稍微瀏覽一下放在架上的書籍與期刊，是一件很有趣的事。你將看到包含許多研究報告的期刊。每一種學術期刊的名稱都有一個圖書編號（call number），就像是一般圖書館的館藏編號。圖書館通常會以期刊名稱的字母順序來整理這些期刊。然而，由於期刊有時候會更換名稱，因此如果仍歸類於原本的名稱下，就會造成一些混亂情況。

學術期刊裡的文章主題是某個學術領域裡的研究，因此，大部分數學期刊裡的文章都是關於最新的數學研究，或是證明的報告。文學期刊裡的文章是對文學作品的評論以及文學批評。而社會學期刊裡的文章則是關於社會學研究的報告。有些期刊涵蓋了一個非常廣泛的領域（例如，社會科學、教育、公共事務等等），且包含了來自整個領域的報告。其他的期刊則專注於一個次領域（例如，家庭、犯罪學、早期兒童教育、或比較政治學等等）。

引用的格式　一篇文章的引用資訊，是找到這篇文章的最重要資訊。在第四章的開頭部分，我們曾經讀過潘培爾（Pampel）對於文化品味、音樂、與吸煙行爲所進行的一項研究。假設你現在想要閱讀這項研究的完整內容，而其引用資訊如下：Pampel, Fred C. 2006. "Socioeconomic Distinction, Cultural Tastes, and Cigarette Smoking." *Social Science Quarterly*, 87(1): 19-35 這項引用資訊告訴你去尋找 2006 年出版的《社會科學季刊》這份學術期刊中的第一期。雖然在引用資訊中，並未提供明確的月份，但卻提供了卷數（第 87 卷）、期數（第 1 期）、以及頁碼（第 319 頁到 335 頁）。

引用文獻的格式有許多種不同的變化，而在內文中最常見的引用格式，就是使用括弧來標明作者的姓氏與出版日期。至於完整的引用資訊，則會出現在單獨的參考書目區。期刊文章、書籍、及其他作品的完整引用資訊，有多種不同的格式，且每一種都有其特定的格式。當你在引用文章時，最好先確認一下指導教授、期刊或其他刊物所要求的格式是哪一種。不過，幾乎所有的格式都包括作者的姓名、文章標題、期刊名稱、卷數和頁數。而在這些基本資訊之外，還會有許多不同的變化。例如，有些格式會包含作者的名字，而有些格式只包含作者姓名的縮寫字母；有些格式會包括所有的作者，有些格式則只會列出第一位作者；有些格式會包含期數和出版日期的資訊，但是有些格式則沒有包括這些資訊（請參閱圖 5.1）。

一般而言，引用的格式可能會相當複雜。在社會科學中各種主題的兩種主要的格式參考工具書，分別是《芝加哥格式手冊》（*Chicago Manual of Style*），其中有將近八十頁的內容在談論參考書目及如何引用的格式；以及《美國心理學協會出版手冊》（*American Psychology Association Publication Manual*），其中有大約六十頁的篇幅在討論格式的問題。而在社會學領域裡，《美國社會學評論》（*American Sociological Review*）之中有兩頁的內容是關於格式的說明，其格式也被廣泛地使用。

圖 5.1 期刊文章的不同文獻引用格式

《美國社會學期刊》（*American Journal of Sociology*）是美國最早期的一份社會學期刊，該份期刊曾經報導過由貝爾曼（Peter Bearman）和布克納（Hannah Bückner）針對貞操宣誓（virginity pledges）（譯註：貞操宣誓所指的是女性宣誓在婚前要守身如玉，要維持處女（virgin）的狀態一直到洞房花燭夜）所進行的一項研究。這篇研究報告出現在 2001 年 1 月那一期（第 4 期）的第 859 頁到 912 頁，而這份期刊從每年三月開始出刊。該期是第 106 卷，也就是這份期刊創立後的第 106 年。以下是引用這篇文章的幾種方式，其中兩種非常普遍的格式是「美國社會學評論」（American Sociological Review, ASR）以及「美國心理學協會」（American Psychological Association, APA）的格式。

ASR 格式

Bearman, Peter and Hannah Bückner. 2001. "Promising the Future: Virginity Pledges and First Intercourse." *American Journal of Sociology* 106: 859-912.

APA 格式

Bearman, P. and Bückner, H. (2001). Promising the future: Virginity pledges and first intercourse. American *Journal of Sociology 106*, 859-912.

其他格式

Bearman, P., and H. Bückner. "Promising the Future: Virginity Pledges and First Intercourse," *American Journal of Sociology* 106 (2001), 859-912.

Bearman, Peter and Hannah Bückner, 2001.
"Promising the future: Virginity pledges and first intercourse." *American Journal of Sociology* 106 (4): 859-912.

Bearman, P. and Bückner, H. (2001). "Promising the Future: Virginity Pledges and First Inter-course." *American Journal of Sociology* 106 (January): 859-912.

Bearman, Peter and Hannah Bückner. 2001.
"Promising the future: Virginity pledges and first intercourse." *American Journal of Sociology* 106 (4): 859-912.

Bearman, P. and H. Bückner. (2001). "Promising the future: Virginity pledges and first inter-course." *American Journal of Sociology* 106, 859-912.

Peter Bearman and Hannah Bückner, "Promising the Future: Virginity Pledges and First Intercourse," *American Journal of Sociology* 106, no. 4 (2001): 859-912.

書籍 書籍能夠傳遞許多種類的資訊、啓發思考、還能娛樂讀者。書籍的種類非常多，包括圖畫書、教科書、短篇故事集、小說、流行小說或非小說、宗教書籍、以及童書等等。而我們在這兒所關心的書籍，是那些含有原始的研究報告、或是匯集許多研究文章的書籍。就像其他類型的書籍，圖書館也會將這些書列入館藏，並指定圖書編目號碼。你可以在圖書館的圖書分類系統中找到這些書籍的引用資訊（例

如，標題、作者、出版社等等）。

把一本報導研究的書籍和其他類型的書籍明確區隔開來，可能是一件相當困難的事。你比較可能在大學的圖書館裡找到此類報導研究的書籍。有一些出版社，像是大學出版社，會專門出版研究報告。然而，在沒有親自去閱讀前，沒有任何一種方法能保證令你輕易辨別出這些與研究報告有關的書籍。有些類型的研究會比其他類型的研究更可能以書籍的方式呈現。例如，由人類學家及歷史學家所進行的研究，其研究報告的長度就很有可能和一本書差不多，而經濟學家或是心理學家的研究報告，則可能比較簡短。話雖如此，有些人類學和歷史學的研究報告仍可能以文章的形式呈現。而有些經濟學與心理學的研究報告，則以書籍的形式出現。在教育、社會工作、社會學、以及政治學這些領域裡，長時間及複雜的研究所產生的結果，可能同時以二到三篇的文章，與書籍的形式呈現。那些牽涉到詳細的臨床或民族誌描述、複雜的理論或哲學討論的研究，通常會以書籍的方式呈現。最後，如果某位作者想要和學術界的同儕及受過教育的民眾進行溝通，那他可能以利用撰寫書籍的方式，縮短學術界的格式和流行的、非小說格式之間的距離。而在書籍當中，找出原本的研究文章，是相當困難的事。因爲當中並無任何一個列出所有研究文章的單一來源。

有三種類別的書籍包含了研究文章或研究報告的集合。第一種類別是被稱作讀本（reader）的教學用書籍，其中可能包括了原本的研究報告。通常，來自於學術期刊、針對同一主題的文章會被蒐集在一起，並重新加以編輯，以使學生能夠更加容易閱讀與理解。第二種類別的書籍，則蒐集了期刊文章、也可能包含原本的研究、或針對一個特定主題的理論性文章。某些此類書籍會包含原本的研究報告，而這些難以找到的研究報告，則依據期刊中的某個專門主題組織安排。在此種書籍的目錄頁裡，會列出所有文章的標題和作者。圖書館把這些書籍和其他的書籍擺放在一起，而且有些圖書館的目錄系統，還會包含文章或章節標題。最後一種類別是年度研究書籍，此種書籍是一種介於學術期刊與文章集合的混合體。此類書籍包含一些在其他地方無法找到的研究報告。而此種類別的書籍一年只出版一次，每年都有個別的卷數。例如《政治社會學研究評論》（*Review of Research in Political Sociology*）以及《比較社會研究》（*Comparative Social Research*）等等，會和其他一般書籍擺在一起。有些年刊類的書籍專精於文獻回顧（例如，《社會學年度回顧》（*Annual Review of Sociology*）以及《人類學年度回顧》（*Annual Review of Anthropology*））。然而，與學術期刊不同的是，關於此類的書籍，並沒有一份完整而詳細的列表。找出這些書籍的唯一方式，就是花大量的時間待在圖書館，或是去詢問一位對某個主題領域已經相當熟悉的圖書館員。

書籍的引用或是參考資訊會比文章的引用資訊更爲簡短。在書籍的引用資訊裡會包括作者姓名、書籍標題、出版日期及地點、以及出版者的名稱等等。

論文　所有獲得博士學位的研究生，都必須完成一份原創性研究的著作，也就是*博士論文*（*dissertation thesis*）。這一份論文會被裝訂好、並放置在博士生取得學位的大學圖書館內。幾乎一半的博士論文，最後都會以書籍或是文章的形式被出版發表。

由於博士論文所寫的內容是關於原創性的研究，因此這些論文可能是非常具有價值的資料來源。有些獲得碩士學位的學生，也會進行原創性研究並完成碩士論文，但是碩士論文很少涉及到嚴謹的研究，而且要找到碩士論文也會比找到未出版的博士論文困難得多。

有些特殊的索引，會列出由公認著名大學的學生所完成的論文。例如，《國際論文摘要》（*Dissertation Abstracts International*）就會列出論文的作者、標題、以及大學。該份期刊索引的組織方式，是依照主題排列，還會附上每一篇論文的摘要。如果學校允許，你可以透過圖書館館際借閱系統，向授予學位的大學借閱大部分的論文。另外一個可行的方法，是向全國論文微縮膠片／複印本中心購買一份論文的影本，像是在密西根大學安納堡分校（University of Michigan, Ann Arbor）裡就有一個收藏全美國大學論文影本的中心。在一些大型的研究圖書館裡，如果之前有人曾經借閱過其他圖書館的論文，那麼圖書館也會保留這些論文的影本。

政府出版品 美國聯邦政府、其他國家的政府、州政府或是地方政府、聯合國、及其他像是世界銀行之類的國際機構，都會出資贊助研究並出版研究報告。許多大學圖書館的館藏都有這些文件或資料，且通常放置在一個特定的「政府出版品」區。這些研究報告很少出現在圖書館的館藏分類系統中，因此，你必須使用特殊的出版品列表與索引，還需要圖書館員的協助，才能夠找到這些研究報告。多數的大學圖書館只收藏那些最常被借閱的出版品及研究報告。

政策報告書及發表的文章 如果你正在進行完整的文獻回顧，你或許還得察看這兩種資料來源。這兩種文獻有些可以在網路上找到，但是大多數都很難取得，除非是受過訓練的專家才有辦法取得這些文獻。研究機構及政策中心（例如，布魯金斯研究院（Brookings Institute）、貧窮研究機構（Institute for Research on Poverty）、美國蘭德公司（Rand Corporation）等等）也會出版一些文章和報告。有些大型的研究圖書館會購買這些文章及報告，並把它們和其他書籍一起放置在架上。要確定某個機構或是中心已經出版過哪些刊物的唯一方法，就是直接寫信去詢問，抑或是向這些機構、中心索取一份已出版報告的清單。

每一年，在各個學術領域（例如，人類學、犯罪司法、地理學、政治學、心理學、以及社會學等等）裡的專業協會，都會舉辦年度會議。在這些會議裡，數千名的研究者齊聚一堂，在此發表、聆聽、或是討論最新研究的口頭報告。大部分的口頭報告都會有書面報告以供與會者索取。那些身爲協會成員、但是無法參加會議的人，也會收到會議的流程，其中列出了每一篇將在會議中發表的文章，並附上這些文章的標題、作者、及作者目前任職的機構。而這些未能參加會議的協會成員，可以直接寫信給作者，並向作者索取一份書面報告的影本。大多數（但並不是全部）在會議中所發表的報告，之後都會以文章的形式出版，有時這些報告也會出現在線上服務中（以供大家進行討論）。

如何有系統地進行文獻回顧

定義並且修正主題　就像你在開始一項研究計畫前，必須先進行規畫，並清楚定義主題和研究問題一樣，你在開始進行文獻回顧之前，也必須有一個定義清楚的、焦點明確的研究問題與計畫。一個好的文獻回顧主題，應該以一個研究問題的方式呈現。例如，「離婚」或「犯罪」都是過於廣泛的主題。因此，較爲恰當的文獻回顧主題應該是「什麼因素有助於繼親家庭的穩定性？」或是「經濟方面的不平等狀況，是否會造成跨國犯罪率升高？」如果你是爲一個研究計畫進行脈絡回顧，此種文獻回顧的範圍，應該比你正在進行檢視的特定研究問題稍微廣泛一些。在研究者尚未檢閱文獻前，通常不會明確決定最後要提出的研究問題。換句話說，文獻回顧能夠幫助研究者將研究問題聚焦。

設計搜尋程序　在爲文獻回顧選擇了一個聚焦且明確的問題之後，下一個步驟就是規畫一個搜尋的策略。你必須決定所要進行的文獻回顧類型、範圍、以及包含在內的資料類型。關鍵在於你必須仔細謹慎、有系統性、有條有理。接下來，你還得爲你的搜尋程序定下界限：你要花多少時間進行文獻回顧、要回溯到多久以前的資料、最少要檢視多少篇研究報告、以及你要走訪多少間圖書館查詢資料等等。

除此之外，你也要決定以何種格式爲每一筆參考資料記錄引用資訊，與如何記下筆記（例如，記在筆記本上、記在 3 乘 5 吋的卡片上、或記錄在電腦檔案裡）。你也應該建立一個檔案夾或是電腦檔案，在裡面你可以放置可能的資料來源、以及新資料來源的想法等等。隨著你的文獻回顧過程不斷地向前進展，你應該要越來越明確地聚焦在一個具體或特定的研究問題，抑或是議題上面。

找出研究報告　找出研究報告的方式，取決於你要尋找的報告或是研究出處的類型。一般而言，你應該使用多種搜尋策略，以抵消單一搜尋方法所可能造成的限制。

學術期刊中的文章　如同前面已經討論過的，大多數的社會研究都發表在學術期刊上。由於目前有數百份不同的學術期刊，每一份又都各自包含數百篇文章，因此，要找出特定的一篇文章，可能是一項極大的挑戰。幸運的是，線上搜尋服務和特殊的刊物，可以幫助我們把這個任務變得簡單一些。

或許你曾經使用過一般刊物的索引，例如「期刊文獻書目資料庫」（*Reader's Guide to Periodical Literature*）諸如此類。許多學術領域都有學術文獻的「摘要」或「索引」（例如，《心理學摘要》（*Psychological Abstracts*）、《社會科學索引》（*Social Sciences Index*）、《社會學摘要》（*Sociological Abstracts*）、以及《老年學摘要》（*Gerontological Abstracts*）等等）。就教育相關的主題而言，「教育資源資訊中心」（Educational Resources Information Center, ERIC）的系統是特別珍貴而有用的。現在已有超過一百種以上如同此類的來源工具供我們使用。利用這些來源工具或線上服務，你就可以使用標題、作者姓名、或主題，尋找相關的文章。

上述這些內容聽起來或許會讓你覺得，你必須做的事就只有找到正確的來源工

具，然後再去尋找某個主題的文章而已。有時候整個過程確實是如此沒錯，但是在某些時刻裡，事情其實更爲複雜。來源工具裡的主題其實都太過廣泛。你所想探討的特定研究問題，通常可能同時符合數個主題領域。因此，你應該仔細查對每一個主題領域。例如，以高中的非法藥物這個主題而言，你或許必須尋找下列這些主題：藥物成癮、藥物濫用、物質濫用、藥物法規、非法藥物、高級中學、以及中等學校。在某個主題領域裡的許多文章，可能與你的文獻回顧毫無關聯。除此之外，從一篇文章發表之後，到出現在來源工具之中，通常會有三個月到十二個月的時間落差。

大型研究導向圖書館會向「美國科學資訊研究所」（Institute for Scientific Information）訂閱《社會科學引用索引》（*Social Science Citation Index, SSCI*），而在這份珍貴的資源裡，包含了超過 1,400 種期刊的資訊。雖然這份索引與其他的索引及摘要相當類似，但卻必須花一些時間學習如何使用它。《社會科學引用索引》一共有四本書。其中一本是來源索引，提供了期刊文章的完整引用資訊。另外三本則是列在來源索引那本書的所有文章，而其組織的方式是依照主題、研究者任職的大學、研究中心、或是被其他文章所引用的作者等等。

你也可以利用作者、文章標題、主題、或是關鍵字來進行線上搜尋。所謂的關鍵字，指的是對某個主題而言相當重要的詞彙，通常也會是標題的一部分。在搜尋的時候，你大概會使用六到八個關鍵字，並考慮到其他數個同義字。不同的電腦程式進行搜尋的方式可能有所差異，且大多數都只尋找在標題或是摘要裡的關鍵字而已。如果你選擇的關鍵字太少、或使用非常狹隘的詞彙，你可能會錯過一些相關的文章。如果你選擇的關鍵字太多、或是使用非常廣泛的詞彙，那麼你有可能會搜尋到大量不相關的文章。學習關鍵字的適當範圍和數量的最好方法，就是不斷地進行嘗試，由錯誤中學習。

數年前，我曾經針對大學生定義*性騷擾*（*sexual harassment*）的方式進行過一項研究（Neuman, 1992）。當時我使用了下列的這些關鍵字：*性騷擾*、*性侵害*（*sexual assault*）、*騷擾*、*性別平等*（*gender equity*）、*兩性平等*（*gender fairness*）、以及*性歧視*（*sex discrimination*）。我後來發現了一些重要的研究，但是在這些研究的標題裡，卻不包含任何一個上述關鍵字。我也曾經嘗試過使用*大學生*及*強暴*這兩個關鍵字，但是卻找到數量極爲龐大的不相關文章，這些文章的數量多到我甚至根本無法大略地完全瀏覽一遍。

除此之外，還有爲數眾多的電腦輔助搜尋資料庫或是系統，可以讓我們加以利用。任何人只要有一部可以連上網際網路的電腦，就能夠搜尋文章索引的集合、許多圖書館的館藏目錄、以及其他全球各地的資訊來源。而上述這些資訊皆可透過網路取得。

所有電腦化的搜尋方法都共享一個類似的邏輯，但是每一種方法仍然有我們必須學習的獨特運作方式。在我的研究裡，我尋找了前七年的資料，並且使用了五個學術文獻的電腦化資料庫：《社會科學索引》、CARL（科羅拉多區域研究圖書館）（*Colorado Area Research Library*）、《社會檔案》（*Sociofile*）、《社會科學引用索引》、以及《心理學指引》（*PsychLit*）。

通常，你可能會在數個不同的來源工具資料庫中找到相同的文章。然而，如果你在搜尋時使用了好幾個不同的來源工具資料庫，你就可能在其中找到一些無法在其他地方找到的文章。因此，我們在此可學到一個極爲重要的教訓：「不要單獨依賴電腦化的文獻搜尋方式、摘要服務、單一領域內的文獻、或是一段任意定義的時期」（Bausell, 1994: 24）。例如，在仔細地研讀相關性最高的文章的參考書目之後，我發現了數個不在任何資料庫裡的、很棒的新資料來源。我的文獻搜尋過程是相當典型的。根據我所進行的關鍵字搜尋，我很快地瀏覽了超過 200 篇文章的標題或摘要。而從這些文章中，我挑選出大約 80 篇文章、報告、以及書藉進一步仔細閱讀。之後，我在這 80 篇文章裡發現了大約 49 篇相當有參考價值的，而且這些文章全部都包含在我後來出版的文章中的參考書目。

學術性書籍　要找到針對某個主題的學術性書籍，可能會是一件相當困難的任務。圖書館目錄系統裡的書籍主題通常不完整、也過於廣泛，因此無法發揮太多作用。除此之外，雖然你也許可以透過館際借書系統來搜尋其他圖書館的館藏資料，但是目錄系統通常只會列出在某個特定的圖書館系統裡的書籍而已。圖書館在爲書籍組織圖書編號時，通常是依據書籍的主題分類。然而，圖書館對於主題的分類方式可能無法反映你感到興趣的主題、或是反映出在某本書裡所討論到的所有主題。圖書館員可以幫助你在其他的圖書館找到你想要的書。例如，「美國國會圖書館全國聯合目錄」（*Library of Congress National Union Catalog*）會列出所有在美國國會圖書館裡的書籍。圖書館員可以進入一些列出其他圖書館館藏書籍的來源資料庫，或者，你也可以使用網路來進行搜尋。由於目前沒有任何一個萬無一失的方法可以幫你找出所有相關的書籍，因此你必須運用多種搜尋方法，像是查閱那些含有書評與文章參考書目的期刊等等。

論文　有一份稱爲《國際論文摘要》（*Dissertation Abstracts International*）的刊物，會列出大多數的論文。就像期刊文章的索引以及摘要一樣，這份刊物也是依據廣泛的主題類別、作者、日期來組織論文。研究者們會檢閱在主題領域內的所有論文標題，而這些標題可能包含他們所感興趣的主題。不幸的是，在你已經找到某份論文的標題及摘要之後，你或許會發現，要取得這份論文，不但要花費許多時間，還可能會有一些額外的金錢支出。

政府出版品　在圖書館裡的「政府出版品」區，通常都會包含這些出版品的特殊清單。《政府出版品每月目錄》（*Monthly Catalog of Government Documents*）是由美國聯邦政府所發行的一份相當有用的出版品索引，而且這份目錄通常可以在網路上取得。這份目錄從 1885 年開始發行迄今，但是，如果你要尋找的是超過十年之前的政府出版品，你應該同時使用其他補充性的來源。這份目錄也有年度索引，而每個月所發行的期刊則包含了主題、標題、以及作者的索引。另外一份有用的資料來源則是《國會聽證會索引》（*Indexes to Congressional Hearings*），其中列出了從 1930 年代晚期至今的所有委員會與主題。《國會記錄》（*Congressional Record*）則包含了美

國國會的辯論過程記錄，法案、投票記錄、以及修法記錄的一覽表。《美國法規大全》（*United States Statutes*）則依據年度及主題，列出了每一條美國聯邦的法規。《聯邦公報》是美國政府所發行的日刊，其中包含了所有美國聯邦機構的條款、規定、公告等等，而且其中也包含了每月和每年的索引。其他的索引則包括了條約以及法規公告等等。其他的政府也有類似的法單，例如，英國政府的《政府出版品索引》（*Government Publications Index*）會列出一年之內所有的政府出版品。《國會報告》（*Parliamentary Papers*）列出了過去二百年來官方的社會和經濟研究。在使用這些特殊的索引時，最好能夠依賴專業圖書館員的協助。不過，在索引中所使用的主題，不一定能完全符合你的特定研究問題。

政策報告書以及發表的文章 政策報告書及在研討會中所發表的文章，是非常難以尋獲的。你或許會看到這些文章被列在已出版研究的參考書目裡、以及在一些來源工具之中。通常，你必須寫信到研究中心去，並且索取其出版刊物的清單，或是取得在專業會議中所發表的文章清單等等。一旦你找到了一篇研究報告，你可以試著寫信給相關的作者或機構，以向他們索取一份完整的研究影本。

如何評估研究文章

在你找到一份已經出版的研究之後，你必須閱讀這份研究並進行評估。一開始，這會是一件困難的工作，但是隨著你越來越熟練，會漸漸變得較爲容易。以下是一些準則，可以幫助你閱讀與評估你所找到的研究報告，並找到適合你寫出自己的研究報告模式。

1. *檢視標題* 一個好的標題必須要明確而具體、在不描述研究結果的情形下指明研究的本質、並避免提出狹隘的是非題。標題會描述研究的主題、可能會提到一或二個主要的變項、也會描述研究的背景或是參與者。以下是一個良好標題的例子：「新加坡國民中學學生的家長參與以及紀律問題的下降」（"Parental Involvement and Reduced Discipline Problems among Junior High School Students in Singapore"）。一個好的標題能夠告訴讀者許多關於研究的資訊，而一個差勁的標題要不是意義模糊不清、就是過度強調技術性細節或術語。以上述的研究爲例，同樣的一篇研究，也可以取一個令人困惑的差勁標題：「針對影響降低分段行爲焦慮因素的三步驟相關性分析」（"A Three-Step Correlation Analysis of Factors That Affect Segmented Behavioral Anxiety Reduction"）。

2. *閱讀摘要* 一個良好的摘要可以總結關於一項研究的重要資訊。良好的摘要能夠說明該項研究的目的、指出使用的方法、並強調重大的發現。此外，良好的摘要也會避免以模糊不清的方式，提及該項研究對於未來的影響或是暗示。在對標題進行過初步篩選過程後，你應該能夠藉由一篇準備充分的摘要，決定一份研究報告的相關性。除了以相關性來進行篩選之外，標題或摘要可以讓你準備好去仔細地檢視一份研究報告。我建議使用一種兩階段的篩選過程。首先利用標題和摘要初步決

定某篇文章的相關性。如果你認爲該篇文章似乎具有相關性，那麼你就快速地瀏覽一下緒論以及結論的部分，以決定這篇研究報告是否要留下來（也就是值得你投入時間去慢慢地、仔細地把整篇文章讀完）。最有可能的情況是，你將會發現只有少數文章符合你的研究目的，許多文章則略爲觸及相關主題而已，這些文章只值得你花一點時間快速地瀏覽並找出一或二個特定的相關細節。另外，特別要注意的是，在你找出這些特定的相關細節時，不要使這些細節脫離原本的脈絡。

3. *閱讀文章*　在閱讀整篇文章之前，你或許會想要先簡略地看過一開始的最初幾段，並且快速地閱讀結論的部分。此種作法能夠讓你對於整篇文章的內容有初步的概念。有一些因素會影響到閱讀一篇學術文章所投入的時間與努力程度、以及整體的報酬率。在以下的三種情況裡，你在閱讀文章時可以達到事半功倍的效果：（1）該篇文章是一篇品質很高的文章，有明確定義的目的、清楚的書寫方式、與平順且合乎邏輯的組織方式，（2）你很明確地聚焦在一個特定的議題或是問題，以及（3）你擁有紮實的理論背景、相當瞭解實質的主題、並熟悉研究方法論。正如同你可以看到的，其實很大的部分取決於讀者自己的事前準備工作。你可以藉由辨識出一項研究的面向（請參閱第二章）、該項研究所使用的理論（請參閱第三章）、以及該項研究所使用的研究方法（請參閱第四章）等方式，發展出良好的讀者準備工作，好讓你能迅速地「評估」（“size up”）一篇文章的價值。除此之外，你還必須知道，不同的作者在寫作時，心中所設定的讀者群是不一樣的。有些作者的寫作對象是科學社群裡一群範圍狹小的、高度專業化的人士；有些作者的寫作對象是廣泛的、橫跨不同領域的學生，以及在數個學術領域裡的學者；或者，有些作者的寫作對象是政策制訂者、議題擁護者、以及應用領域的專業人員等等。

當你在閱讀一篇高度相關的文章時，應該從緒論的部分開始讀起。緒論部分有三個目的：（1）介紹一個廣泛的主題，且將這個廣泛的主題平順地轉化成爲一個明確具體的研究問題，而這個研究問題將會是該項研究的主要焦點；（2）確立研究問題的重要性（在擴展知識、連結過去的研究、或處理一個應用的議題等各方面）；以及（3）概略描繪出理論架構，並定義主要的概念。有時，一篇文章會把緒論與脈絡文獻回顧混合在一起，不過在其他時候，文獻回顧會單獨放在另一個部分。

爲了寫出一篇良好的文獻回顧，你必須有所取捨、面面俱到、勇於批判、並且要能不斷改進。有所取捨的意思是，你不必列出關於某個主題的所有文章，只須列出最具相關性的文章即可。而面面俱到的意思是，你應該把具有高度相關的過去研究放入文獻回顧裡，並不遺漏任何重要的文章。除此之外，你也不能對過去的研究照單全收，而應以批判性的方式進行評估，換句話說，你必須對某些特定研究的細節作出評論，並且評估這些細節與你自己研究的相關性有多高。最後，直到你完成研究之前，你不會完全瞭解你研究裡的一切，因此，即使你已經完成了文獻回顧，你還是應該準備好，隨時對其內容進行微調、甚至完全重寫。

你應該在你的文獻回顧裡包含最近的研究。根據這些研究的大小與複雜度，或許你必須把理論、方法、發現、及評估等各部分區隔開來。例如，你可能會檢閱理論性議題及爭議、調查先前的研究者所使用過的方法、總結研究發現並突顯出任何

的漏洞，抑或是不一致的地方。對過去的研究進行評估，可以幫助你爲進行此項研究，提出具重要性並合理的理由。

依據一篇文章中所使用的研究取向，在文獻回顧之後的部分，可能便是說明假設或研究方法的部分。這些部分會概略說明特定的資料來源，或者是蒐集資料的方法、描述測量各種變項的方式、是否使用了抽樣方法、以及如果有使用抽樣方法，關於抽樣過程的細節。你可能會發現這部分的書寫方式較緊密，且充滿了許多技術性的細節。這些部分在量化研究中，會比在質化研究中還要來得更長。

在談論研究方法的部分後，則是關於研究結果的部分。如果某項研究是一個量化研究，在這個部分就應該不只是提供許多的統計圖表，或相關係數與百分比例而已，應該還要討論這些圖表和資料所顯示的意義。如果是一項質化研究，這個部分就應該不只列出引文或直接描述而已。呈現資料時的組織方式，通常是由描繪一個廣泛的範圍開始，然後再進入比較複雜的部分以及具體的發現。資料的呈現通常包括針對核心發現的直接討論，並且註明這些發現的重要性。在量化研究裡，其實沒有必要討論在某個圖表裡的每一個細節，只要註明重大的發現與任何意料之外的、或是不尋常的發現就足夠了。在一篇好文章裡，作者將會引導讀者閱讀資料、指出研究的內容、並且展現出所有資料細節。在質化研究裡，資料的組織方式通常是講述一個故事、或呈現出一種推理的思考方式。讀者跟隨著作者所講述的故事，但仍然可以自由地詢問與故事有關的問題。

在某些文章裡，作者會把討論的部分和結果的部分結合在一起。而在其他的文章，這兩個部分則是各自分開的。在討論的部分裡，其內容其實不只是單純的描述，而是詳細說明研究結果對於過去的發現、理論、或是應用議題所具有的含意。討論部分也可能包含由文獻回顧來建立過往發現的說明，和對於特定研究問題的意義與暗示。除此之外，這個部分也可能包括對任何意外所發現的評論。

大多數的研究者會把研究的方法論限制放在討論部分，而內容通常包括特定的測量方式、抽樣、個案、地點。或其他因素會以何種方式限制研究發現的可推論性、抑或是仍存在其他解釋法的可能性。讀者們會預期研究者能完全地坦誠以及公開。在一篇好的文章裡，作者會進行自我批判，並表現出自己能夠察覺該項研究所具有的缺陷。

在你已閱讀過討論與結果的部分之後，你應該要再讀一次該篇文章的結論／總結部分。一篇好的結論／總結會回顧研究問題、重大發現、以及具有意義的意外結果。除此之外，一篇好的結論也會概略描繪出對於未來的可能影響及建議後續的研究方向。你或許會想尋找可能包含額外研究細節的附錄、並檢視一下參考書目的部分。一篇文章的參考書目能夠提供你一些相關研究或者是理論陳述的線索。

閱讀並以批判方式來評估學術文章，需要專注力及時間，也必須多加練習才能夠熟能生巧。雖然學術文章須經過同儕審查，且有一定程度的退稿比例，但是文章的品質仍然有極大差異。有些學術文章可能包含許多錯誤、鬆散的邏輯、以及大量的漏洞。你必須注意的是，文章的標題及緒論的部分，可能無法與研究結果裡的特定細節完全吻合。作者未必總會描述所有發現。當一名讀者有著清楚聚焦的目的、並相當仔細地閱讀一篇文章時，有可能會在研究發現中注意到新的細節。舉例而言，

作者可能不會提到在一份統計圖表裡相當明顯的重要結果，或過於強調某些微小的、較不重要的結果。當你在評估一篇文章時，請特別注意該篇文章所提及的研究是以何種方式進行、文章的各個部分在邏輯上如何整合、以及文章之結論是否真的是由所有研究的發現所推導出來的。

記筆記的方法

當你在蒐集相關的研究文獻時，你或許會覺得快要被大量的資訊所淹沒了，因此，你必須有系統地記筆記。舊式的記筆記方法，是把筆記寫在索引卡片上，然後更動並整理這些筆記卡片，把卡片一疊疊地放好，同時尋找這些筆記內容之間的關聯、發展出一份報告或是論文的綱要。雖然這個方法目前仍然有用，但時至今日，大多數人會選擇使用文書處理軟體，以便於記筆記和整理筆記，並蒐集許多文章的影本或是書面版本。

當你發現了新的資料來源，你或許會想要爲筆記卡片或電腦文件建立兩種文件夾的類型：一種是*資料來源文件夾*（*source file*）、另外一種則是*內容文件夾*（*content file*）。你應該把每個資料來源的*所有*參考書目資訊都放在資料來源文件夾裡，即使你或許不會用到其中的一些資訊也一樣。不要遺漏完整的參考書目引用資訊裡的任何細節，像是頁數或是第二作者的姓名等等。如果你遺漏了某些細節，之後有可能會悔不當初。刪除掉一筆你不會用到的資料來源，會比你後來試圖在一個你已發現的資料來源中找出你需要的、但卻遺漏掉的參考書目資訊來得容易多了。我建議你建立兩種不同的資料來源文件夾，或把一個大型的文件夾分割成兩個部分：*已閱讀文件夾*（*have file*）與*待處理文件夾*（*potential file*）。放在已閱讀文件夾裡的，是那些你已經找到的、並已經記過內容筆記的資料。而放在待處理文件夾裡的，則是一些你還沒有仔細追蹤或是閱讀的線索提示，以及可能的新資料。每當你偶然發現了一份新的資料、或是一篇新文章的參考書目時，你都可以把這些資訊增加到待處理文件夾當中。而在接近完成一份報告時，待處理文件夾的內容應該會逐漸減少，最後通通消失，而已閱讀文件夾裡的內容，則會在之後成爲你的參考書目。

在內容文件夾裡，包含了來自於一份資料的實質資訊，而且是你有興趣的資訊。其中通常包括重大發現、方法論的細節、概念的定義、或是耐人尋味的引文。如果你直接引述一份資料裡的語句、或從中擷取一些具體的資訊，你必須記下這些引述的語句或是資訊出現的確切頁碼。除此之外，你也可以在每個內容文件夾的名稱加入像是作者、日期之類的關鍵資料與來源資訊，以便把不同的內容文件夾串連。

應該記錄的資訊　你必須決定要記下關於一篇文章、一本書、或是其他資料的何種資訊。寧可盡量多記下一些資訊，也不要記下太少資訊。一般而言，你應該記下被檢驗的假設、重大概念的測量方式、主要的發現、研究的基本設計、所使用的團體或樣本、以及對於後續研究的建議（請參閱圖 5.2）。比較明智的作法，是去檢視該份報告的參考書目，並記下一些你可以增加到自己研究中的可用資料。

圖 5.2 文章筆記的範例

參考書目的完整引用資訊（資料來源文件夾）

Bearman, Peter, and Hannah Bückner. 2001. "Promising the Future: Virginity Pledges and First Intercourse." *American Journal of Sociology* 106: 859-912. (January, issue no. 4)

筆記卡（內容文件夾）

貝爾曼（Bearman）以及布克納（Bückner）2001

主題：青少年懷孕與性行為，宣誓／承諾，貞操，第一次性行為，南方浸信會教友，認同運動

自從 1993 年以來，南方浸信會就在青少年之間發起一項運動，要求青少年公開宣誓在結婚之前都要維持貞操。到目前為止，已經有超過二百五十萬名青少年作出這項宣誓。本項研究檢視了這項宣誓是否會影響青少年發生性行為的時機，以及作出宣誓的青少年與未作出宣誓的青少年之間是否有任何差異。這項運動的批評者對這項運動感到不自在，因為支持宣誓的人通常拒絕接受性教育、對於婚姻抱持過度浪漫化的看法、且堅守傳統的性別角色。

假設

根據限制從事禁忌行為機會的社會控制力量的強弱，青少年會刻意從事一些成人們可以享受、但卻禁止青少年去做的行為。出生於較為放任的非傳統家庭中的青少年，比較有可能會從事禁忌的行為（性行為）。而出生於傳統家庭並與家長關係較為密切的青少年，會延遲發生性行為的年齡。與家庭之外的「認同運動」有密切關係的青少年，則會依據這些運動所教導的規範，修正自己的行為。

方法

本項研究的資料來自於一份 1994 年到 1995 年之間，針對美國公立或私立學校裡就讀 7 年級到 12 年級的青少年，所進行的全國健康調查。在 141 所學校裡，總共有九萬名學生完成了問卷調查。在這些接受問卷調查的九萬名學生中，有二萬名學生之後再完成第二次的問卷調查。問卷的詢問內容關於宣誓、宗教的重要性、以及性行為等等。

發現

本項研究發現，作出宣誓的青少年確實大幅度地延遲了發生第一次性行為的時間。然而，宣誓所產生的效果，會根據青少年的年齡而有所差異。除此之外，宣誓只有在某些社會脈絡下才能發揮作用（也就是說，至少宣誓在某方面而言算是一項社會規範時）。作出宣誓的人通常對宗教比較虔誠、在生理方面比較晚熟、而且通常來自於傳統的社會與家庭背景。

把所有相關的文章或是報告通通影印下來，不但可以節省你記筆記的時間，也能夠確保你擁有一份完整報告。除此之外，你也可以在影本上標注筆記。不過，關於這個作法，你必須考量到幾個問題。首先，如果你所進行的是大型的文獻搜尋，把所有的文章或報告通通影印下來，將會所費不貲。其次，你必須小心並遵守相關的著作權法。美國的著作權法僅允許人民以個人研究用途的名義影印整份文件。第三，記得要記錄或是影印完整的文章，包括所有的引用資訊。第四，整理一大疊的

文章可能會是一件相當累人的工作，尤其是當你想要使用同一篇文章的數個不同部分時。最後，除非你仔細地標註或是筆記做得很好，否則，你之後可能必須重新讀過整篇文章。

整理筆記　在你已經蒐集了大量的參考資料及筆記之後，你必須有一個整理這些資料的方法。其中一個方法，是先快速地瀏覽筆記，以把各種不同的研究或特定的發現分門別類，並在腦海中建立一個心智地圖以標示這些研究，或發現是以何種方式連結在一起的。在決定最後要使用哪一種整理方法前，你應該多嘗試幾種不同的組織整理方法。整理資料是一種會隨著不斷練習而進步的技能。例如，把筆記放進代表一般主題的類別裡，或是畫出圖表來比較不同的報告對於相同問題的說明或陳述，抑或是標記出任何相同或是不同之處等等。

在整理筆記的過程中，你會發現某些參考資料或筆記無法放進任何一種類別之中。在遇到這種情況時，你應該把這些資料或是筆記視爲不相關，將它們丟棄。你也可能會發現一些漏洞、相關的領域或主題，但是你卻還沒檢視過這些領域或主題。當你遇到這種情況時，你就必須重新回到圖書館，查詢相關的資料。

最好的整理方法，其實取決於文獻回顧所要達到的目的。如果是*脈絡回顧*，那就代表著你必須以一個特定的研究問題爲中心，組織安排新近的研究報告。如果你進行的是*歷史回顧*，這就表示你必須依據主要的主題與出版日期，組織安排你所找到的各項研究。如果是*整合式回顧*，那麼你在組織安排各項研究時，就必須圍繞著某一個領域裡的一個核心，共同發現及被驗證的主要假設。若是*方法論回顧*，就必須依照主題組織安排各項研究，且在每個主題裡，必須依據研究設計或是使用的研究方法組織安排各項相關的研究。至於*理論回顧*，必須依據理論以及主要的思想家，組織安排各項研究。

規畫以及撰寫文獻回顧

良好的文獻回顧需要完善的規畫與清楚的撰寫方式，而且必須不斷地重覆修改。所有良好寫作的規則（例如，清楚的組織架構、引言和結論、段落之間的起承轉合等等），都可以應用於撰寫文獻回顧上。當你在撰寫文獻回顧時，要把你的目的時時放在心上，並以清楚而有效的方式進行溝通。你必須利用整篇文獻回顧的組織方式，來傳達或溝通這篇文獻回顧所要達成的目的。*錯誤*的撰寫方式，就是只列出一連串的研究報告和每篇報告的研究發現摘要而已。

此種方式根本無法傳達你所要表達的目的，因爲以此種方式所撰寫成的文獻回顧，讀起來就像只把一堆筆記串連在一起而已。每當我看到此種文獻回顧時，都會覺得該篇文獻回顧的作者只是在敷衍了事，而且在撰寫這篇文獻回顧時，完全忽略了一個相當重要的組織步驟。撰寫文獻回顧的*正確*方式，是去整合並組織共同的研究發現。有一種被廣泛接受的撰寫方式，就是先處理最重要的那些概念，然後以合乎邏輯的方式連結共同的陳述或發現，再標記不一致的差異處或缺點（請參閱範例 5.3 不佳以及良好文獻回顧的範例）。

範例 5.3

不佳以及良好文獻回顧的範例

不佳文獻回顧的範例

性騷擾會造成許多影響與後果。亞當斯、柯特克、以及帕德吉特（Adams, Kottke, and Padgitt, 1983）發現有些女學生說她們會避免選修某些課程、或和某些教授一起工作，因為她們擔心會有被性騷擾的風險。作者也發現到，男學生與女學生對於性騷擾的反應是不同的。該項研究是針對一千名男性和女性研究生，以及大學生所進行的調查研究。班森與湯姆生（Benson and Thomson, 1982）在《社會問題》（*Social Problems*）上所發表的研究，則列出了許多由性騷擾所造成的問題。而在狄傑奇與韋納（Dziech and Weiner, 1990）合著的《好色教授》（*The Lecherous Professor*）這本廣受好評的書裡，兩位作者亦列出一長串性騷擾的受害者曾遭遇的困境。

不同的研究者會採取不同的方式來研究這個主題。杭特與麥克利蘭（Hunter and McClelland, 1991）曾經針對一所小型博雅學院裡的大學生，進行一項研究。他們挑選了三百名學生作為樣本，並給這些學生一些圖卡，圖卡上面畫有各種性騷擾受害者的不同反應與情境。賈區克以及弗瑞茲（Jaschik and Fretz, 1991）放了一段影片給一所位於美國中西部大學裡的九十位女學生看，影片內容是關於一位教學助理遇到性騷擾的過程，而這也是典型的性騷擾案例。在研究者還未把影片中的行為稱為性騷擾之前，幾乎沒有任何一位女學生把影片中的內容稱為性騷擾，但在研究者詢問這些女學生影片中的行為是否就是性騷擾時，有高達百分之九十八的學生同意這個說法。韋伯－柏汀以及羅西（Weber-Burdin and Rossi, 1982）以麻州大學的學生為對象，複製了一個之前針對性騷擾所進行的研究。在這項研究中，兩位研究者要求 59 位學生為 40 個假設的情境進行評分。雷利、卡本特、杜爾、以及巴特列特（Reilley, Carpenter, Dull, and Bartlett, 1982）等人，則針對加州大學聖塔芭芭拉分校裡的 250 名女性大學生、150 名男性大學生、以及 52 名教職員工進行了一項研究。這三個樣本群體（其中兩個群體是大學生，而另外一個群體則是教職員工）都完成了一份問卷，而在這份問卷中，填答者必須對各種性騷擾情境的插圖進行評分。帕波維奇等人（Popovich et al., 1986）則創造了一個關於九種性騷擾情境的量表，然後把一所中型大學裡的 209 名大學生分成 15 人到 25 人不等的團體來進行研究。他們後來發現在大學生當中，對於性騷擾的定義有許多意見不同和混淆的情況。

良好文獻回顧的範例

性騷擾受害者會承受各種程度不同的影響或結果，範圍從自尊低落、失去自信，到從社會互動中退縮、改變生涯目標、以及憂鬱沮喪等等（Adams et al., 1983; Benson and Thomson, 1982; Dziech and Weiner, 1990）。例如，亞當斯等人就注意到，有百分之十三的女學生說她們會避免選修某些課程、或是與某些教授共事，因為她們認為會有被性騷擾的風險。

研究者採取了數種不同的方式來研究校園中的性騷擾問題。除了調查研究之外，許多研究者也曾嘗試使用圖卡、或是呈現假設情境的方式（Hunter and McClelland, 1991; Jaschik and Fretz, 1991; Popovich et al., 1986; Reilley et al., 1982; Rossi and Anderson, 1982; Valentine-French and Radtke, 1989; Weber-Burdin and Rossi, 1982）。受害者的言語回應方式與情境因素，似乎會影響觀察者是否將某種行為認定是性騷擾。在把不適當的行為貼上性騷擾標籤這方面，似乎仍有些混淆的狀況。舉例而言，賈區克及弗瑞茲（Jaschik and Fretz, 1991）發現，在觀看過一段關於教學助理被性騷擾的典型案例的影片之後，只有百分之三的女學生會在一開始時就把影片中的行為稱為性騷擾。相反地，大多數的女學生會認為那是「性別歧視」、「粗魯無禮」、「不專業的」、或是「人品低下的」。而當她們被問到影片中的行為是否為性騷擾時，則有高達百分之九十八的女學生同意這樣的說法。至於羅斯可等人（Roscoe et al., 1987）也報告了類似的、在認定性騷擾方面所遭遇的困難。

如何在社會研究中使用網路

網際網路已經徹底地改變了社會研究者進行研究的方式。在大約十多年以前，幾乎沒有什麼社會研究者使用網路。然而，到了今日，所有的社會研究者都經常使用網路幫助他們回顧文獻、與其他研究者進行溝通、搜尋其他的資訊。而網際網路本身也在不斷地持續擴展以及改變當中。而人們對於網路的看法毀譽參半，網路也非如某些人最初所認知的萬靈丹。網際網路提供了嶄新的、快速的、以及重要的尋找資訊之法，但仍然只能算是許多工具中的一種而已。然而在社會研究中，使用網路有其優點，當然也有其缺點。

優點

1. *網際網路容易使用、迅速、便宜*。在許多地方都可以連結並使用網路，而這項近乎免費的資源，也讓人們幾乎能夠由任何地方找到資料。不論是地區的公共圖書館、家中、實驗室或教室、咖啡店、或任何有電腦可以連上網路之處，網際網路一天 24 小時、一個星期 7 天都在運作當中。只要有最低限度的訓練，大多數的人都能很快地學會如何搜尋並取得資訊。而不過幾年前，人們還必須特地到大型研究圖書館一趟，才能取得這些資訊。以利用電腦系統的方式來搜尋大量的資訊，會比用人工的方式進行搜尋，還來得更爲容易及快速。網際網路大大地擴展了資料來源的數量與種類。除此之外，一旦找到了資料，研究者通常可以用電子檔案的方式儲存、或是直接利用手邊的印表機將這些資料列印出來。

2. *網際網路上有「連結」（"links"），可以提供我們找到並連結其他資訊來源的額外方法*。網站、首頁、及其他網路資源網頁，都會提供許多連結，讓使用者可以直接點擊連結的指標（通常是一個按鈕或特別標注的文字），就能連結到相關的網站或資源，取得我們想要的資訊。此種情形讓使用者可以連結更多的資訊，也讓使用者能夠取得交互參照的資料。在一個由相關來源所組成的網絡中，連結會鑲嵌在其中的一個來源，讓使用者藉此連結到其他的相關來源。

3. *網際網路大幅地增加了資訊在全球流通的速度，並產生一種「民主化」效果*。網際網路讓資訊（例如，文件、新聞、資料、以及照片）可以在遙遠的距離與不同的國界之間，迅速地進行傳輸。在十年之前，要取得某些報告，可能必須等上一個星期或一個月的時間，而且還可能需要花一筆錢。時至今日，你卻可以在幾秒之內便取得這些報告，而且完全免費。網路幾乎不會限制什麼人能夠將資料放置於網路上、或什麼樣的資料才能放置上網路。而這也意味著，那些原本無法出版或散佈資料的人，現在很輕易地就能夠出版或是散佈任何資料。由於網際網路的開放性，也強化了普世主義的規範。

4. *網際網路讓人們可以取得範圍廣泛的資訊來源，其中有些資訊的格式是動態的且相當有趣的*。你可以取得一份以黑白文件格式呈現的報告，就如同在傳統的學術期刊及資料中一樣。或者是一份包含明亮的色彩、圖表、動態影片、照片、甚至是聲音或影片片段的報告。作者與其他的資訊創作者可以在呈現方式上盡情地發揮

他們的創意。

缺點

1. *對於能夠放置在網際網路上的資料，沒有任何品質方面的控管機制*。與標準學術刊物不同的是，網路上的資料不必經過同儕審查或者是任何的審查過程，任何人都能把任何東西放在一個網站上。這些資訊可能品質低劣、未註明出處、具有強烈偏見、憑空捏造、或全然是欺騙。一旦你在網路上找到資料，你必須具有一些技巧，才能分辨出「垃圾」和可用資料。你也必須小心謹慎地對待網頁，就你對待在街上從別人手中拿到的宣傳單一樣。因爲其內容可能只是一個瘋子的胡言亂語，但也可能含有高價值資訊。另外一個較不嚴重的問題，是那些在網站上常見的引人注意的花招，好比鮮豔的色彩、音樂、或動態圖像等等。可能會使毫無戒心的使用者分心，因而未注意到較爲嚴肅的內容。因爲這些使用者可能誤以爲炫目的花招就是高水準的資訊。除此之外，網際網路比較適合迅速地瀏覽及短時間的注意，卻不適合緩慢地、審慎地、及仔細地閱讀與研究其內容（請參閱資訊箱 5.3 網站：網路瀏覽者請小心！）。

2. *許多極佳的資料以及一些具有關鍵性的資源資料，仍然無法在網際網路上取得*。與一般普遍想法相反的是，其實網路並未讓所有人都能免費取得所有資訊。一般而言，免費的資訊都是有所限制的，只有付費使用者才能取得較爲完整的資訊。

3. *在網際網路上尋找資料，可能必須耗費大量的時間*。要找出特定的資料來源，並不是一件容易的事。目前網路上的數種搜尋引擎（例如，Google、Bing、Yahoo、Altavista、Lycos、以及 AskJeeves.com 等等）的運作方式都有些微不同之處，而且可能產生差異極大的結果。我在五分鐘之內利用四種不同的搜尋引擎來搜尋同樣的

資訊箱 5.3　網站：網路瀏覽者請小心！

由於網際網路快速普及，而人們也越來越依賴利用網路取得各種資訊，因而帶來許多便利之處。由於網際網路可說是形同不受管制，因此幾乎任何人都能建立一個網站，並在網站上發表任何意見。在 2000 年，有超過六千萬名美國居民在網路上搜尋與健康相關的資訊。而在網際網路的使用者當中，有超過百分之七十的人認為，他們在網路上所找到的資訊，將會影響他們對於治療方式的決定與選擇。一項針對在網路上可取得的健康資訊所進行的研究（Berland et al., 2001）發現，網路上的健康資訊通常是不完整或不正確的。研究者使用了十個英語的搜尋引擎及四個西班牙語的搜尋引擎，搜尋以下四個詞彙：「乳癌」、「兒童氣喘」、「憂鬱症」、以及「肥胖」。他們發現在與健康資訊有關的網頁上，所提供的背景資訊中，只有不到四分之一的資訊是有根據、相關的資訊。

在上述這項研究中，有三十四名醫生評估了二十五個健康資訊網站的品質，而他們所得到的結論是，只有不到一半的網站提供關於某個主題充分且完全正確的資訊。研究者們還發現到，在多數時候，某個網站的某部分資訊，與同一個網站裡的其他資訊、或是同樣主題裡的資訊之間互相矛盾。他們也發現到，在是否提供完整資訊來源與記錄方面，不同的網站之間也有極大的差異。平均而言，只有大約百分之六十五的健康資訊網站，提供了關於作者和資料日期的正確資訊。

詞彙：「褫奪公權」（voter disenfranchisement）。接著我檢視了每種搜尋引擎所找到的最前面三項的結果。每種搜尋引擎都會找到一個或多個其他引擎所未找到的網站結果。只有兩個網站出現在超過一個以上的搜尋引擎結果裡，所有其他的網站都只出現在個別的搜尋引擎所產生的結果裡。而在這兩個出現在超過一個以上的搜尋引擎前三名結果裡的網站，其中之一的連結竟然是失效的。顯而易見地，你必須使用多種搜尋引擎，且需要檢視第一頁之外的結果。大多數的搜尋引擎只是單純地去尋找，關於某個網頁的簡短描述中的特定字詞。此外，搜尋引擎也可能找到數萬筆資料，而這些資料量實在多到沒有任何人能夠逐一檢視的地步。除此之外，某些網站之所以會出現在最前面的搜尋結果，可能只是因爲在這些網站的簡短描述裡，包含了搜尋關鍵字的各種不同說法。而那筆對你而言「最佳」的網路資料，可能被深埋在搜尋結果裡的第 150 筆資料。

4. *網路資料可能「不穩定」，而且難以證明*。你可以進行網路搜尋並找到一些提供有用資訊的網頁。然而，當你在一個星期之後要連結到同樣的網頁時，卻可能發現，其中好幾個網頁已經消失不見了。你必須注意到某一個特定網頁所在的「統一資源定位器」（uniform resource locator, url）或者是「位址」（“address”）（通常是以 http:// 開頭）。網頁位址會指向位於某處的一部電腦裡的一個電子檔案。一般的期刊文章會被收藏在圖書館的書架上，或是微縮膠片上，時間可以長達數十年之久，並且讓任何人都能夠取得與閱讀。然而，和期刊不同的是，網頁可能會在短時間內完全消失無蹤。如此一來，若要輕易地確認某人的網路參考資料、驗證一份文件裡的引文、或是回溯原始資料等等，就會變成是一件幾乎不可能的事。除此之外，我們也可以很容易地複製、修改、或扭曲一份資料，然後再產生該份資料的各種不同版本。例如，某個有心人士可以擅自修改一段文字或是一張圖片，然後再建立一個新的網頁來散佈這些錯誤的資訊。此種情況引起了各種與著作權保護和來源資料的真實性有關的議題。

你必須花費時間及多加練習，才能瞭解網際網路、網路用語、以及辨識有價值的網路資訊的各種方法。目前幾乎沒有任何規則可幫助你在網路上找到那些包含有用的、及真實資訊的最佳網站。就進行研究的目的而言，來自於大學、研究機構、或政府機構的資料，其可信度通常會高於那些來路不明的個人首頁資料、或是由商業組織或支持某些政治抑或是社會議題的團體所贊助的網站的資料。除了可能會轉移網址或消失外，許多網頁與網路資料都沒有提供完整的引用資訊。品質較好的網路資料來源則會提供較爲完整的引用資訊，像是作者、日期、地點等等資訊。

社會研究中的倫理

我們現在要接著討論在你設計一項研究之前所必須處理的第二個重大議題。雖然如同你在第四章中所看到的，不同的研究取向會以不同的方式來處理價值議題，但社會研究仍有一個倫理道德的面向。所有的研究取向都承認研究具有倫理的面

向。直到你正式進行研究時，你才能完全地體會研究中所遭遇到的倫理困境。然而，等到研究進行一半時才去思考倫理問題，其實爲時已晚。當你在設計一項研究時，你必須準備並思考研究中的倫理議題。如此一來，你才能在研究設計中建立健全的倫理實踐作法。

倫理規範與其他的研究者都能夠提供相關的指引，但是合乎道德的行爲最終還是取決於個別的研究者。即使當研究參與者不瞭解、或不關心研究倫理時，你也有道德上及專業上的義務去遵守研究倫理。不可諱言的，確實有許多研究參與者幾乎完全不在意是否要保護自身的隱私，以及他人的權利。[4]

倫理議題意指在我們考量如何以適當的方式進行研究時，所遭遇的重要議題、兩難困境、以及衝突爭執。倫理界定了什麼是正當的、什麼是不正當的行爲，或是「合乎道德」的研究程序應該包含什麼作法。雖然幾乎沒有什麼倫理方面的絕對規範，但卻有許多大家都能認同的原則。不過，這些原則在實務上可能會產生一些衝突。許多的倫理議題都會要求你在下列兩種價值之間取得平衡：其中一個是對於科學知識的追求，另一個則是研究對象與社會上其他人所擁有的權利。你必須在潛在的利益（像是增進對於社會生活的瞭解、提升制訂決策的品質、或是幫助研究參與者等等）以及可能的代價（像是失去尊嚴、自尊、隱私、或民主自由等等）之間作出權衡與取捨。

進行研究的倫理標準，可能會比許多組織（例如，收款代理機構、警政部門、廣告商等等）裡的標準還要更爲嚴格。在進行專業的社會研究時，你必須瞭解適當的研究技術（例如，抽樣技術），並且敏銳地知覺到研究中的倫理議題。

個別研究者

倫理問題始於研究者，也結束於研究者。身爲研究者，你個人的道德信念是對抗違反倫理行爲的最佳防範措施。在進行一項研究之前、期間、及之後，你都有機會也*有義務*去反思你的研究行動，並要時時捫心自問是否心安理得。研究是否合乎倫理，取決於個別研究者的誠信正直和價值觀。「如果我們認真嚴肅地看待價值觀的問題，那麼價值觀不應該在被明確地表達之後，就被閒置在一旁。相反地，價值觀應該指引社會學家的行動。這些價值觀決定了什麼人應該被調查、爲了什麼目的、以及研究是爲誰服務」（Sagarin, 1973: 63）。

信守研究倫理的理由

既然大多數社會研究者皆真誠地關懷別人，那麼爲什麼還會有研究者做出不顧倫理道德的行爲呢？撇開少數心懷不軌的人，多數不合倫理的行爲，都是導因於研究者缺乏警覺性、以及承受壓力而不得不抄捷徑，因而踰越倫理道德的規範。許多研究者面對莫大的壓力，這些壓力包括建立事業、發表研究論文、提升知識、獲得聲望、讓家人和朋友留下深刻印象、以及保住工作等等。合乎倫理的研究要花費更多的時間才能完成、研究花費較大、更爲複雜、且更有可能在中途就被迫結束。除此之外，形諸於文字的倫理標準都是一些模糊不清的原則。在許多情況之下，研究

者都有可能會做出不合倫理的行爲，但是被逮到的機會卻微乎其微。

此外，也沒有人會因爲你遵守研究倫理並且正當行事而褒揚你。違反倫理的研究者如果被查獲，就必須面對來自公眾的羞辱、事業可能因而告終、還可能面臨訴訟。但是遵守倫理的研究者卻不會得到任何讚美。大部分的研究者在接受專業訓練的過程中、擔任專業角色時、以及在與其他研究者的私人接觸中，會逐漸地將倫理行爲內化。除此之外，科學社群的誠實與公開的規範，也強化了倫理行爲。那些真心想成爲專業研究者的人、認同科學精神的人、經常與認真的研究者互動的人，最有可能在行動時恪遵倫理規範。

科學的不當行為　研究社群反對科學的不當行爲，其中包括研究造假與抄襲。當一名研究者僞造、扭曲資料或蒐集資料的方法，抑或抄襲他人著作時，就犯了所謂**科學的不當行為**（scientific misconduct）。科學的不當行爲還包括在進行或報告研究時，嚴重地偏離科學社群裡被廣爲接受的慣例作法。研究機構及大學都訂立了相關的政策與程序，以偵測不當行爲、向科學社群和贊助機構舉發不當行爲、並處罰那些從事不當行爲的研究者（例如減薪或解聘等等）。[5]

研究造假（research fraud）所指的是研究者憑空捏造一些並非真正蒐集而來的資料、或是刻意以錯誤的方式說明研究的進行方式。雖然研究造假的情況並不多見，但卻要非常慎重地處理。最廣爲人知的研究造假案例，就是英國的教育心理學之父伯特爵士（Sir Cyril Burt）。當伯特於 1971 年過世時，是一位備受尊崇的研究者，這是因爲他以雙胞胎進行研究，並顯示出智力的基因基礎，因而聲名大噪。然而在 1976 年，他卻被人發現其研究資料以及共同研究者的姓名都是僞造的。不幸的是，當這起造假案件東窗事發時，整個科學社群已經被誤導了將近三十年之久。

抄襲（*plagiarism*）也是一種造假。當某位研究者剽竊了他人的想法或著作、或是在引用時未註明出處時，都算是抄襲。有一種特殊型式的抄襲，是剽竊另一位研究者、助理、或是學生的著作，並將之當作自己的著作發表。上述這些行爲都嚴重地違反了倫理標準。[6]我們將在第十六章更進一步地討論抄襲問題。

不合倫理但是合法　有些行爲可能是不合倫理的，但卻並未違法。此種存在於合法行爲以及合乎倫理行爲之間的差異，可用下面這個抄襲的案例來加以說明。美國的社會學協會曾經證實，一本由東新墨西哥大學（Eastern New Mexico University）的一位校長於 1988 年所出版的書籍，其書中大篇幅地引用了一份由塔夫茲大學（Tufts University）的一名社會學教授於 1978 年所寫的論文內容，但卻沒有任何的附註。此種複製的行爲並不算*違法*行爲，這是因爲該名社會學家的論文並沒有向美國政府申請著作權，因此這位校長的作法並未違反著作權法。話雖如此，根據專業行爲的標準，

科學的不當行為　某些從事研究造假、抄襲、或其他違反倫理行為的研究者所做出的行為，這些行為嚴重違反了科學社群在進行研究與報告研究方面所建立的慣例作法。

研究造假　一種違反倫理的行為類型，意即研究者憑空捏造或創造出錯誤的資料，抑或是以錯誤方式報告研究步驟。

圖 5.3　社會研究中法律及道德行為的分類

參考書目的完整引用資訊（資料來源文件夾）

法律層面	倫理層面	
	合乎倫理	*不合倫理*
合法	既合乎倫理又合法	合法但是不合倫理
不合法	不合法，但是合乎倫理	既不合倫理也不合法

這種行爲顯而易見地是不合乎倫理的。[7]（請參閱圖 5.3）

權力

研究者與研究參與者之間的關係，涉及到權力以及信任。相對於研究參與者以及助理，實驗者、調查主持人、或研究者，都是握有權力的一方。而此種權力關係和信任，是由證照、專業知識、訓練、以及科學在現代社會中的角色所賦予的。部分倫理方面的議題正是導因於濫用此種權力和信任所致。

研究者擁有進行研究的權力，但此種權力也伴隨著引導、保護、以及關心研究對象利益的相關責任。例如，一位醫生被發現在未徵得當事人同意的情況下，對三十三名婦女進行了實驗性的婦科手術。此種行爲不但是不道德的，也違反了他人對他的信任。這些婦女信任這位醫生，但是他卻濫用了病人、醫療社群、以及社會對他的信任。[8]

如果你想要尋求倫理方面的指引，你可以求助於下列這些資源：專業的同僚、倫理諮詢委員會、在大學或某些機構裡的機構審查委員會，或是以人群爲受試者的審查委員會、專業協會的倫理規章、及探討研究倫理的著作。

與研究參與者有關的倫理議題

你是否曾經在某項研究中擔任過研究參與者？如果有，你受到什麼樣的對待？與任何議題不同的是，研究倫理的相關討論，都把焦點放在對於研究參與者可能造成的副作用上。進行合乎倫理的研究，意謂著我們必須在增進知識與不擾亂他人生活這兩種價值之間取得平衡。如果研究參與者擁有不受干擾的絕對權利，那麼大部分的實徵研究就不可能進行。而假使研究者擁有進行探究的絕對權利，則可能會侵犯研究參與者的基本人權。因此，此處所要處理的道德議題是，在什麼時候（如果真的有的話）研究者可以有正當理由讓研究對象冒險、讓研究對象感到難堪、侵犯研究對象的隱私、或是對研究對象造成某種傷害呢？

法律及倫理規範已經訂出了一些明確的禁止事項，其中包括絕對不能對研究參與者造成不必要或不可逆的傷害；盡可能事先取得自願性的同意書；以及若非必要，

絕對不施以羞辱、貶低研究對象，或者散佈那些爲了研究目的而蒐集的，關於特定個人的有害資訊。這些都只是最低的要求標準，也都仍有不同的詮釋空間（例如，在特定的情境之下，什麼可以算是「不必要」的？）。

對研究參與者加以保護的起源　在越來越多以科學之名而違反基本人權的案例被揭露之後，社會大眾也開始關注研究受試者所受的待遇。違反基本人權最爲惡名昭彰的例子，就是德國納粹的研究者，在猶太人及其他人身上所進行的「醫學實驗」。在這些實驗中，研究科學家以進行科學研究的名義，對實驗對象施加了恐怖的酷刑。例如，有些人被放入冰冷的水中，看看多久之後才會死去；其他人則被蓄意地餓死；而一些兒童的四肢則被截下，並且移植到其他人身上。[9]

此種違反人權的案例，並不只發生在德國，而且也不只發生在過去。塔斯克基梅毒研究（Tuskegee Syphilis Study），也被稱作壞血事件（Bad Blood），就在納粹集中營被關閉將近三十年之後，於美國發生。一直到 1970 年代，因爲報紙的報導，才讓這樁醜聞爆發出來。事情起源於美國公共健康服務部所贊助的一項研究。在此項研究中，阿拉巴馬州內一群窮困而未受過教育的黑人罹患了梅毒，並因爲沒有接受治療而死於此疾病。然而，研究者卻藉此機會研究那些在罹患梅毒末期所出現的嚴重生理失能情況。雖然此項研究是在 1929 年開始的，而當時許多人還不知道可以用盤尼西林來治療梅毒。但是，在此種治療梅毒的方法已經普及許久之後，此項研究竟然還持續進行。雖然研究者們以不道德的方式來對待受試者，但他們卻仍然能夠持續發表研究結果，且時間長達四十年之久。此項研究最後終於在 1972 年結束了，但是美國總統一直到 1997 年才承認該項錯誤，並且對於那些研究中的參與者與受害者正式道歉。[10]

不幸的是，壞血事件的醜聞並非特例。在冷戰期間，美國政府經常爲了軍事及政治的目的，在研究倫理的原則方面讓步妥協。在 1995 年，有一些報告指出，美國政府曾經在 1940 年代晚期，批准對不知情的人們注射放射性物質。而在 1950 年代，美國政府曾經警告過柯達（Eastman Kodak）和其他的軟片製造商關於原子彈試爆的核子落塵所可能造成的影響，以避免發生軟片成像模糊的情況。然而，美國政府卻未針對居住在試爆區域附近的民眾提出健康方面的警告。1960 年代時，美國軍方提供 LSD（一種迷幻藥）給不知情的士兵，因而造成嚴重的精神性傷害。時至今日，上述這些行爲都被廣泛地認爲違反了兩項最基本的倫理原則：避免生理傷害，以及獲得知情後的同意。[11]

生理傷害、心理虐待、以及觸法風險　社會研究可能在生理方面、心理方面、法律方面、以及經濟方面對研究參與者造成傷害，並因而影響個人的事業或是收入。即使是在生物醫學研究這種介入情形較多的研究裡，生理傷害也是相當罕見的。在不同類型的研究中（例如，在實驗研究及田野研究中），所可能產生的特定傷害類型也都各自不同。研究者必須清楚地知道所有類型的傷害，並時時刻刻努力地把各種傷害減少到最低程度。[12]

生理傷害 核心的倫理原則之一，就是研究者絕對不能對研究參與者造成任何生理傷害，這意味著我們在開始進行一項研究之前，必須先考量可能的風險，其中包括基本的安全注意事項（安全的建築物、傢俱、以及設備等等）。如果研究牽涉到壓力，而我們預期研究參與者或助理有可能會受傷，甚至是有遭受攻擊的危險時，我們就必須事先進行篩選，並且排除高危險性的受試者（像是患有心臟疾病者、患有精神疾病者、或是罹患癲癇者）。此外，如果有任何人因爲參與研究的關係而受到傷害，我們應該要負起道德上與法律上的責任。而這也意味著，倘若我們無法保證研究受試者的人身安全，我們就必須立刻中止研究（請參閱在範例 5.4 倫理爭議的三個案例裡的辛巴度研究）。

心理虐待、壓力、或是失去自尊 雖然造成生理傷害的風險並不多見，社會研究者卻有可能讓人們置身於充滿壓力的、令人難堪的、造成焦慮的、或令人不愉快的情境中。爲了要瞭解人們在真實生活的、造成高度焦慮的情境中會如何反應，社會研究者會讓研究參與者置身於造成心理不適或壓力的真實情境中。知名的米爾葛蘭（Milgram）服從研究中所涉及的倫理議題，目前仍然引起許多爭議（請參閱範例 5.4）。有些人認爲，米爾葛蘭所採取的預防措施、與此項研究所獲得的知識，其價值高過於研究參與者所經歷的壓力以及可能的心理傷害。而其他人則相信，極端壓力以及造成永久傷害的風險，是遠遠超過一般人所能承受的。

有些研究者曾經讓研究參與者感受到高度的焦慮或不適感，比方像讓研究參與者觀看一些駭人的照片；謊稱男學生擁有強烈的女性人格特質；欺騙學生說他們被當掉了；創造一個令人高度恐懼的情境（例如，讓煙霧飄進一個門被鎖上的房間）；要求研究參與者去傷害其他人；讓人們置身於一個因爲社會壓力而必須否認自己信念的情境；以及強迫研究參與者說謊、欺騙、或是偷竊等等。[13] 研究助人行爲的研究者，可能會讓研究參與者置身於緊急的情境中，以觀察他們是否會協助「受害者們」。例如，匹利文以及同事（Piliavin and associates, 1969）就曾經利用某人突然倒地不起的騙人情境，來研究地鐵內的助人行爲。而在這個田野實驗中，地鐵的乘客們都不知道這是一個實驗，也皆非自願參與此項實驗。

敏感的研究者也會知覺到對於研究參與者的自尊所可能造成的傷害。例如，沃斯特爾（Walster, 1965）曾意欲觀察，女性自我價值的感受是否會影響到她們在戀愛方面的偏好。她先對一群女大學生實施人格測驗，然後給予假的回饋。她告訴一些研究參與者說她們缺乏想像力以及創造力，接著，有一位假裝也是研究參與者的帥氣男研究生，開始與這些女學生展開交談。這位男研究生裝作對其中一位女生非常有興趣，並約她一起出去吃晚餐。沃斯特爾想要測量這位女學生對於這位男研究生的浪漫吸引力。而在實驗結束之後，這位女學生才知道根本沒有晚餐約會，而那位男研究生也只是在進行實驗而已，並不是真的對她動心。雖然研究者後來對所有的研究參與者都進行了說明，這些研究參與者仍可能會有自尊心受損的感受，也可能會承受心理上的傷害。[14]

範例　5.4

倫理爭議的三個案例

米爾葛蘭的「服從研究」（obedience study）（Milgram, 1963, 1965, 1974）企圖透過檢視服從權力的社會壓力強度，來發現在納粹統治下的恐怖大屠殺是如何發生的。在簽署了「知情同意表」之後，受試者經過被操縱的選擇程序而被指派為「老師」，而實驗者的同謀則擔任「學生」。老師要負責測試學生對於單字表的記憶，並在學生犯錯時增加電擊的強度。學生被安置在鄰近的房間，所以老師可以聽得見學生，但卻不能看到學生。電擊設備清楚地標示出逐漸增加的電壓值。當學生犯錯且老師打開電擊開關時，學生會發出哀叫聲，就像是承受了巨大的痛苦。研究者也會待在現場，還會對老師提出一些評論，像是「你必須繼續下去」之類的。根據米爾葛蘭的報告，「受試者被觀察到會冒汗、顫抖、結巴、咬嘴唇、呻吟、並用指甲掐自己的肉。這些都是本實驗中的典型反應，而不是異常的反應」（Milgram, 1963: 375）。受到震驚的程度到達危險地步的受試者比例，遠遠超過研究者原本預期的比例。對於本項研究中所使用的欺騙行為、以及受試者所經驗的極端情緒壓力，都引起了社會大眾對於倫理問題的關切。

在韓福瑞（Laud Humphreys）的「茶室交易研究」（tearoom trade study）（一項關於男同性戀者在公共廁所裡進行性行為的研究）裡，韓福瑞假裝是一個「把風皇后」（watchqueen，又可譯為「放哨的兔仔」，意指在警察接近時發出警告的人，在此是指研究者一邊進行窺探、一邊幫忙把風），趁機觀察了將近一百位男性從事性行為的過程。當受試者回到自己的車上時，韓福瑞也在其後跟隨他們，並秘密地記下受試者的車牌號碼。之後，韓福瑞又喬裝成一名市場研究者，順利地從警方的登記資料那兒得到受試者的姓名與住址。一年之後，在偽裝的情況之下，韓福瑞利用一個關於健康調查的騙人故事，在受試者的家中對他們進行了訪談。韓福瑞小心翼翼地把受試者的姓名收藏在保險箱中，並將所有能夠辨識出受試者姓名的資料通通銷毀。他大幅度地增進了我們對於那些經常出入於「茶室」的同性戀者的知識，也顛覆了過去對於這些同性戀者的錯誤想法。然而，這項研究亦不斷受到爭議，其中包括受試者從未同意參與這項研究；研究者使用了欺騙的手法；受試者名單可能會被用來敲詐勒索受試者、害他們失去婚姻、或告發他們的犯罪行為等等。

而在「辛巴度的監獄實驗」（Zimbardo prison experiment）中（Zimbardo, 192, 1973; Zimbardo et al., 1973, 1974），男學生被分成兩組扮演不同的角色：警衛與囚犯。在實驗開始之前，研究者先針對自願參與實驗的學生實施人格測驗，只有屬於「正常」範圍內的學生才能被挑選出來加入實驗。自願者簽署參加為期兩週的實驗，扮演囚犯的學生則被告知他們會受到監視，而且也會被暫時剝奪一些公民權利，但是不會受到任何身體方面的傷害。在設於史丹福大學一棟大樓地下室的模擬監獄裡，囚犯被去除了任何個人化的特徵（一律穿著標準的制服，並且只用號碼來稱呼他們），而警衛則進行軍事化管理（穿著制服、配備警棍以及反光太陽眼鏡）。警衛被告知必須維持合理程度的秩序，並採取八小時的輪班制，至於囚犯則一天二十四小時都被關在監獄裡。出乎意料地，這些自願者後來竟然變得過度投入於自己的角色之中。囚犯變得被動而混亂，警衛則變得具有攻擊性、獨斷專制、而且失去人性。到了第六天時，辛巴度基於倫理的考量，取消了這個實驗。因為如果繼續下去，對於受試者造成永久心理傷害、甚至是生理傷害的風險實在是太大了。

只有當研究者能事先對可能引起的焦慮或不適感做好預防措施、而且擁有豐富經驗時，才可以考慮進行一些可能會引起壓力或是焦慮的研究。在規劃此種研究時，研究者應該要先請教那些曾經進行過類似研究的研究者，以及心理健康方面的專家、篩選掉高風險的群體（例如，那些有情緒問題的人、或心臟衰弱的人）、並安排

好在出現危險情境時的緊急介入措施，抑或是終止研究的方式。研究者都必須在研究之前取得充分知情的同意（稍後將再進一步討論），並在取得同意之後，立刻向受試者進行詳細的說明。

核心倫理原則之一，就是研究者絕對不能在研究參與者身上製造出*非必要的*壓力。所謂非必要的，意指產生超出預期效果所需要的最低限量，或那些與研究目的無直接關係的、不合理的壓力。對於最低限量的認識與瞭解，其實是來自於反覆嘗試的豐富經驗。研究者展開實驗時，最好能先從過小的壓力開始，寧可在發現缺乏效果之後再來修正，也不要一開始便製造過高的壓力。如果壓力的程度可能造成長期效果，研究者應該要進行後續的追蹤，並免費提供心理諮商。另外一個危險，則是研究者對於研究參與者的態度，可能會變得冷酷無情或是具有操縱性。如果所進行的實驗會對受試者造成心理方面的傷害，研究者之後會感到內疚及後悔。當所進行的實驗會讓研究參與者置身於焦慮情境中時，有道德感的研究者通常都會覺得過意不去。

法律傷害　身爲研究者，我們有責任保護研究參與者免於增加被逮捕的風險。如果參與一項研究可能會增加研究參與者被逮捕的風險，那麼這樣的情況將會破壞人們對於社會科學研究的信任，造成未來的研究參與者不再願意參與任何研究。進行某些類型的研究之前，研究者或許可以先取得某些執法單位的豁免許可。例如，美國司法部有時候會提供書面的豁免證明，給予研究犯罪行爲的研究者。然而，正如同本章開頭部分的幫派研究所指出的，這些措施對於研究者的保護其實相當有限，研究者本身同樣也必須極爲謹慎小心。

對於韓福瑞在 1975 年所進行的「茶室交易研究」（請參閱範例 5.4）的批評之一，就是其中存在可能的法律傷害。而在紐澤西的負所得稅實驗（*New Jersey Negative Income Tax Experiment*）中，有些研究參與者接到了收入補貼。然而，研究者們並未監督這些參與者是否也接到了公共福利的支票。一名當地的檢察官要求研究者交出參與者的資料，以確認是否有「福利詐騙」的情形。換句話說，研究參與者由於加入了這個實驗而面臨觸法的風險。雖然這個衝突到最後終於解決，但這件事也說明了研究者在設計一項研究時，必須注意到潛在的法律議題。

當研究者在蒐集資料時發現了非法行動，就會產生另外一個相關的倫理議題。此時我們就必須衡量下列兩者之間孰輕孰重：其一是維護研究者與受試者之間的關係、以及對於未來的研究者所可能帶來的利益；另外一個則是保護無辜人士、避免他們遭受到任何傷害。例如，在他對於警察所進行的田野研究中，馬能（Van Maanen, 1982: 114-115）提及他曾經看到警察毆打民眾，並親眼目睹警察從事非法行爲，與採取不當的程序。但他卻說：「關於這些惱人事件及在事件發生過後……我遵從了警察的慣例：保持緘默。」

田野研究者經常會面臨困難的倫理抉擇。例如，在研究一所精神治療機構時，泰勒（Taylor, 1987）發現在該機構裡的員工會虐待及辱罵病患。他有兩個選擇：放棄研究、要求警方介入調查；或是保持沉默，並且繼續進行數個月的研究、對外發

表他的研究發現、然後發起終止虐待行爲的運動。在衡量了整個狀況後，他選擇了第二條路，而他現在則是成爲一名爲精神治療機構的病患爭取權利的行動份子。

在針對紐約市的餐廳所進行的一項研究中，保護機密資訊的議題使得類似的倫理兩難困境變得更爲複雜。當時有一名社會學研究生正在對服務生進行參與觀察的研究。在研究過程中，研究的現場，也就是一家餐廳，被大火燒毀了，而且警方懷疑火災的起因是人爲縱火。當地的司法單位要求這名研究者交出田野筆記，並打算要向他質詢有關餐廳裡的活動。這名研究者有兩個選擇：與調查單位合作，違反參與者對他的信任、保密、以及基本的研究倫理；或者，他也可以堅守機密並且遵守研究倫理，但是就必須面臨藐視法庭以及阻礙司法的處分，包括罰款以及坐牢服刑。他想要遵守研究倫理，卻又不想承受牢獄之災。在多年的法律纏訟之後，整個僵局最後獲得解決，也就是研究者有限度地配合調查，而司決裁決也維護了田野筆記的機密性。然而，整起事件花了好幾年的時間才得以解決，而這名研究者也承受了巨大的財務與個人損失。[15]

觀察非法行爲也可能是一項研究計畫中的重要主題。如果研究者暗中秘密地進行觀察並記錄下非法行爲，然後把相關資料提供給執法單位，那麼就違反了關於研究參與者的倫理規範，也不利於未來的研究。但是如果研究者不向警方舉發非法行爲，就等同於間接地縱容犯罪行爲發生，也可能被當成共犯而遭到起訴。研究者究竟是一個追求知識的專家、還是一名獨立的秘密告發者呢？

對於研究參與者的其他傷害　研究參與者亦可能面臨其他類型的風險。例如，一項要求受訪者去回憶不愉快事件的調查訪談，可能會對受訪者造成焦慮不安及不舒服的感受。我們必須警覺到任何對研究參與者可能造成的傷害、考量可能的預防措施、並在潛在的傷害以及可能的效益之間權衡輕重。研究參與者有可能因爲參與了某項研究，因而使他們的事業或收入受到負面的影響。例如，假設有一個研究者對員工進行了調查，而其結論則是管理人員的表現不佳。由於研究者公開發表了這項研究結果，因而使得這名管理人員被解聘了。或者，假設有一位研究者的研究對象是那些接受社會福利維生的人。結果，由於該項研究的發現，使得其中某些人失去獲得社會福利的資格，生活品質也因此下降。在這樣的案例中，研究者所該負起的責任是什麼？我們必須考量到研究結果對於研究對象可能造成的影響與後果。然而，這些問題其實並無固定或標準的答案。我們必須仔細評估每個不同的案例、審慎權衡可能造成的傷害以及潛在的益處、承擔起各種決定的責任。

欺騙　是否曾經有人爲了要讓你去做某些事而告訴你一件半真半假的事、或是對你說謊？當時你的感受如何？在社會研究中有一條非常重要的倫理原則，那就是**自願同意原則**（principle of voluntary consent），也就是絕對不強迫任何人參與研究。另外一條相關的倫理原則，則是不要對研究參與者說謊，除非是爲了合法正當的研究理由，而不得不如此做時。另外一項非常重要的倫理規範，則是研究參與者應該要明確

自願同意原則　一項倫理原則，也就是除非人們明確地而且自由地同意加入一項研究，否則他們就不該參與該項研究。

地同意參與一項研究。當我們使用了欺騙的手段、將研究加以僞裝、或使用隱匿的研究方法時，個人不參與研究的權利，就會成爲一個非常關鍵的議題。[16]

社會研究者有時會在田野研究或實驗研究中欺騙參與者、或是對參與者說謊。我們必須是爲了正當之方法論方面的理由，才能扭曲、隱藏我們的行動或是真正的意圖。也就是說，如果參與者知道了真正的目的，他們有可能會修正自己的行爲，使得我們無法瞭解他們真正行爲的時候；或是如果研究者說出實情，可能無法進入研究現場時，研究者才能欺騙或說謊。但是，如果我們能在不使用欺騙手段的情況下完成同樣的事，就絕對不可採用欺騙的方法。

只有在基於明確方法論之目的下，欺騙才能被接受。但即使如此，我們也不能超越必要的最低程度。如果我們使用了欺騙，我們應該要取得知情之後的同意，絕對不能向參與者謊報研究中的風險，並應該在研究結束之後，立刻向參與者進行詳盡的說明。我們可以描述研究中的基本程序，而只隱瞞與研究相關的部分資訊。

知情的同意 另外一條基本的倫理原則是：絕對不能強迫任何人參與研究，所有的參與行爲都*必須*是自願的。光是取得參與者的同意是不足夠的，人們還必須知道他們所參與的研究是什麼。唯有如此，人們才能作出知情之後的決定。當研究參與者閱讀並簽署一份**知情的同意**（informed consent）文件時（此種文件是一種同意參與研究的書面證明，表示人們已經瞭解研究程序的一些基本細節），他們就能夠更加瞭解他們自身的權益、以及他們所參與的是什麼樣的研究。

美國聯邦政府並未強制要求所有以人類爲受試者的研究，都必須取得知情的同意。然而，研究者還是應該取得書面同意書，除非是在經過**機構審查委員會**（institutional review board, IRB）的判斷之後（請參閱之後討論機構審查委員會的部分），認爲研究者有充分的理由可以不必這麼做時（例如，隱匿的田野研究、使用次級資料等等），才可以省略這個步驟。

知情的同意書必須提供明確的相關資訊（請參閱資訊箱 5.4 知情的同意）。[17] 一般而言，在知情的同意書中，只要概略地說明研究程序或是研究問題，以及使用資料的方式，就已經足夠了。在由辛格（Singer, 1978）所進行的一項研究中，有一組隨機挑選之調查研究的受試者，收到一份詳細的知情同意書，而另外一組則沒有收到這份同意書。她發現到在兩組受試者之間的回覆率方面，並沒有顯著的差異。如果真要說有什麼差異的話，就是那些拒絕簽署同意書的人們比較可能會用猜測的方式來回答問題、或者乾脆回答「沒意見」。而根據她們對文獻所進行的分析，辛格與同事（Singer and colleagues, 1995）發現，在問到高度敏感的主題時，若是研究者向受試者保證資料的機密性，則能夠適度地提高受試者回答問題的意願。而

知情的同意 一項通常是書面的聲明，向研究參與者解釋一項研究的相關細節，並在研究開始之前，請求參與者們自願加入研究的同意書。

機構審查委員會 美國聯邦法律要求在大學、醫院、以及研究機構裡所必須設置的一種委員會，以確保那些以人類為受試者的研究能以一種負責任的、合乎倫理的方式來進行。而在研究開始進行之前，該委員會負責事先檢視研究的各種細節。

資訊箱 5.4　知情的同意

知情的同意書包含了下列數項內容：

1. 關於研究目的與研究程序的簡短描述，其中包括研究的預期進行時間。
2. 對於參與研究可能產生的任何風險，或是不適感的說明。
3. 對於匿名隱私以及記錄保密性的保證。
4. 說明研究者的身份，說明何處可以取得關於研究參與者的權利以及關於研究的各種問題的資訊。
5. 明確地說明參與研究完全出於自願，參與者可以隨時退出研究，而不必承擔任何罰則。
6. 說明可能會使用的其他替代研究程序。
7. 說明參與者能夠得到的任何報酬或是補償，以及研究中的參與者數目。
8. 同意提供一份研究發現的摘要給研究參與者。

在其他的情況下，不論研究者再怎麼保證資料的機密性，也不會影響受試者回答問題的方式，或回答問題的意願。

對於大部分的調查研究、田野研究、以及次級資料研究而言，參與者簽署知情同意書並不是必要的。但在實驗研究中，通常都會要求受試者必須先簽署知情同意書。至於在文件研究及電話訪談中，則不可能取得知情的同意書。一般的原則是，造成可能傷害的風險越大，就越需要徵得書面同意書。總結而言，有許多理由支持研究者去徵得知情後的同意，但反對此種作法的理由卻寥寥可數。

隱密觀察　在調查研究與實驗研究中，徵得知情的同意或許是相當簡單的事。但是有一些田野研究者認爲，當研究者所觀察的是真實生活的田野環境時，徵得知情的同意是比較不適合的作法。此外，這些田野研究者也表示，除非他們進行隱密觀察，否則，他們根本無法進入田野現場或進行田野研究。在過去，田野研究者會使用隱密觀察的技術，例如像是假裝自己有酗酒問題，然後加入一個協助人們戒酒的團體。如此一來，他們才能夠研究這個團體。田野研究者有三項選擇，這些選擇都模糊了存在於知情的同意及並非完全知情的默許間的界線。范恩（Fine, 1980）借用了情報工作的術語，並指出三種不同的僞裝方式：深度僞裝（deep cover）（研究者絕口不提與研究相關的事情，並在行動上表現得和參與者幾乎完全一模一樣）、輕度僞裝（shallow cover）（研究者讓參與者們知道研究正在進行當中，但對於研究細節則語帶保留）、以及明白的僞裝（explicit cover）（研究者清楚而完整地揭露自己的目的，並請求參與者的允許）。

有些人支持隱密觀察，並認爲田野研究應該可以免除取得知情同意的義務（Herrera, 1999）。而其理由之一，即是在田野研究中，知情的同意不切實際、而且也可能擾亂研究的過程。除此之外，由於取得知情的同意可能會擾亂原本正在進行中的活動，因而妨礙參與者或研究現場，也因此而造成傷害。此種推理方式的錯誤之處在於，確保參與者尊嚴的道德原則，其重要性應該要勝過研究者在實際上的方

便或是權宜之策。換句話說，此種推理方式的出發點是自私的，認爲進行研究的價值高於維護誠信和隱私的價值。此種推理方式假定研究者比參與者更有能力判斷研究中的風險。然而，道德及倫理規範都告誡我們，必須尊重研究對象的自由與自主權，且須讓他們自己作出決定。一旦研究參與者知道，研究者爲了研究的目的而在未經授權同意的情況下，侵犯了他們的「隱私」時，他們或許就不會保持自然的狀態，而且也可能覺得不舒服。

另外一個支持隱密觀察的理由，則是在人們的溝通及日常事務中，充滿了各種隱密的活動。日常的活動都會涉及到某些數量的隱密活動，其中有許多的「人群觀察者」（"people watchers"）與沒有惡意的竊聽者。在日常生活中，有許多的零售商店、執法單位、或安全人員等等，都會做出隱密的及欺騙的行爲，這樣的行爲相當普遍，而人們也幾乎都會預期遇到這類行爲。既然此行爲是在人們預期之中、而且也不會造成傷害，那麼社會研究者爲什麼必須要與眾不同呢？像是「每個人都這麼做」以及「反正這種事遲早會發生」之類的藉口，並不能當作讓自己免於某些義務的合理說辭。此處的真正議題，是進行研究的道德倫理規範。或許在某些社會裡，偷窺、監視、及利用秘密告發者的情況，正在與日俱增當中，但難道此種情況能使這些行爲在道德方面被合理化爲正當行爲、並且也能確保個人的隱私嗎？我們應該把這些行爲當作在社會研究中確保誠信與信任的模範嗎？隱密監視案件不斷增加，可能會使民眾懷疑、不信任、不願意合作的情況日漸嚴重。缺乏知情的同意，是倫理上的灰色地帶，而且許多人也相信，與事先徵得知情同意的情況相較下，若研究者不先徵得知情同意，那麼此種作法在道德倫理方面所遭遇的風險，極有可能造成更大的傷害。

關於隱密研究這個議題，目前仍處於爭議不斷的狀況，亦有許多研究者認爲所有的隱密研究都是不道德的。[18] 美國人類學協會的倫理規範譴責隱密研究是「不切實際且不受歡迎的」作法。有些人則認爲在某些特定情況之下，可以接受隱密研究是合乎倫理的作法。然而，即使是這些人也主張只有在無法進行公開觀察之情況下，才能使用隱密觀察。除此之外，如果非得要進行隱密研究，我們也應該在事後告知研究參與者，並且給予他們表達關切事項的機會。

欺騙與隱密研究可能會增加大眾對於社會研究的不信任以及懷疑，也會傷害大眾對於社會研究的尊重。田野研究中的僞裝或是欺騙的行爲，就好比在專制和極權社會裡的秘密警察，或是告密者。對於經常成爲研究對象的人們而言，欺騙的行爲會增加他們的不信任感。在下面這個案例，欺騙使得人們減少幫助他人的行爲。1973 年時，曾有一名學生在華盛頓大學西雅圖校區遭到槍擊，但是在校園內來來往往的學生們，竟然沒有人試圖去幫助這名學生。後來才發現，許多圍觀者之所以沒有伸出援手，並不是因爲他們沒有同情心，而是因爲他們認爲這場槍擊案件只不過是一項實驗的一部分。所以他們認爲那只是演戲、而不是真實的。[19]

特殊族群以及新的不平等

特殊族群以及強迫 有些族群或是研究參與者無法給予真正的、自願性的知情後

的同意。**特殊族群**（special populations）可能缺乏必要的能力、或可能受到間接的脅迫。學生、囚犯、員工、軍人、無家可歸的遊民、接受社會福利救濟者、兒童、或心智障礙者可能會同意參與研究，但他們可能沒有充分的能力來作出決定。或因爲他們想要得到一些好處（例如較高的成績、提早假釋、升遷、或額外的服務等等），所以才不得不同意參與研究。

以「無行爲能力」者（如兒童、心智障礙者等等）作爲研究對象是不道德的，除非我們符合下列兩種條件：獲得法定監護人的書面同意，以及我們遵守所有不傷害研究對象的倫理原則。例如，我們想要進行一項調查，主題是高中學生之間的吸煙與藥物／酒精濫用的問題。假如這項研究是在學校裡進行，就必須獲得校方的同意。此外，我們也必須取得所有未成年參與者的家長同意書才行。同樣的，我們最好也能徵得每一位學生的同意才進行研究。

強迫人們參與研究是不道德的，其中也包括提供一些他們無法以其他方式取得的特殊利益，當作誘使他們參與研究的條件。例如，指揮官命令士兵加入研究、教授要求學生必須成爲研究受試者才能及格、以及老闆要求員工參與一項調查以作爲繼續雇用的條件等等，都是不道德的行爲。即使由研究者以外的人（例如，老闆）來強迫人們（例如，員工）參與研究，也都是不合乎道德的行爲。

要決定受試者是否因爲受到強迫而參與研究，是一個相當複雜的議題，而我們必須審慎地評估每一個案例。例如，有一個研究者對一個被判刑的犯人提供了兩個選擇：繼續坐牢、或加入一個實驗性的改過遷善計畫。這名被判刑的犯人或許並不相信這個計畫會有什麼成效或益處，但是研究者卻深信這個計畫將可以幫助這名犯人。這就是一個牽涉到強迫的案例。但這名研究者必須判斷，這項研究對於受試者與社會的益處，是否高過於禁止強迫行爲的倫理原則。

教授有時候也會要求修社會科學課程的學生，必須加入研究計畫。這是一種特殊的強迫案例。已經有人提出三種論點，來支持此種要求學生加入研究計畫的作法：（1）若要以其他方法來找到研究參與者，將會相當困難、而費用也會高到非常嚇人的地步；（2）以學生充當參與者的研究所創造出來的知識，對於將來的學生和社會都有益處；以及（3）由於學生直接體驗到真實的研究情境，因此將可學習更多關於研究的知識。上述的三種論點中，只有第三個論點可以用來支持有限度的強迫。然而，只有當強迫的行爲具有明確的教育目標、學生可以從中得到關於研究的經驗、以及沒有違反任何其他的倫理原則時，強迫的行爲才可以被接受。[20]

製造新的不平等　另外一種可能的傷害，則是發生在當某一群人們因爲參與了一個研究，因而無法得到某項服務或利益時。例如，假設你現在發現了一種新的治療方式，可以治療一種像是愛滋病（acquired immune deficiency syndrome, AIDS）（又稱爲後天免疫缺乏症候群）之類的可怕疾病。爲了要瞭解新療法的效果如何，你刻意讓某些人接受新

特殊族群　特殊種類的研究參與者，這些參與者可能由於年紀、監禁、潛在的強迫，或因為未具有完整的生理、心理、情緒、以及其他的能力，因之可能缺乏完整的自由或是警覺性，來對參與某項研究給予自願的同意。

療法，而其他人則只有施予安慰劑或是無任何作用的藥丸。這項研究是被設計以證明這種新藥是否具有療效，但是那些被施予安慰劑的參與者，卻有可能因而死亡。當然，在研究者對此種新藥的療效有更多的認識之前，服用新藥的參與者也有可能會死亡。對於那些在研究中被隨機分配到控制組的人而言，我們爲了要瞭解更多關於新藥的療效、因此不給予他們可能拯救生命的治療方法，這樣的作法合乎倫理嗎？

我們可以用下列三種方式，減少研究參與者之間新的不平等狀況。首先，那些並未接受「新的、改良過的」治療方式的參與者，可以持續接受到目前爲止最佳的、被廣爲接受的治療方式。換句話說，所有的參與者都可以獲得必要的協助，而且每個人都可以獲得在新的療法進行測試之前的最佳療法。如此一來，就能確保沒有人會因而處於絕對不利的狀況，只是暫時地處於相對不利的狀況。第二，我們可以使用**交叉設計**（crossover designs）的方式，也就是在研究中第一階段裡的控制組（也就是沒有接受新療法的那些參與者），在第二階段中成爲實驗組並接受新的療法。而原先在第一階段中的實驗組，則在第二階段中成爲控制組。最後，我們必須仔細謹慎地、持續不斷地監控結果。如果在研究的早期階段就發現新的療法具有高度的效果，我們就應該讓所有人都接受此種新的療法。除此之外，在與醫藥治療有關、或是可能造成生理傷害的高風險研究中，研究者或許可以改用動物或其他的替代物，取代人類以進行研究。

隱私、匿名性、以及保密 如果在你不知情的情況下，你個人生活的私密細節被公布周知，你會有什麼感受？由於社會研究者會爲了研究社會行爲而侵犯研究參與者的隱私，因此他們必須採取一些預防措施，以保護研究參與者的個人隱私。

隱私 當調查研究者以一種會揭露私密細節的方式去探究信念、背景、以及行爲時，便會侵犯個人的隱私。實驗研究者有時會使用雙面透視鏡，或是隱藏的麥克風來「暗中監視」研究參與者。即使人們已經被告知他們正在接受研究，他們也不清楚實驗者所要觀察的是什麼。田野研究者可能會去觀察人們行爲中極爲私密的面向，或是偷聽人們的對話。在田野實驗及民族誌田野研究中，研究者可能在未事先提出警告的情況下，就侵犯了他人的隱私。當韓福瑞（Humphreys, 1975）在同性戀者進行性行爲的公共廁所外面充當「把風皇后」時，他在沒有知會這些研究參與者的情況下，觀察到了極爲私密的行爲。而當匹利文以及同事（Piliavin and colleagues, 1969）故意讓人們假裝在地鐵中昏倒、以研究助人行爲時，對於那些在地鐵車廂裡的人們而言，他們搭乘交通工具的隱私就受到了侵犯。人們經常會在公共場所（例如，在候車室裡、走在街道上時、在教室裡等等）成爲被研究的對象，但是有些「公共」場所比其他的場所更爲私密（例如，想想看研究者利用潛望鏡觀察那些以爲只有自己一人獨自在公廁裡的人們）。[21]

重視倫理的研究者，只在必要的最低程度內、

交叉設計 一種實驗研究中的研究設計類型，在此種研究設計當中，所有的組別都會接受相同的實驗處遇，但是卻是在不同的時間點接受這些實驗處遇。如此一來，所有的受試者都會遇到相同的不適感以及益處，而且也不會造成不平等的情況。

只爲了正當的研究理由，才會去侵犯他人的隱私。除此之外，重視倫理的研究者也會保護研究參與者的相關資訊，而不會把這些資訊公諸於大眾。

在某些情境中，隱私是受到法律保障的。曾經因爲發生一件侵犯隱私的案例，因而促使國會通過了一條聯邦法律。在 1954 年的「維奇塔陪審團研究」中（*Wichita Jury Study*），芝加哥大學法律學院的研究者記錄下陪審團的討論過程，以檢視陪審團審議的團體歷程。雖然該項研究的發現相當重要，而研究者也採取了預防措施，但仍引來國會介入調查。並於 1956 年通過一條法律，明令禁止爲了任何目的而對大小陪審團進行「竊聽、窺探」（"bugging"），即使在獲得陪審團員同意的情況下也不例外。[22]

匿名　在蒐集了相關資訊之後，研究者以不揭露參與者身份的方式來保護其隱私，而這種作法有二種形式：匿名性與保密。**匿名性**（anonymity）意指使研究參與者維持匿名或是沒有名字的狀態。例如，一位田野研究者提供了關於一名特定個人的社會背景，但是卻使用虛構的姓名以及地址，並更改了一些個人的特徵。如此一來，這個人的身份就受到保護，也沒有人知道這個人的名字、或這個人是誰。調查和實驗研究者會盡快地銷毀研究參與者的姓名或住址，並只以編號來指稱這些參與者，以維持匿名性。如果研究者在使用郵寄調查時，在問卷上加上編號以得知有哪些人沒有回覆問卷，那麼受訪者的隱匿性就沒有受到完整的保護。而在同樣本多次研究中，由於相同的個人會被持續地進行追蹤及研究，因此較不可能達到匿名性。相同地，歷史研究者也會在歷史研究或是文件研究中，使用具體的姓名。如果原始的資訊是來自於公開來源，他們或許可以這麼做，但是如果原本的資訊來源是不對外公開的，他們就必須取得文件擁有者的書面同意，才能使用具體的姓名。

要維護研究參與者的匿名性是一件困難的事。在《大社會中的小城鎮》（*Small Town in Mass Society*）（Vidich and Bensman, 1968）這本書中，有一項關於一個小鎮「春之鎮」（"Springdale"）的研究，雖然這個名稱是虛構的，但是人們仍然可以輕易地辨識出這座小鎮以及其中的特定人物。結果，鎮上的居民對於研究者描述他們的方式感到不滿，並發動一場遊行來嘲諷研究者。而在印第安那州曼西區（Muncie, Indiana）的著名「中鎮」（Middletown）研究中，人們也在其中辨認出自己所居住的小鎮。當研究者利用虛構的資訊來保護個人身份時，卻往往會在真正研究的事物，及向他人報告的資訊之間造成落差。而此種情形也引起人們質疑到底何者才是研究者的真正發現、而何者又是研究者所捏造的。

匿名性　讓研究參與者保持匿名狀態的倫理保護措施，也就是參與者的身分受到保護，因此不會被揭露出來，也不會被其他人所得知。

保密　一種基於倫理方面的考量，而對研究對象所應該給予的保護措施，其方法是秘密地保管研究資料，或不把研究資料公諸大眾；或是在公布資訊時，不讓其他人將特定個人與特定的回應連結在一起。研究者可利用只以集體方式呈現之資料的方式（例如，百分比、平均數等等），以達到保密的目的。

保密　即使我們無法在研究中維持匿名性時，我們也應該盡可能保密。匿名性所指的是保護特定個人的身份免於被外界知曉，而**保密**（confidentiality）的意思是，雖然我們可能在資訊中附上姓名，但是

我們必須秘密地保管這些資訊，或不把這些資訊向大眾公開。在公布資料時，我們絕對不能讓人們把特定的個人和這些資訊連結在一起。我們只能以集體的形式（例如，百分比、平均數等等）公開呈現研究結果。

雖然匿名性和保密兩者通常一起出現，我們仍然可以做到匿名但是不保密，或是保密但不匿名。當我們公布了關於一個特定個人的細節，卻保留這個人的姓名以及一些或許可以辨識出這個人的關鍵細節時，就是做到匿名但是不保密。而如果我們不公開個人的資料，但是私底下把個人的姓名和一些特定個人的資料連結在一起時，就是做到保密但不匿名。

研究者已經發展出詳盡細緻的程序，以保護研究參與者的身分不會被公諸於世。像是以匿名的方式提出受試者對於各種問題的答覆；請一位不屬於研究者及參與者之外的第三方，保管參與者的名單，並將名單加以編碼，只有這位第三方的保管者握有解開這些編碼的工具；或是利用隨機回覆的技術（在第十章中將有更進一步的討論）等等。由於過去曾經發生過一些濫用的情形，因此也讓我們知道這些措施或許是必要的。舉例而言，狄恩納以及克雷恩多（Diener and Crandall, 1978: 70）曾經報導過，在 1950 年代，美國國務院以及聯邦調查局曾經要求知名的金賽性學研究交出曾經參與該項研究人士的研究記錄。金賽性學研究機構拒絕遵從政府的要求，並威脅寧可銷毀所有的研究記錄，也不願意公布絲毫的研究資訊。到了最後，政府單位終於退讓。研究者所肩負的道德與倫理責任，使得這些研究者寧可銷毀記錄，也不願意把這些記錄交給政府官員。

保密或許也能保護研究參與者免於受到生理方面的傷害。例如，我曾經遇過一研究者，這名研究者曾經研究過在極權社會裡的秘密警察的內部工作情形。如果他把提供相關情報的告密者身份公布出來，這些人可能會被處死或監禁。因此，爲了保護這些研究參與者，這名研究者所有的筆記都以密碼的方式來撰寫，並把所有記錄都秘密地鎖起來藏在某處。雖然他住在美國，他的人身安全仍然遭受外國政府的威脅，而且他也數次發現，有人曾試圖侵入他的辦公室。而在別的情境中，其他原則的重要性則可能會優先於保護機密的重要性。

有些研究者爲了合乎倫理而付出高昂的個人代價。例如，雖然華盛頓州立大學社會學博士班學生史凱斯（Rick Scarce）從未因爲違反法律而被控告或是被判刑，同時也謹守美國社會學協會所訂立的倫理原則。但是，他仍然因爲藐視法庭而在華盛頓州的史坡堪市（Spokane）監獄中待了十六週之久。他之所以會被監禁入獄，是因爲他拒絕在大陪審團之前作證，也不願違反社會研究資料的保密原則。史凱斯在這之前一直在研究激進的動物解放團體，並已出版過一本關於這個主題的著作。他曾經訪談過一位研究參與者，而這位參與者被懷疑曾帶領一個團體侵入多所動物收容所（animal facilities），造成了高達十五萬美元的損失。有二名法官拒絕接受社會研究資料保密性的說法，因而判決史凱斯藐視法庭，且必須入監服刑。[23]

參與者的資訊是個人資產 如果你基於研究的目的而免費提供自己的相關資訊，那麼你是否會失去對於這些資訊所擁有的全部權利？這些資訊是否可被用以傷害

你？研究參與者同意提供關於自己的知識，並讓研究者分析這些知識。，而這些資訊之後可以被用於數種不同的目的上，其中包括一些違反受試者利益的行動。大型企業每天都在蒐集、購買、販賣、分析、以及交換關於人們的各種資訊。私人企業與政府機構都會使用這些關於購買習慣、個人品味、消費模式、信用評等、投票傾向、網路瀏覽習慣等等的資訊。資訊是一種個人資產的形式。就像其他的「智慧」財產（著作權、軟體、專利權等等）一樣、而且與大多數實體財產不同的是，資訊在進行交換之後，其價值仍然會繼續存在。

多數人都是在幾乎沒有、或是根本沒有任何補償的情況之下，把他們的時間和資訊提供給研究者。然而，對於隱私的關切、以及資訊的蒐集問題，使我們可以合理地把個人資訊當作是私人資產。如果個人資訊確實是私人資產，個人當然有權保留、販售、或是讓渡這些資訊。如果有人使用這些個人資訊時，是以一種參與者若在完全知情的情況下，將會反對的方式，此種情況就會引起極大的倫理爭議。例如，有一群堅定的不吸煙者參與了一項關於他們的習慣以及心理特徵的研究。有一家市場調查公司取得了這些資訊，而有一家煙草公司則要求這家市場調查公司設計一項活動，來向不吸煙者進行促銷活動。如果這些不吸煙者事先曾被告知他們所提供的資訊用途，他們或許便會選擇不加入這項研究。重視倫理的研究者可用一些方法以增加對於參與者的保護，像是提供一份研究發現給所有的研究參與者，並在知情的同意書裡，向參與者說明這些發現的所有用途，且徵得其同意等等。

關於研究參與者的資料是由誰掌控的議題，與在第四章當中所描述的社會科學研究取向有關。實證主義認爲，資訊的蒐集和運用應該要由專家進行，不該讓研究參與者或是一般人來決定。而詮釋取向及批判取向的社會科學則基於不同的理念，對於研究對象在研究過程中的融入和參與、以及研究資料與研究發現的運用方式，各自提出不同主張。[24]

對於研究參與者的明文保障　美國聯邦政府與其他國家的政府，都有保護研究參與者，及其權利的相關規定和法律。在美國，此種法律方面的限制載明於由「美國衛生部」及「保障免於研究風險的人群服務辦公室」所頒布的法令規章。雖然這只是美國聯邦政府裡的一個機構，大多數的研究者與其他的政府機構，都會向此機構尋求指引。目前美國政府的法規是從 1966 年實施、並於 1971 年擴大適用的「公共衛生服務」政策所演變而來。美國《國家研究法》（National Research Act, 1974）設立了「國家保護生醫與行爲研究中人類受試者委員會」（National Commission for the Protection of Human Subjects in Biomedical and Behavioral Research）。此委員會大幅度地擴展各項相關法規，也要求在大部分的社會研究中都要取得知情的同意，而且該委員會也明訂由各研究機構及大學來承擔維護倫理標準的責任。美國衛生部與人群服務辦公室在 1981 年頒布了相關的規章，至今仍然有效。除此之外，關於科學不當行爲以及保護資料機密性的規定，則於 1989 年加以擴大增修。

聯邦政府的法令規章所依循的是生物醫學的模式，並保護受試者免於遭受生理方面的傷害。其他的規定則要求在所有的研究機構、學院、和大學之中，都要由「機

構審查委員會」審查所有會以人們作爲受試者的研究。機構審查委員會的成員，主要是由研究者及社群成員所組成。也有一些類似的委員會負責監督在研究中使用動物作爲研究對象的方式。而機構審查委員會也會管理、監督、與審查所有研究程序對於人類受試者所造成的影響，並且要求研究者遵守倫理守則。除此之外，當一項研究首次提出申請，機構審查委員會亦會審查研究的程序。教育測驗、「正常的教育工作」、大多數的調查、大多數對於公共行爲的觀察、及對於無法辨識個人資訊的現存資料所進行的研究等等，則都可以不必經過機構審查委員會的審查程序。[25]

倫理與科學社群

醫生、律師、諮商師、與其他的專業人士，都有**倫理規範**（code of ethics）以及同儕審查委員會或是證照制度。這些規範明訂了專業的標準，當在實務上出現問題時，也可以當作指引的方針。[26] 社會研究者在提供服務時並未收取費用、只接受有限的倫理訓練、也很少獲得證照。然而，社會研究者仍然把倫理的議題融入於研究當中，因爲這才是在道德方面與社會方面負責任的作法。而此種作法也能幫助社會研究不至於被指控爲對社會漠不關心、或是在佔人們的便宜。全世界的專業社會科學協會都有自己的一套倫理規範，這些規範明確地說明了合宜的以及不合宜的行爲，且代表了專業人士對於倫理的共識。研究者或許不會同意所有的倫理規範，倫理規則也可能有不同的詮釋空間。但是，研究者們都被預期要維護倫理規範，因爲這是他們身爲專業社群的一分子所應盡的義務。

研究倫理規範可以追溯到**紐倫堡規範**（Nuremberg code），這是第二次世界大戰之後，同盟國在紐倫堡軍事法庭審判納粹黨人的戰爭罪行時，所採用的規範。紐倫堡規範原先是爲了因應在納粹集中營裡的殘酷實驗而發展出來的，此項規範概略地描述了倫理原則，以及人們作爲研究受試者所擁有的權利。紐倫堡規範裡所陳述的原則，主要集中於醫學實驗方面，後來則成爲社會研究倫理規範的基礎。類似的人權規範，例如 1984 年的聯合國世界人權宣言、1964 年的赫爾辛基宣言等等，也都爲社會研究者帶來了許多啓示。[27]（請參閱資訊箱 5.5 合乎倫理的社會研究之基本原則）

自 1960 年代與 1970 年代開始，許多專業社會科學協會（例如，美國心理學協會、美國人類學協會、美國政治學協會、以及美國社會學協會等等）開始制訂並採用了倫理規範。專業社會科學協會設立了各種委員會來審查倫理規範及探查可能的違反情形，但是並未嚴格地執行這些規範。對於輕微的違規情形，其懲處幾乎不過是寄給研究者一封糾正信函而已。如果沒有違反法律，那麼對於嚴重違反倫理規範的案例之最大懲罰，就是對此情形詳加記錄並公諸於眾。此種負面的消息可能導致研究者失

倫理規範 由專業組織所發展出來的原則及指引方針，其目的在於指引研究的進行方式，並澄清在合乎倫理以及不合倫理的行為之間的界線。

紐倫堡規範 第二次世界大戰後的戰爭罪行審判時，所採用關於道德及倫理行為的國際規範。此規範是針對納粹黨人所進行的殘忍醫學實驗所提出的因應之道，而且也是以人們作為受試者的研究倫理規範的開端。

資訊箱5.5　合乎倫理的社會研究的基本原則

- 正式承認倫理責任應該由個別研究者承擔。
- 不能為了個人利益而剝削研究參與者或是學生。
- 極力建議或要求獲得某種形式的知情後的同意。
- 尊重所有對於隱私、保密、以及匿名性的保證。
- 絕對不能強迫或是羞辱研究參與者。
- 只有在必要時才使用欺騙，而且之後必須進行相關的說明。
- 使用適合於研究主題的研究方法。
- 偵測並排除任何可能會對受試者不利的影響。
- 事先預想研究本身或公布結果之後可能造成的影響。
- 確認提供研究資金的贊助者是誰。
- 在進行比較研究時，應該要與地主國共同合作。
- 研究設計的細節必須和研究結果一起公布。
- 對於結果的詮釋必須與資料一致。
- 在方法論上要採取高標準，並力求精確。
- 不可進行秘密的研究。

去工作、學術期刊拒絕刊登其研究發現、而該名研究者也會被禁止接受研究的資金補助——換句話說，就是把該名研究者從專業研究者的社群當中驅逐出去。

倫理規範的作用並不僅止於使思考過程更有系統性以及提供指引而已，還能夠幫助大學以及其他機構捍衛合乎倫理的研究。舉例而言，在 1994 年時，有一位研究者在訪談了二十四位工作人員以及進行了觀察之後，證實了「密爾瓦基公設辯護律師辦公室」（Milwaukee Public Defenders Office）裡的工作人員超時工作的情況非常嚴重，因而無法爲窮困的人們提供有效的法律辯護。在得知這項發現之後，該辦公室的高層官員立刻與大學聯絡，並且要求知道哪些工作人員曾經與研究者談過話，似乎暗示著將會對這些員工採取報復行動。這所大學引述了被廣泛接受的保護人類研究參與者的倫理規範，以保護這名研究者並且拒絕提供相關的資料。[28]

倫理及研究的贊助者

告密　你可能會找到一份爲贊助者進行研究的工作，而贊助者可能是和你簽約進行研究的一名雇主、一個政府機構、或是一家私人公司。當贊助者提供經費進行一項研究，尤其是一項應用研究時，便可能產生特殊的倫理問題。你可能會被要求在倫理或專業的研究標準方面作出妥協或讓步，以作爲取得合約或繼續雇用的條件。而這也意味著你必須先設定倫理界線，一旦超出你的界線時，就必須拒絕贊助者所提出的要求。當你面對著不合理的要求，你有三種基本的選擇：效忠於組織或是較大的團體、脫離這個情境、抑或提出反對意見。[29] 這三種選擇所代表的意義分別是，

屈服於贊助者、離職、或成爲一名**告密者**（whistle-blower）。你必須選擇你自己的行動方案，但最好能在簽約之前就先考慮到倫理的議題，並在事前表達自己所關切的事項。

告密舉發可能是件吃力不討好的工作。其中牽涉到三方面：發現違反倫理的不當行爲的研究者、外在的機構或媒體、以及聘雇組織中的監督者。研究者必須確認某項行爲已經嚴重地違反倫理規範，且該行爲是由組織所默許進行的。在窮盡內部管道仍無法解決此項議題之後，研究者才轉而訴諸外部人士的力量。外部人士對於這個問題不一定有興趣，也不一定有能力幫忙。外部人士可能有自己優先考量的事項（不要破壞這個組織的外在形象、或不要過度炒作這個問題等等），而這些優先事項可能不同於研究者最爲關切的議題（終結違反倫理的行爲）。監督者或是管理者則可能設法醜化或是懲罰那些揭發問題，以及不忠於組織的人（請參閱範例 5.5 一位告密者的故事）。正如同弗瑞希特－施瑞德（Frechette-Schrader, 1994: 78）所指出的：「告密舉發的行爲是一種特殊的違抗組織行爲。或者，更確切地說，是一種忠於更高尚的原則、而非忠於雇主的行爲。」即使在最好的情況下，要解決一項議題仍然曠日費時，也會造成情緒上的緊張狀態。爲了要讓行爲合乎道德，告密者必須有作出犧牲的心理準備：像是失去工作與升遷的機會、減薪或是被迫調職、被同事唾棄、抑或是惹來法律方面的麻煩等等。而且，就算有人告密舉發違反倫理的行爲，也不保證做正確的事就能改變這些違反倫理的行爲、或保護研究者免於被報復的風險。

獲得特定的發現 假使贊助者以直接或間接的方式告訴你，你應該要獲得某種特定的研究結果時，你應該要怎麼做？如果進行研究的前提是必須得到特定的研究結果，重視倫理的研究者就會拒絕參與研究。所有的研究都應該在未對研究之發現設下限制的情況下進行。例如，有一個調查組織獲得一份合約，要爲一個購物中心的公會進行研究。這個購物中心的公會，正和一個打算在購物中心舉行示威活動的政治團體進行法律訴訟。在調查組織內的一個採訪者反對在問卷中的許多問題，因爲他認爲這些問題是無效的，而且過於偏袒購物中心公會。在與一家報社聯絡並公開這些帶有偏見的問題之後，這名採訪者就被開除了。然而，過了幾年之後，在一場告密者的訴訟當中，這名採訪者獲得超過六萬美元的賠償金，以作爲補償他被雇主積欠的薪資、精神賠償、以及對於該調查組織的懲罰性賠償金。[30]

另一個可能有人施加壓力以獲得特定發現的例子，則是在於教育測驗的領域中。測量美國兒童學業成就的標準化測驗已經飽受抨擊。例如，在美國大約百分之九十的學區裡的兒童，在標準化測驗所得到的分數都是「平均以上」，而此種現象被稱爲*烏比岡湖效應*（*Lake Wobegon effect*）。這個名稱是以烏比岡湖這個神秘的小鎮來命名的。根據廣播節目主持人凱勒（Garrison Keillor）的說法，在這個小鎮內「所有的兒童都在平均以上」。造成此種狀況的

告密者 一個發現在組織裡有違反倫理或是不合法行為、對這些行為提出異議、並透過組織的管道以試圖阻止這些行為的人。此人有可能會失敗並因而被懲罰，但是在失敗後，此人仍然可能持續在組織外部說出對於違反倫理及非法行為的反對意見。

範例　5.5

一位告密者的故事

法蘭克林（David Franklin）是一名擁有博士學位的微生物學家，他被華納－朗伯特（Warner-Lambert）公司雇用作為與醫學界間的聯絡人。他的工作是取得醫生的信任，並提供他們科學資訊，以販售藥品給他們。在他受訓的過程中，他被要求對某項藥物的療效作出與實際情況不合的宣稱，並被告知該如何迴避法律和倫理規則，以增加藥物的銷售量。除此之外，某些藥物明明在研究中只顯示出一點點效果，他卻被要求誇大這些研究結果，並隱瞞副作用的報告。當他向他的主管表明憂慮之意，亦提出一些已發表的報告，證明這些藥物具有危險副作用時，他的主管卻完全不理會他的意見。他觀察到這家公司支付了數萬美元的金額給醫生，讓這些醫生為藥物的效果背書，或充當一些醫學文章的作者，而這些文章事實上都是由這家公司的行銷部門所撰寫的。他認為這家公司的行為已違法，可能對人們造成危害。他在正式工作僅僅四個月之後就辭職了，但是他卻被威脅不准揭露任何關於公司的事情。之後，他前後總共花了七年的時間，才解決他告發公司不法行為的法律案件。

*資料來源：*Excerpt from Melody Petersen, "Doctor Explains Why He Blew the Whistle," *New York Times* (March 12, 2003)

主要原因，是因爲研究者把現在學生的分數，拿去和許多年前的學生的分數標準進行比較。因爲這些研究者面臨了來自於教師、學校校長、督學、以及學校董事會的壓力，要求研究者必須獲得特定的結果，好讓這些人能夠向家長與選民報告，說他們學區裡的學童成績都是在「平均以上」。[31]

對於如何進行研究的限制　不論是直接地或間接地（例如，限制研究經費），贊助者是否有可能利用定義什麼可以被研究、或利用限制可以使用的研究技術之方式，來對研究設下限制？贊助者可合法地對研究中使用的研究技術設定一些條件（例如，調查或是實驗），並限制研究的費用。然而，我們依舊必須遵守一般廣爲接受的研究方法。我們應該針對在既有經費數目下所能完成的事項，進行務實的評估。

此種關於各種限制的議題，在**簽約研究**（contract research）中屢見不鮮。而所謂的簽約研究，是指一家公司或一個政府機構要求研究者從事一個特定的研究計畫。在簽約研究當中，研究者可能必須在品質及成本之間進行權衡與取捨。阿布特（Abt, 1979）是一家大型私人社會研究公司——阿布特學會（Abt Associates）——的總裁，他認爲要根據研究的實際成本來取得合約，是件困難的事。一旦研究開始進行，我們可能必須重新設計研究計畫以降低成本。但是，合約的簽訂程序使得在研究中途要作出變更變得困難重重。我們可能會發現，因爲合約的關係，我們被迫使用較不理想的研究程序或方法。因此，我們就會面臨一個兩難的困境：履行合約並進行一項品質低劣的研究，或是不履行合約，因而連帶損失金錢以及未來的工作機會。

簽約研究　一種由政府機構、基金會、以及公司等等所贊助（支付經費）的一種應用研究。在進行此種研究時，研究者同意針對贊助者所提出的研究問題來進行一項研究，並且以固定的價格以及在截止期限之前完成研究。

如果你無法堅守一般認可的研究標準，你就應

該拒絕繼續從事一項研究。如果贊助者要求使用偏差的樣本或是誘導性的問題，爲了堅守研究倫理，你就必須堅決地拒絕繼續合作。如果合法的研究顯示出，贊助者所偏好的想法或計畫是一個拙劣的行動方案，你可能甚至得預期，若不是結束聘僱關係、就是面臨違反專業研究標準的壓力。長遠而論，研究者、贊助者、科學社群、以及整個社會都將因爲違反正規的研究作法而受到傷害。你必須作出抉擇，多加三思自己是否要當個「雇工」（"hired hand"），無條件地給予贊助者任何他想要的東西，即使在倫理上是錯誤的；或選擇當一位專業人士，爲了更高尚的道德原則而去教導、指引、甚至反對贊助者的意見。

如果贊助者對於使用研究發現或對於真相根本不感興趣，我們應該要問：爲什麼這些贊助者想要進行社會研究？答案是此種贊助者根本不把社會研究當成是獲得知識或真相的工具，而是將社會研究僅當作一種掩護，好讓他們可以將其決策與作法加以合理化。而且這是他們以其他方法無法輕易辦到的。這些贊助者其實是以欺騙的方式在濫用研究者身爲認真的、值得信任的專業人士的地位，並試圖「利用」（"cash in"）科學研究者所擁有的誠實與正直的聲譽。當此種情況發生時，合乎倫理的行動方案就需揭露並終止濫用的惡行。

隱瞞研究發現 如果在你進行一項研究之後，研究發現對於贊助者不利、或是贊助者拒絕發表研究結果，你應該怎麼辦？這並非罕見的情況。例如，有一名社會學家爲威斯康辛州樂透委員會進行了一項研究，主題是州政府公辦賭博所造成的影響。在這名社會學家完成報告、但尚未對大眾發表前，該委員會要求她刪掉那些描述賭博會造成許多負面社會影響的章節，並要求她刪除創辦社會服務工作以幫助強迫性賭徒的建議。這名研究者陷於一個非常艱難的處境，她得決定哪一種倫理價值必須優先考量：是要幫這些支付研究經費的贊助者掩飾真相呢？還是要把真相公諸於世，然後一肩挑起所有後果？[32] 有一名羅馬天主教的神父，曾經調查過美國主教對於官方教會政策感到不滿的情況，但是他的上司卻命令他隱瞞研究發現，並且銷毀所有的問卷。不過，他後來決定辭去他已擔任長達二十四年之久的神職工作，將研究結果公布。[33] 由此可見，爲了要堅守倫理原則，研究者必須付出高昂的個人與經濟代價。

政府機構可能會隱瞞那些抵觸官方政策或是會使高層官員難堪的科學資訊。此外，受雇於政府機構的研究者，若把這些資訊公開，也可能遭到報復。例如，有一位社會研究者受雇於美國人口普查局，針對美國對於伊拉克發動波灣戰爭所造成的死亡人數進行研究。之後，這名研究者報告表示，政府官員爲了政治理由而隱瞞研究發現。這名政府機構想要解雇的研究者報告，提供統計數據的美國政府機構刻意拖延，並忽視高死亡率的研究發現。在這項資訊得以對外發表之前，必須經過一名官派機構首長的同意。這名研究者指控這位官派首長一心一意只想維護政府的外交政策。在另外一個例子裡，美國國防部下令銷毀數項研究，因爲這些研究顯示出美軍中大約有百分之十的男同性戀或女同性戀者，而軍方在禁止同性戀從軍方面，並未提供任何協助。2005 年時，美國白宮先是威脅要解僱沒沒無聞的司法統計局局長，而最後則是將他降級，因爲他公布了在執法方面的種族歧視資料。司法統計局

每年都會提出許多關於犯罪模式、藥物濫用、警政策略、囚犯人口等等的報告。這些資料詳盡地記載了，當人們被警方攔下來臨檢時，警方對於不同種族的處理態度有著極大差異。因此，政府高層強烈要求把此種關於種族歧視的資料從報告中刪除。這份資料是來自於 2002 年，對八萬名民眾所進行的訪談結果。這份資料顯示出，在該年度當中，全國各地被警方攔下臨檢的白人、黑人、以及西班牙人司機比例大致相同，大約皆爲百分之九左右。然而，一旦被警方攔下臨檢，接下來會發生的狀況，就因爲司機的種族差異而有天壤之別。[34]

在接受贊助的研究中，我們可以在研究*開始之前*，就和贊助者協商公布研究發現的條件，並簽訂合約作爲證明。在沒有此種事前保證的情況下，就貿然地進行研究，是不明智的舉動。雖然一些較欠缺倫理顧忌的競爭對手可能會這麼做。我們也可以選擇承受贊助者的批評與敵意，並忽視贊助者的抗議，直接公布研究發現。不過，大多數的研究者都比較偏好第一種選擇，因爲如果他們選擇了第二種作法，便可能嚇跑未來的贊助者。

社會研究者有時會限制或是延遲發表研究發現，以保護提供消息人士的身份、維持進入研究現場的管道、保住他們的工作、維護自己或家庭成員的人身安全。[35]這是一種比較不會造成困擾的自我審查行爲，因爲這並非由外在力量所強加於研究者身上的，而是由熟悉研究、對於可能的後果相當瞭解的研究者所進行的。研究者必須擔負起研究的最終責任。通常，研究皆可善加利用許多不同的資源，但同時也會面臨許多相互競爭的壓力（請參閱資訊箱 5.6 評估研究中常見的誤用類型）。

隱匿真正的贊助者　如果不透露贊助者身份的話，此種行爲是否合乎倫理呢？例如，有一家墮胎診所贊助了一項研究，研究主題是反對墮胎的宗教團體的態度。我們必須在下列的倫理價值間做出權衡取捨：把贊助者的身份公開讓受試者知道、違反贊助者希望保密的要求並且公布研究結果、以及受試者合作意願降低的可能性。

資訊箱 5.6　評估研究中常見的誤用類型

- 提出「錯誤的」研究問題（例如，在使用公式化問題較適合的時候，卻提出總結性的是非題，或提出一些把主要的利害關係人排除在外的問題）。
- 在決策已經制訂之後，才實施評估研究，只是利用評估研究來當作延遲或是合理化該項決策的理由。
- 要求採用不適合於該項方案評估任務的研究設計或是資料蒐集技術。
- 干涉研究設計或資料蒐集的過程，以確保得到想要的特定結果。
- 當評估研究的結果明確地顯示某項方案毫無效果時，卻仍然繼續實施該項方案；或是當評估研究的結果明確地顯示某項方案成效斐然時，卻終止進行該項方案。
- 隱瞞／刪除正面結果以終止／縮減某項方案，或是隱瞞／刪除負面結果以繼續／擴展某項方案。

資料來源：修訂自 Stevens and Dial（1994），兩位作者同時也提供了誤用的範例。

如果研究者發表了研究結果，研究者在倫理方面便有義務揭露出真正的贊助者是誰。至於另一個倫理議題，也就是研究者是否應該向受試者透露真正的贊助者身份的議題，則比較沒有一致的共識。普雷瑟與同事（Presser and colleagues, 1992）發現，在進行調查研究時，受試者所回覆的答案，會因爲贊助者的身份而有所不同。如果受試者相信，某項調查研究是由一家在某項議題上採取強硬立場的報社所進行，那麼這名受試者較不會否定該報社在此項議題上的公開立場。如果受試者相信某項調查研究是由一個中立的學術組織所贊助，則比較不會有上述的問題。向受試者告知贊助者的身份，是合乎倫理的作法。除非研究者在方法論方面有充分的理由可以不這麼做。

女性主義社群主義研究倫理

有一些採取詮釋取向社會科學或是批判取向社會科學（請參閱第四章）立場的研究者，認爲大部分的倫理爭論、倫理規範、以及審查委員會都是不適當的、而且是根植於實證主義的假定而來。他們提議採取一種女性主義社群主義模式的研究倫理（feminist communitarian model of research ethics），以作爲傳統研究倫理的替代選擇。因爲他們認爲傳統研究倫理的基礎是正式的程序、對於成本與效益的理性權衡、以及道德上與善相關的抽象原則。他們主張「道德任務不能被化約爲專業的倫理」（Christians, 2003: 232）。由於這些研究者認同參與式行動研究（participatory action research）（請參閱第二章）的理念，因此他們主張研究參與者應該也有權利決定研究的進行方式，並應該積極地涉入研究的進行過程。倫理應該能夠反映研究的最終目的，意即在研究參與者本身的日常生活經驗中，賦予他們權能，促進以及提升人類的自由。

女性主義社群主義模式所倚賴的是三項道德原則。首先，合乎倫理的研究是意義曖昧的（multivocal）。換句話說，合乎倫理的研究承認人類經驗的多樣性，並包含了此種多樣性，而且此種研究是由以下前提開始的：所有的人類生活，都是處於一個由社會建構的性別、種族、階級、以及宗教脈絡裡。人們生活在多重社群裡，每一個人也都有某些重要的事情可以和他人分享。其次，合乎倫理的研究，要求研究者投入一場關於道德關懷的對話，而此種對話中所使用的詞彙，都是以參與者的日常生活經驗來表達。研究者必須投入及參與這些持續進行的道德辯論與討論。這些辯論與討論都發生在他們想要研究的人們所處的社群之中，而且研究者不應該把自己的抽象法制權利或是原則強行施加於研究參與者身上。第三，當研究者及參與者在研究過程中以開放、平等的方式進行互動時，此種情況將會揭露出權力關係，以及產生一些能夠促進更多反思以及相互瞭解的社會評論。到了最後，研究者以及參與者之間的合作關係將會逐漸浮現，而在此種關係之中，「對於隱私的侵犯、知情的同意、以及欺騙行爲等等，都將不再是議題了」（Christians, 2003: 234）。

女性主義社群模式的研究倫理目前仍在發展初期，且尚未被實施在任何研究當

中。儘管如此，此種模式批判了研究倫理的主流取向，認爲主流取向過度地傾向於形式—法制主義（formal-legalistic）、以程序爲基礎、而且抽象。此外，此種新的模式也突顯出社會科學中的不同取向（如同在第四章中所描述的）與研究倫理中的道德議題產生關聯性的方式。

結論

本章是從社會研究的一般基礎過渡到研究設計種種細節的銜接章節。我們在本章中討論了兩個重要的議題：文獻回顧與倫理議題。這兩個議題都是在設計一項研究時的一部分準備工作，且兩者都涉及到將你的研究放置於大型研究者社群的脈絡之中、以及將一個特定具體的研究與較大的議題連結起來。

在第一章裡，我們曾討論過科學對於社會的獨特貢獻，以及社會研究成爲獲得關於社會世界知識的方式。社會研究的觀點與技術，可以成爲瞭解世界的強大工具。然而，伴隨著此種力量而來的，是不可逃避的責任——不論是對你自己、你的贊助者、研究者的社群、以及更大的社會等等。在某些時候，這些不同的責任可能、而且也確實會與彼此產生衝突。

說到底，最後仍是你個人必須決定是否要去進行合乎倫理的研究、去維護和捍衛你所採取的社會科學取向的原則、以及去要求別人的行爲也要合乎倫理。社會研究所產生的知識真實性、以及此種知識的善用或是誤用，全都取決於像你這樣的個別研究者。因此，研究者們必須反思自己的行動，並仔細思考將社會研究融入於社會的方式。

在下一章，我們將會檢視基本的研究設計方法以及相關議題。不論是在質化研究或量化研究裡，研究者們都會遇到這些方法以及議題。

問題回顧

1. 文獻回顧的四個主要目標爲何？
2. 何種文獻的出處是最容易找到的？而何種文獻的出處則是最難找到的？
3. 你要如何找到一份博士論文？
4. 如何區分良好的文獻回顧以及不佳的文獻回顧？
5. 在社會研究中使用網際網路的主要優點和缺點爲何？
6. 避免在研究中出現違反倫理行爲的基本防範措施爲何？
7. 欺騙或是強迫個人參與研究的作法，與自願同意原則有何衝突之處？
8. 請解釋在米爾葛蘭、韓福瑞、及辛巴度等人的例子中，所牽涉到的倫理議題。
9. 何謂知情的同意？知情的同意如何保護研究中的受試者？
10. 匿名性以及保密之間有何差異？

注解

1. See Hunt (1997) and Hunter and associates (1982).
2. From Hargens (1988).
3. Based on Hargens (1991).
4. See Reynolds (1979:56–57) and Sieber (1993).
5. See research fraud discussion Broad and Wade (1982), Diener and Crandall (1978:154–158), and Weinstein (1979). Also see Hearnshaw (1979) and Wade (1976) on Cyril Burt. Kusserow (1989) and the September 1, 1989, issue of the National Institutes of Health weekly *Guide* summarize some recent scientific misconduct issues.
6. See "Noted Harvard Psychiatrist Resigns Post after Faculty Group Finds He Plagiarized," *Chronicle of Higher Education* (December 7, 1988).
7. See Blum (1989) and D'Antonio (1989) on this case of plagiarism.
8. See "Doctor Is Accused of 'Immoral' Tests,'" *New York Times* (December 9, 1988). For a more general discussion of power and trust, see Reynolds (1979:32).
9. Lifton (1986) provided an account of Nazi medical experimentation.
10. See Jones (1981) and Mitchell (1997) on the Bad Blood case.
11. Diener and Crandall (1978:128) discuss these examples.
12. See Warwick (1982) on types of harm to research participants. See Reynolds (1979:62–68) on rates of harm in biomedical research. Kelman (1982) discusses different types of harm from different types of research.
13. College counselors report that anxiety and low self-esteem over dating are major problems among college women (Diener and Crandall, 1978:21–22). Also see Kidder and Judd (1986:481–484).
14. See Dooley (1984:330) and Kidder and Judd (1986: 477–484).
15. See Hallowell (1985) and "Threat to Confidentiality of Fieldnotes," *ASA Footnotes*, 12:6.
16. For more on the general issue of the right not to be researched, see Barnes (1979), Boruch (1982), Moore (1973), and Sagarin (1973).
17. Informed consent requirements and regulations are discussed in detail in Maloney (1984). Also see Capron (1982) and Diener and Crandall (1978:64–66).
18. The debate over covert research is discussed in Denzin and Erikson (1982), Homan (1980), and Sieber (1982). Also see Miller and Tewksbury (2000), especially Sections 1 and 4.3.
19. See Diener and Crandall (1978:87) and Warwick (1982:112).
20. See Diener and Crandall (1978:173–177) and Kidder and Judd (1986:469).
21. See Boruch (1982), Caplan (1982), Katz (1972), and Vaughan (1967) on privacy.
22. For more on the Wichita Jury Study, see Dooley (1984:338–339), Gray (1982), Robertson (1982), Tropp (1982:391), and Vaughan (1967).
23. See Monaghan (1993a, 1993b, 1993c).
24. See Gustavsen (1986).
25. IRBs are discussed in Maloney (1984) and Chadwick and associates (1984:20). See Taylor (1994) for an international survey of ethical standards.
26. See Abbott (1988), Brint (1994), and Freidson (1986, 1994) on professionals.
27. See Beecher (1970:227–228) and Reynolds (1979:28–31, 428–441).
28. See "UW Protects Dissertation Sources," *Capital Times* (December 19, 1994):4. Greenwald (1992: 585–586) remarked, "Sociology stands out among the learned professions as critical of the authority of established institutions such as government or large business firms" and in its provision to "explicitly state the shortcoming of methodologies and the openness of findings to varying interpretations."
29. See Hirschman (1970) on loyalty, exit, and voice. Also see Rubin (1983:24–40) on ethical issues in applied research.

30.Additional discussion can be found in Schmeling and Miller (1988).

31.See Fiske (1989), Koretz (1988), and Weiss and Gruber (1987) on educational statistics.

32.See "State Sought, Got Author's Changes in Lottery Report," *Capital Times* (July 28, 1989), p. 21.

33.See Chambers (1986).

34.See Dale W. Nelson, "Analyst: War Death Counts Falsified," *Wisconsin State Journal* (April 14, 1992:3A; "Ex-Official Says Pentagon Dumped Findings on Gays," *Capital Times* (April 1, 1993); and Eric Lichtblau, "Profiling Report Leads to a Clash and a Demotion, " *New York Times* (August 24, 2005).

35.See Adler and Adler (1993).

第 6 章

研究設計的策略

三角檢證法
研究的質化以及量化取向
質化設計議題
量化設計譯題
結論

> 實質的問題因而必須被轉譯為社會探究的詞彙……經由作出各種選擇以及某種在工作方面的合宜次序，你將可以獲得一種思考的方式，並使你有能力去回答一個研究問題。
>
> ——Robert Alford, *The Craft of Inquiry*, p. 25

在 1995 年時，僅僅數天之內，就有超過七百人死於芝加哥的熱浪之中。新聞報導以官員們都無法解釋為什麼會發生這樁意外事件。就在此事件發生之後沒多久，大眾與媒體對這起災難的相關討論突然就消失無蹤了。克林能柏格（Klinenberg, 2002）在一項研究中，針對這起「極端的事件」進行了「社會剖析」（"social autopsy"），而他所使用的是社會學研究的工具，例如民族誌田野工作、訪談、詳細調查檔案文件（報紙、統計報告、各種不同的記錄、以及地圖等等）、以及對於統計資料所進行的分析。這項研究的設計目的是要回答下列這些問題：為什麼這麼多人在這麼短的時間內死去？這起災難又是如何發生的？他利用社會研究來剖析這起事件，並揭露出造成這起事件背後的社會、政治、以及經濟因素。這項研究讓我們得知這起災難發生的原因和過程，而這個例子也顯現出我們應該如何設計一項社會研究，以回答一個尚未被探討的、或是被忽略的重要問題（在短短數天之內，數百人無辜死去的真正原因）。

本章把焦點放在一些議題上，這些議題是關於如何設計一項研究、以及發展一項策略以在研究過程之中爲你提供指引。根據你所進行的主要是量化或質化的研究，你在設計與進行研究時的策略也會有所變化。在進行量化研究時，你必須先詳盡地進行規畫，才能蒐集或分析資料。而你或許會想要提出的問題，則是如何才能夠產生一個邏輯嚴謹的設計，以精確地定義並測量所有的變項、選擇具有代表性的樣本、蒐集資料、以及進行統計分析呢？如果你要進行的是質化研究，你就必須試著讓自己完全地沉浸在某種範圍的資料之中。而且在整個蒐集資料的過程中，還必

三角檢證法　一種認為從多重角度觀察事物可以增加正確性的概念。

須非常留意各種新的洞見。你會想提出的問題，則是你要如何捕捉動態社會生活中的豐富性、構造、以及感受。當然，你也可以混合量化研究和質化研究的特色，以在兩者之間截長補短。混合的研究取向雖然有其優點，但也會增加研究的複雜度，而且比較耗費時間。我們可以在下一個即將要討論的三角檢證法的部分，清楚看到混合式研究取向所具備的優點。

三角檢證法

調查員與水手會從不同的位置進行觀察，以測量物體之間的距離。利用從數個不同的角度和位置觀察某個物體的方式，調查員及水手就能夠更加確定某個物體的真正位置（請參閱圖 6.1）。社會研究者也會採用一種被稱爲**三角檢證法**（triangulation）的類似程序。在社會研究中，我們所依循的原則是，當我們由多種不同觀點進行觀察的時候，會比只從單一觀點進行觀察，得到更多的資訊。

社會研究者會使用數種不同類型的三角檢證法（請參閱資訊箱 6.1 四種不同類型三角檢證法的範例）。最常見的一種類型是*測量的三角檢證法*（*triangulation of measure*），意爲我們會使用多種不同的方式，來測量相同的現象。例如，假設你想要知道關於某個人的健康狀況。首先，你要求此人完成一份有多項選擇題的問卷。接著，你進行一項開放式的非正式訪談。除此之外，你也向某位與這個人同住的同伴或是照顧者詢問此人的健康狀況。你也可以去訪談這個人的醫生，並一起檢視他或是她的病歷，以及各種健康檢查的結果。相較於只使用一種單一測量方法，當你使用多種不同的測量方式，尤其若是每種測量方式都得到相當類似的結果時，你對於測量結果的正確性就會更有信心。而你在這些不同的測量結果中，所看到的差異之處，也可以激勵你更進一步地去探討其他問題。

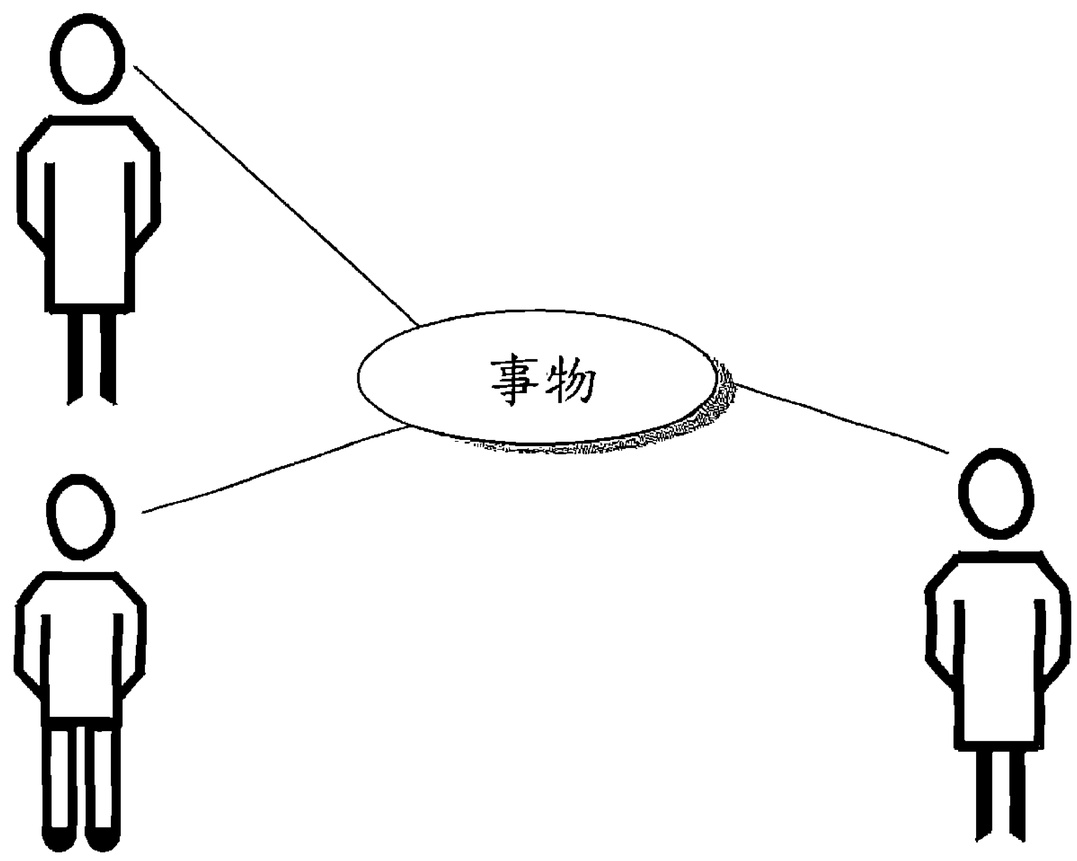


圖 6.1　三角檢證法：從不同的觀點進行觀察

資訊箱 6.1　四種不同類型三角檢證法的範例

主題

美國流行電影中的暴力情節與暴力畫面的數量。

測量　建立三種測量暴力的量化指標：頻率（例如，謀殺或是鬥毆場景出現的次數）、強度（例如，尖叫的音量大小及持續時間、在臉部以及肢體動作顯示出痛苦的強弱程度）、以及在電影中明白露骨地展現暴力的程度（例如，顯現出一具淌著血的屍體、被截肢的身體部位、受傷部位的鏡頭特寫等等）。

觀察者　請五個不同的人獨立地觀賞、評估、並記錄在十部最受歡迎的美國電影中，其暴力形式與程度。

理論　比較一下女性主義理論、功能論、以及象徵互動論如何解釋在電影中常見的暴力形式、成因、以及所造成的社會影響。

方法　針對十部流行電影進行一項內容分析、測量受試者對於每部電影中的暴力情節或畫面的反應、調查電影觀眾對於影片中暴力情節或是畫面的態度。並在電影播放過程中與一結束後，對於觀眾的行為進行田野觀察。

觀察者的三角檢證法（*triangulation of observers*）是第一種類型的變化型態。在許多的研究裡，我們會獨自進行訪談，或擔任各種事件及行爲的單一觀察者。如此一來，任何單一觀察者所可能造成的限制（例如，缺乏某種領域內的技術、對某個議題抱持偏見、對某些細節視而不見等等），就會變成整個研究的限制。多個不同的觀察者可以帶來各種不同的觀點、背景、以及社會特徵，也因此減少上述的限制。例如，有二名觀察者彼此進行互動，並在一所安親班裡一起觀察十個五歲兒童的行爲。其中一名觀察者是六十歲的白人男性小兒科醫師，並且具備在大型城市的醫院中長達二十五年的工作經驗。另一名觀察者是個三十一歲、育有二名子女的西班牙裔女性，並具備在一個小鎮的國小裡教書六年的經驗。這兩名不同的觀察者可能會觀察並記錄下不同的資料。因此，把這兩名觀察者所看到及體驗到的資料結合起來，會得到一幅比僅僅依賴單一觀察者還要更加完整的圖像。

理論的三角檢證法（*triangulation of theory*）則是要求我們使用多種不同理論觀點，以規畫一項研究或是詮釋資料。每一種理論觀點都有其假定與概念，因此每一個理論觀點就像是一面透鏡，讓我們可以透過這面透鏡來觀察社會世界。例如，在關於銀行內部工作關係的一項研究，研究者可以使用衝突理論，以強調其中的權力差異及不平等的現象。這項研究也可以突顯出，由於職權地位的差異（例如，經理和出納員），因而造成薪資及工作環境方面的不平等現象。這項研究同時也會揭露出在社會背景方面的相關差異：擁有企管碩士學位的中年白人男性經理，以及僅擁有學士學位的年輕黑人女性出納員。接下來，研究者可以應用理性選擇理論，把焦點集中於個人用來將個人利益最大化的決策制訂過程以及理性策略。此種觀點突顯出銀行經理如何根據客戶的貸款金額、或存款帳戶的金額數量，而調整他花費在每位不同客戶身上的時間和精神。此種觀點也可以更清楚地讓我們觀察到，銀行出納

員如何根據她所認爲每個不同的主管幫助她升遷的可能性，而在不同的主管身上投資不同的時間和精力。每一種不同的觀點都能夠爲研究提供指引：每一種觀點都能夠指出不同的相關資料、提供一套概念、並幫助研究者詮釋資料的意義與重要性。

方法的三角檢證法（*triangulation of method*）則是融合了質化、量化的研究取向及資料。大多數的研究者只專精於其中一種取向，但是這兩種取向之間，有著彼此不同、但可互補的優點。一項融合了兩種取向的研究，通常會比較豐富、也比較全面。我們可以用數種方式來結合這兩種取向[1]，例如依序使用這兩種取向，也就是先使用其中一種、然後再使用另外一種；以平行並進或同時使用的方式來結合這兩種取向。在本章開頭部分所提到的研究中，克林能柏格就把對於死亡人數的量化資料之統計分析、訪談與文件分析結合在一起（請參閱範例 6.1 採用多重方法的研究）。

研究的質化以及量化取向

在所有研究裡，我們努力地以系統性的方式蒐集實徵資料，並檢視資料模式，好讓我們能夠更清楚地瞭解與解釋社會生活。然而，在不同研究取向間的差異，卻會造成溝通和理解方面的錯誤。其實這兩種研究取向間可以互相啓發瞭解，只是必須多花一些時間與努力，才能掌握這兩種研究取向、看出彼此之間如何互補。接下來，我們就要看看一些造成差異的來源。

首先，造成差異的第一個來源，是來自於資料的本質。*軟性資料*（*soft data*）（也就是文字、句子、照片、符號等等）主導了質化研究的策略以及蒐集資料的技術，而這些策略與技術都與量化取向所使用的*硬性資料*（*hard data*）（也就是以數字形式呈現的資料）大相逕庭。這些差異可能使得量化研究中所使用的研究工具，不適合用在質化研究中、或是與質化研究毫無關聯，反之亦然。

另一個存在於質化和量化研究之間的差異，則是源自於與研究過程有關的原則、以及關於社會生活的假定。質化研究與量化研究各自的原則，造成了不同的「研究語言」，而且各自強調的重點也完全不同。在量化研究裡，我們比較依賴實證主義原則，使用一種充滿各種變項與假設的語言，我們強調的重點在於精確地測量各種變項及驗證假設。而在質化研究裡，我們則比較依賴來自於詮釋取向或是批評取向社會科學的原則。我們所使用的是一種充滿了「個案與脈絡」、以及文化意義的語言。強調重點在於對特定具體的個案進行詳盡檢視，而這些個案都在社會生活的自然流動之中發生。耐人尋味之處在於，採取質化取向的女性社會研究者的人數，遠高於男性的社會研究者的人數。[2]

質化與量化研究之間的第三個差異，則來自於我們在研究中所試圖要完成的事物。「良好研究的核心」——不論是量化研究或是質化研究——「就是一道謎題和一個想法」（Abbott, 2003: xi）。在所有研究中，我們都試圖解決一道謎題或是回答一個問題，但根據我們所使用的研究取向，我們會以不同的方式來解決謎題或回答問題。在本章一開始的部分所提及的熱浪研究當中，克林能柏格（Klinenberg, 2002）

範例　6.1

採用多重方法的研究

在一項關於美國境內多重種族身分的研究中，李以及賓恩（Lee and Bean, 2007）混合了量化以及質化的研究取向。他們觀察到，自 1970 年以來，由於移民的人數不斷地增加，因此使美國社會更加多元化。而且在 2000 年時，美國的人口普查首度開放讓人們可以把自己歸類為多重種族。此種在種族方面的新多樣性，與長久以來的單一種族類別與主流的白人和黑人的種族二分法，形成了強烈的對比。李以及賓恩詢問了那些多重種族的人們，當他們在勾選單一種族或是多重種族的身分類別時，是否覺得比較自由、或覺得受到許多的限制。兩位研究者也詢問了這些受訪者，當他們在人口普查的表格上勾選多重種族這個類別時，這究竟是一種象徵性的行動、或是真實地反映出自己在日常生活中屬於多重種族的狀況。在這項研究的量化部分，作者們以統計方法分析了在 2,000 份人口普查資料裡，把自己歸類為多重種族的人們之數量，及其不同種族混血的情況。而在這項研究的質化部分，作者們與來自加州北部及南部的多重種族成人們，進行了四十六次的半結構式深度訪談。在這些訪談裡，李和賓恩詢問這些受訪者如何與為什麼認為自己的身分是多重種族、此種身分認同是否會隨著時間，或隨情境脈絡而改變、以及一些與種族或族群有關的語言使用問題，和其他的實際活動。作者們訪談了各種由亞洲人、白人、拉丁美洲人、以及黑人所混血而成的多重種族成人。根據這些訪談的結果，李和賓恩發現，與其他不同種族類別混血的人們相比之下，多重種族的黑人比較不會把自己稱作是具有多重種族的人。此種自我限制的情形，與美國歷史上的模式一致，也就是即使一個人僅僅只有部分的黑人血統，一般的美國大眾也會把這個人認定為黑人。至於具有白人與亞洲人或拉美洲人的混血血統、或是拉丁美洲和亞洲的混血血統的人，在對自己的種族認同方面，則比較具有彈性。由於大眾的普遍看法，以及亞洲人與拉丁美洲人在外表上的狹隘刻板定義，因此有些混合了亞洲與白人、或是拉丁美洲與白人血統的人，也會認為自己是白人。至於其他混合了白人與亞洲、以及白人和拉丁美洲血統的人，則說他們對於自己的混血血統感到自豪，即使這對他們的日常生活造成一些影響也無所謂。此外，作者們也發現，人們不會一直堅持認同某一種種族身分，而是會在不同的情境中，宣稱自己具有不同的種族或族群背景。在把量化以及質化的發現結合起來之後，李和賓恩認為，拉丁美洲人及亞洲人的種族團體界線消失的速度比黑人還快。他們的結論是，一種新的黑人及非黑人之間的區隔方式已經逐漸浮現，並漸漸取代老舊的白人與黑人之間的區隔方式。然而，相對於其他的種族類別，黑人仍然處於較為弱勢的地位。

想知道爲什麼會有這麼多人在熱浪來襲時喪生。但是他也想知道這些受害者是怎麼死去的，以及爲什麼某些階層的人們受到很大的影響，而其他階層的人們卻沒有受到太大的影響。在量化研究裡，我們通常要證明某項關係或假設爲真或是爲假。因此，我們會把焦點放在從許多個案所發現的結果或者是效果上面。在考驗一個假設時，可能不僅只有單純地證明該假設爲真或是爲假的答案。更爲常見的情況是，此種過程還包括去瞭解某個假設在某些個案或情境裡成立、但在其他個案或是情境裡卻不成立的情況。在前述的熱浪研究中，克林能柏格想知道人們的社會階級是否會影響最後的結果，也就是一個人的社會階級對於他或她在熱浪中死亡的可能性，是否造成任何影響。因此，利用量化的資料，他比較了在熱浪中死亡的大約七百個受害者、以及其他數千個未受影響的人們在社會階級方面的差異，以考驗社會階級和死亡率之間的關係。

在許多質化研究中，我們通常會產生新的假設，並且會描述在少數個案裡的因果機制或過程的相關細節。讓我們再回頭看看前面提過的熱浪研究，在這項研究裡，克林能柏格（Klinenberg, 2002）考驗了關於階級與死亡率的現有假設。而當他仔細地審視了造成某些人死亡、而其他人卻未受影響的機制時，他也發展出數個新的假設。他得知高死亡率通常發生在飽受貧困和犯罪活動所苦的地區。男性的死亡人數高於女性，而黑人的死亡人數也高於拉丁美洲人或是白人。藉由親自走訪數個不同的低收入地區、以及直接訪談許多人的方式，他辨識出都市隔離的機制（mechanisms of urban isolation），而該機制可以解釋爲什麼相同社會階級的人們，在熱浪中的存活率會有如此巨大差異的原因。他檢視了年紀較大的黑人男性的社會情況，並發現地區環境是最爲關鍵的因果機制。他同時檢視了那些造成 1990 年代中期芝加哥地區的社會情境以及地區環境的外在力量。

量化研究及質化研究間的第四個差異，是這兩者各自有涇渭分明的「邏輯」以及進行研究的路徑。在量化研究中，我們所採用的是具有系統性的邏輯，並依循一種線性的研究路徑。而在質化研究中，邏輯則是形成於持續進行的實務之中，我們所依循的是一種非線性的研究路徑。在下一個部分，我們將逐一檢視這兩種不同的邏輯與研究路徑。

重建邏輯以及實務中的邏輯

我們學習以及討論研究的方式，通常會依循兩種邏輯的其中之一。[3] 這兩種邏輯總結了我們的研究策略的清晰程度、編碼程度、以及標準化的程度。在具體的研究中，我們通常會混合這兩種邏輯，但是在不同的研究中，這兩種邏輯的比例，會有極大差異。

重建邏輯（reconstructed logic）強調使用一種清晰明白的研究程序。重建邏輯是將進行實際研究時的許多混亂細節，加以「重新建構」或是重新敘述，成爲理想化的、擁有標準做法以及連貫的原則、術語、和規則的正式步驟。你可以把重建邏輯想成關於應該如何進行高品質研究的一種「純淨的模型」。依循這種邏輯來進行研究，就像是完全精確無誤地按照一份已印刷出版的食譜來烹調食物一樣。因此，進行簡單隨機抽樣（將在第七章中進一步討論）的方法，不但是直接明白的，也依循著一個清楚的漸進程序。

至於**實務中的邏輯**（logic in practice），則顯得較爲雜亂無章，也比較接近於進行研究時的實際具體作法。在實務中的邏輯所包含的建議，是來自於在具體實際生活中進行研究的實際活動，而非來自一組重新敘述過的、理想的規則。此種邏輯相當仰賴那些積極主動的、經驗豐富的研究者所共有的「主觀判斷」（"judgment calls"）以及「行業裡的訣竅」

重建邏輯 一種研究的邏輯，主要是將研究知識以及實務加以重新組織、標準化、系統化，使其成為明確的規則、正式的程序及技術。重建邏輯是量化研究的特徵。

實務中的邏輯 一種研究的邏輯，主要依據學徒制的模式，分享關於實務議題與特定具體經驗的隱性知識。實務中的邏輯是質化研究的特徵。

（“tricks of the trade”）。我們學習此種邏輯的最好方式，就是閱讀大量的研究與擔任研究者的學徒，並學習那些經驗豐富的研究者以非正式的方法所流傳下來的民間智慧。這彷彿是在沒有書面食譜的情況下烹調食物——在這邊加點原料、不停地攪拌，直到某些部分「看起來對了」、不斷地在烹調過程中調整火候等等，一直到我們煮出了某種特定的氣味或是味道爲止。

你可以在量化研究報告裡的研究方法部分，清楚地看到重建邏輯。對比之下，在質化研究的報告中，你可能無法看到任何關於研究方法的討論（此種情況在歷史比較研究中相當常見）、或是發現研究方法和一個特定研究的個人自傳式敘述融合在一起（此種情況在田野研究中相當常見）。質化研究並不會因爲缺乏標準方法而比較不具說服力。然而，對於新手而言，若想要精通質化研究，通常需要花費比較多的時間、並採取一種不同的思考方式。

線性以及非線性路徑

所謂的路徑只是一種代稱，用來指稱一連串要做的事情：例如你應該先完成什麼事、或是你曾經到過什麼地方、接下來會發生什麼事等等。你可以依循一條筆直的、許多人走過的、而且標記清楚的路徑。此種路徑沿途有著明顯的路標指示前進的方向，之前也已經有許多人走過。或者，你也可以選擇依循另一條路徑，這條路徑會蜿蜒曲折地帶領你進入幾乎沒有人去過的未知領域。此種路徑幾乎沒有任何標誌，所以當你在往前邁進時，可能會偏離路徑而走到路線的邊緣，而且還可能往後倒退了一些距離，之後再度往前推進。

當我們使用**線性研究路徑**（linear research path），我們依循一套順序固定的步驟，像是引導我們朝著單一方向往上移動的樓梯。藉由依循線性路徑，我們以一種直接的、狹窄的、以及筆直的方式，朝向結論移動前進。此種直接引導我們完成任務的路徑，是西歐與北美文化中的主流取向，而且非常廣泛地用於量化研究中。對比之下，**非線性研究路徑**（nonlinear research path）則要求我們在每個步驟連續地通過重重的關卡。我們可能會先往前移動、往後移動、甚至是往側邊移動，然後才能再度有所進展。因此，此種路徑比較像是螺旋狀的、而不是筆直的樓梯。我們會逐漸地往上移動，但卻是以緩慢並間接的方式在移動。隨著每一次的循環或重覆，我們或許會蒐集到新的資料，也得到新的洞見。

習慣於直接的、線性路徑的人，通常會對比較間接的、循環的路徑感到不耐煩。雖然非線性路徑並非意味著一種毫無章法的、定義不明的混亂狀況，但是循環路徑在表面上看來，似乎是沒有效率且不嚴謹的。而那些習慣於非線性路徑的人，則通常覺得被線性路徑悶住以及「困住」（“boxed in”）了。對於他們而言，線性路徑感覺起來比較不自然或是僵硬死板。他們認爲線性路徑會使他們無法自

線性研究路徑　以一種清楚的、合乎邏輯的、逐步的直線方式進展之研究，通常使用於量化研究當中。

非線性研究路徑　以一種循環的、反覆的、或是來來回回的模式進展的研究，通常被使用於質化研究當中。

然地發揮創意及隨機應變。

每一種路徑都有各自的優點。線性路徑的優點是具有邏輯性、易於遵循、且較有效率。至於非線性路徑，則可能在形成對整體環境的真正感受、捕捉意義的微妙層面、將資訊的分散部分整合在一起、以及轉換觀點等各方面極爲有效。每一種路徑都有其本身的專業紀律及嚴謹之處。線性路徑借用了自然科學中重視邏輯與精確性的取向，而非線性路徑則借用了人文學科裡的一些手法（例如，隱喻、類比、主題、動機、以及反諷等等），並適合於像是轉譯語言之類的任務。因爲在轉譯語言的過程裡，意義的幽微層面、微妙的含意、或是脈絡的區隔等等，都是相當重要的（關於以圖像方式來呈現兩種不同的路徑，請參閱圖 6.2）。

客觀以及誠信

我們都會盡量努力在研究活動中保持公平、誠實、真誠、以及不帶偏見的態度。然而，在創造知識的過程中，包括社會研究在內，我們也都可能出現帶有偏見、不誠實、或是不道德的時刻。這兩種主要的研究取向，各自以不同的方法來處理例如減少困難和確保誠實以及真誠研究之類的議題。

在質化研究中，我們通常會試圖獲取關於研究情境深入的、第一手的知識。因此，我們不會讓自己遠離我們所研究的人們或事件。獲取對於某個情境的深入瞭解，並非意味我們可以武斷地加入個人的意見、以草率的態度蒐集資料、或是選擇性地使用證據來支持我們的偏見。更確切地說，我們必須充分利用個人的深刻理解能力、內在的感受、以及生命的觀點來理解社會生活。我們「走在」介於親密與疏離之間的「微妙界線」（“walk a fine line”），並將正直及誠實視爲最重要的事物。有一些技術或許可以幫助我們順利地走在這條界線上，而其中之一是對於我們本身的看法、先入爲主的成見、以及之前的假定都保持高度敏銳的狀態。然後將這些看法、成見、或是假定都「放入括弧」（“bracket”），也就是暫時先懸置在一旁。如此一來，我們就可以不受這些事物的限制而看得更加清楚。這樣做並不是要去隱藏或否認我們所抱持的假定、觀點、以及價值觀，相反地，我們認爲最佳的作法就是承認這些事物的存在、並以坦白而開放的態度來面對它們。如此一來，我們就可以看出這些事物可能會以何種方式影響我們。在我們對於研究情境的參與、處理研究中的人們與任何浮現的相關議題等各方面，我們都盡力保持直率而坦白的態度。而且我們不只是在進行研究時要抱持這種直率而坦白的態度，在報告研究的發現時，也同樣要抱持這樣的態度。

個別研究者的個人開放態度以及誠信，對於質化研究而言是非常重要的。對比之下，在量化研究中，我們所強調的是中立和客觀。在量化研究裡，我們依賴複製原則、堅持標準化的程序、以數字來進行測量、並用統計技術分析資料。[4] 就某種方面來說，我們盡力減少、或是排除在量化研究中的人爲主觀因素。正如同波特（Porter, 1995: 7, 74）所主張的：

> 理想上，專門知識應該是機械化與客觀化的……以具體的特定技術為基礎……此種客觀性的理想既具有政治性質、同時也具有科學性質。客觀所

圖 6.2　表示線性以及非線性路徑的圖像

> 指的是法治、而不是人治，因而意味著將公共的標準置於個人的利益以及偏見之上。

在量化研究中的誠信議題，則反映了自然科學的取向。在量化研究中，誠信意味著使用明確且客觀的術語（像是以精確並中立的詞彙作出陳述）、使用已被廣爲驗證的標準技術、以及在進行測量時使用可以複製以及客觀的數字等等。

量化社會研究的檢核方式和自然科學的特徵相當一致：明確而且標準的程序、精確的數字測量、以及複製。對比之下，質化研究的檢核方式則比較依賴可靠的、可以信賴的研究者，以及研究者本身的誠信、自律、以及可信賴性。[5] 在質化研究中其他四種形式的檢核方式，與量化研究中的客觀程序或多或少是相當類似的。[6]

第一種檢核的方式，是研究者已經仔細地評估過各種不同形式的證據，並已經確認過證據之間的一致性。例如，一位田野研究者傾聽並且記錄下一位學生所說的：「史密斯教授把一塊橡皮擦丟到瓊斯教授身上」。這位研究者必須小心翼翼地思考這項證據，其中也包括考量其他人對於這起事件的說法。這位田野研究者也要尋找支持這起事件的證據，並檢查內部的一致性。研究者也會詢問這名學生對這起事件的資訊是否爲第一手資訊。換句話說，意即這名學生是否真的親眼目睹這起事件，而且必須瞭解這名學生的感受或是私利是否有可能導致他刻意說謊（例如，這名學生不喜歡史密斯教授）。

第二種檢核方式，則來自於在大多數質化研究中爲數眾多的詳盡筆記。除了關於

證據的詳盡逐字描述之外，其他的文件紀錄還包括了資料來源、研究者的評論以及引文、照片、影片、地圖、圖表、釋義、以及計數等等。數量龐大的資訊、豐富的多樣性、以及環環相扣並相互補充的文件紀錄，都有助於檢核研究與資料的真實性。

第三種檢核的方式，則是來自於其他的觀察者。雖然大部分的質化研究者都是獨自一人進行研究，但卻有許多其他人知道或瞭解研究中所提出的證據。例如，我們研究了在某個特定背景環境中的人們，而這些人們到今日依然健在。於是，其他的研究者就能夠走訪相同的背景環境，並且和相同的人們交談。我們所研究的人們也能夠閱讀研究的細節，並證實研究的內容、或對研究的內容提出問題。同樣地，歷史比較研究者會引述歷史文件、檔案來源、或是視覺資料。只要研究者提出精確的引述資訊，留下詳細的「審核蹤跡」（“audit trail”），那麼其他人就能夠查核這些參考資料，證實資料來源的真實性。

第四種檢核真實性的方式，則是來自於我們公開地發表研究結果的方式。在量化研究裡，當我們在撰寫研究報告時，必須嚴格遵守一種標準的格式。在其中，我們會詳盡地解釋我們如何依循著公認的程序來進行研究。我們會描述研究的每一個步驟，以圖表、圖形、或表格呈現量化資料，也會開放資料檔案讓其他人能夠再次分析。除此之外，我們也願意回答任何與研究有關的問題。而在質化研究內，我們無法在研究報告裡公開地展示或分享堆積如山的詳細筆記、訪談錄音、照片、或是原始的來源資料，因爲這些資料或記錄可能塞滿一整個房間！相反地，我們把相互關聯的細節「編織成網」（“spin a web”），並使用彼此緊密交互參照的資料。透過我們的寫作呈現，我們提供充分的結構與細節，以在讀者們的內心產生一種「身歷其境」（“I-was-there”）的感受。我們在提供豐富而具體的描述時，也以地圖、照片、逐字的引述等方式，加以輔助增補。如此一來，我們就能夠傳達對於某個背景環境的深入而詳盡的知識，也能夠在讀者心中建立一種共同的熟悉感受。一位技巧熟練的質化研究者，能夠善加利用堆積如山的實徵證據，而重新創造出曾經存在的視覺影像、說話聲、氣味、各種聲音、緊張感受、以及整體的氣氛。

事先規畫以及意料之外的研究問題

我們可以用許多不同的方式來開始進行一項研究，但通常第一步都是從選擇一個主題開始。[7] 然而，關於應該如何選擇一個適切的主題，其實並沒有一定的公式可循。不論我們是經驗豐富的老手、或是剛開始進行研究的新手，最好的方式就是挑選一個我們覺得有興趣的事物。選擇主題的方法有很多種（請參閱資訊箱 6.2 研究主題的來源）。我們或許會先從一個主題開始，但這個主題的範圍通常會太過龐大，而且也僅僅是一個起始點。因此，我們還須縮小範圍，使該主題轉變爲一個聚焦的研究問題。依據我們的研究主要是質化或量化研究，聚焦的方式也會有所差異。對於某些主題來說，這兩種取向的研究都相當適合。例如，如果我們想要研究貧困問題，我們可以檢視政府的統計資料、進行調查、進行民族誌田野研究、或完成一項歷史比較分析。然而，有些主題比較適合進行質化研究（例如，人們如何透過參與

資訊箱 6.2　研究主題的來源

1. *個人經驗*　你可以根據曾經發生在你身上、或是你知道的事物來選擇一個主題。例如，當你趁暑假在一間工廠打工時，當地的工會發動了一場罷工活動。雖然你對勞資雙方都沒有什麼強烈的認同感，但仍然被迫選邊站。你注意到緊張對峙的情勢逐漸升高，而資方和勞方對於彼此的敵意也越來越強烈。根據這次的經驗，你可以把工會或是有組織的勞工當作研究的主題。
2. *對媒體報導的事物的好奇心*　有時候，在你讀過一篇報紙或雜誌的文章、或是看完一個電視節目後，在心中會留下許多的疑問。你所讀到的內容在你心中造成了疑問，抑或使你想要複製看看別人的研究所發現的事物。例如，你在《新聞週刊》上讀到了一篇關於街友的文章，但其實你不太瞭解街友的狀況、他們為什麼會無家可歸、以及此種現象是否一直是社會問題等等。因此，你可以依據你的好奇心，把街友當作一個研究主題。
3. *某個領域中的知識現況*　能夠擴展知識範圍的新研究發現及理論，都能促進基礎研究的發展。當理論性的解釋被詳盡地加以闡述或擴展時，為了使整個領域能夠向前進步，就必須解答某些議題或是問題。隨著研究者逐漸辨識出一些新的議題、並針對這些議題進行研究，知識就能不斷地向前進展。例如，你讀到一篇文章，內容是關於人們對於死刑所抱持的態度，也發現大多數的研究指出，死刑的支持者基本上都認為犯罪者身上有一種與生俱來的邪惡特質。你能注意到目前為止沒有任何一個研究者檢視過，在那些教導此種關於內在邪惡特質觀念的宗教團體中的成員們是否也支持死刑，況且也沒有人具體地標示出這些宗教團體的地理位置。你對於這個領域的知識，產生了一個研究計畫的主題：不同地區的人們對於死刑以及宗教的看法。
4. *解決問題*　應用研究的主題通常是從一個有待解決的問題開始。例如，假設你是一名宿舍輔導員，而你份內的工作之一，是去幫助大一新生在彼此之間建立友誼。因此，你所遇到的問題，就可以轉化成為研究的主題，也就是大學新生之間如何建立友誼的過程。
5. *社會獎金*（*social premiums*）　這是由辛格頓與同事（Singleton and colleagues, 1988: 68）所提出的一個詞彙，意指某些主題很「熱門」（"hot"）、或是能夠提供機會。例如，你讀到一些報導，因而得知目前有許多單位提供大量資金，讓研究者可以針對療養院進行相關的研究，但是對此種研究感興趣的人卻不多。由於你正好需要一份工作，因此，你可以試著以療養院當作研究主題，並且申請相關的資金補助。
6. *個人的價值觀*　有些人極為忠誠地堅守一些宗教、政治、或是社會價值觀。例如，你強烈相信種族平等的理念，而每當你聽到有關於種族歧視的消息時，就會覺得義憤填膺。由於你具有強烈的個人信念，因此你可以試著把種族歧視當作研究的主題。
7. *日常生活*　在日常生活當中的古老俗諺、小說、歌曲、統計數據、以及其他人所說的話（尤其是那些與你意見不同的人），都可以找到潛在的研究主題。例如，你曾經聽說過，主場優勢在籃球比賽中是非常重要的。因此，你可以把主場優勢當作是你的研究主題。

青少年次文化的方式來重塑自我認同），而其他的主題則比較適合進行量化研究（例如，過去五十年來對死刑的相關輿論轉變；以及個人對於這項議題的意見，是否會

受到相關議題的意見所影響、或者受到某些主題在新聞媒體上的曝光率所影響）。

大多數的質化研究都是從一個模糊的、或定義鬆散的主題開始。在研究進行過程中，具體明確的主題會慢慢地浮現出來，並可能因爲新證據的出現而改變研究方向。在第五章開頭部分所提及的范卡德胥的研究（Venkatesh, 2008），正是此種情況的例子。原本一開始時，他想要研究的主題是在都市貧民區住宅計畫中的貧困情形，但後來的研究焦點卻轉移到販賣毒品的幫派。當我們在蒐集資料時，我們仍然持續地聚焦在一個特定的研究問題上面。范卡德胥逐漸把幫派活動這個主題聚焦爲更尖銳的問題：在低收入住宅計畫中的幫派份子，是以何種方式、以及爲什麼維持一種地下經濟狀態，並且還爲住宅計畫中的住戶提供保護與服務？

質化研究所具備的彈性，鼓勵我們在整個研究過程中，持續地進行聚焦。某個意料之外的問題，可能只有在研究的過程中，才會變得越來越清楚。在我們蒐集了一些資料並開始進行初步分析之後，我們才能聚焦且更明確地定義研究問題。在許多質化研究裡，只有當我們完全地沉浸在資料裡之後，那些最重要的議題和最有趣的問題，才會變得清楚可見。我們對於意料之外的想法、資料、以及議題，必須保持著開放的態度。在研究進行的初期階段，我們應該要定期地重新評估我們的焦點，並且隨著準備好改變研究方向以及依循新證據的線索。同時，我們也必須練習自制以及紀律。如果我們無止盡地一直在改變研究的焦點，我們可能會永遠也無法完成一項研究。正如同大多數的事物一樣，取得某種平衡點確實是必要的。

典型的質化研究問題包括：某個特定的狀況或社會情境是如何產生的？人群、事件、以及狀況是以何種方式維持某種情境的？此種情境在改變、發展、或是結束時，會經歷哪些過程？另外一種類型的研究問題，則是試圖確認現存的信念或假定〔例如，麥可德莫特（McDermott, 2006）曾進行過一項關於亞特蘭大及波士頓地區的勞動階級社區的研究。那麼，在與其他種族的人們相處時，南部與北部的白人在行爲上是否有所差異，就像麥可德莫特在研究中所發現的一樣？〕至於最後一種類型的研究問題，則是試圖去發現新的想法。[8]

進行量化研究時，在我們完成研究計畫之前，我們會先完成一個具體的規畫步驟。也就是把一個主題的範圍縮小，成爲一個聚焦的研究問題。在發展出一個可以進行考驗的假設過程中（稍後將再進行討論），將問題聚焦是其中一個重要步驟，因爲這麼做可以在你開始蒐集任何資料前，爲你的研究設計提供引導。[9]

在質化研究裡，我們可以利用資料來幫助我們縮小焦點。在量化研究中，我們必須在沒有任何資料的情況下進行聚焦。而且必須使用其他的技術。在挑選了一個主題之後，我們必須問自己：我們對於這個主題最感興趣的部分是什麼？對於一個我們所知甚少的主題，我們首先必須閱讀與這個主題相關的一些研究，以獲得相關的背景知識。閱讀研究文獻，可以激發許多和如何聚焦研究問題相關的想法。

在大多數的量化研究中，研究問題所指的是在少數幾個變項之間的關係，而這意味著，當我們試圖把主題聚焦成研究問題時（請參閱資訊箱 6.3 將主題聚焦成爲研究問題的技巧），應該把所有的變項清楚地列出來。例如，像是「什麼是導致離婚的因素？」這樣的問題，就不是一個好的研究問題。比較好的問法，應該是「結婚

資訊箱 6.8　將主題聚焦成為研究問題的技巧

1. *檢視文獻*　已出版的文章是尋找研究問題的絕佳來源，因為這些文章通常都具有適當程度的具體性，也會在下列各方面幫助我們找到適合的研究問題：
   a. 完全精確地複製一項先前的研究，或是稍微做一些變化。
   b. 探索在先前的研究中所找到的意外發現。
   c. 依循作者在文章結尾部分對於未來研究方向的建議。
   d. 將現存的解釋或理論延伸到新的主題或是環境。
   e. 挑戰該項研究的發現，或者試圖拒絕一項關係。
   f. 具體說明某個中介的過程，並且思考任何具有連結性的關係。
2. *與其他人討論想法*
   a. 向那些熟知你的研究主題的人們詢問一些問題，並請教他們的意見和想法。
   b. 找到那些在這個研究主題上面和你意見相左的人，並且和他們討論可能的研究問題。
3. *應用於特定的脈絡*
   a. 將焦點集中於一個特定的歷史時期或時間區段。
   b. 將主題縮小、集中在一個特定的社會或是地理單位上。
   c. 思考哪些次團體或是哪些類別的人們／單位和你的研究有關，以及在這些人事物之間是否存在著任何差異。
4. *明確定義研究目標或是想要獲得的結果。*
   a. 你所提出的研究問題適合於探索性研究、解釋性研究、或是描述性研究？
   b. 你的研究屬於應用研究或基礎研究？

的年齡是否與離婚有關？」。在第二個研究問題裡有兩個變項：結婚年齡、以及是否離婚（請參閱範例 6.2 不佳以及良好研究問題的範例）。

個人經驗可以幫助我們找到研究的主題。或許就和佩吉爾（Pager, 2007）所進行的研究一樣，我們也會因爲個人的經驗，而把剛被釋放出來的犯人當作研究主題。我們可能會在數十本著作與數百篇文章裡，讀到一些有前科的犯人又再度因爲犯案而入監服刑的情形、以及有關於緩刑的問題。一個聚焦的研究問題可能如下：「一個具有非暴力犯罪記錄的人，是否會比一個沒有任何前科的人更難以找到工作？」這個研究問題在犯罪記錄的類型、及犯罪記錄所可能造成的影響方面，是比較具體而明確的。這個問題把焦點集中於下列兩個變項上面：某個人是否有犯罪記錄、與這個人是否能找到工作。一種常見的研究問題類型，是去詢問在數個因素裡，哪一個因素會對結果造成最明顯的影響。就像佩吉爾一樣，我們或許會問，種族類別（黑人以及白人）、和一個人是否有犯罪記錄，會對此人找到工作的機會造成什麼樣的影響？種族的不同是否會造成任何差異？曾經入獄服刑是否會造成差異？在對工作機會的影響方面，這兩個因素是各自獨立運作、互相抵消、還是互相增強？

範例　6.2

不佳以及良好研究問題的範例

不佳的研究問題

無法以實徵方式加以檢驗，非科學性的問題

- 墮胎應該合法化嗎？
- 執行死刑是對的嗎？

過於廣泛的主題，不是研究問題

- 酒精及藥物濫用的治療
- 性能力與老化

只是變項的組合，不是研究問題

- 死刑與種族歧視
- 都市衰敗與幫派

過於模糊以及模稜兩可

- 警方對於違法行為有何影響？
- 如何預防兒童受虐？

必須要再更加明確一些

- 兒童受虐的案件是否增加了？
- 貧窮如何影響兒童？
- 成長於貧窮之中的兒童，會經歷什麼其他兒童所不會經歷的問題？

良好的研究問題

探索性的問題

- 在過去十年之中，威斯康辛州是否出現了不同形式的虐待兒童案例？

描述性的問題

- 在父母曾經離婚的家庭裡，是否會比家庭完整、父母從未離婚的家庭，更容易出現以暴力、或在性方面虐待兒童的情況？
- 在貧困家庭中長大的兒童，是否會比其他兒童更容易出現醫療、學習、社會與情緒調適方面的困難？

解釋性的問題

- 在離婚過程中所經歷的情緒不安全感，是否會增加離婚父母在肢體上虐待其子女的可能性？
- 缺乏進行預防治療的資金，是否為造成貧窮家庭的兒童經歷其嚴重醫療問題的主要原因？

我們也必須詳細說明我們能夠把一個研究問題的答案加以推論的**母群**（universe）所有的研究問題及研究都可以應用於某種類別的人們、組織、或是其他單位。這裡所謂的母群，是指研究問題所涵蓋的所有單位、或我們。可以加以推論的單位。例如，在佩吉爾（Pager, 2007）的研究裡，他所研究的單位

母群　某項關係或假設所涵蓋、解釋的完整類別、或是所有的單位。

是個人，更明確地說，是年輕的白人與黑人男性。因此，我們可以將他的研究發現加以推論的母群，就包括了所有屬於這兩個種族類別的二十幾歲之美國男性。

當我們把一個主題聚焦凝鍊爲一個研究問題、並設計好一項研究之後，我們還必須考量實際的限制。設計一個完美無缺的研究計畫，是一項有趣的學術活動，但如果我們想要實行一項研究，就必須在研究設計中，把實際上的限制也列入考量。主要的限制包括時間、成本、取得資源的管道、管理機構的同意、倫理議題、以及專業知識等等。如果我們只有五個禮拜的時間、而每個禮拜只有十個小時可用來進行研究，但要解答的研究問題卻需要花上長達兩年的時間，我們就必須把研究問題縮小，以符合實際的限制。

時間一直都是必須列入考量的重要限制，然而，要正確評估一項研究所必須花費的時間，卻是一件極爲困難的事。具體的研究問題、使用的研究技術、研究的複雜程度、以及我們計畫要蒐集的資料數量以及種類，都會影響到該項研究所必須花費的時間。若你想要得到關於研究所必須花費的時間的良好評估，經驗豐富的研究者將會是你最佳的請益對象。

成本是另一個限制，而我們有時會因爲牽涉到高昂的費用，因而無法回答某些研究問題。例如，假設我們的研究問題是，球迷們是否會因爲支持的隊伍在球季中大勝，而對隊伍的吉祥物產生強烈正向感受、或因爲支持的隊伍在球季中輸球而對吉祥物產生負面的感受。若是要花費長達十年的時間去調查美國境內所有不同球類運動的所有隊伍，以便於我們回答這個問題的話，這項研究所需花費的時間和金錢就會非常地龐大且驚人。因此，我們可以把焦點縮小，只去研究某一項球類運動（例如，橄欖球）在大學裡所舉辦的比賽，以及在三個球季中僅僅四所大學的學生球迷們。就和時間這個限制一樣，經驗豐富的研究者，比較能夠正確地評估進行某項研究所需花費的成本。

取得資源的管道與方式也是一種常見的限制。資源包括專業知識、特殊設備、以及資訊等等。例如，如果我們的研究問題是關於在許多國家內的竊盜率，以及家庭收入的話，這種研究問題幾乎無法回答。因爲在許多國家裡，根本沒有人蒐集過關於竊盜以及收入的資料。或者，此種資料根本就無法取得。其他的研究問題則可能要徵得管理機構的同意（例如，當你要調閱某人的就診記錄時）、或是牽涉到違反基本的倫理原則（例如，對某個人說謊、並讓此人陷於險境）。我們身爲研究者所擁有的專業知識或是背景，也會成爲一項限制。在回答某些研究問題時，會牽涉到使用某些蒐集資料的技術、統計方法、對於某種外國語言的知識、抑或是一些我們可能並未擁有的技巧。除非我們能夠獲得必要的訓練、或能以付費方式得到他人所提供的相關服務，否則，某些研究問題可能顯得不切實際。

簡而言之，質化研究以及量化研究擁有許多共同之處，但是在數個設計的議題上卻有所差異，例如像邏輯、研究路徑、檢核的模式、產生研究問題的方式等等（請參閱表 6.1）。除此之外，不同的研究取向會有不同的「語言」，也會著重截然不同的研究設計特點。我們將在下一個部分詳細地討論這些議題。

表 6.1　量化研究與質化研究

量化研究	質化研究
研究者所考驗的是在研究一開始的時候所陳述的假設	一旦沉浸在資料裡之後，研究者便會去捕捉並發現意義
概念是以明確的變項形式呈現	概念是以主題、動機、概論、以及分類的形式呈現
在開始蒐集資料之前，就已經以系統化的方式建立好測量工具，並將測量工具加以標準化	測量工具以一種特別的方式建立，且經常針對個別的情境或是研究者而設計
資料的形式是由精確測量所得到的數字資料	資料的形式是從文件、觀察、以及謄寫本中，所得到的文字與意象
理論主要強調因果關係，而且是演繹的	理論可以是因果的或非因果的，而且通常是歸納的
程序為標準化，且經常被複製	研究程序是個殊的，而且幾乎不能複製
進行分析的方式是利用統計、表格、或是圖表，並討論它們與假設之間所呈現的關係	進行分析的方式是由證據中提取出主題或概論，並組織資料以呈現一幅一致且連貫的圖像

質化設計議題

個案以及脈絡的語言

大部分的質化研究都會涉及一種個案與脈絡的語言、採用*拼湊*（*bricolage*）（將在本章稍後部分討論）、審視在其社會脈絡裡的社會過程以及個案、並且研究在特定社會文化環境裡的詮釋或意義。我們從多種不同的觀點來審視社會生活，並解釋人們以何種方式建構認同。只有在非常罕見的情況下，我們才會使用變項、考驗假設、或創造數字形式的精確測量工具。

大多數的質化研究都是建立在下面這個假定之上：社會生活的某些領域在本質上是質化的。由於這個原因，質化資料並非不精確的或是有缺陷的。相反地，質化資料其實非常有意義。我們並非試圖把流動、活躍的社會生活轉變成爲變項或數字。更確切地說，我們是從我們所研究的人們那裡，借用他們的想法及觀點，並將這些想法或觀點，放置於一個流動的自然情境之中。我們所檢視的不是各種變項，而是動機、主題、區別、觀點。我們的方法通常是歸納式的，而且仰賴於某種形式的*紮根理論*（*grounded theory*）（我們已在第三章當中，討論過紮根理論的相關概念）。

質化資料從表面上看或許是軟性的、模糊而不具體的、而且是難以捉摸的。然而，這並非意味著我們因此無法捕捉質化資料。我們可以利用下列這些方式來蒐集質化的資料：詳細記錄真實的事件、記錄真實的人們所說的話（像是字詞、手勢、以及音調等等）、觀察特定的行爲、檢視書面的文件、以及研究視覺意象等等。上述

這些資料都是社會世界中明確而具體的面向。當我們仔細地審視人們與社會事件的照片或是影片時，我們就正觀看著「硬性」的實體證據。[10] 與在量化研究中所使用的關於態度、社會壓力、智力等等的數字測量工具相較之下，這些質化證據同樣也是「硬性」的以及實體的證據。

紮根理論

在質化研究裡，我們可能在蒐集資料的過程中發展出理論。而此種主要是歸納式的方法，意味著我們是從資料裡建立理論、或是將理論紮根於資料之中。紮根理論增加了彈性，並允許資料與理論兩者之間進行互動。而此種過程也有助於我們對於意料之外的發現保持開放態度。換句話說，如果我們發現了某項新奇而令人興奮的事物時，即使是在研究計畫進行的過程當中，我們仍然可以改變研究方向，甚至放棄原來的研究問題。[11]

我們是以進行比較的方式來建立理論的。例如，我們觀察了一個事件（像是一名警察面對一名被攔下的超速駕駛人），接著，我們可能會仔細思考一些問題，並且尋找一些類似以及差異之處。當我們看著一名警察時，我們可能會問：警察在執法之前，是否總是會先通報車子的車牌號碼？在通報車子的位置之後，警察是否會要求駕駛人下車，還是有時會隨意地走向車子，並與坐在車裡的駕駛人交談？當我們在蒐集資料與形成理論這兩者之間往返時，新的理論性問題可能會逐漸成形，並暗示了未來的觀察方向。如此一來，我們就可以修正新資料，以回答那些經由思考先前的資料而形成的理論性問題。

在紮根理論裡，我們由特定而具體的觀察結果，逐步進展到能夠組織觀察資料的較大概念，然後再繼續建立能夠連結這些概念的原則或是主題。與其他形成理論的方法相較之下，紮根理論通常比較不那麼抽象，也比較接近具體的觀察結果或特定的事件。此種利用歸納法由資料發展成理論的方式，使資料和理論之間產生了強烈的連結。然而，這同時也可能是一項缺點，因爲如此一來，若要在許多互異的情境之間連結各種概念以及原則，就會變得相當困難，並且可能使創造更廣泛抽象知識的概念發展過程，變得較爲緩慢。爲了要抵消此項缺點，我們必須熟知在其他的研究中所發展出來的概念及理論，以在適當時運用這些共享的概念，並記下任何的相似處及相異處。如此一來，我們就能建立跨研究的相互連結性，發展出更具推論性的知識。

脈絡是關鍵

在質化研究裡，我們通常會強調社會脈絡，因爲社會行動、事件、或陳述的意義，都必須由其發生的脈絡來決定。如果我們把一個事件、社會行動、或是對話從其發生的脈絡當中抽離出來，就很容易扭曲其意義、並改變其社會重要性。

社會脈絡包括時間脈絡（某件事物發生的時間）、空間脈絡（某件事物發生的地點）、情緒脈絡（對於發生某件事的感受）、以及社會文化脈絡（某件事物發生的社會情境及文化氛圍）。例如，假設某個社會活動（一場牌局、性行爲、或是意見不合

的情況）在深夜時分發生於一座大都市中一條低收入地區的街上，而在一個種族與經濟極端不平等的文化氛圍裡，這是個易於發生藥物濫用、恐懼與憤怒、暴力犯罪、以及賣淫等等行爲的背景環境。假設同樣的活動於大白天時發生在一個富裕的市郊社區中一棟豪宅的後院裡，在一種放鬆與悠閒的社會情境中，充滿了信任和情緒的親密性，並且在既定的富裕以及特權的文化氛圍裡。由於脈絡會明顯地影響活動的社會意義，因此，由於具有不同的脈絡意義，相同的活動或是行爲，就可能帶來完全不同的結果。

在量化研究中，我們很少會認爲脈絡具有任何重要性可言。我們通常會去除脈絡，把脈絡視爲「混亂的」或只是「干擾」而已，只專注於精確的計數或數字測量結果之上。因此，質化研究認爲不可或缺的重要事物，在量化研究中可能被視爲只是不相干的干擾而已。例如，假設某項量化研究計算了在不同的時間點或文化中的投票率，那麼質化研究者可能會思考在不同的脈絡裡，投票所代表的意義是什麼。質化研究者可能會根據發生某種行爲的社會脈絡，以不同的方式來看待相同的行爲（例如，投票給某位總統候選人）（請參閱範例 6.3 脈絡對於意義的重要性的範例）。

脈絡不僅止於社會事件、行爲、以及陳述而已，同時也包括實體的物體。一把手槍可能是一件藝術作品、某種娛樂活動的一部分、從事暴力犯罪的關鍵要素、一項證明父母不負責任的證據、一種協助自殺的工具、或是一種維護社會和平與保護社區的手段。每一種意義都取決於脈絡而定，因此，若不把周圍的脈絡列入考量範圍，我們就無法正確地爲一個物體賦予意義。

拼湊

所謂的*拼湊者*（*bricoleur*）（譯註：此字原爲法文，意指擅長利用手邊現有的材料來敲敲打打、從事各種零碎工作或是完成任務的人。此種人雖然通曉許多領域、也具備許多技巧與能力，但通常是雜而不精。除了譯爲「拼湊者」之外，亦可譯爲「修補匠」、「藝術工匠」、「零雜工」、「能工巧匠」等等），是指一個熟悉各種不同領域的人，而且也能夠利用各式各樣的來源，並且用手邊任何可以得到的東西，拼湊出自己所要的東西。[12] **拼湊**（bricolage）的技術包括以自己的雙手來進行工作，並用一種實際可行的、技巧熟練的、以及有創造力的方式結合各種零星瑣碎的物品，以完成一項工作或任務。一名成功的*拼湊者*擁有對於材料的深入知識、深藏不露的技巧、以及彈性地結合或創造物品的能力。典型的拼湊者通常是具有高度創造能力而技巧熟練的工匠、修理工、或是萬事通（jack-of-all-trades）。

進行質化研究時，研究者必須視情況的需要，而用到許多不同技巧、材料、以及方法。此種情況通常發生在我們無法預測該使用什麼技巧、材料、或是方法時。而此種混合不同來源的材料、應用不同的方法、以及把零碎東西組合成一個整體的過程，正可比擬爲一個利用手邊任何現有的材料而創造或修補許多事物的能工巧匠。

拼湊 一種即興創作的活動，主要是善加利用手邊可以得到的各種不同材料，並以一種具有創意的方式使用這些材料，以完成一項實際的任務。

範例　6.3

脈絡對於意義的重要性範例

「在全國性選舉中投票」這件事，在不同的脈絡之下會有不同的意義：

1. *一黨專政的政體、沒有其他候選人，而且人民依法必須投票。*未參與投票的人將會被警方記下姓名，並被懷疑是反政府的破壞份子。這些人可能會因為沒有投票被處以罰鍰、甚至可能丟掉工作。
2. *一個處於叛軍與執政者之間產生激烈暴力衝突的國家。*其投票行為充滿了危險，因為雙方中任一方的武裝士兵都可能射殺那些他們懷疑背叛己方的人。投票的結果將會賦予其中一方權力，並戲劇性地重建整個社會。任何年滿十六歲的人都具有投票權。
3. *人們可以在十幾個不同的政黨之間作出抉擇，而且這些政黨擁有大致相等的實力，各自代表非常不同的價值觀與政策。*每一個政黨都有相當大的組織，也擁有自己的報紙、社交俱樂部、以及社區裡的動員人士。選舉日是國定假日，當天全國人民通通不必上班。在投票時，選民只要攜帶身分證件到許多地區性的投票所，就可以參與投票。投票本身無記名，年滿十八歲者即擁有投票權。
4. *只有那些擁有固定工作、而且年滿二十一歲的白人男性才可以參與公開的投票。*家庭、朋友、以及鄰居都可以看到彼此投票的過程。不同政黨間所提出的政策沒有明顯的差異，相反地，政黨是和種族或宗教團體綁在一起的，也是個人的種族—宗教認同的一部分。種族與宗教的認同是非常強烈的，此種認同會影響到一個人的居住地、工作地、以及結婚對象等等。在進行投票之前，會先舉行由種族和宗教團體所舉辦的大型遊行，以及時間長達一周的社區活動。
5. *某一個政黨獨攬大權，受到來自於一或二個非常小型、力量薄弱的其他政黨挑戰。*在過去的六十年間，這個獨大的政黨透過貪污、賄賂、恐嚇威脅的手段而把持了權力，並且得到社會中各種組織領導者的支持（例如，宗教團體、教育機構、企業、工會、以及大眾媒體等等）。任何人在政府機構中的工作（例如，每一名警察、郵局員工、學校教師、以及清潔隊員等等），都由執政的政黨來決定。
6. *有兩個政黨可讓人民選擇，但是這兩個政黨之間幾乎沒有任何差別。*人民在選擇候選人的時候，主要是依據電視上的廣告來作判斷。候選人接受來自於富裕的人們或是強大組織的政治獻金，並用這些金錢刊登廣告。投票只不過是意義模糊的公民義務，沒什麼人認真看待投票這件事。選舉是在工作日當中舉行的。如果想要投票，人們必須符合許多條件，而且在選舉日之前的數個星期就必須先進行登記。最近才移民的人、任何曾經犯罪而被逮捕過的人，都被禁止參與投票。

個案以及過程

我們可以把所有的實徵社會研究區分爲兩大類：*個案研究*（*case study*）（只有一個或少數的個案）、或是*跨個案研究*（*cross-case*）（包含了許多不同的個案）。[13] 大多數的質化研究都是採取一種「個案導向的方式，把個案、而非變項當作最重要的焦點」（Ragin, 1992a: 5）。因此，我們會檢視少數幾個個案的許多面向，此種針對少數個案進行密集、深入研究的作法，取代了在量化研究中對眾多個案進行廣泛的、表面層次研究的典型作法。通常，個案導向的分析所強調的是在「混亂」的自然環境中所同時發生的偶然事件（換句話說，就是在同一時間與地點中，許多特定因素以及事件同時發生的狀況）。量化研究的典型作法，是對數量龐大的個案進行精確的

測量，但在質化研究中，我們不採取此種作法，而是針對少數的個案進行研究，以獲得深入的知識及細微的洞見。

針對少數個案所進行的研究，通常會產生複雜的解釋或詮釋，而其呈現的形式則是一個逐漸展開的情節，或一個具有敘事性質的故事，而其內容則與特定的人們或是具體事件有關。此種情況使得時間的流逝成爲構成整體解釋中一個不可或缺的部分。通常，事件的次序會變成強調的重點：什麼事情是先發生的、次發生的、之後發生的，以此類推。此種把焦點放在過程上的作法，有助於揭露某項議題演變的過程、某個衝突出現的來龍去脈、或是某種社會關係形成的經過。

詮釋

進行*詮釋*（*interpret*）意味著爲事物賦予重要性或一致的意義。在量化研究中，意義來自於使用數字（例如，百分比例、或是統計係數），而我們會去解釋這些數值資料與假設之間的關係。質化研究很少使用到數字表格，而在呈現資料時，唯一可能會用到的視覺呈現方式，或許只有地圖、照片、或顯示出各種概念間的關係圖示而已。在討論各種概念所具備的重要性時，我們會把資料交織於討論之中。而資料是以文字的形式呈現，包括引文或對於特定事件的描述等等。任何數字形式的資訊，都只是用來補充文字證據的不足之處。

質化研究賦予資料意義、把資料加以轉譯、或使資料變得明白易懂。我們由研究對象所擁有的觀點開始，再找出他們觀看世界以及定義情境的方式。我們試著瞭解各種事件、行爲、以及活動對於研究對象的意義。爲了要進行質化的詮釋，我們首先必須瞭解各種事物對於研究對象所產生的意義。[14]

人們在創造出社會活動與行爲時，背後都有進行這些活動及行爲的個人理由或動機，而這就是**第一級詮釋**（first-order interpretation）。當我們發現並重新建構第一級詮釋時，就會演變成**第二級詮釋**（second-order interpretation），這是因爲我們是以局外人的身分發現曾經發生的事情。在第二級詮釋裡，我們會從資料中引導出隱藏的一致性或是意義感。意義只有在與一組更大的、其他的意義相互參照時，才能發展出來，而不能憑空產生。在第二級詮釋中，我們把被研究的人類行動放入相關的「行爲流」（"stream of behavior"）或是事件之中，也就是該行動的情境脈絡。

假如我們採取一種非常嚴格的詮釋取向，我們或許會在第二級詮釋就停止了。也就是說，在我們瞭解某種行動對於研究對象的重要性與意義後，就不再繼續深究下去。然而，大部分的質化研究者還會更深入地探究，因爲他們想要針對第二級詮釋進行更進一步的推論，或把第二級詮釋與一個理論或是更廣泛的知識連結在一起。如此一來，他們就邁向了一個更廣闊的詮釋層次：**第三級詮釋**（third-order interpretation）。換句話說，研究者會爲資料賦予一般理論的重要性。

第一級詮釋　由被研究的人們的觀點所得到的詮釋。

第二級詮釋　由進行研究的研究者的觀點所得到的質化詮釋。

第三級詮釋　由研究報告的讀者所提出的質化詮釋。

由於質化研究的主要目的及結果，通常是詮釋在情境脈絡中的社會意義。因此，你必須記住，這三個步驟或詮釋的層級，有助於爲你提供一個組織研究過程的方法。

量化設計議題

變項以及假設的語言

變異以及變項 簡單地說，一個**變項**（variable）就是一個會變化的概念。在量化研究中，我們所使用的是一種變項以及變項間關係的語言。

在第三章裡，我們曾經討論過兩種類型的概念：其中一種是用來指稱某一種固定的現象（例如，科層體制的理想型），而另一種則會在數量、強度、或是總量等各方面產生變化（例如，接受教育的程度）。變項正是上述第二種類型的概念，並且是測量此種概念的指標。

一個變項必須有二個以上的數值，而且一旦我們開始注意到變項的存在，就會發現變項幾乎是無所不在的。例如，性別就是一種變項，有兩種可能的值：男性或女性。婚姻狀態也是一種變項，而這個變項的值可能是未婚、已婚、離婚、或寡居。犯罪的形式也可以是一種變項，其可能的值包括搶劫、闖空門、竊盜、謀殺等等。家庭收入也是一種變項，其數值的範圍則可以從零到數十億元之多。一個人對於墮胎的態度也是一種變項，而其可能範圍則從強烈贊成墮胎合法化，到強烈認同胎兒生命的神聖性，因此極力反對墮胎等等。

一個變項的數值或是類別，就是它的**屬性**（attributes）。變項和屬性很容易被混淆在一起，而之所以會產生此種混淆的情況，是因爲一個變項的屬性本身也可以是一個獨立的個別變項，而且在定義上只有稍微變化。此種情況來自於會產生變化的概念之間的區別，與在這些概念之間的不同條件。例如，「男性」並不是一個變項，因爲這個詞所描述的是性別這個變項中的一個類別。也就是說，「男性」是「性別」這個變項的屬性。然而，另外一個相關的概念，「男子氣概的程度」，則是一個變項。這個變項所描述的是在一個文化中，對於某一組與男性化概念有關的信念、傾向、以及行爲的依附強度。同樣地，「已婚」也不是一個變項，而只是「婚姻狀態」這個變項的屬性。而一些相關的概念，像是「結婚年數」或是「對於婚姻承諾的程度」，則都是變項。上面提到的第三個例子裡，「搶劫」並不是一個變項，而是「犯罪型式」這個變項的屬性。然而，「搶劫次數」、「搶劫率」、「一次搶劫所獲得的物品數量」、以及「搶劫的型式」等等概念，都因爲本身有所變化、而且具有一定範圍的值，因此都可以被當作變項。

在量化研究中，我們會利用變項的語言，把所有的概念加以重新定義。正如同上述關於變項和屬性的例子所顯示的，此種重新定義的過程，通常只需稍微修改原本的定義就可以了。如同在第三章中

變項　一個概念或此概念的實徵測量方式，可能有多重不同的值。

屬性　一個變項的類別或是水準。

所提到的，概念是理論的基石，因爲概念可以組織我們對於社會世界的思考方式。具有謹慎定義的清楚概念，在理論中是不可或缺的。

變項的類型　由於我們的焦點在於變項間的因果關係，因此我們通常會先從一個結果開始，然後回頭尋找造成此種結果的原因。我們可以依據變項在因果關係或因果鍊中的位置，把變項分成不同的類型。所謂的**自變項**（independent variable），就是一種原因變項、或是一種作用於其他事物的力量或條件。而本身是由另一個變項所造成的效果、結果、或產物的變項，就被稱作**依變項**（dependent variable）。自變項「獨立於」先前對它產生作用的任何因素，而依變項則會受到自變項這個原因所影響。

要決定一個變項到底是自變項還是依變項，並非總是件容易的事。以下兩個問題可以幫助我們確認某個變項是否爲自變項。首先，這個變項在時間上是否比其他變項更早出現？自變項出現的時間必須早於任何其他類型的變項。第二，如果兩個變項同時發生，其中一個變項是否會對另一個變項造成影響？自變項會對其他的變項造成影響、或是對其他的變項產生作用。我們通常會以依變項表達研究主題與研究問題，因爲依變項通常是我們想要解釋的現象。例如，在檢視美國德州達拉斯地區犯罪率上升的成因時，我們通常會把達拉斯地區的犯罪率當作是依變項。

在一個簡單的因果關係裡，僅僅需要一個自變項和一個依變項。而第三種變項的類型，也就是**中介變項**（intervening variable），則會出現在比較複雜的因果關係中。中介變項出現在自變項和依變項之間，因而有助於顯示存在於自變項和依變項兩者之間的連結或機制。如同在第三章中已經提過的，知識的進展不只是仰賴詳細記錄各種因果關係而已，也有賴於具體指明造成此種因果關係的各種機制。就某方面而言，中介變項相對於原本的自變項，就像是一個依變項，對於原本的依變項，則像是一個自變項。

例如，法國的社會學家涂爾幹曾經發展出一個關於自殺的理論，這個理論詳細說明了在婚姻狀態以及自殺率之間的因果關係。涂爾幹找到一些證據，證明了已婚者的自殺可能性比單身者還低。他相信這是因爲已婚者有比較高的社會整合（social integration）（也就是一種歸屬於某個團體或家庭的感覺）。他認爲造成某種自殺形式的主要原因，正是人們缺乏一種歸屬於某團體的感受。因此，他所提出的自殺論可由三個變項所組成的一種關係來加以重新敘述：婚姻狀態（自變項）會影響社會整合的程度（中介變項），而社會整合的程度則會影響自殺行爲（依變項）。詳細說明整個因果鍊，可使理論中的各種連結變得更爲清楚，也能夠幫助研究者考驗複雜的解釋。[15]

自變項　一種變項的類型，此種變項會在一個因果假設中對一個依變項產生作用或是結果。

依變項　在一個因果假設中，由一個自變項所引起的效果或是結果的變項。

中介變項　一種在邏輯上或時間上出現於自變項之後、依變項之前的變項，而且自變項和依變項之間的因果關係，是透過此種變項來運作的。

簡單理論只有一個依變項和一個自變項，複雜理論則可能包含了十幾個變項，其中有許多不同的自變項、中介變項、及依變項。例如，一個有關犯罪行爲（依變項）的理論指出四個自變項：個人的

經濟困境、容易犯罪的機會、身爲一個默許犯罪行爲的偏差次團體的成員、以及缺乏對於犯罪行爲的處罰。一個多重原因的解釋通常會明確指出哪一個自變項所造成的效果最爲顯著。

一個複雜的理論性解釋通常會包含一連串環環相扣的中介變項。例如，家庭瓦解造成兒童自尊降低、自尊降低會引起憂鬱、憂鬱會造成學校成績不佳、學校成績不佳則減少找到好工作的機會、而找不到好工作則導致成年後的收入較低。這一連串的變項分別是家庭瓦解（自變項）、兒童時期自尊低落（中介變項）、憂鬱（中介變項）、學校成績（中介變項）、工作機會（中介變項）、以及成年後的收入（依變項）。

即使探討相同議題的兩個理論，也可能在自變項的數量方面有所差異。除此之外，不同的理論可能在自變項及依變項兩者上看法一致，但在中介變項或因果機制方面意見分歧。例如，有兩個理論都提到家庭瓦解會導致成年後的收入降低，但兩者所提出的原因卻各自不同。其中一個理論主張，家庭瓦解會鼓勵兒童加入偏差的同儕團體，而此種團體並不認同工作和節儉等等社會規範。另外一個理論則強調家庭瓦解會造成兒童時期的憂鬱以及學業表現不佳。在這個理論裡，憂鬱與教育程度不足則會直接導致較差的工作表現以及較低的收入。

在一項研究裡，我們通常只會考驗因果鍊中的一個或是少數的部分。例如，一項檢視六個變項的研究計畫，可能是從一個有二十四個變項的大型複雜理論中篩選出來的。與一個大型理論間的明確連結性，則會強化並澄清一個研究計畫的重要性。

因果理論以及假設

假設以及因果關係　**因果假設**（causal hypothesis）是一種有待檢驗的命題，或是關於兩個變項間之關係暫時性、試驗性陳述。假設是關於社會世界運作方式所作出的猜測，而且是以一種價值中立的形式來陳述的。柯林傑（Kerlinger, 1979: 35）曾經指出：

> 只要瞭解假設的內容、以及建構假設的方法，就可以知道假設在科學研究中的重要性，遠高於假設在表面上所呈現出來的表象。假設具有一個深層的、極為重要的目的——也就是讓人們跳脫自己……假設是促進知識發展的強大工具。因為雖然假設是由人們所構思而成的，但假設卻可以在不受到人類的價值與信念影響的情況下，被加以檢驗並且證明其為正確或不正確。

一個因果假設具有五項特徵（請參閱資訊箱 6.4 因果假設的五項特徵）。舉例而言，如果有一個假設的內容是參加宗教活動可以減少離婚的機率，那麼我們可以把這個假設重新敘述爲下列的預測：經常參加宗教活動的夫妻，其離婚率會比很少參加宗教活動的夫妻來得低。我們可以利用實徵證據來考驗這項預測。我們應該以合乎邏輯的方式，將這個假設連結到一個研究問題，以及一個更爲廣泛的理論。畢竟，我

因果假設　一種因果解釋或是命題的陳述，在該陳述中至少會有一個自變項和一個依變項，而且尚未以實徵的方式加以驗證過。

資訊箱 6.4　因果假設的五項特徵

1. 因果假設至少有兩個變項。
2. 因果假設表達了變項之間的因果關係。
3. 因果假設可以被描述為一項預測、或是未來預期出現的結果。
4. 因果假設在邏輯上連結到一個研究問題以及理論。
5. 因果假設具有可否證性，換句話說，因果假設能夠用實徵證據加以考驗，並顯示其真假。

們考驗假設的目的，就是爲了回答研究問題，或找到支持某個理論的實徵證據。在邏輯上必然爲真的陳述，或是無法透過實徵觀察回答的問題（例如，什麼是「好的生活」？上帝是否真的存在？），都不能當作科學的假設。

我們可以用數種方式陳述因果假設。有時候我們會使用「*導致*」（*cause*）這個詞，但這並不是必要的。舉例而言，我們可以用十種不同的方式，來描述在參加宗教活動及離婚率降低之間的因果假設（請參閱範例 6.4 陳述因果關係的方式）。

在科學研究裡，當我們談論到考驗假設時，會避免使用「*證實*」（*proved*）這個詞彙。在新聞媒體、法庭、以及廣告中，都會使用「*證實*」（*proof*）這個詞，但是研究科學家幾乎不會使用這個詞。陪審團可能會說某項證據「證實」了某人是有罪的，或者，某個電視廣告可能會說：「研究證實我們的阿斯匹靈能夠在最短的時間之內治好頭痛」，然而，這些都不是科學研究的語言。在科學裡，我們承認知識是暫時性的，而且知識的創造是一個持續進行的過程，應該避免過早提出定論。「證實」這個詞暗示著已成定局、絕對的確定性、或某件事物已經不需要更進一步的調查了。對於謹慎的科學世界而言，這個詞實在是太過強烈了。我們或許會說某項證據支持或確認某個假設，但我們不會說某項證據證實了某個假設。即使在數百項研究都已

範例　6.4

陳述因果關係的方式

- 參與宗教活動*導致*（*causes*）離婚的可能性降低。
- 參與宗教活動*促使*（*leads to*）離婚的可能性降低。
- 參與宗教活動與離婚可能性降低*有關*（*is related to*）。
- 參與宗教活動*影響了*（*influences*）離婚可能性降低的情況。
- 參與宗教活動與離婚可能性的降低*有所關聯*（*is associated with*）。
- 參與宗教活動*產生了*（*produces*）降低離婚可能性的效果。
- 參與宗教活動*造成*（*results in*）離婚的可能性降低。
- *如果*（*If*）人們參與宗教活動，*那麼*（*then*）離婚的可能性就會降低。
- 宗教活動的參與程度*越高*（*the higher*），離婚的可能性就*越低*（*the lower*）。
- 參與宗教活動*減少了*（*reduces*）離婚的可能性。

經顯示了完全相同的結果之後，例如像是抽煙與肺癌間的關聯性，科學家仍然不會說我們已經百分之百地證實了這項關聯性。相反地，我們可以說，有數量極爲龐大的證據、或到目前爲止的所有研究，都支持某項假設、或是與某項假設一致。科學家從未排除發現一些可能否定過往發現之新證據的可能性，而且科學家們也不願意中斷未來的研究、或停止探索中介的機制。有些人們原以爲已經證實爲真的關係，後來卻被發現是錯誤的，而此種例子在歷史上可以說是不勝枚舉。在指稱邏輯的或是數學的關係時，我們可以使用「證實」這個詞，就如同我們在數學方面的證明，但我們卻不能在實徵研究中使用證明或證實之類的詞彙。

檢驗以及精進假設　知識的進展很少來自對於單一假設的一次檢驗。事實上，如果研究者只專注在檢驗某個假設的單一研究上，就有可能會得到一幅關於研究過程的扭曲圖像。隨著科學社群中許多不同領域的研究者檢驗各種相異的假設，知識將會隨著時間經過而逐步進展。知識會在許多假設之間不停轉換並且去在蕪存菁的過程之中緩慢地成長進展。每一個假設都代表了對於一個依變項的一種解釋。如果證據不支持某些假設，這些假設就會逐漸地被研究者們排除在考量範圍之外。而那些獲得證據所支持的假設，則會繼續保留下來讓研究者們去進行爭論。理論家與研究者會不斷地創造出新的假設來挑戰那些已經得到證據支持的舊假設（請參閱圖 6.3）。我們可以從圖 6.3 的內容看到，在 2010 年，只有三個假設仍在被爭論當中，而且有可能成爲最佳的解釋，但從 1970 年到 2010 年之間，原本研究者們列入考量之中的假設共有十一個，但隨著時間經過，其中有八個假設在一個或多個檢驗的過程當中被拒絕了。

科學家是個抱持著懷疑態度的團體，因此，某一項研究的結果支持某個假設，並不足以讓科學家們接受這個假設。可複製性的原則主張，在一個假設獲得廣泛接受之前，必須經過數次的檢驗並得到一致而重覆的支持。另外一個能夠強化對於某個假設的信心的方法，就是考驗該假設所依據的理論中的相關因果連結。

身爲科學家，我們所能接受的是擁有最多實徵證據支持的最堅強假設，並把該假設當作當時的最佳解釋。當我們針對某個假設考驗了越多可能的替代假設時，我們對於該假設的信心就會越加強烈。有一些假設的考驗方式被稱爲**決定性實驗**（crucial experiments）或決定性研究，在此種類型研究中，

> 有兩個或是更多個對於某些現象的可能替代性解釋，每一種解釋都符合某種特定的實徵資料。而決定性實驗的設計目的，就是其產生的結果只能夠由這些解釋中的唯一一個加以說明，而唯一有效的解釋就是「最佳的解釋」（Kaplan, 1964: 151-152）。

因此，雖然科學家們很少進行決定性實驗，但是決定性實驗卻是對理論的重要考驗方式。在決定性實驗中，來自於兩個不同理論的假設彼此正面對決，且其中之一將會被淘汰出局。雖然決定性實驗相當罕見，但一旦進行決定性實驗，就會帶來重大而且深遠的影響。

決定性實驗　一種針對相同現象的不同解釋進行直接比較以及評估的實驗，其設計目的在於決定何種解釋才是最佳的解釋。

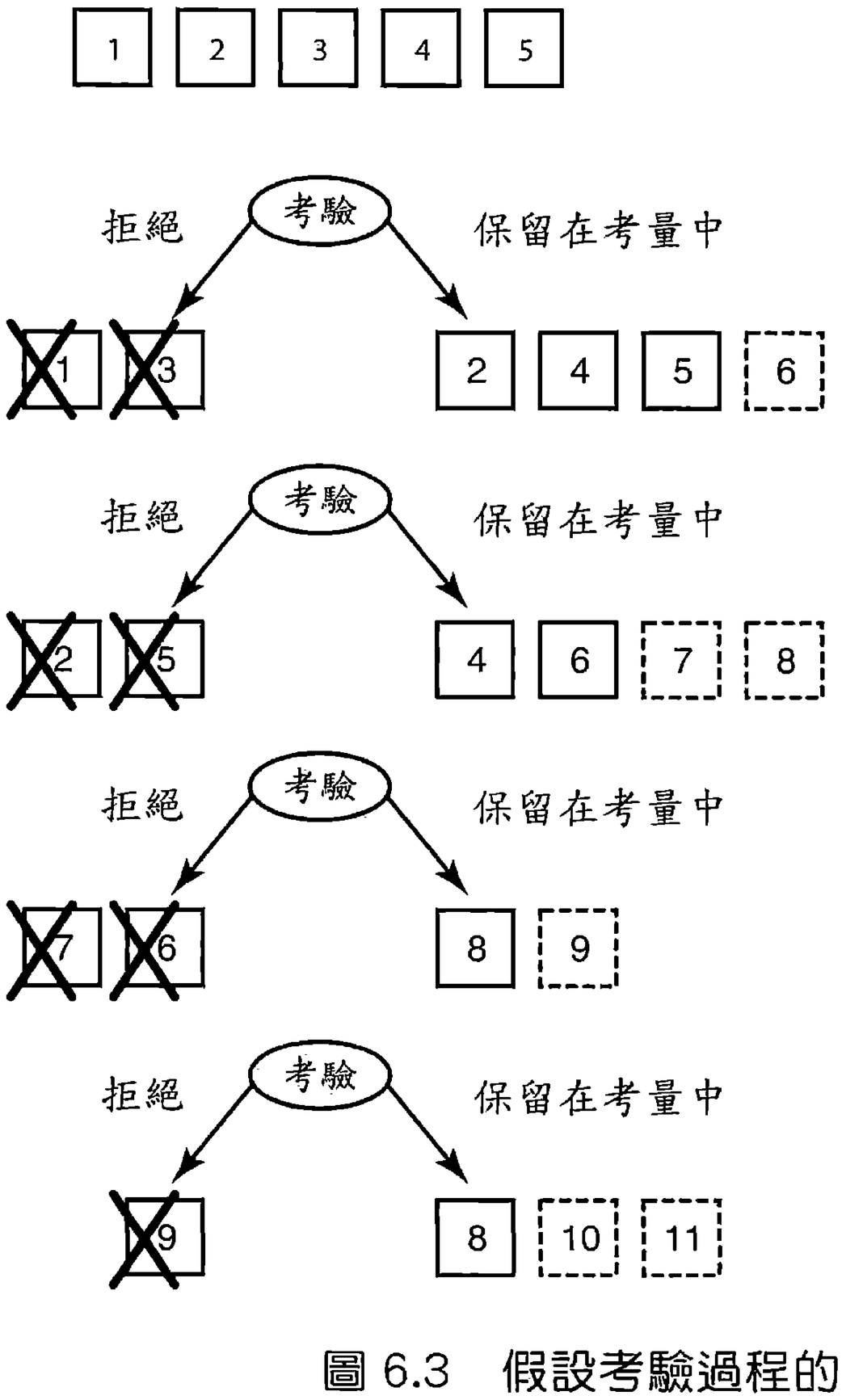


圖 6.3　假設考驗過程的運作方式

假設的類型　假設是一連串因果鍊中的環節，並被用來考驗變項之間關係的方向與強度。當一個假設擊敗了其競爭對手，就支持了研究者所提出的解釋。在假設考驗過程中的一個較爲奇特的面向，就是研究者對於支持某個假設的證據、以及反對某個假設的證據，有著截然不同的對待方式。更爲確切地說，研究者其實比較重視負面的證據。此種認爲負面證據在評估一個假設時具有決定性關鍵的想法，來自於**否定假設的邏輯**（logic of disconfirming hypotheses）[16]，此種邏輯與巴柏的證僞原理的觀念（請參閱第四章裡關於實證主義的內容）以及使用虛無假設（請參閱本章稍後部分的內容）之間有著密切關聯性。

否定假設的邏輯　這是虛無假設背後的邏輯概念，其基礎在於以下這個觀念：在證明某種關係的存在時，找到支持性的實徵證據並不是理想的作法，相較於蒐集支持性證據的作法，若是先假設沒有任何關係存在、並且考驗此種假設，則更能夠為某種關係存在的可能性提供謹慎而間接的支持性證據。

回想一下之前我們關於「證明」或「證實」的相關討論內容。在先前的討論內容裡曾經提過，我們從來不會去證實一個假設，然而，我們卻可以駁斥一個假設。在擁有支持性證據的情況之下，我們只能說某個假設仍然具有可能性、或是這個假設仍然可以列入考慮。負面證據則更爲重要。在負面證據出現之後，某個理論就變成「有污點的」

（“tarnished”）或「失去光彩了」（“soiled”），這是因爲假設應該是可以作出正確預測的，但是負面與否定的證據則顯示出這些預測是錯誤的。由於各種不同的替代假設都可能作出完全相同的預測，因此，正面的或是肯定的證據就不是那麼重要。當我們找到某個預測的肯定性證據，或許可以使某個解釋的地位暫時高於其他的替代解釋，但是其他的替代解釋也可能擁有各自的肯定性證據。

舉例而言，假設有一個人拿著一把雨傘站在街角，並聲稱他的雨傘可以保護他不被從天而降的大象砸中。他也有證據支持他提出的這個雨傘可以保護他的假設，也就是在他把雨傘打開的期間內，他完全沒有被任何一頭從天而降的大象砸中。然而，這樣的支持性證據其實是相當薄弱的，而且這項證據也能支持另外一個替代的假設：大象不會從天上掉下來。這兩個假設都預測了這個人不會被從天而降的大象砸中，然而，只要出現一項負面證據——一頭大象真的從天而降、砸中他和他的雨傘、並把兩者都壓扁了——就可以完完全全地摧毀這個人所提出的假設。

我們可以用兩種方式來考驗假設：以一種直截了當的方式，或是以虛無假設的方式。許多的量化研究者，尤其是實驗研究者，會根據否定假設的邏輯，以**虛無假設**（null hypothesis）的方式提出假設。這些研究者會尋找證據，好讓他們決定要接受或拒絕虛無假設。大多數人在談論假設時，都把假設當作預測某種關係的方式，而虛無假設則是恰巧相反，因爲虛無假設所預測的是沒有任何關係存在。例如，莎拉相信住在學校宿舍裡的人，會比住在學校外面而且通勤上學的人得到更高的成績。因此，她的虛無假設就是在居住地點與成績之間沒有任何關係。研究者在使用這個虛無假設時，也會使用相對應的**對立假設**（alternative hypothesis）或實驗假設。對立假設主張有關係存在，因此，莎拉的對立假設如下：學生住在校內宿舍會對成績有正面的影響。

對於大部分的人來說，此種使用虛無假設的作法，看起來似乎是在思考假設考驗時的後退方式。使用虛無假設的作法，來自於我們想要發現關係的這個假定。正因爲我們內心渴望找到關係，所以我們在設計假設考驗的過程時，必須把確定發現關係的標準訂得很高。當我們使用虛無假設時，我們直接考驗的只有虛無假設。如果證據支持或引導我們接受虛無假設，我們就可以下結論，認定接受檢驗的關係是不存在的，而這也意味著對立假設是錯誤的。另一方面，如果我們找到了拒絕虛無假設的證據，那麼對立假設仍然有成立的可能性。我們無法直接證實對立假設，因此，我們只能藉由考驗虛無假設的方式，讓對立假設可以繼續維持接受討論與挑戰的狀態。當我們在肯定性證據中加入虛無假設考驗的結果，主張對立假設成立的論點便會隨著時間經過而變得越來越強而有力。

如果這些有關虛無假設的討論讓你覺得暈頭轉向，至少你要記得，科學社群是極爲謹慎的。畢竟，科學社群的主要活動就是在創造真正的、經過驗證的真理。因此，直到堆積如山的證據顯示某項因果關係爲真之前，科學社群都寧可認爲此因果關係是錯誤或是不存在的。此種態度很類似英裔美國人的法律觀念：在被證明有罪之前，每個人都是清白的。

虛無假設　一種假設，此種假設主張，某個自變項對於某個依變項不會造成任何顯著的效果。

對立假設　與虛無假設成對出現的一種假設，此種假設主張，某個自變項會對某個依變項造成顯著的效果。

我們假定、或刻意裝作虛無假設是正確的，直到有*合理的懷疑*（*reasonable doubt*）暗示我們可能是另一種情況爲止。當我們在使用虛無假設時，我們也可以使用專家爲此種思考方式所設計的特定統計考驗方法（例如像 *t* 考驗或是 *F* 考驗）。因此，如果統計考驗的結果指出，虛無假設有 99%的機率是錯誤的，那麼我們就能合理懷疑虛無假設的正確性。此即當我們說統計考驗的結果讓我們「在.01 的顯著水準下拒絕虛無假設」時，所指的真正意思了（我們將在第十二章中，進一步討論統計的顯著水準）。

另外一種假設的類型是**雙載假設**（double-barreled hypothesis）[17]，此種類型的假設顯示出含糊不清的思考方式、並且造成不必要的混淆情況，所以應該極力避免。雙載假設意指在一個假設之中同時放入兩種各自不同的關係。例如，我們認爲在某個地區中，貧窮與青少年高度集中的現象，會造成財物犯罪案件增加。上述的說法就是一種雙載假設，而在這個雙載假設裡，我們所指的可能是下列兩種說法之中的任一種：貧窮或青少年高度集中的現象之一，會造成財產犯罪；或只有當貧窮以及青少年高度集中的兩種現象結合時，才會造成財產犯罪率上升。如果我們以「兩者其中之一」的假設來進行考驗，而且只有一個自變項會產生效果時，則假設考驗的結果就會含糊不清。例如，如果證據顯示，貧窮會導致犯罪、但是青少年集中的現象不會導致犯罪，那麼我們的假設是否獲得支持呢？如果我們所考驗的是兩者結合才會產生效果的假設，那麼我們真正所指的意思是，只有當貧窮與青少年集中的現象同時發生時，才會導致財產犯罪，而其中任何一項單獨出現時，則不會產生此種效果。如果我們要考驗的是此種結合兩者的假設，就不算是雙載假設。我們必須非常清楚而且明確地陳述此種結合兩者的假設。而此種結合兩者的假設，則被稱作*交互作用效果*（*interaction effect*）（稍後將會進一步討論交互作用效果，同時請參閱圖 6.4）

因果解釋中的可能錯誤

在爲任何理論（例如，因果的、詮釋的、網絡的理論）發展一個良好解釋的過程中，必須避免犯下一些常見的邏輯錯誤。這些錯誤可能發生在剛開始進行研究時、在詮釋及分析量化資料時、或是在蒐集以及分析質化資料時。像這樣的錯誤一般被稱作*謬誤*（*fallacies*）或是*錯誤的解釋*（*false explanations*）。從表面上來看，這些錯誤似乎是正確而合理的，但只要更爲深入地調查，就可以看出隱藏在表面之下的嚴重問題。

雙載假設　一種易於造成混淆且設計不良的假設，由於在此種假設中總共有兩個自變項，因而無法釐清到底是其中一個自變項、或是兩個自變項一起產生效果。

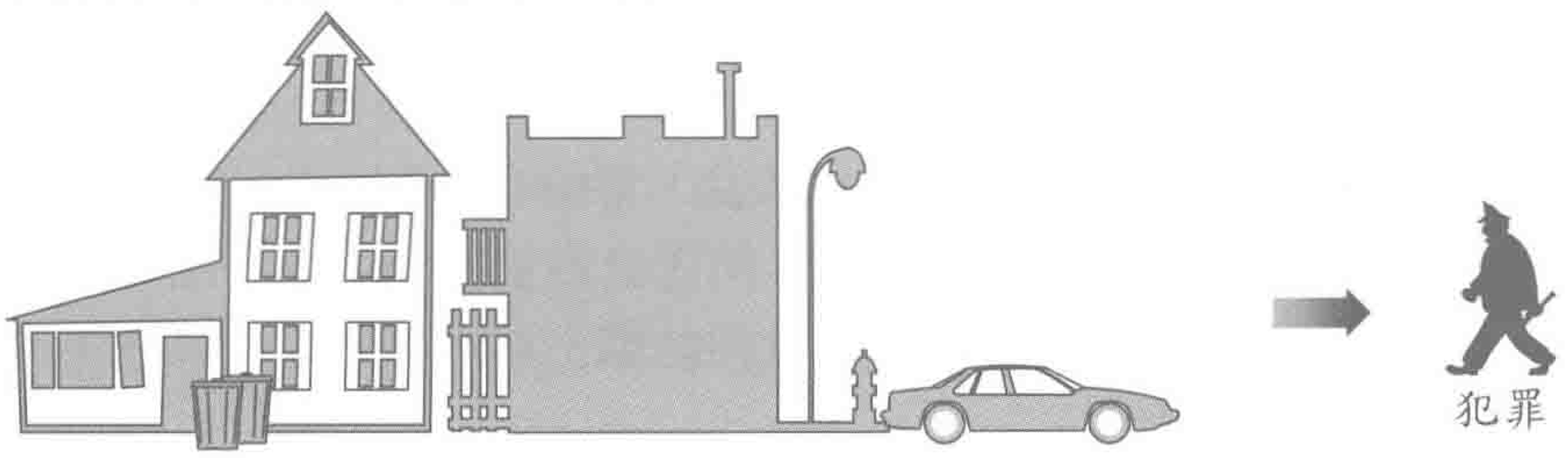
套套邏輯　一種解釋上的錯誤，在此種錯誤中，原因（自變項）及結果（依變項）實際上是完全相同的，或者是彼此的重新陳述而已，因此使得某種明顯的因果關係在定義上為真。

套套邏輯　所謂的**套套邏輯**（tautology）是指一種循環推理的形式。我們表面上好像在說一件新的事情，但事實上只是反覆地談論同樣的事物，使得某個陳述*在定義上看似爲真*（*true by definition*）。我們無法用實徵資料來驗證套套邏輯。例如，我聽到一

假設： 貧窮和青少年集中在某個地區的現象會造成犯罪率上升

雙載假設： 這可能是指下列三種情況中的任何一種

貧窮

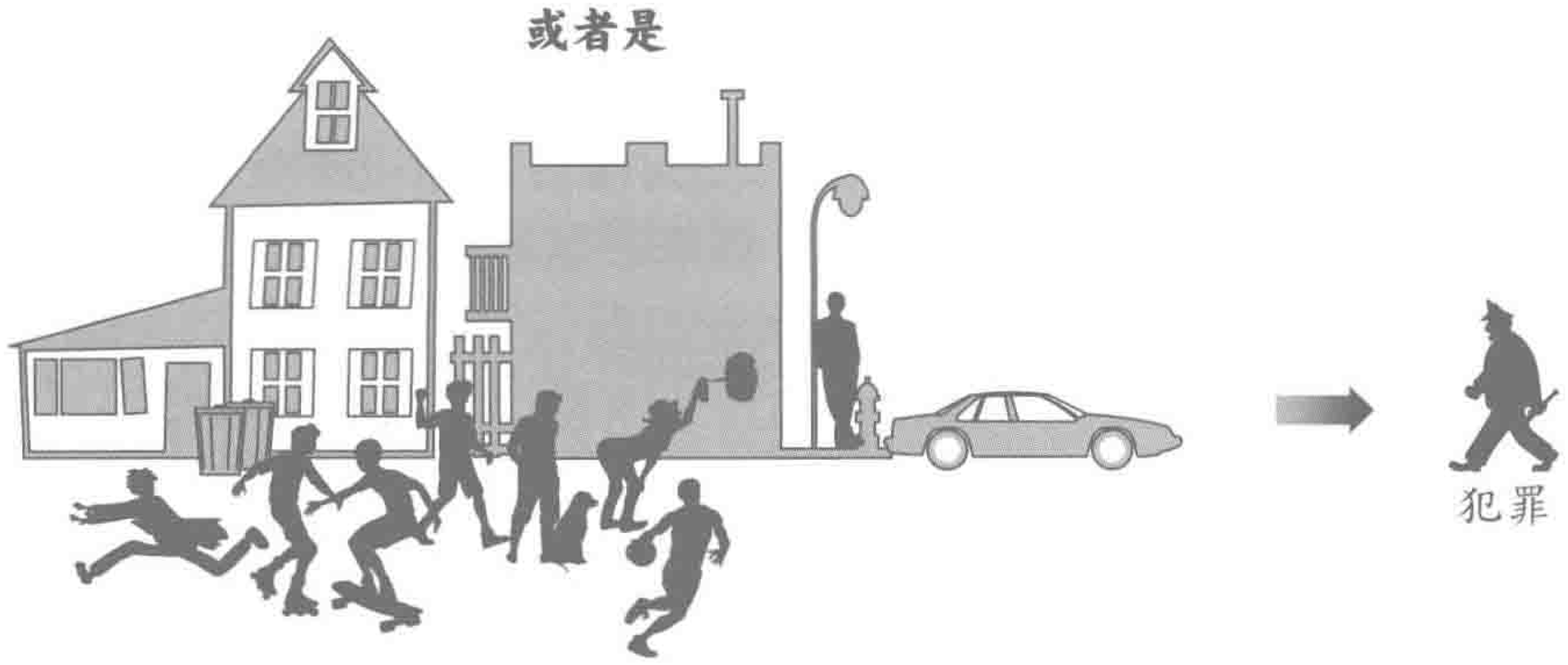


或者是

青少年高度集中

或者是

貧窮以及青少年高度集中



交互作用效果： 這意味著事件的結合

貧窮以及
青少年高度集中
兩種情況同時存在

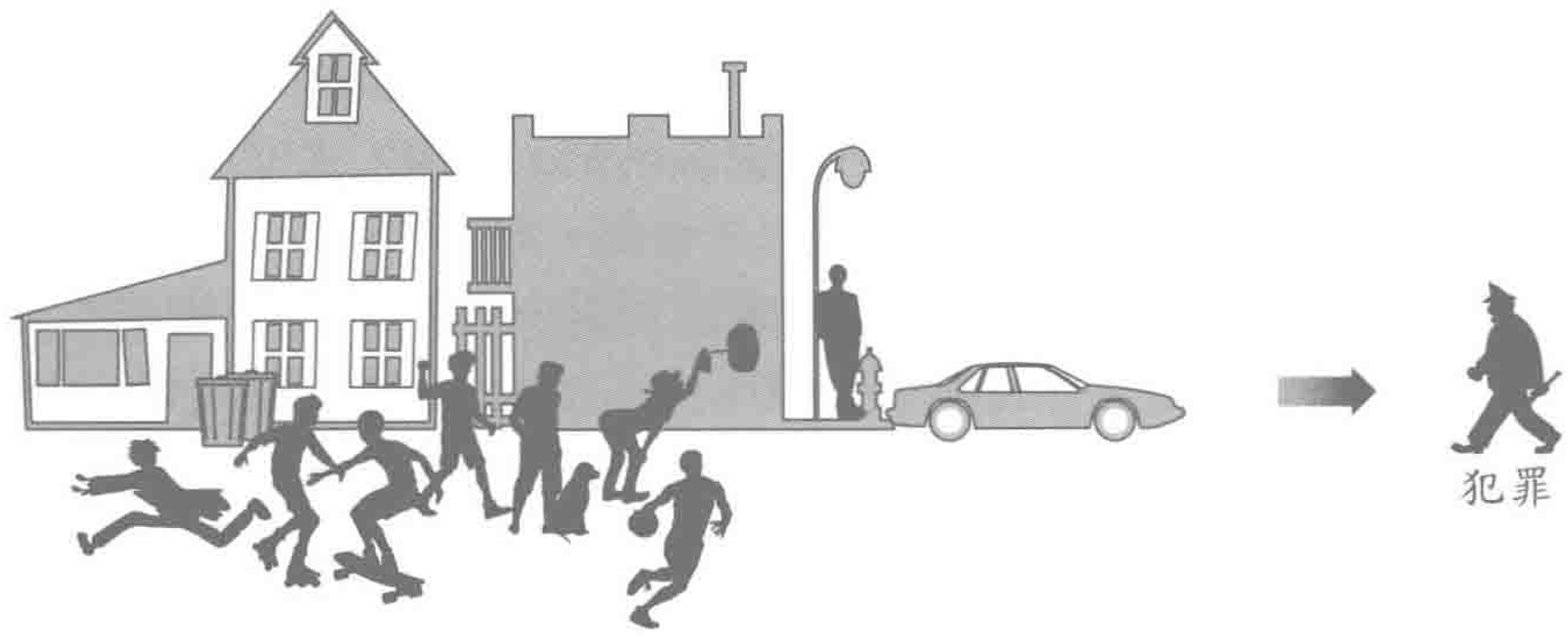


圖 6.4　雙載假設與交互作用效果

則新聞報導，內容是關於一位美國國會議員鼓吹一條把更多 14 歲與 15 歲的罪犯送到成人法庭的新犯罪法律。當被問到他爲什麼只是一味地強調嚴刑峻法、而不好好地討論如何預防犯罪時，這位議員的回答是，違法的人在受到重罰之後，將學到慘痛的教訓：犯罪是得不償失的。如此一來，就可以預防犯罪了。因此，他相信預防

犯罪的唯一方法，就是以嚴刑峻法來重懲犯罪的人。當我聽到這則消息時，我覺得這位議員的說法聽起來有點怪怪的。於是，我重新審視了這個議員的論點，然後瞭解到這就是一種套套邏輯（也就是說，這個論點已經犯了邏輯上的錯誤）。基本上，這位議員的說法是懲罰犯罪可以預防犯罪，這是因爲他已經把「預防」重新定義爲與「懲罰」完全相同了。就邏輯上來說，當他說懲罰可以產生預防的效果時，他其實是將重罰及預防這二件事情畫上等號了。政客或許會利用循環推理混淆大眾的視聽，但社會研究者卻必須學習如何看穿此種把戲，並避免類似的曲解。

範例 所謂的保守份子，指的是那些抱持特定態度、信仰、以及價值觀的人（要求政府減少干預、對高收入者不予課稅、建立強大的軍隊、在公立學校教導宗教課程、廢除反歧視法案等等）。如果有人說，要求政府減少對各種事務的干預、建立強大的軍隊之類的訴求，將會*導致*保守主義，這種說法便是套套邏輯。在較寬鬆的日常用法中，我們可以說：「莎莉是一個保守份子，*因爲*她相信政府應該減少對人民的干預」。這句話在表面上看來，似乎是一段因果陳述，事實上並不是。這些態度是把莎莉歸類爲保守份子的*理由*（*reason*），但這些態度並非造成她抱持保守主義的*原因*（*cause*）。她的態度*是*保守主義，因此上面提到的那句陳述就定義上而言是真實的。但我們永遠也無法找到證據來證明，這些態度與保守主義之間毫無關聯。

目的論 **目的論**（teleology）意指某件事是由一個的最終目的或目標所引導的。目的論可能有兩種形式。首先，目的論常讓人聯想到一些特殊事件，而這些事件的發生原因，乃是由於「上帝的旨意」或某種支配一切的、神秘的、無形的、不可知的力量。換句話說，這些事件之所以會發生，完全是因爲上帝或是一種無形的、不可知的強大力量，已經預先決定了這些事必定會發生。當我們說某件事之所以會發生，是因爲此爲某種全能內在神靈「自然顯現」的一部分時，我們所說的就是一種目的論。因此，若我們說某個社會是因爲「國家的精神」或「天意」而朝向某個方向發展，也是一種目的論。類似的目的論說法則把人性當作原因，例如，「犯罪事件發生的原因，乃是因爲人性使然」，這種說法也是一種目的論。目的論也出現在與歷史有關的理論中，例如，有些人認爲，我們正邁向一個「理想的社會」或是烏托邦、並以此種想法來解釋當今所發生的各種事件。這種說法亦犯了目的論的錯誤。目的論也會出現在功能性的論點中。當我們說家庭演變成某種形式（例如，核心家庭），是因爲核心家庭符合社會系統對於社會延續的「需求」時，就是一種目的論。在邏輯上而言，以上的論點意味著社會系統延續到未來的功能性需求，是造成我們今日所見到的家庭形式的主要原因。然而，我們根本不可能測量這個原因，也不可能以實徵的方式來檢驗目的論。

目的論違反了因果關係中的時間順序條件，而且在目的論中，沒有真正的自變項。因爲當中的「原因因素」是極爲模糊不清、遙遠、而且不可見的。許多人會把目標動機（goal motivation）（也就是對

目的論 一種解釋上的錯誤，在此種錯誤中，因果關係無法以實徵的方式加以檢驗，因為就時間順序而言，原因並未出現在結果之前，或原因是一種模糊不清的、過於籠統的力量，無法以實徵的方式來進行測量。

於一種尚未發生的事物的渴望）與目的論混爲一談。我或許會說，一項目標導致了一個行動。例如，我想要在某一門課得到 A 的目標，促使我得到了好成績。我所清楚意識到的目標或渴望，可以是一個合理的原因，而不是一種目的論。爲了清楚地展現此種關係，我必須描述一下因果鍊。首先，我們可以在某個時間點用實徵的方式來測量我的心理狀態（例如，目標、渴望、或是抱負等等）。其次，我們可以把我的心理狀態和未來的事件（這些事件可能會發生、也可能不會發生）進行比較，像是在某一門課程中得到特定的分數等等。我的心理狀態可以是一個動機，促使我從事某些行爲，好比認真念書等等（中介變項）。此種努力念書的行爲能夠增加一個未來事件（在某一門課程得到特定的成績）發生的機會。有意識的人類目標不同於上帝的旨意、一個社會的精神、或者是系統需求等等，因爲我們無法以實徵方式測量這些事物、這些事物沒有固定存在的時刻、而且這些事物對於發生的事情，總是擁有一套自圓其說的解釋，因而無法加以駁斥。

範例　「核心家庭是西方工業社會中的主流家庭形式，因爲此種形式具有幫助社會延續的功能。」這句來自於結構功能論的陳述，是一句無法驗證的目的論陳述。這句話說的是「社會的延續」*導致了*「家庭形式的發展」。然而，我們能觀察到一個社會是否成功延續下來的唯一方式，是在這個事實發生之後，或是在已經發展出一種家庭形式之後的結果。以下爲另一個目的論所陳述的例子：「因爲美國註定成爲主要的世界強權，因此我們發現在十九世紀早期，有成千上萬的移民進入這塊尚未被開發的新大陸。」這句話的意思是，「成爲主要的世界強權」這件事（發生於 1920 年到 1945 年之間）導致了「大群人們往西部移民」的現象（發生於 1850 年到 1890 年之間）。這句陳述使用了「註定」這個模稜兩可的詞彙，而這個詞彙就像其他類似的詞彙一樣（例如，「上帝的旨意」），我們根本無法觀察其間的因果關係。

生態謬誤　**生態謬誤**（ecological fallacy）是來自於分析單位之間的錯誤搭配情況。此種錯誤意指，在我們擁有實徵證據的單位、以及我們想作出一般陳述的單位之間，出現了無法互相配合的狀況。歸根結底，發生此種錯誤的原因，乃是來自於不精確的推理方式，和對於證據作出過度的推論。當我們以*較高的*（*higher*）或是*集合的*（*aggregated*）分析單位來蒐集資料、但卻想以這些資料解釋*較低的*（*lower*）或是*分散的*（*disaggregated*）單位時，便是犯了生態謬誤。此種狀況之所以被認爲是一種謬誤，是因爲在某個層次的分析單位中所發生的事，不見得也會發生在不同層次的分析單位之中。[18] 因此，當我們蒐集了大型集合體（例如，各種組織、整個國家）的資料、並以這些資料針對個體的行爲作出結論時，我們就會犯下生態謬誤。爲了避免犯下此種錯誤，我們必須確保我們在某個解釋中所使用的分析單位，與我們蒐集資料時所使用的分析單位是完全相同的、或是非常接近的（請參閱範例 6.5 生態謬誤）。

生態謬誤　一種解釋上的錯誤，在此種錯誤中，關於大尺度的分析單位間的關聯性實徵資料被過度推論，並被當作是說明較小單位間的關係的證據。

範例　大約有四萬五千人住在湯姆鎮（Tomsville）

範例　6.5

生態謬誤

許多研究者都批評涂爾幹著名的自殺研究（[1897]1957）犯了生態謬誤，因為他把群體資料當作彷彿是個人層次的資料一樣。在這項研究裡，涂爾幹比較了十九世紀西歐國家中的清教徒以及天主教徒之間的自殺率，並將觀察到的差異解釋為是來自於這兩種宗教中的信仰與行為方面的不同。涂爾幹認為，清教徒的自殺率比天主教徒來得高，是因為清教徒較傾向於個人主義，而且社會整合程度較低。涂爾幹以及早期的研究者所擁有的只有區域性的資料。由於人們傾向於和有相同宗教信仰的人住在一起，因此涂爾幹利用了群體層次的資料（也就是宗教）來解釋個體的行為。

之後，有兩位研究者（van Poppel and Day, 1996）只以在其他地區所發現的個人層次的資料，重新檢視十九世紀的自殺率。他們比較了死亡記錄，也檢視了正式的死亡原因與宗教信仰，但是他們所得到的結果卻與涂爾幹不同。顯而易見的，在十九世紀時，地方官員是用不同方式來記錄不同宗教信仰者的死亡原因。在記錄天主教徒的死亡原因時，因為天主教在道德方面強烈禁止自殺，所以官員們常將死亡原因標註為「死因不明」。涂爾幹的大型理論可能是正確的，但是他用以檢驗理論的證據卻很薄弱，因為他使用的是在群體層次所蒐集到的資料，然而卻試圖以此解釋個人的行動。

與瓊安鎮（Joansville）。在湯姆鎮裡，高收入的居民比例很高，在湯姆鎮有超過一半以上的家庭，其年收入都超過十六萬美金。而且湯姆鎮登記的摩托車數量，也多於任何其他規模類似的小鎮。至於瓊恩鎮裡則有很多窮苦人家，其中有一半以上的家庭生活於貧窮線以下。而且瓊安鎮登記的摩托車數量，也少於任何其他規模類似的小鎮。但是，如果我們單獨依據上述的資訊，就說有錢人比較可能擁有摩托車、或是這項證據顯示出家庭收入和擁有摩托車之間有關，那就是一種*謬誤*。理由是我們不知道在湯姆鎮以及瓊安鎮裡的哪些家庭擁有摩托車。我們知道的僅有以這兩個城鎮爲整體的兩個變項，也就是平均收入和摩托車總數。我們在觀察變項時的分析單位，是以每個城鎮作爲一個整體來觀察的。或許在湯姆鎮裡，所有中低收入的家庭都擁有摩托車，但是沒有任何一個高收入的家庭擁有摩托車。或者，也有可能在瓊安鎮裡的一個富裕家庭以及五個窮困家庭都擁有摩托車。若要作出關於家庭擁有的摩托車數量與家庭收入間的關係陳述，我們就必須以家庭爲單位來蒐集資料、而不是以整體城鎮爲單位來蒐集資料，如此才能作出正確的陳述。

化約論　另外一個也是由分析單位之間的錯誤搭配，以及對於證據作出不正確推理所造成的問題，便是**化約論**（reductionism），也稱爲*未等化的謬誤*（*fallacy of nonequivalence*）（請參閱範例 6.6 化約論的錯誤）。當研究者使用來自於特定個人的證據來解釋鉅觀層次的事件時，就會發生此種錯誤。也就是說，當某個人所觀察的是*較低的*或*分散的*分析單位、但卻以此觀察資料來對*較高的*或是*集合的*分析單位之運作方式作出陳述時，就會發生此種錯誤。從某種意義來看，化約論就像是生態謬誤

化約論　一種解釋上的錯誤，在此種錯誤中，關於小型尺度分析單位之間關係的實徵資料，被過度地推論，並當作說明大型單位之間關係的證據。

範例　6.6

化約論的錯誤

假設你拿起一本書，並在這本書裡讀到了以下內容：

> *在 1960 年代的民權運動時期，美國的種族關係出現了戲劇性的變化。隨著全國各地的法律與法庭判決出現變化，多數族群、也就是白人們，對於其他種族的態度轉變成具有更大的包容力。原本一些在法律上及形式上只開放給白人的機會——在像是住宅計畫、工作、學校教育、投票權等等領域內——也開放給其他種族的群眾。從 1955 年的「布朗判例」(Brown vs. Board of Education)、到 1964 年的《民權法案》(Civil Rights Act)、以及從 1966 到 1968 年間的「向貧窮宣戰」(War on Poverty)運動等等，一個嶄新的、引人注目的未來前景席捲了整個美國。而此種現象正是來自美國最重要的民權運動領袖，馬丁路德金恩博士的遠見、奉獻、以及行動。*

在本段文章裡，依變項為「在 10 到 13 年期間之內，美國種族關係所發生的重大變化」，而自變項則是「金恩博士的遠見及行動」。

如果你熟知美國民權運動時期的歷史，那麼你就會發現其中有一個問題，也就是將整個民權運動與該運動的成功完全歸結於單獨的個人身上。當然，我們不能否認金恩博士確實發揮了重大的影響力，並幫助建立及引導了民權運動。但是，在上述的推論中，「運動」的概念卻消失不見了。換句話說，作為一個因果力量的社會政治運動的概念，被化約為這個運動的主要領袖，原本明確可見的社會現象——一項重大的運動——卻變得模糊不清。其中被忽略的，是投入於推動共同目標的成千上萬人們的行動（遊行、法庭案件、演說、祈禱大會、靜坐抗議、暴動、請願、被警方毆打等等）、以及對這些人的回應。這個運動的中心思想、群眾動員、政治、組織、以及策略等等，全都付之闕如。可能對此運動發生影響的其他相關鉅觀層次的歷史事件（例如，對於越戰的抗議、甘迺迪總統被刺殺身亡後的全國民心變化、非裔美國人的分離主義政治、非裔美國人的遷移現象等等），也同樣被忽視了。

此種錯誤並不僅僅發生於歷史解釋之中。許多人進行思考時，僅想到個人的行動，並且也會有一種個人主義偏見，這有時也被稱作是*方法論的個人主義*（*methodological individualism*）。此種情況在極度重視個人主義的美國文化中尤其常見。此種思考方法的錯誤之處，在於完全漠視分析單位的差異、或超越個人之外的其他力量。此種化約論錯誤，把對於原因的解釋轉移到一個相當低的分析單位。如果依照此種模式思考，我們甚至可以繼續將個人的行為化約成身體裡的生物過程、微觀層次的神經化學活動、或次原子的層次等等。

由於多數人所處的「社會世界」的焦點是地域性的、當下的情境，以及他們和其他一小群人們的互動。因此，人們對現實的日常感知，會鼓勵他們將社會潮流或是事件當成個人行動或是心理過程。因而人們往往對較為抽象的、鉅觀層次的實體——像是社會力量、過程、組織、機構制度、運動、或結構等等——視而不見。社會學的核心概念，就是所有的社會行動都不能被化約為單獨的個體。在他的經典著作《自殺論》中，涂爾幹極力反對方法論的個人主義，並且也詳細說明了，較大的、未被辨識的社會力量，能夠解釋更高度個人化的、私人的行動。

這種錯誤搭配的鏡像反射。當某人所擁有的是關於個人行爲方式的資料、但卻想以此資料談論鉅觀層次單位的動態變化時，就犯下了化約論的錯誤。當個人層次的資料比較容易取得，就可能會犯下此種錯誤。除此之外，鉅觀層次單位的運作方式也比較抽象而且模糊。李柏森（Lieberson）認爲此種錯誤會導致不一致、矛盾、以及混淆的情況。他（Lieberson, 1985: 108, 113-114）擲地有聲地說：

> 在決定關於較高層次運作過程的命題的有效性時，較低層次之間的關聯性是毫不相關的。事實上，我們根本無法從較低層次的分析中，獲得任何對於較高層次結構的有用理解……如果我們對於較高層次的過程以及事件有興趣，那純粹是因為我們瞭解到，較高層次的結構具有明顯的特質，而這些特質無法單純由次單位的資料中歸納得到。

正如同生態謬誤一樣，爲了避免犯下化約論的錯誤，我們必須確定在我們解釋當中所使用的分析單位非常接近於我們所擁有的實徵證據分析單位。當我們未能正確地思考分析單位、或是無法讓資料契合於理論時，我們就可能犯下生態謬誤或化約論的錯誤。這兩種錯誤都是來自於只擁有適合於研究問題的資料、但卻針對這些資料進行了嚴重的過度推論。

對於我們以實徵方式進行研究之外的分析單位，我們確實可能會去作出假定。因此，對於個人的研究，是以下列這個假定當作基礎的：個人是在一組社會制度之中做出行動的。當我們進行社會機構的研究時，是以關於個人行爲的假定作爲基礎的。我們知道，許多微觀層次的單位結合在一起而形成鉅觀層次的單位。危險之處在於，我們很容易在不知不覺的情況下，使用微觀單位（例如，個人）的行爲來解釋鉅觀單位（例如，社會機構）的行動。發生在某一層次的分析單位之間的事物，在其他不同層次的分析單位之間，不一定總是成立。作爲一個學術領域，社會學所依賴的信念是，在個人之外存在著一個有明顯區別層次的社會實在。關於這個層次的解釋，必須使用到超越個人之外的資料以及理論。我們不能把存在於鉅觀單位的各種原因、力量、結構、或是過程，化約爲個人的行爲。

範例 爲什麼會發生第一次世界大戰？你或許已經聽過，1914 年時，一名塞爾維亞人槍殺了奧匈帝國的大公，因而引發了第一次世界大戰。這種說法就是一種化約論。沒錯，這次的暗殺事件確實是引起第一次世界大戰的原因之一，但是在各國之間的鉅觀政治事件，也就是戰爭，不能被僅僅化約爲單一個人的特定行動。如果此種化約論成立的話，那麼我們也可以說，因爲這名殺手的鬧鐘沒有故障，並在那天早上準時把他叫醒，所以才會發生第一次世界大戰。如果當時那個鬧鐘故障了，就不會發生槍殺事件了。所以，根本是這個鬧鐘引起了第一次世界大戰！當然，第一次世界大戰的真正成因，其實遠比上述的情形更爲複雜，而且也是由於許多社會、政治、以及經濟力量在歷史的特定時間點匯聚所造成的。特定個人的行動雖然也扮演了一定的角色，但與這些鉅觀力量相較之下，根本是小巫見大巫。個人會影響事件，而且在與大規模的社會力量及組織結合之後，這些事件最終會影響其他事件並且推動國家。然而，只有單一個人的行動，並不成什麼氣候。因此，即使當時並未發生這起槍殺事件，第一次世界大戰仍然可能會於大致相同的時間點爆發。

虛假關係 當我們說在兩個變項之間的關係是*虛假的*（*spurious*）時，意思是指此種關係是錯誤的、是一種幻象。如果我們認爲自己已經找到一個虛假關係，我們通常會覺得很興奮。因爲我們可以證明，這個世界比表面上看起來的還要更複雜難解。

由於兩個變項之間的任何關聯性，都有可能是虛假的，因此，當我們發現某兩個變項間有所關聯時，我們必須要謹慎小心。在進一步調查探究之後，此種關聯性可能並非某種因果關係的基礎。此種關聯性或許只是一種幻象，就像是在炎熱的天氣裡，路上出現一坑水池的幻象一樣。

當兩個變項之間有所關聯，但是卻不具有因果關係時，就是所謂的**虛假關係**（spuriousness）。這是由於某個尚未被發現的第三個因素，才是真正的原因（請參閱範例 6.7 虛假關係，以及範例 6.8 夜明燈與虛假關係）。第三個變項是明顯的自變項與依變項的共同原因，並能解釋後二者間的關聯性。就因果關係的條件而言，尚未被發現的第三個因素，代表著一個解釋力更爲強大的替代性解釋。

你要如何辨別一個關係到底是不是虛假關係呢？你又要如何發現那個神秘的第三個因素可能會是什麼呢？你可能需要使用統計技術（在本書後面的章節會有進一步的討論）來考驗某個關係是否爲虛假關係。在使用這些統計技術時，你需要有一

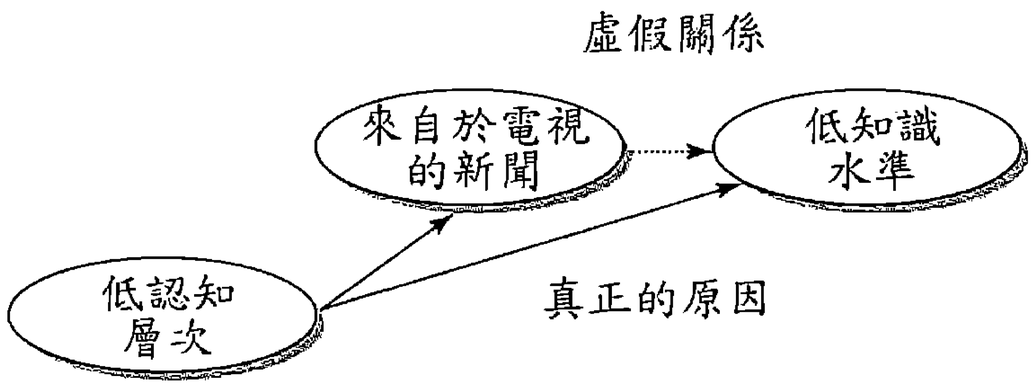
虛假關係　一種表面上很明顯、但其實卻是虛幻的因果關係，因為某一個尚未被發現的、或是一開始隱而未顯的因素，才是真正產生影響的變項。這個尚未被發現的因素對於自變項及依變項都具有因果影響，而且產生了後兩者之間存在著關係的錯誤印象。

範例　6.7

虛假關係

在他們對於新聞媒體的研究當中，紐曼與同事（Neuman and colleagues, 1992）發現了在新聞來源的類別及知識之間，存在著相關性。偏好從電視獲得新聞消息的人，在知識方面會少於那些從印刷媒體獲得新聞消息的人。此種相關性通常被詮釋為資訊的「愚民化」（“dumbing down”）。換句話說，人們幾乎無法從電視新聞當中獲得任何知識。

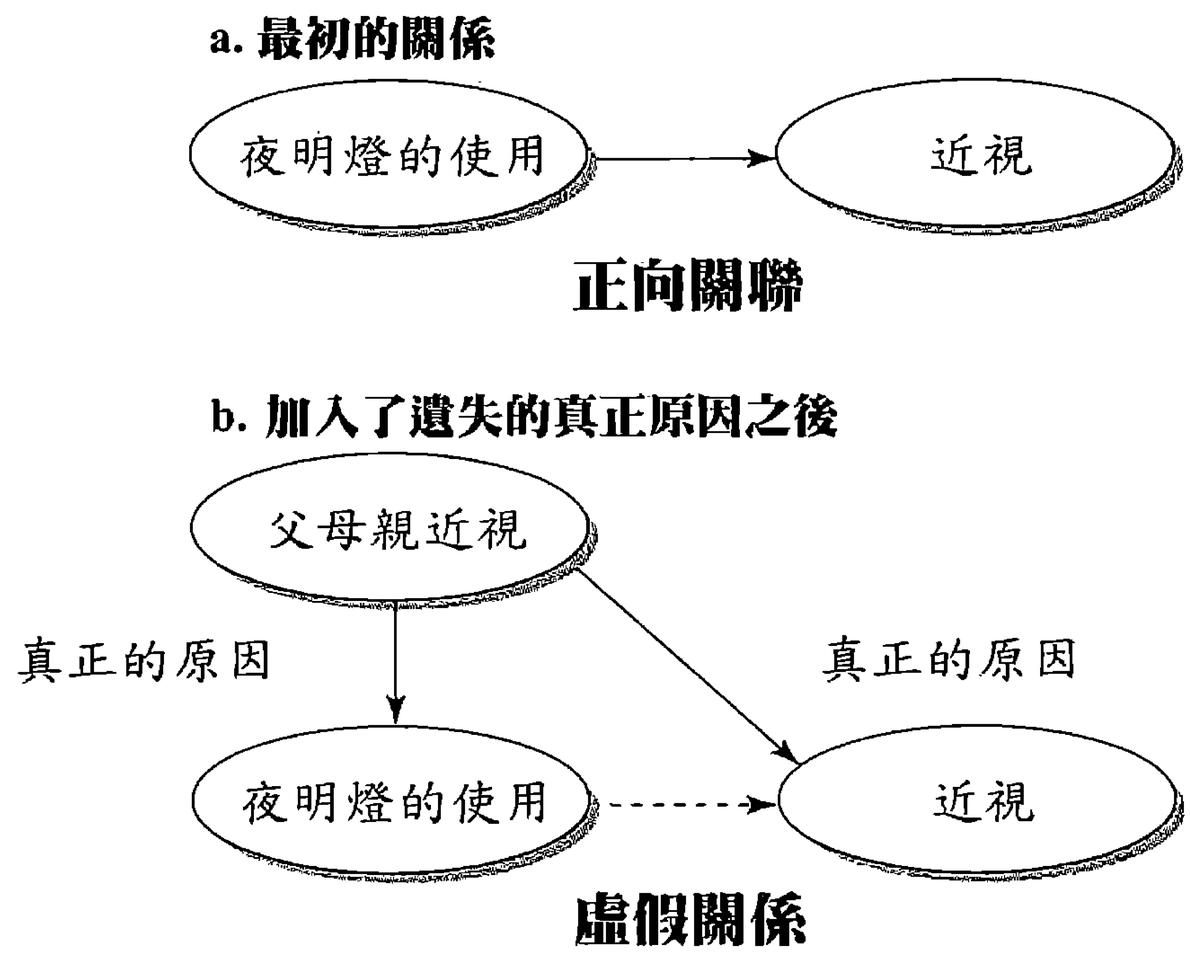
作者們發現，此種關係其實是一種虛假關係。「我們能夠證明，存在於電視新聞偏好及較低的知識分數間的整個關係，只是一種虛假關係」（p. 113）。作者們發現一個一開始未被發現的第三個變項，能夠同時解釋對於電視新聞的偏好、與關於時事知識的水準。作者們提到：「我們發現，真正造成此種『電視就是問題』效果的原因，是那些認知技巧較低的人們，偏好從電視獲得新聞消息」（p. 98）。在這個關係裡，「認知技巧」（“cognitive skills”）正是那個遺失或是隱藏的變項。作者們把「認知技巧」定義為一個人使用理智及運用抽象概念的能力。換句話說，那些覺得難以處理抽象、複雜資訊的人，會轉向電視新聞獲取資訊。其他人或許也會使用這些高度聳動的、具有娛樂性質的電視新聞來源，但他們其實較少使用電視新聞，並會經常以對於認知技巧要求較高的、資訊內容較為豐富的印刷來源作為補充。資訊技巧較為貧乏的人，通常對於時事的認識較少，也比較無法掌握其他需要用到抽象思考或處理複雜資訊的主題。



範例　6.8

夜明燈與虛假關係

許多年以來，研究者們在使用夜明燈與近視的兒童這兩者之間觀察到強烈的正向關係。許多人認為夜明燈或多或少造成了兒童的視力受損問題（如下圖所示）。其他的研究者則認為，在使用夜明燈和造成近視這個因果連結之間，已經沒有其他可能的原因或是變項。然而，一項在 1999 年所進行的研究，提供了這個問題的答案。該項研究發現，本身有近視問題的父母，比較有可能會使用夜明燈，而他們也會把視力不佳的問題遺傳給他們的孩子。該項研究也發現，一旦把父母親的視力加入這項解釋中，則在夜明燈的使用與近視之間，就找不到任何連結了（請參考下圖的 b）。因此，一開始的因果連結其實是被誤導的、或者也可以說是一種虛假關係（來自於 *New York Times*, May 22, 2001）。



資料來源：“Vital Signs: Update; New Study Vindicates Night Lights” from *The New York Times*, Health Section, 5/22/2001 Issue, Page(s) 6.

個理論，或至少要有關於可能的第三個因素的猜測想法。事實上，虛假關係是以一些你已經在使用的常識邏輯當作基礎的。例如，你已經知道在冷氣的使用量與冰淇淋的銷售量之間，存在著某種關聯性。如果你測量每天的冷氣使用量以及冰淇淋銷售量，你將會發現強烈的相關性。在冰淇淋銷售量越多的那些日子，冷氣的使用量也會跟著攀升。但是，你也知道吃冰淇淋這件事並不會導致人們打開冷氣。相反地，第三個變項，也就是炎熱的天氣，才是影響這兩件事的真正原因。你可以測量每日的溫度、冰淇淋的銷售量、以及冷氣的使用量，以驗證這個想法。而在社會研究裡，一些對立的理論可以幫助我們在許多主題之中，找出相關的第三個因素（例如，犯罪的成因、造成戰爭或是虐待兒童的原因）。

範例　有些人主張，使用非法毒品會導致自殺、輟學、以及暴力的行爲。支持「毒品就是問題」此種立場的人，指出在施打毒品和自殺、輟學、從事暴力行爲之間，存在著正向的相關。支持此種觀點的人們主張，終結毒品的使用將可以大幅度地減

少自殺、輟學、以及暴力行爲。其他人則主張，許多人之所以會施打毒品，是由於他們本身有一些情緒問題、或是他們所居住的社區具有高度失序混亂的情況（例如，高失業率、不穩定的家庭、高犯罪率、寥寥可數的社區服務、缺乏禮儀教養等等）。本身有情緒問題、或住在失序社區的人，也比較可能會自殺、輟學、以及從事暴力行爲。換句話說，如果能夠減少情緒問題與社區混亂的情況，就能夠大幅度地減少施打毒品、輟學、自殺、以及暴力行爲等等問題。單獨只減少施打毒品這個問題，只能產生有限的效果。因爲此種作法忽視了根本的原因，而這個根本的原因並不是毒品。「毒品就是問題」的論點是虛假的，因爲一開始存在於施打非法毒品、以及此種立場支持者所指證的問題之間的關係，其實是受到誤導的。情緒問題與社區的混亂失序，才是真正的、而且經常被忽視的原因變項。

我們現在可以把注意力從那些在因果關係中應該要避免的錯誤上移開，移向其他與假設有關的議題。表 6.2 回顧了這些重大的錯誤，而圖 6.5 則以圖例來表示這些錯誤。

從研究問題到假設

從一個廣泛的主題轉化成一個假設，是一件困難的事，但是將一個仔細規畫的研究問題轉化成爲一個假設，卻容易得多。一個好的研究問題本身已內含假設，除此之外，假設也是對於研究問題的暫時性解答。

思考一下以下這個研究問題的範例：「結婚年齡是否與離婚有所關聯？」這個問

表 6.2　解釋中的重大錯誤摘要

錯誤類型	簡短定義	範例
套套邏輯	某種關係在定義上為真，而且涉及到循環推論。	造成貧窮的原因是因為沒有錢。
目的論	原因是一種不恰當的意圖，或是在時間次序上出現錯置的情況。	人們在宗教儀式中結婚，是因為社會希望他們這麼做。
生態謬誤	實徵觀察結果的層次過高，不適合用來解釋所陳述的因果關係。	紐約的犯罪率很高。瓊安住在紐約，因此，她很有可能偷了我的手錶。
化約論	實徵觀察結果的層次過低，不適合用來解釋所陳述的因果關係。	因為史蒂夫剛丟了工作，而且沒有買新車，因此這個國家必然已進入了長期的經濟衰退。
虛假關係	一個尚未被發現的第三個變項，才是影響自變項與依變項的真正原因。	頭髮的長度和喜歡的電視節目類型有關。短髮的人偏好觀看美式足球，而長髮的人則偏好愛情故事。（尚未被發現的變項：性別）

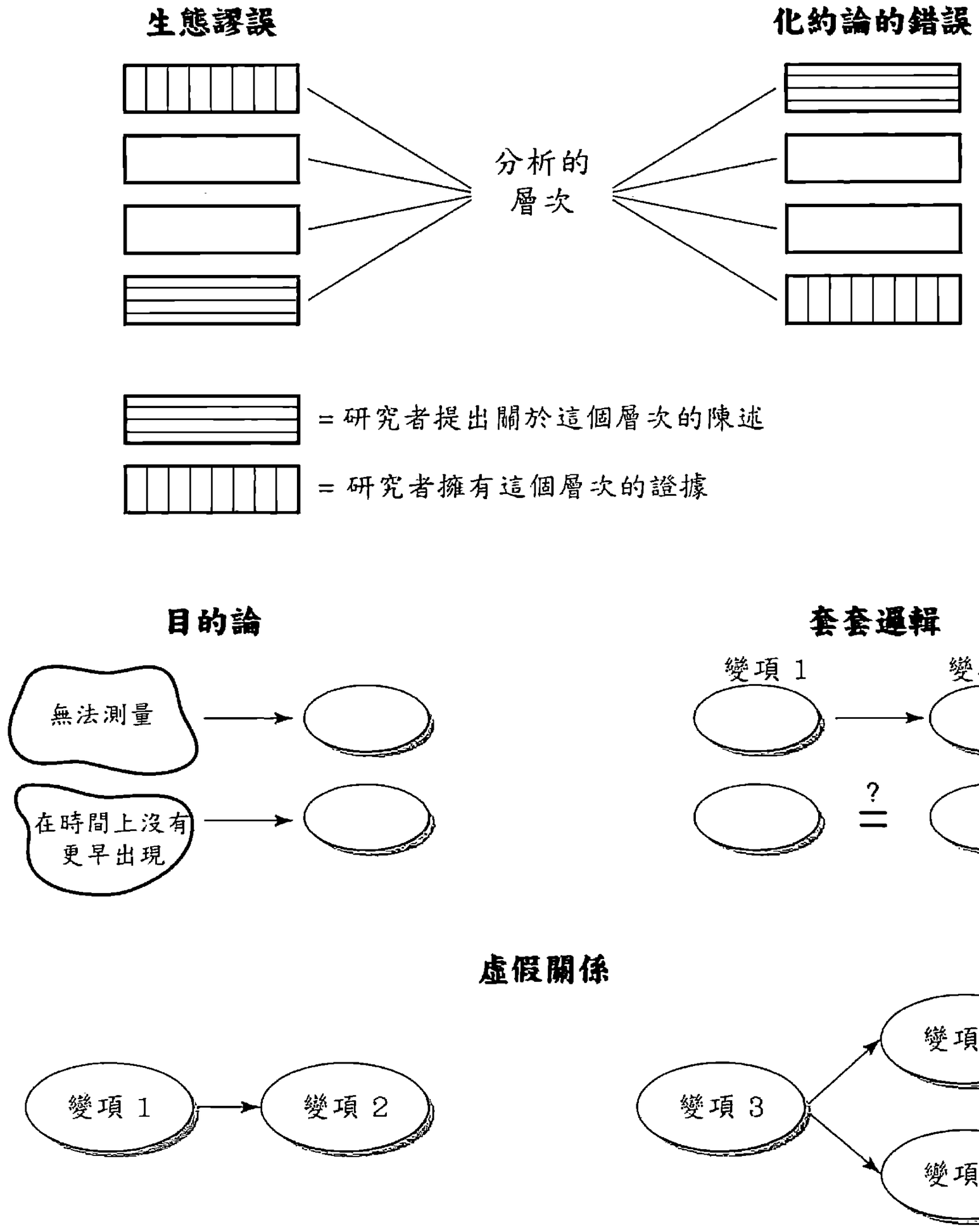


圖 6.5　在解釋時應該避免的五種錯誤

題有兩個變項：「結婚年齡」與「離婚」。爲了要發展出一個假設，我們必須決定何者是自變項。在這個研究問題中，自變項是結婚年齡，因爲就邏輯上而言，結婚必定早於離婚。我們或許會詢問此項關係的方向爲何。根據這個研究問題而發展出的可能假設如下：「結婚年齡越低，這段婚姻以離婚收場的可能性就越高」。這個假設回答了研究問題，也作出了預測。請注意，我們也可以重新形成假設，並使這個假設更爲聚焦：「結婚時年齡越小的夫妻，是否越有可能離婚？」

我們可以針對同一個研究問題，提出數個不同的假設。因此，針對同樣的研究問題，我們所提出的假設也有可能是：「在結婚時男女雙方的年齡差距越小，這段婚姻以離婚收場的可能性就越低。」在這個新的假設裡，我們是以不同的方式說明結婚年齡這個變項。

我們也可以提出一個假設，詳細說明某個關係在某些條件下成立，但在其他條件下則不然。正如同李柏森（Lieberson, 1985: 065）所說的：「爲了要評估某個特定

的因果命題的效用，清楚明確地陳述該命題能成立的條件，是一件重要的事。」例如，有個假設的內容是：「在結婚時，結婚對象的年紀越小，這段婚姻以離婚收場的可能性就越高，除非夫妻雙方都是某種緊密團結的傳統宗教社群的成員，因爲在此種社群中，早婚是一種常態。」

形成研究問題與假設，不一定要以固定的步驟或是階段來進行。我們可以先形成一個暫時性、試驗性的研究問題，然後發展出一些可能的假設。這些假設可以幫助我們更加精確地陳述研究問題。在這個過程中，研究問題和假設之間可以不斷地交互作用，而我們也必須發揮自己的創造力。

你可能覺得好奇的是，在從一個主題轉換成一個可以考驗的假設時，理論應該如何與這個過程結合在一起。回想一下，在第三章裡我們曾經提過，理論其實有許多種不同的形式。我們可以利用一般的理論性議題當作主題的來源。理論提供我們一些可以轉換成變項的概念，也提供了我們推論的方式或機制，幫助我們把變項連結在一起，以形成研究問題。假設可以回答研究問題，同時也是一個由理論發展而成的、尚未經過考驗的命題。我們可以在一個抽象的、概念性的層次上表達假設，或以一種較爲具體的、可測量的形式，重新陳述這個假設。特定研究的範例或許有助於更加具體說明研究過程的各個部分。關於三個量化研究的範例，請參閱圖表 6.1 的內容，而關於兩個質化研究的範例，請參閱圖表 6.2 的內容。

結論

你在這一章所讀到的內容，是在開始進行一項研究之前所必須知道的一些基礎知識。你已經看到，存在於質化及量化取向之間的差異，會以何種方式引導我們以不同的方法準備一項研究。在所有類型的研究中，你必須將一個主題聚焦成爲一個更具體的、焦點更爲明確的研究問題。這兩種進行研究的主要取向，在研究的形式與進行的次序方面，會引導你作出截然不同的決定。而在關於何時以及如何聚焦研究問題方面，也各自有不同的解答。最有效的取向其實仍然取決於你所選擇的主題、你的目的、研究結果的預定用途、你對於社會科學所採取的傾向、你個人的假定與信念等等。

一般而言，量化研究所採取的是一種線性的路徑，並且強調客觀性。在量化研究中，你將會使用明確的、標準化的程序，以及一種因果解釋。量化研究所使用的是一種變項和假設的語言，此種語言被廣泛地應用於許多科學領域中，而其基礎則來自於實證主義傳統。量化研究的過程通常是演繹的，而且在蒐集資料之前，會有一系列清楚明確的步驟：把主題聚焦成一個焦點更明確的研究問題、把模糊不清的理論性概念轉變成更加精確的變項、並發展出一個或多個可以進行考驗的假設。在實際進行研究時，你可能會在不同的步驟之間來來回回地移動，但整體的過程大致上是朝向一個單一的、線性的方向流動。除此之外，你也應該特別地小心，盡量避免在發展出假設以及因果解釋的過程中出現邏輯方面的錯誤。

圖表 6.1 量化研究的範例

研究引文以及標題	李吉威以及艾瑞克森（Ridgeway and Erickson, 2000），「創造以及散播關於地位的看法」（"Creating and Spreading Status Beliefs"）	穆錫克、威爾森、以及畢能（Musick, Wilson, and Bynum, 2000），「種族及正式擔任志工工作：階級與宗教的差別效果」（"Race and Formal Volunteering: The Differential Effects of Class and Religion"）	巴洛、巴洛、以及奇里寇斯（Barlow, Barlow, and Chiricos, 1995），「媒體中的經濟狀況與犯罪意識型態」（Economic Conditions and Ideologies of Crimes in the Media）
使用的方法論技術	實驗	調查	內容分析
主題	人們發展出關於他人社會地位的看法之過程	白人與黑人成人擔任志工的比例	美國大眾媒體對於犯罪者的描述
研究問題	在個人進行互動時，那些影響互動方式的外在、結構性因素，是否會影響人們將來對於某種類別裡整體成員的看法？	在黑人及白人可以取得的資源中，哪些資源能夠解釋為何黑人較少參與志工活動的現象？	經濟狀況是否會影響媒體描述犯罪者的方式？
考驗的主要假設	當互動有限，且在互動時某種類別的成員會運用更多技巧時，人們會被「教導」對不同類別的人們作出地位之間的區別。	社會階級以及宗教會影響黑人與白人在擔任志工方面是否會有所差異。	當經濟狀況較差時，媒體對於犯罪的扭曲報導，會以較為負面的方式（譴責罪犯）來描述犯罪者。
主要的自變項	人們在與某個類別裡的成員進行互動時，此種互動是否顯示出該類別成員擁有較為優越或是較為低劣的工作技巧	社會階級、宗教參與、種族	1953 到 1982 年之間數年的失業率
主要的依變項	人們是否發展出一種不平等的看法，並將此種看法推論到該類別裡的全體成員	某人是否曾在五種中的任何一個組織（宗教、教育、政治、勞工、年長者、或是地區性組織）擔任過志工	是否發生扭曲報導的現象，測量方式是評估在數年之間，媒體關注程度（在《時代》雜誌上的文章）以及犯罪統計數據之間不一致的狀況
分析單位	個別的大學生	個別的成人	媒體報導
母群	所有的個人	美國境內所有的成年白人及黑人	所有的美國大眾媒體報導

圖表 6.2　質化研究的範例

研究引文以及標題	呂及范恩（Lu and Fine, 1995），「民族正統性的呈現：中國食物作為一種社會成就」（"The Presentation of Ethnic Authenticity: Chinese Food as a Social Accomplishment"）	莫洛契、弗登伯格、以及鮑森（Molotch, Freudenburg, and Paulsen, 2000），「歷史會重覆，但是如何重覆？城市特性、都市傳統、以及地方的成就」（History Repeats Itself, but How? City Character, Urban Tradition, and the Accomplishment of Place）
使用的方法論技術	田野研究	歷史比較研究
主題	在美國，民族文化在可被接受的範圍內所展現出來的方式，以及民族文化配置文化資源的方式	城市發展出一種特殊的都市「特性」的方式
研究問題	中國餐館如何呈現食物以平衡正統性、並且滿足非中國人的美國顧客？	為什麼聖塔芭芭拉（Santa Barbara）和凡圖拉（Ventura）這兩個位於加州的城市在表面上看起來非常相似，但卻發展出非常不同的特性？
紮根理論	民族餐廳會入境隨俗地把供應的食物美國化，以符合當地民眾的口味，但也建立一種正統性的印象。這是一種妥協的過程，既要滿足顧客的期望與口味方面的傳統，也要滿足一種異國的而且正統的飲食體驗。	作者們使用了兩個概念：「拼湊」（"lash up"）（許多因素之間的交互作用）與「結構」（過往事件對於後續事件造成限制）來詳細說明特性及傳統。經濟、政治、文化、以及社會因素會結合在一起，創造出獨特的文化經濟地區。視脈絡的不同，相似的力量可能會產生完全相反的結果。
拼貼	作者觀察並訪談了四家中國餐館，也利用了過往研究的證據。	作者們利用歷史記錄、地圖、照片、官方統計資訊、以及訪談等等。除了經濟和社會情況，他們也檢視了自願性的協會及實體的材料。
過程	餐館作出一些修正和改變，以符合可獲得的食材、市場立基、以及當地顧客的文化與食物品味。	兩個城市的情況，在回應石油與公路方面造成了不同的經濟發展狀況。凡圖拉以石油產業為基礎，形成了一個雇用許多當地居民的大型工業。聖塔芭芭拉則是限制石油和公路兩者的發展，反而專注於創造觀光產業。
脈絡	中國餐館，尤其是位於喬治亞州雅典地區的四家中國餐館	過去一百年來的加州海岸中部地區

在質化研究中，你可能會依循一種非線性的路徑，並把強調的重點放在逐漸熟悉一個自然情境或是特定文化歷史脈絡的各種細節上。質化研究比較沒有標準化的程序或明確清楚的步驟，因此你必須經常爲某個情境或是研究而隨機應變發明一些適合的技術。質化研究所使用的是一種個案與情境脈絡的語言，此種語言會引導你對特定的個案或過程進行詳盡的調查以追尋真實性。在質化研究中，關於研究規畫與研究設計方面的決定，很少與資料蒐集階段完全明確地區隔開來，而是在資料蒐集階段的早期不斷地繼續發展。事實上，在進行質化研究時，你所採取的是一種歸納的過程，此種過程鼓勵你依據你從資料中所得到的知識，以一種緩慢的、有彈性的方式，朝向一個具體的焦點逐漸演變。紮根理論會從你對於資料和脈絡的持續反思中，逐漸地浮現出來。

質化研究及量化研究之間的區別，經常被過度誇大，也過於死板地被一分爲二。某種取向的支持者，通常會以自己的假定與標準，評判另外一種取向的研究。量化研究者強烈要求知道所使用的變項以及被考驗的假設，而質化研究者則強烈反對此種把人性轉化爲冰冷數字的作法。一個受過良好訓練而且謹慎的社會研究者，將能從不同的角度去瞭解並欣賞每種研究取向，而且也承認每個取向都有其優點及缺點。社會研究的終極目標，是更加清楚地瞭解與解釋社會世界，而要達成這個目標，則有賴於社會研究者對於每種研究取向的內容，有更爲深入的瞭解。

問題回顧

1. 當我們說質化研究比較偏向於實務中的邏輯、而不是重建邏輯時，其涵義爲何？
2. 當我們說質化研究所依循的是非線性路徑時，這是什麼意思？非線性路徑能夠在何處展現其價值？
3. 請描述自變項、依變項、以及中介變項之間的差異。
4. 爲什麼我們在社會研究中無法「證實」研究結果？
5. 請選擇一個你感興趣的主題，並發展出兩個研究問題。針對每一個研究問題，詳細說明其分析單位與母群。
6. 如果研究者使用了否定假設的邏輯，那麼這位研究者會用到哪兩種假設？爲什麼負面證據的力量比較強大？
7. 請將下列敘述重新陳述爲一個有自變項及依變項的假設：「一個人在一年之中開車的里程數，會影響這個人到加油站的次數，而且在這兩個變項之間，有一種正面的單向關係。」
8. 試比較量化研究者和質化研究者在處理個人偏見以及信任研究者這個議題上所使用的方式。
9. 質化研究者以及量化研究者如何運用理論？
10. 試解釋質化研究者詮釋資料的方式。請分別以第一級詮釋、第二級詮釋、以及第三級詮釋的區別進行說明。

注解

1. See Tashakkori and Teddlie (1998).
2. Ward and Grant (1985) and Grant and colleagues (1987) analyzed research in sociology journals and suggested that journals with a higher proportion of qualitative research articles address gender topics but that studies of gender are not themselves more likely to be qualitative.
3. See Kaplan (1964:3–11) for a discussion.
4. On the issue of using quantitative, statistical techniques as a substitute for trust, see Collins (1984), Porter (1995), and Smith and Heshusius (2004).
5. For discussion, see Schwandt (1997), Swanborn (1996), and Tashakkori and Teddlie (1998:90–93).
6. For examples of checking, see Agar (1980) and Becker (1970c).
7. Problem choice and topic selection are discussed in Campbell and associates (1982) and Zuckerman (1978).
8. See Flick (1998:51).
9. *Exceptions* are secondary data analysis and existing statistics research. In working with them, a quantitative researcher often focuses the research question and develops a specific hypothesis to test after she or he examines the available data
10. See Ball and Smith (1992) and Harper (1994).
11. For place of theory in qualitative research, see Hammersley (1995).
12. See Harper (1987:9, 74–75) and Schwandt (1997:10–11).
13. See Gerring (2007:20) and George and Bennett (2005).
14. See Blee and Billings (1986), Ricoeur (1970), and Schneider (1987) on the interpretation of text in qualitative research.
15. See Lieberson (1985:185–187) for a discussion of basic and superficial variables in a set of causal linkages. Davis (1985) and Stinchcombe (1968) provide good gen-eral introductions to making linkages among variables in social theory.
16. The logic of disconfirming hypothesis is discussed in Singleton and associates (1988:56–60).
17. See Bailey (1987:43) for a discussion of this term.
18. The general problem of aggregating observation and making causal inferences is discussed in somewhat technical terms in Blalock (1982:237–264) and in Hannan (1985). O'Brien (1992) argues that the ecological fallacy is one of a whole group of logical fallacies in which levels and units of analysis are confused and over-generalized.

第 7 章

質化與量化測量

測量的必要性
量化以及質化測量
測量的過程
信度與效度
量化測量指引
量表與指數
結論

簡而言之，測量本身並不是目的。只有從一種工具論的觀點來看，我們才能夠真正瞭解到測量的科學價值。在此種觀點中，我們想要知道的是測量所要服務的目的是什麼、測量在科學情境中被要求去扮演什麼樣的角色、以及測量在研究中發揮了什麼樣的功能。

——Abraham Kaplan, *The Conduct of Inquiry*, p. 171

哪些人應該被歸類為窮人？貧窮的情況又有多嚴重？在 1960 年代，為了要回答上述的這兩個問題，美國政府官員利用所謂的貧窮線來衡量貧窮的狀況，而新的社會福利計畫則是要提供協助給這些被定義為貧窮的民眾（像是學校教育、健康照護、以及居住方面的協助等等）。一開始的時候，這些官員的想法是，當一個家庭無法購買足夠的食物以使家庭成員獲得足夠的營養時，就是屬於非常貧窮的狀況。當時的一些研究顯示出，低收入的人們必須把收入的三分之一花費在食物上。政府官員們走訪了許多雜貨店，並且計算了一個家庭必須花費多少錢來取得一些低成本的食物，然後把這些食物價格的總金額乘以三，就得到所謂的貧窮線的標準。從那時候開始，這個數字已經根據通貨膨脹的狀況而加以調整過。當布雷狄（Brady, 2003: 730）檢閱了從 1990 到 2001 之間的出版刊物時，他發現在美國有百分之六十九點八的研究，都是使用官方所訂立的比例。然而，為數眾多的研究卻發現到，美國官方測量貧窮的方式有重大的缺陷。當美國國家科學研究委員會（National Research Council）於 1995 年檢視了此種測量方式時，委員會中的成員們宣稱此種測量方式已經過時，並且認為此種方式不應再被繼續使用下去。他們認為此種測量貧窮的方式只是武斷、任意地指定一個收入水準，而且「此種測量方式會模糊了在不同群體之間、以及在不同地理脈絡之間的貧窮範圍的差異，而且在描繪趨勢的演變方面，只能提供一幅不正確的圖像」（Brady, 2003: 718）。美國官方測量貧窮的方式無法捕捉貧窮現象的複雜本質，而且也未把新的家庭情境、新的福利計畫、稅制的改變、以及新的生活支出等等事物列入考量。使此種混亂狀況更為雪上加霜的是，由於每個國家用來測量貧窮的方式都各自不同，因此我們根本無法比較美國以及其他國家在減少貧窮方面的變化情形。由於所有測量

貧窮的方法都已經有所改善、而且將會把數量更為龐大的人們歸類為貧窮人口，因此，幾乎沒有任何的政府官員想要改變現有的測量方式。

測量的必要性

身爲研究者，我們每天都會遇到各種不同的測量方式，像是測量智力的「史丹佛—比奈智力測驗」（Stanford Binet IQ test）、測量種族隔離狀況的「相異指數」（the index of dissimilarity）、或者是測量犯罪數量的「統一犯罪報告」（uniform crime report）等等。我們需要各種測量的方式來考驗假設、評估解釋、爲理論提供實徵支持、或者是研究應用的議題等等。我們測量社會生活中某種層面的變化程度或者是範圍（像是自尊、政治權力、疏離、或者是種族歧視等等各種不同層面）的方法，正是本章的中心主題。不論是量化研究或者是質化研究，都會用到測量，但是量化研究者最爲關心測量的問題。在*量化研究*中，測量本身就是研究過程中的一個明確獨立的步驟，而且其發生時間早於蒐集資料的階段。量化測量擁有特殊的專門用語以及一整套的技術，這是因爲量化測量的目標，就是要精確地捕捉實徵社會世界的細節，並且用數字來表達我們所發現的事物。

而在*質化研究*裡，我們會用數字以外的其他替代事物來進行測量，而且測量本身比較不像是一個獨立的研究步驟。由於質化研究的過程是比較偏向歸納式的，因此我們會在蒐集資料的過程當中，同時進行測量並且創造出新的概念。

進行測量並非是一種我們可以快速跳過的、神秘難解又充滿技術性的議題（像是拉出一條捲尺來測量某個物體的長度、或者是把一個物體放在磅秤上以測量重量一樣）。測量能夠把我們對於社會世界所觀察到的、所思考的事物，緊密地與我們在其中所發現的事物連結在一起。品質不佳的測量方式，可能很快地就毀了一項原本相當優秀的研究。測量在日常生活中也有重大的影響力，舉例而言，心理學家以及其他專家就一直爲了智力的意義以及測量方式而爭論不休。我們使用智力「測驗」來測量一個人在學校裡、在應徵工作方面、以及在關於種族或是其他遺傳優點方面的智力表現。但是，智力到底是什麼？大多數的智力「測驗」所測量的，只是分析推理的能力而已（例如，一個人以抽象方式進行思考的能力、或者是以合乎邏輯的方式進行推論的能力）。然而，我們都承認智力有許多不同的類型，像是藝術的、實際的、機械的、以及創意的。有些人甚至認爲還有更多種不同類型的智力，像是社交與人際的、情緒的、身體動覺的、音樂的、或者是空間的智力。如果智力具有多重的形式、但是我們卻只有狹隘地測量其中一種形式的話，我們就會在許多事物方面畫地自限，像是學校提供教育以及促進學習的方式；我們選擇、評估、以及晉升員工的方式；以及整體社會重視多重人類能力的方式。

就如同本章開頭部分所指出的，我們測量貧窮的方式，會決定人們是否有資格能夠得到許多社會福利計畫的協助（例如住宅補助、食物津貼、健康照護、兒童照

顧等等）。有些人認爲，如果人們無法負擔食物以預防營養不良的狀況，這些人就該被歸類爲窮人。其他人則認爲，*貧窮*的定義應該是指年收入少於全國平均收入（中位數）的一半。另外還有人認爲，*貧窮*的定義應該是指某人所賺得的收入少於能夠維持最低生活條件的「基本生活工資」（"living wage"），而基本生活工資的計算方式，是指要達到在健康、安全、適度衛生、住宅、衣著、飲食、交通等等方面的最低基本標準所需要的收入。在測量貧窮方面的不同方式與決定，將會嚴重地影響到數百萬人的日常生活情況。

我們在日常生活中會使用到許多種不同的測量方式，例如，今天早上我起床之後，站到體重計上面去測量體重，看看我的減重計畫是否發揮了作用。然後，我瞄了溫度計一眼，以決定今天是否要加穿一件外套。接著，我坐進我的車子並且檢查了一下油量表，確定還有足夠的油可以讓我開到學校去。在我開車時，我隨時注意車上的速度表，以免因爲超速而被警察開罰單。到了早上八點爲止，我已經測量了體重、溫度、油量、以及速度，而這些都是關於實體世界的測量結果。這些精確的、發展完善的日常生活測量工具，在自然科學中是基本而且必要的。

我們在日常生活中對於非實體世界的測量方式，通常比較不是那麼地精確。當我們說某間餐廳的食物很棒、帕布羅真的很聰明、凱倫對生命抱持著負面的態度、強森有很強烈的偏見、或者是昨晚看的電影裡有太多暴力成分時，也是在進行某種形式的測量。像是「有強烈的偏見」或者是「有太多暴力成分」之類的日常判斷，都是草率而且不精確的。

測量工具也能夠擴展我們的感官知覺能力。天文學家或者是生物學家會使用望遠鏡或是顯微鏡來擴展人類天生的視力範圍。測量工具能幫助我們看見原本無法看到的事物，而且也讓我們能夠觀察到一些理論預測其存在、但是過去卻無法看見或者是未知的事物。例如，我們可能無法用我們天生的感官去看見或是感受到磁力。磁力來自於一個關於實體世界的理論。我們可以用間接的方式來看見磁力的作用，像是小鐵片會被磁鐵所吸引過去之類的情況。磁鐵讓我們可以「看見」或者是測量磁場。與我們天生的感官相比之下，科學測量工具更加靈敏、比較不容易因爲特定的觀察者而有所變化、而且也能夠產生更爲精確的資訊。我們都承認，溫度計能夠比觸覺提供更爲明確的、精準的溫度資訊。同樣地，比起我們用手抱起一個五歲女孩並且覺得她是「沉重」或是「輕盈」，一個好的體重計能夠提供關於這個女孩更爲具體、一致、而且精確的體重資訊。

在我們能夠進行測量之前，我們必須對於我們感興趣的事物有非常清楚的概念。這是非常關鍵的原則，因爲測量可以把我們腦中的想法、與我們在實徵世界中所做的事物連結在一起，好讓這些想法變得清晰可見。自然科學家會用到許多理論，而且他們會創造出各種測量方式來「看見」非常微小的東西（分子或者是昆蟲的器官）、或者是非常巨大的東西（巨大的地質陸塊或者是星球），而這些東西都是無法透過一般的感官來觀察的。因此，所有的研究者都在不斷地創造出新的測量方式或者是工具。[1]

我們或許可以輕易地看見在社會研究中所測量的年齡、性別、以及種族等事物（例如，年齡所造成的皺紋、不同性別的身體特徵、膚色、以及眼睛的形狀等等），

但是社會世界的許多層面（例如，態度、意識型態、離婚率、偏差行爲、社會角色等等）是非常難以直接進行觀察的。正如同自然科學家會針對「無形的」分子以及重力而創造出一些間接的測量方式一樣，社會科學家也會創造出不同的測量方式，來測量社會世界中那些難以直接觀察的部分。

量化以及質化測量

在所有的社會研究中，不論是質化研究或者是量化研究，我們都會把資料與想法或是概念連結在一起。我們可以把在一項研究中的資料想成是某個概念的實徵表現。測量能夠把資料連結到概念，然而，依據我們的資料以及研究取向主要是量化或者是質化取向，測量的過程也會因而有所不同。有三項特徵可以區別量化與質化取向之間的不同測量方式。

第一項差異是時間的安排。在量化研究裡，我們會在規畫階段裡去思考變項、並且把變項轉化爲特定的行動。規畫階段早於、並且獨立於資料的蒐集以及分析階段。而在質化研究裡，我們會在資料蒐集階段裡同時進行測量的工作。

第二項差異則與資料本身有關。在量化研究中，我們所使用的技術能夠產生數字形式的資料。通常，這樣的過程是來自於以演繹的方式將抽象概念轉變成特定的資料蒐集技術、以及這些技術所產生的精確數字資訊。數字資料代表了一種統一的、標準化的、以及簡潔的方法，以實徵的方式來呈現抽象的概念。在質化研究中，資料有時候會以數字的形式呈現，但是，在大多數時候，資料是書面的或是口說的文字、行動、聲音、符號、實體物體、或者是視覺影像（例如，地圖、照片、影片等等）。與量化研究不同的是，質化研究不會把所有的觀察結果都轉化成爲一種像是數字之類的單一、共同的媒介，而是讓資料維持著各式各樣非標準化的外形、大小、以及形式。量化研究所使用的數字資料可以把資訊轉化成爲一種標準的、濃縮的格式，而質化資料的特性則是數量龐大的、種類多元的、以及非標準化的。

第三項差異則牽涉到我們把概念與資料連結在一起的方法。在量化研究裡，在我們真正去蒐集資料之前，我們會去深入並且仔細地思考各種概念。我們會審慎地選擇測量技術，以銜接抽象概念以及實徵資料。當然，在我們蒐集並且檢視過資料之後，我們不會就此停止思考或者是不再繼續發展新的想法，而是會以清楚的、仔細思考過的概念作爲出發點，並且去考量我們應該要以何種方式來測量這些概念。

在質化研究裡，我們也會在蒐集資料之前先仔細地思考各種概念。然而，我們所使用的許多概念，都是在資料蒐集的過程之中或是之後所發展出來以及修正的。我們會同時去重新檢視以及思考資料與概念，並且會讓這兩者之間產生互動。當我們在蒐集資料時，我們同時也在對資料進行深入的反思，並且產生新的想法，而這些新的想法則會爲我們指出方向，並且暗示新的測量方法。接下來，這些新的測量方法則會影響我們蒐集其他額外資料的方式。簡而言之，我們是在一個持續進行中的、互動的過程中去連結概念以及資料。

總結來說，在量化研究中，我們會在正式去蒐集資料之前，先仔細思考測量的方式，並且作出相關的決定。而量化研究的資料是以一種標準化的、統一的格式來呈現的，也就是數字。對比之下，在質化研究中，我們大部分的思考過程以及與測量有關的決定，都是發生在蒐集資料的過程當中，而且資料的呈現形式是相當多元的。

測量的過程

當我們在進行測量時，我們把在我們腦海中的無形概念、想法、或者是構念，與我們用來觀察實徵世界的技術、過程、或者是程序連結在一起。[2] 在量化研究裡，我們傾向於從抽象的想法開始，並且以實徵資料作爲結束。而在質化研究裡，我們在蒐集資料時，會把資料以及想法混合在一起。然而，在特定的具體研究中，真正的情況其實是比較混亂的，而且也會比上述的一般性陳述更具有互動的特質。

我們在測量時會用到兩種主要的過程：概念化以及操作化。**概念化**（conceptualization）所指的是先選擇一個抽象的構念，然後爲這個構念賦予一個概念性的或者是理論性的定義，以使這個構念更爲凝鍊。所謂的**概念性定義**（conceptual definition）是指以連結到其他想法或是構念的具體文字或是理論性詞彙，來陳述在你腦海中的想法。目前沒有任何神奇的方法，可以把一個構念轉化成爲一個精確的概念性定義。在這個轉化的過程中，會牽涉到進行仔細的思考、直接的觀察、向他人請教、閱讀別人所說過的事物、以及嘗試可能的定義等等。

一個好的定義應該要有一個清楚的、明確的、而且具體的意義，不能有任何模稜兩可或者是模糊不清的地方。有時候，概念化是一種具有高度創造力的過程，而且能夠產生新的洞見。有些學術文章專門在討論如何針對一些關鍵概念進行概念化的過程。梅爾賓（Melbin, 1978）把*夜晚*轉化爲邊界的概念，吉布斯（Gibbs, 1989）分析了*恐怖主義*這個概念的意義，而波爾以及卡利（Ball and Curry, 1995）則討論了*街頭混混*的意義爲何。關鍵的重點在於，我們必須對概念擁有清楚的、明確的定義，才能夠發展出健全的解釋。

某一個構念可能會有數種不同的定義，而且人們可能會對這些定義產生爭論。此外，概念性定義會與理論架構連結在一起。例如，一名衝突理論家可能會把*社會階級*定義爲社會上的某一群人所擁有的或者是缺乏的權力或是財產。而一名結構功能論者，則可能把*社會階級*定義爲共享一種社會地位、生活型態、或者是主觀認同的一群個體。雖然人們會對不同的定義產生爭議，但是我們身爲研究者，就應該總是清楚地說明我們正在使用的是哪一種定義。

有一些構念（例如：疏離）是極爲抽象而且複雜的。這些構念之中包含了一些較低層次的概念（例如：無力感），而這些較低層次的概念則可以被轉化成爲更加具體明確的概念（例如：一種無法改變自

概念化　針對抽象想法或者是概念發展出清楚的、嚴謹的、以及系統的概念性定義的過程。

概念性定義　針對一個構念所賦予的仔細而且系統性的定義，並且清楚地以書面文字的方式來呈現該定義。

已居住地點的無奈感受）。其他的構念則是具體而且簡單的（例如：年齡）。我們必須注意到一個構念的複雜以及抽象程度。例如，在爲一個像是*年齡*（例如，從出生之後已經活了幾年）之類的具體構念下定義時，就會比爲一個像是*士氣*之類的複雜、抽象概念下定義還要容易地多。

在我們能夠進行測量之前，我們必須精確地區別我們所真正感興趣的事物、以及其他近似的事物。這其實是一種常識而已。如果我們不知道所要追尋的事物是什麼的話，那麼我們要怎麼去進行測量呢？例如，除非一名生物學家先知道什麼是癌細胞、擁有一部顯微鏡、而且也能區別顯微鏡中所看到的細胞與非細胞物質之間的差異之處，否則的話，這名生物學家根本無法去觀察癌細胞。測量的過程所牽涉到的不僅僅是擁有測量的工具而已（例如，一部顯微鏡）。在測量的過程中，我們需要有三樣事物：構念、測量工具、以及辨識出我們尋找的目標的能力。[3]

舉例而言，假設我現在想要測量教師士氣的話，我首先必須爲*教師士氣*（*teacher morale*）一詞下定義。*士氣*這個構念的意義是什麼？作爲一個變項構念，士氣可能會有數種不同的值：高或低、優或劣。接著，我必須爲我的構念創造出一種測量方式，測量的形式可能是進行問卷調查、檢視學校裡的文件記錄、或者是對教師們進行觀察。最後，在問卷調查的回答中、學校的記錄裡、或者是觀察結果中，我還必須把士氣和其他的事物區隔開來。

社會研究者的工作比自然科學家的工作還要更爲困難，因爲社會測量牽涉到與人們交談、或者是觀察他們的行爲。與星球、細胞、或者是化學物質不同的是，人們所給的答案以及他們的行動可能會是模稜兩可而且不精確的。人們可能會因爲意識到自己正在被問問題或是被觀察，因而做出一些與平常不同的反應。因此，社會研究者面臨了雙重的負擔：首先，他們必須擁有一個清楚的構念、良好的測量方式、以及辨識出自己所追尋的事物的能力；其次，他們也必須盡力去測量流動的以及令人困惑的社會生活，而且社會生活還可能因爲人們意識到研究者正在試圖去測量他們，因而產生一些變化。

我要如何發展出*教師士氣*的概念性定義、或者至少是一個讓我可以開始進行研究的暫時性定義呢？我從自己對於士氣的日常理解開始：也就是某一種像是「人們對於事物的感受」之類的模糊想法。我詢問了一些朋友，看看他們是如何定義士氣這個構念的。我也去查了完整本的字典以及同義字辭典，而這些書則告訴我一些定義或是同義字，像是「信心、精神、熱誠、愉悅、團隊精神、對於某種事物的心理狀態」等等。我也造訪了圖書館，並且搜尋了關於士氣或者是教師士氣的研究文獻，看看其他人是怎麼下定義的。如果某個人已經提出了一個極佳的定義，我可能會直接借用這個定義（當然，我會在研究中註明引用的出處）。如果我無法找到一個適合研究目的的定義，那麼我可能會轉而求助於團體行爲、個人心理狀態、以及相關的理論，以尋求一些想法或概念。當我蒐集了各式各樣的定義、許多定義的不同部分、以及相關的觀念時，我就會開始逐漸地看出核心觀念的邊界爲何。

現在，我已經找到了許多定義，所以必須把這些定義整理一下並且去蕪存菁。大部分的定義都主張士氣是一種精神、感覺、對於某件事物的心理狀態、或者是一

種團體感覺。我把這個構念中的兩個極端區隔出來，這麼做可以幫助我把這個概念轉變成爲一個變項。高士氣牽涉到信念、樂觀、愉悅、團結的感覺、以及爲了共同的目標而願意忍受困境。低士氣則是剛好相反，也就是缺乏信心、悲觀、憂鬱、孤立、自私、以及不願意爲了他人而付出努力。

由於我所感到興趣的是*教師*士氣，所以我必須先瞭解教師們，才能對他們詳細地說明這個構念。可行的策略之一，是針對高或是低教師士氣的範例列出一張表。高教師士氣包括了提到一些關於學校的正面事項、不會抱怨做額外的工作、或者是喜歡和學生在一起等等。低教師士氣則包括了經常抱怨、除非必要否則不出席學校的活動、或者是尋找其他的工作等等。

士氣牽涉到一種對於其他事物的感覺，而一個人會對某種事物擁有士氣。我列出了各式各樣教師們可能會有感覺的事物（例如，學生、家長、薪資、學校行政、其他的教師、教師專業等等），而在這個過程中，發生了一個在發展定義的過程中經常出現的議題。教師士氣是否有數種不同的類型？或者，上述的這些事物都只是單一構念的不同面向而已？由於目前仍然沒有完美的答案，因此我必須決定，士氣是否意指一種擁有多種不同部分的單一而且廣泛的感覺，或者是數種明顯不同的感覺。

除此之外，我的構念所能適用的分析單位是什麼？是一個團體或者是個人？士氣是否爲個人或是團體（例如：學校）的特徵、或者是兩者都有？爲了配合我的研究目的，所以我決定將士氣定義爲適用於團體，而這也意味著我的分析單位是一個團體，也就是在一所學校裡的所有教師。

我還必須把我有興趣的構念和其他的相關想法區別開來。我對於教師士氣的構念，與其他的相關概念有何相似之處或者是相異之處？例如，*士氣*是否有別於*情緒*（*mood*）？我認爲與士氣相較之下，情緒是比較個人化的、而且也是比較暫時性的感受。同樣地，士氣也與樂觀或是悲觀不同，因爲後兩者都是個人對於未來的看法或是展望。士氣是一種團體的感覺，其中可能包括了對於未來的正面或是負面感受，以及相關的信念與感覺。

在概念化的過程裡，我們必須去仔細思考某一個構念的各種可能意義。到目前爲止，我已經知道教師士氣是一種心理狀態或是感覺，其範圍是由高（樂觀的、愉悅的）到低（悲觀的、沮喪的）；士氣有數個不同的面向（關於學生的、關於其他教師的）；士氣是一種團體的特質；而且士氣會持續數個月之久。比起我剛開始要進行研究時，現在我心裡對於我想要測量的事物已經有了一幅更爲清晰具體的圖像。如果我沒有先經過概念化的過程，我就可能只會試著去測量我一開始的想法：「人們對於事物的感覺」。

然而，即使已經經過了概念化的過程，仍然還有些模糊不清的地方。爲了要完成概念化的過程，我們有必要去設定邊界。我必須清楚明確地決定我想要包括在內的以及排除在外的事物。例如，哪些人才能算是教師？教師是否包括了提供指引的諮商者、校長、體育教練、以及圖書館員？而實習老師、兼職教師、或者是代課教師算不算是教師呢？*教師*這個詞彙是否包括任何一個以教學維生的人，即使這個人並未受聘於學校裡呢（例如，一位企業的訓練師、指導學徒的在職指導者、訓練民眾的醫院醫

操作化　從一個構念的概念性定義轉化成為具體活動或是測量方法的過程，而這些具體活動或是測量方法可以讓研究者以實徵的方式去進行觀察。

操作性定義　一個變項，其內容是在實徵世界中進行測量或是指出該變項的具體行動。

師）？即使我把我對教師的定義限制爲在學校裡從事教學工作的人，那麼什麼樣的機構才能算是學校？學校可能包括托兒所、訓練醫院、大學的研究所、幫助人們參加標準化測驗的營利機構、訓練狗的學校、教導學生如何打籃球的夏令營、以及教導人們如何駕駛半拖車或卡車的職業學校等等。

有些人直覺地認爲，教師就是指一位全職的、受過專業訓練的學校雇員，負責教導一到十二年級的學生，並且在一天當中會花費大部分時間在教室裡與學生相處。其他人則是採取一種法律上的、或者是政府官方的定義，此種定義會把一些已取得教師證照的人包括在內，即使他們並且在教室中進行教學；而且此種定義也會把一些未取得教師證照的人排除在外，即使他們正在教室裡與學生共處。重點在於，概念化的過程要求我必須以非常清楚的方式來進行思考。在我可以開始進行測量之前，我必須確切地知道我自己對於*教師*以及*士氣*的定義是什麼。我必須以一種非常清楚而且明確的、並且能夠讓別人理解的方式來陳述我自己的想法。

操作化（operationalization）能夠把概念性定義連結到一組測量的技術或者是程序，也就是一個構念的**操作性定義**（operational definition）（換句話說，就是依據明確具體的操作或是行動過程來陳述的定義）。操作性定義可以是一份調查的問卷；在一個田野情境中去觀察事件的方法；一種在大眾媒體中測量象徵性內容的方法；或者是反映、記錄、或是呈現在概念性定義中所表達的抽象構念的任何一種過程。

我們通常可以用數種方法來測量一個構念，而且有些方法會比其他方法更佳、也更切合實際。關鍵的重點在於，我們必須使測量方式符合於特定的概念性定義，而且必須考量到我們所面臨的所有實際的限制（例如：時間、金錢、可以找到的參與者等等）。我們可以從頭開始發展出一種新的測量方法、或者是使用其他的研究者正在使用的測量方法（請參閱資訊箱 7.1 想出測量方法的五個建議）。

操作化的過程連結了理論的語言以及實徵測量方法的語言。理論中有許多抽象的概念、假定、定義、以及因果關係。對比之下，實徵測量方法則是發生於具體而真實的情境中的具體行動，而且這些情境是由真正的人們以及事件所構成的。測量方法是針對我們所從事的操作或是行動而創造出來的，其目的則是在於指明某個構念在於具體、可觀察的實在中是否確實存在。

量化的概念化以及操作化

量化的測量是以一種直線前進的順序來進行的：首先是概念化、接著是操作化、然後是應用操作性定義或者是蒐集資料。我們必須嚴謹地將抽象觀念連結到能夠產生以數字形式來表達精確資訊的測量程序。而達到此種目的的方法之一，就是利用對應規則或者是輔助理論（auxiliary theory）。對應規則的目的，是在於把構念的概念性定義連結到測量構念的具體操作過程。[4]

資訊箱 7.1　**構想出測量方法的五個建議**

1. *牢記概念性定義*　不論任何一種測量方法，其基本的原則就是一定要符合在研究中所使用的構念的特定概念性定義。
2. *保持開放的心胸*　不要被單一的測量方法或是某種類型的測量方法給困住了。發揮你的創意，並且經常尋找更好的測量方法。除此之外，也要極力避免卡普蘭（Kaplan, 1964: 28）所說的「工具法則」（"law of the instrument"），意思是指固執地只用一種測量工具來測量所有的問題。
3. *向其他人借用想法*　只要適當地引述出處，你就不必害怕從其他研究者那兒借用想法。在其他的研究中可以發現一些關於測量方法的好點子，或者也可以將其他的測量方法加以修正。
4. *遇上困難是正常的*　當你試著去測量你想知道的變項時，經常會遇到邏輯上以及實務上的問題。有時候，只要事先仔細地思考以及規畫，就可以預見某些問題的發生、甚至還可以避免這些問題。
5. *不要忘記你的分析單位*　你的測量方法應該要符合你研究中的分析單位，並且允許你將結果推論到感興趣的母群。

對應規則（rules of correspondence）是一種邏輯性的陳述，其內容是一個指標如何對應到一個抽象構念的方式。例如，某條對應規則主張，只要某個人符合了十項具體的陳述，這就可以當作證據，讓我們大致上同意這個人具有強烈的反對女性主義的態度。此種輔助理論或許可以解釋指標以及構念之間是如何以及爲什麼會連結在一起的。卡明斯以及札勒（Carmines and Zeller, 1979: 11）提到：「對於社會研究而言，詳細說明概念以及指標之間關係的輔助理論，其重要性等同於連結不同概念的實質理論。」例如，或許我們想要測量疏離，而我們對於疏離的定義有四個部分，每一個部分個別屬於不同的生活領域：家庭關係、工作關係、與社區的關係、以及和朋友的關係。輔助理論可能會詳細地說明，在每一種不同的生活領域內的特定行爲或是感覺，可以當作證明疏離的堅實證據。該理論告訴我們，在工作的領域內，一個人產生疏離現象的指標，是指這個人覺得無法掌控某些事情的時候，其中包括了他的工作時間、地點、以及同事；他在工作時所做的事；或者是他必須多快地完成工作等等。

對應規則　一種標準，研究者利用這些標準而將抽象構念連結到實徵社會實在中的測量操作。

圖 7.1 清楚地顯示出了連結一個理論以及假設裡的兩個變項的測量過程。我們必須考量到三個層次：概念的、操作的、以及實徵的。[5] 在最爲抽象的層次，我們所感興趣的或許是在兩個構念之間的因果關係，或者是**概念性假設**（conceptual hypothesis）。而在操作性定義的層次，我們感興趣的是去考驗一個**實徵假設**（empirical hypothesis），以決定兩個指標之間的相關性程度有多高。而我們

概念性假設　一種假設的類型，此種假設以抽象的、概念性的詞彙來表達變項以及變項之間的關係。

實徵假設　一種假設的類型，在此種假設中，研究者以具體的實徵詞彙來表達變項，並且以一種可以觀察的、實徵的方式來表達被測量的指標之間的關聯性。

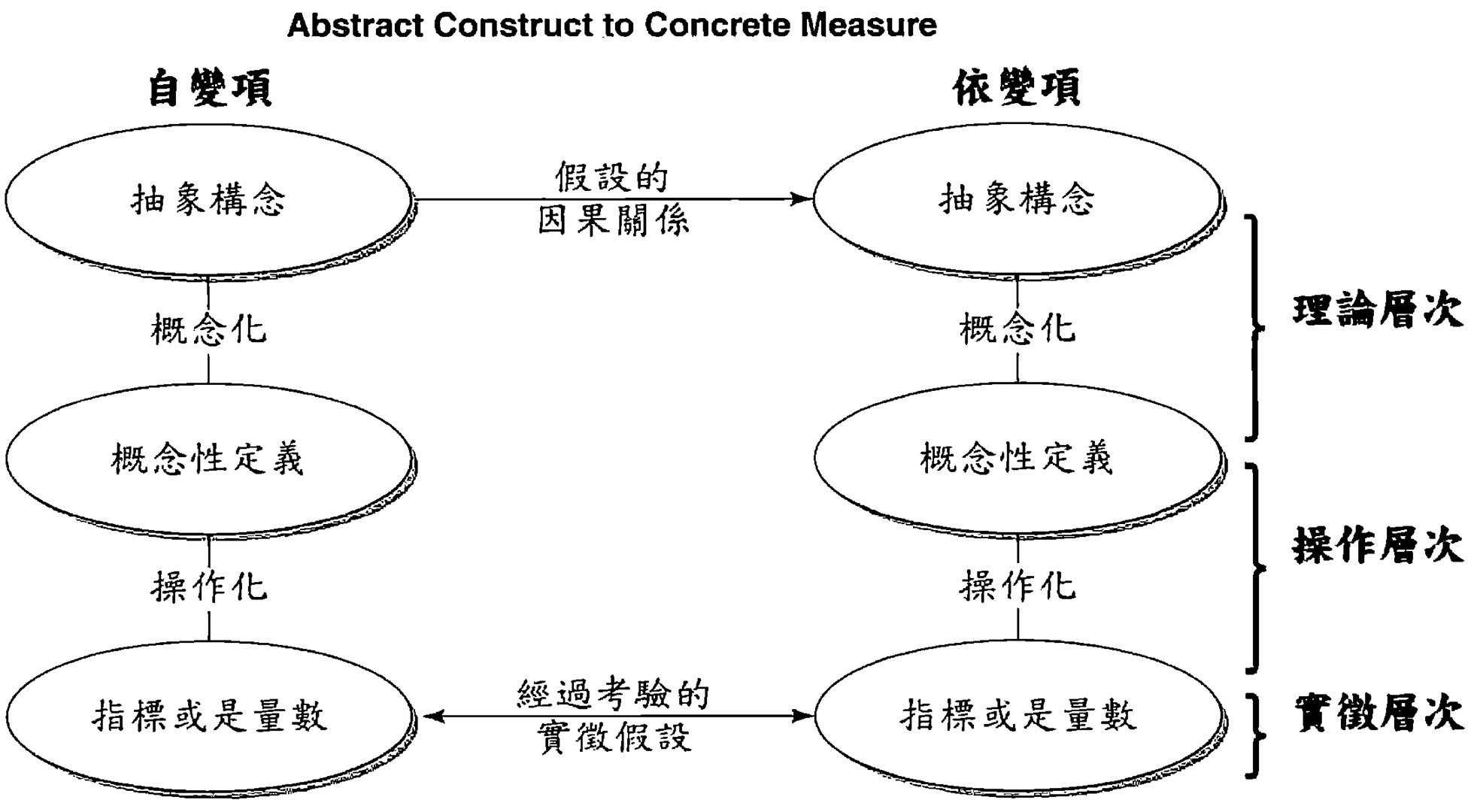


圖 7.1　概念化以及操作化

就是在這個層次上去考量相關程度、統計數據、問卷等等事物。至於第三個層次，則是實際社會世界的實徵實在。當我們把操作性指標（例如：問卷裡的題目）與一個構念（例如：疏離）連結在一起時，我們就是在捕捉實際社會世界裡正在發生的事物，並且將這些事物連結回到概念層次。

當我們在進行測量時，我們會把這三個層次連結在一起，並且以演繹的方式從抽象層次往具體層次移動。首先，我們先將一個變項概念化，也就是爲這個變項賦予一個清楚的概念性定義。接下來，我們會將這個變項操作化，也就是爲這個變項發展出一個操作性定義或者是一組指標。最後，我們會應用這些指標來蒐集資料以及考驗實徵假設。

讓我們回到剛剛提過的範例上。我應該如何爲我的教師士氣構念賦予一個操作性定義呢？首先，我先去閱讀其他人所出版的研究報告，並且看看是否已經有一個良好的指標。如果沒有任何現存的指標，我就必須從頭開始去創造一個指標。由於士氣是一種心理狀態或是感覺，所以我只能間接地透過人們的言語以及行動來測量士氣。我可能會爲教師們設計一份問卷，並且詢問他們對於我所下的定義中的各種士氣面向的感覺。我或許也會去實地造訪學校，並且去觀察教師們在教師休息室裡、在與學生互動時、以及在參與學校活動時的狀況。我可能會使用學校裡關於教師行爲的人事記錄，以尋找一些可能與士氣有關的陳述（例如：缺席、要求應徵其他工作的推薦信、工作表現報告等等）。我也可能會去調查學生、學校行政人員、以及其他人，以發現他們對於教師士氣的看法。不論我所選擇的是哪一種指標，我都會在發展指標的過程中（例如：寫下具體的問卷問題）進一步地修正我的概念性定義。

對於每一個變項而言，概念化以及操作化都是必要的。在先前的例子裡，士氣是一個變項，而不是一個假設。士氣可能是一個由其他事物所引起的依變項，也有

可能是一個引起其他事物的自變項。至於士氣到底是依變項或是自變項，則是取決於我們的理論性解釋。

質化的概念化以及操作化

概念化　在質化研究裡，我們不會在研究過程的早期就把抽象想法轉化爲理論性的定義，而是在蒐集資料以及分析資料的過程中去精鍊初步的「可行觀念」（"working ideas"）。*概念化*所指的是當我們努力地去「理解」或者是組織資料以及我們對於資料的初步想法時，形成一致的理論性概念的過程。

當我們去蒐集並且分析質化資料的時候，我們同時也會發展出新的概念、形成主要構念的定義、並且思考這些構念之間的關係。到了最後，我們會把概念以及構念連結在一起，以創造出理論性的關係。我們會在檢視資料（例如：田野筆記、照片以及地圖、歷史文件等等）時形成並且精鍊構念，而且我們會提出關於資料的理論性問題（例如，這算是階級衝突的個案嗎？事件的結果是什麼？會與之前有所不同嗎？爲什麼此種情況發生在這兒，而不是發生在別處？）。

我們必須用文字來表達清楚而且明確的定義、以及對於特定行動的描述，這些描述可以連結到其他想法，並且與資料緊密地結合在一起。在質化研究裡，概念化的過程主要是來自於資料。

操作化　在質化研究裡，操作化通常早於概念化（請參閱圖 7.3），並且提出歸納式的測量過程（至於演繹式的測量過程，請參閱圖 7.2）。當我們在進行觀察或者是蒐集資料時，我們可能會從初步的「可行觀念」中創造出概念性定義。在質化研究中，我們進行操作化的方式，不是把精鍊過的概念性定義轉化成爲測量操作，而是去描述特定觀察結果以及關於資料的想法如何促成可行的觀念，而這些可行的觀念則是概念性定義的基礎。

因此，質化研究裡的操作化主要是牽涉到發展出一項描述，以說明我們在進行觀察時如何使用可行觀念的方式。操作化描述了我們獲得特定觀察結果或者是蒐集資料的方式，以及在資料演變成爲抽象構念的過程中，我們努力去瞭解資料的方式。因此，質化的操作化比較像是一種在事情發生之後所提出的描述，而不是一種事前規畫的技術。

正如量化的操作化有可能會偏離嚴謹的演繹過程一樣，質化研究者汲取靈感的來源也可能超越了從某個特定研究環境所得到的資料。質化的操作化過程，包括了使用先前已經存在的技術以及概念，並且把這些技術以及概念與在資料蒐集過程中所浮現的技術以及概念混合在一起。

范特西亞（Fantasia, 1988）針對勞工的抗爭運動所進行的田野研究，正可以用來說明質化的操作化過程。范特西亞以*團結的文化*（*cultures of solidarity*）當作一個中心的構念。他把這個構念與下列兩種觀念連結在一起：其中之一是與充滿衝突的職場關係，另外一個則是在非管理階層的員工之間逐漸升高的階級意識。他把團結的文化定義爲一種由勞工所創造出來的文化表達方式，而且此種文化會隨著時間經

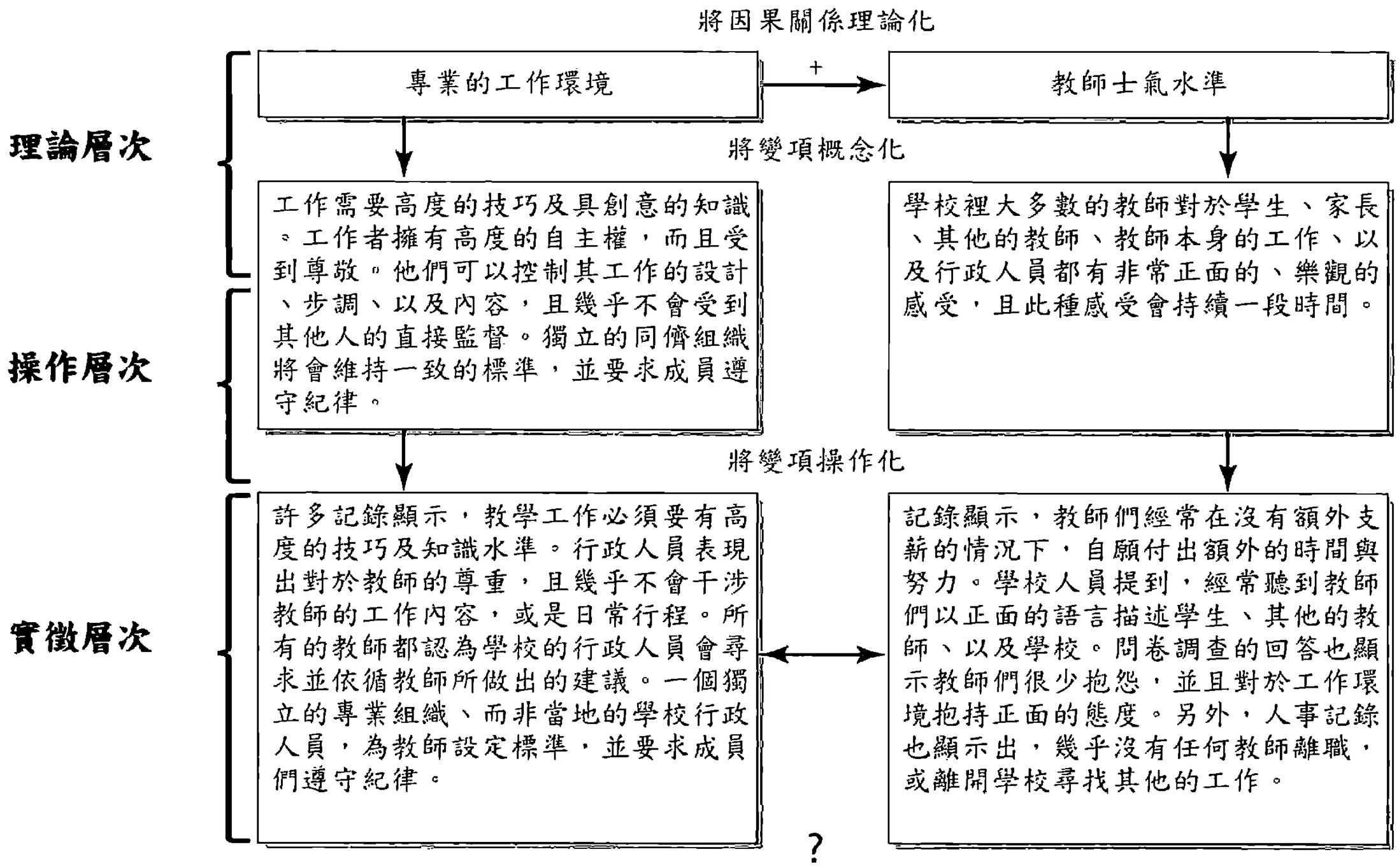


圖 7.2　假設的演繹測量過程的範例：專業的工作環境會提升教師士氣水準

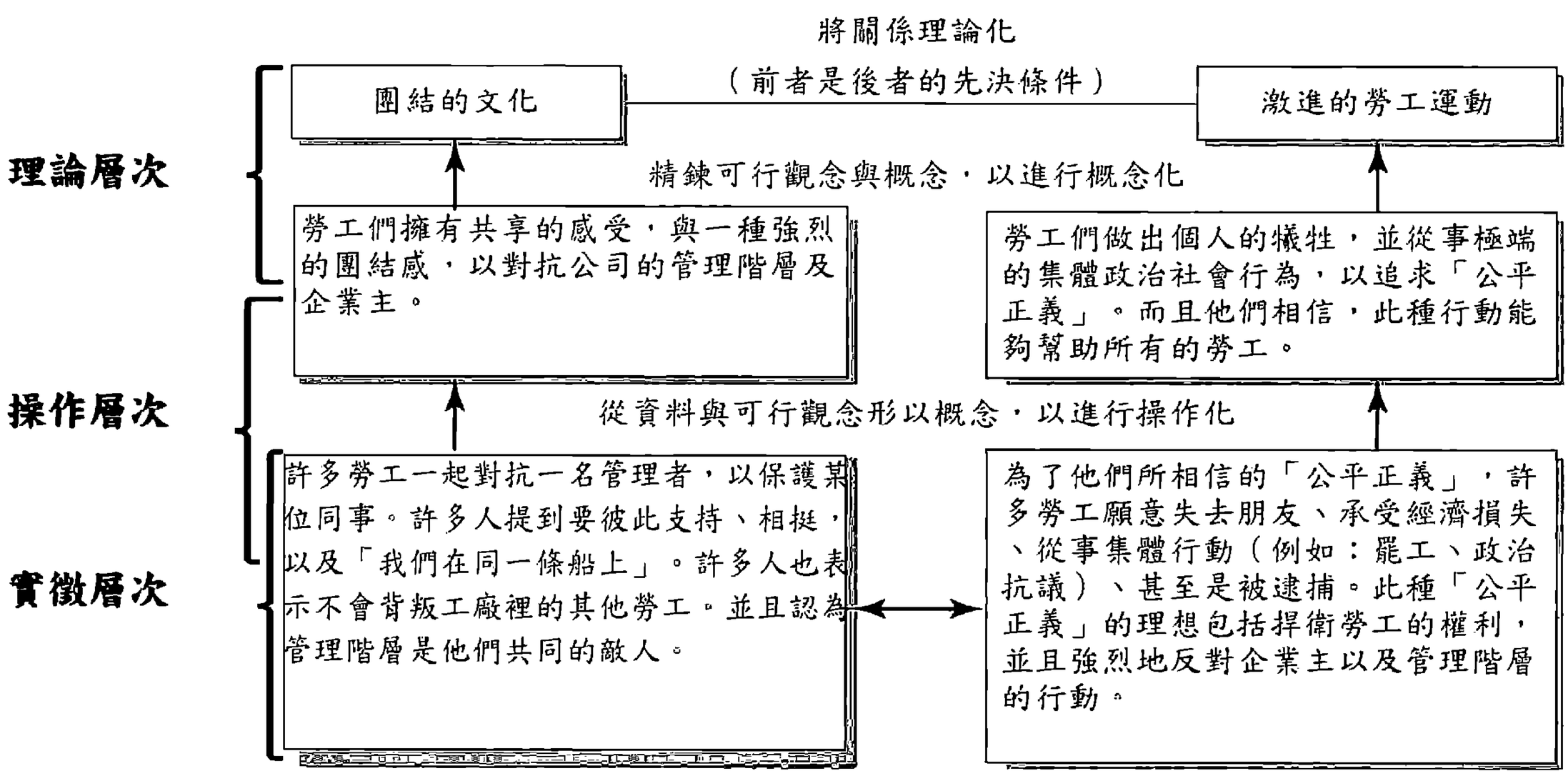


圖 7.3　假設的歸納測量過程的範例：激進的勞工運動可能會發生在勞工們創造出一種團結文化的職場裡

過而在特定的地點逐漸演變。隨著時間經過，勞工會發展出共享的感覺以及一種團結感，以對抗管理階層以及企業主。這是一種具有互動性質的過程，而且在這個過程中，勞工們會緩慢地發展出共同的想法、理解、以及行動。此種情況「比較不像是一種脫離現實的心理態度，而像是一種可以進行實徵調查的、範圍較爲廣泛的實踐以及計畫」（Fantasia: 14）。

個案化　在質化研究中發展出個案的過程。

爲了將這個構念操作化，范特西亞描述了他蒐集資料的方式。他利用資料來說明他的構念，並且也解釋了他對於資料的想法。他描述了他在蒐集資料時所採取的具體行動（例如：他在一間特定的工廠裡工作、參加了一場記者會、並且也訪談了一些人）。他也把資料鉅細靡遺地展現給我們看（例如，爲了佐證他的構念，他描述了許多具體事件，像是呈現數幀地圖以指明人們在一場與工頭的對抗行動中所站的位置、重述了在一間工廠裡發生的事件次序、詳細說明管理階層所採取的行動、以及覆述個別勞工所說過的話等等）。當他進行反思並且試著去瞭解他的經驗、以及利用現存的想法來發展出新的想法時，他也讓我們得以一窺他的思考過程。

個案化　在質化研究中，想法以及證據是相互依賴的，此種情況尤其適用於個案研究分析。個案並不是某種特定的、事先建立好的實徵單位，或者是與資料無關的理論類別。事實上，個案是由資料以及理論來加以定義的。藉由分析某種情境的方式，研究者就能夠把資料組織起來，並且同時把想法應用於創造或者是說明一個個案。製作或者是創造一個個案的過程被稱爲**個案化**（casing），此種過程可以把資料和理論結合在一起。決定把什麼事物當作是個案，可以解決在研究者所觀察到的事物、以及研究者對於這些事物的想法之間的緊張或是衝突的情況。「個案化被視爲是一種方法論的步驟，可能會發生在研究過程中的任何階段，但是特別容易出現在研究計畫的開端以及結束的時候」（Ragin, 1992b: 218）。

信度與效度

我們所有身爲研究者的人，都想要獲得信度以及效度，而這兩者正是所有測量中最重要的中心議題。信度與效度兩者都能夠把測量方式和構念連結在一起。雖然我們不可能達到完美的信度與效度，但這仍然是我們應該努力去追求的理想。信度與效度之所以如此重要而突出，正是因爲我們的構念經常是模稜兩可、模糊籠統、而且無法觀察的。信度與效度這兩種重要的概念，都能夠幫助我們建立研究發現的真實性、可靠性、或者是可信度。除此之外，這兩個詞彙都具有多重的意義，然而，當本書使用這兩個詞彙時，所指的是與測量有關的、值得追求的各種層面。

*信度*的意思是指可靠性（dependability）或者是一致性（consistency），換句話說，信度是指在完全相同或者是非常相似的條件之下，同樣的事情會重覆或者再次發生。與信度相反的狀況，就是由於測量本身的問題而造成不規律的、不穩定的、或者是不一致的結果。至於*效度*所指的就是真實性（truthfulness），也就是某個觀念

與實際狀況「契合」（"fit"）的程度有多高。若是研究缺乏效度的話，就意味著我們所用來分析社會世界的觀念，無法契合於在真正的社會世界裡所實際發生的事物。簡單地說，效度所處理的問題，就是我們關於社會實在的構念以及測量社會實在的方式之間的契合程度有多高。

所有的研究者都想要進行可信的以及有效的測量，但是除了在一般層次的基本想法上有一些共識之外，質化以及量化研究者們看待信度以及效度的方式其實是有所差異的。

量化研究中的信度與效度

***信度*　測量信度**（measurement reliability）所指的是，一個指標所產生的數字結果不會因爲測量結果或者是測量工具本身的特性而有所變化。舉例而言，我站上我家裡的體重計，並且得知了我的體重數字。我離開體重計，然後一再地站上去測量體重。如果我每次站上去量體重時，所量到的體重都是一樣的數字，那麼這個體重計就是可信的，不過，當然前提是在我每次量體重時沒有在吃東西、喝東西、或者是改變衣著等等。至於一個不可信的體重計，則是每次都顯示不同的體重數字，即使我「真正」的體重並未改變。另外一個例子則是我車上的速度表。如果我在一條平坦的道路上以緩慢而穩定的速度開著車，但是速度表上的指針卻在兩端之間來回地跳動，那麼這個速度表就不是一個可信的指標，也無法測量出我開車的速度有多快。事實上，信度一共有三種不同的類型。[6]

信度的三種類型

1. **穩定信度**（stability reliability）是指長期的信度。穩定信度所處理的是下列的這個問題：在不同的時段裡，某種測量方法是否能夠得到同樣的結果？剛剛提到的體重計的例子，就是屬於穩定信度。我們可以利用測試與再測試（test-retest）的方法來驗證一個指標在穩定信度方面的程度有多高。在驗證的過程裡，我們必須以同一個指標針對同一群人來進行重覆測試。如果我們所測量的事物是穩定的，而且該項指標也具有穩定信度，那麼我們每次都會得到相同的結果。測試與再測試方法的變化形式之一，是採用另一種替代的測試形式，但是此種替代形式必須與原本的測試形式非常相似。例如，我提出了一個假設，內容是關於性別以及在大學自助餐廳裡坐位子的模式。我觀察依變項（坐位子的模式）的方法，是在三個小時的時間內去觀察並且記錄坐在桌旁的男學生以及女學生的人數，並且標註誰坐在第一個位置、誰坐在第二個位置、誰坐在第三個位置等等，以此類推。如果在我進行觀察的期間，我覺得疲倦、分心、或者是忘記做記錄，因而在三個小時的觀察時間快要結束時錯失了許多人的話，那麼我的指標的穩定信度就不算高。

測量信度　某個變項的測量方式所具有的可靠性或一致性。

穩定信度　長期的測量信度，換句話說，假定在被測量的事物本身並未改變的情況下，某個測量方法在不同的時段，都能得到一致的結果。

2. **代表性信度**（representative reliability）指的是橫跨各個次母群或者是不同類型個案的信度。代表性信度所處理的是下列這個問題：當某個指標被應用到不同的群體時，是否能夠產生相同的結果？如果一個指標在被應用到不同的次母群時（例如，不同的階級、種族、性別、年齡群體等等），能夠對同一個構念產生相同的結果，那麼這個指標的代表性信度就很高。例如，我向某人詢問了他或是她的年齡。如果二十幾歲的人在回答我的問題時誇大了他們的實際年齡，而五十幾歲的人則是少報了他們的實際年齡，那麼這項指標的代表性效度就會很低。若想要達到代表性效度，我所用的測量方法就必須能夠獲得每個年齡群體的正確資訊才行。

次母群分析（*subpopulation analysis*）可以確認某項指標是否具有代表性信度。此種分析法會橫跨不同的次母群或者是次群體來比較該項指標，並且使用關於這些次母群或者是次群體的獨立知識。例如，我想要測試某份問卷裡的一個題目的代表性信度，而這個問題所詢問的是某個人的教育程度。我進行了次母群分析來看看這個問題對於男性和女性是否都同樣適用。我向男性和女性都詢問了這個問題，並且得到了獨立的資訊（例如，檢查學校的記錄），並且也檢查了他們的答案，看看男性和女性在回答這個問題時是否有同等的錯誤率。如果男性和女性的錯誤率是完全相同的話，那麼這個問題就具有高度的代表性信度。

3. **等值信度**（equivalence reliability）的應用時機，是當研究者使用了**多重指標**（multiple indicators）的時候，也就是說，研究者使用了多種特定的測量方式（例如，在一份問卷中的數道題目都是在測量同一個構念）來測量一個構念。等值信度所處理的是下列的這個問題：某種測量方法是否能夠從不同的指標中產生一致的結果？如果有數個不同的指標都是在測量同一個構念，那麼一個可信的測量方法，將可以在所有的指標中都獲得相同的結果。

我們可以利用*折半法*（*split-half method*）來驗證等值信度。在使用此種方法時，我們會把同一個構念的許多不同指標分成兩組，通常是利用隨機的過程來進行分配，然後確認這兩組指標是否都得到了相同的結果。例如，我在一份問卷中設計了十四道題目，而且所有的題目都是在測量大學生的政治保守主義程度。如果我的指標（例如：問卷裡的題目）具有等值信度的話，那麼我就可以利用隨機的方式把這些題目分成兩組，每組各有七個問題、而且也都能得到相同的結果。例如，我使用了第一組的七個問題，並且發現某一班裡的五十名主修商業的學生的保守程度，是另外一班五十名主修教育的學生的兩倍。而我利用第二組的七個問題來進行測量時，也得到了相同的結果。除此之外，特殊的統計量數（例如：克朗巴哈的 α 係數）（Cronbach's alpha）也能用來判定等值信度。當有數個觀察者、評判者、或者是資訊編碼者（在第十一章裡會進一步解釋）時，我們

代表性信度　跨群體的測量信度，也就是指某個測量方法能夠針對不同的社會群體產生一致的結果。

等值信度　橫跨不同指標的測量信度。意思是指當某一種測量方法使用了不同的指標、而且假定所有的指標所測量的都是相同的構念時，所得到的結果都是一致的。

多重指標　使用多重程序或是數種特別的測量方法，來為某個變項的水準提供實徵證據。

也可以使用一種特殊類型的等值信度，也就是編碼者間信度（intercoder reliability）。在某種意義上，每一位觀察者都是一個指標。因此，如果所有的觀察者、評判者、或者是編碼者彼此之間意見一致的話，某項測量方法就是可信的。此種測量方法在內容分析的研究中，是一種常見的信度類型。例如，我請六名學生去觀察在自助餐廳裡的坐位子的模式。如果這六名學生都同樣地擅長於進行觀察以及記錄，我就可以把這六名學生所得到的資訊結合在一起，以成爲一個單一的可靠量數。但是如果其中有一或兩名學生較爲懶散、粗心大意、或者是輕浮草率的話，我所得到的量數的信度就會變得較低。測量編碼者間信度的方式，是讓數位不同的編碼者去觀察完全相同的事物，並且比較其量數。例如，我請三位編碼者在三天中的相同時間裡，各自獨立地去對坐位子的模式進行編碼。接下來，我把他們所記錄下來的觀察結果進行比較。如果他們之間彼此意見一致的話，我就可以有信心地認爲我的測量方法具有編碼者間信度。除此之外，也有一些特別的統計技術可以測量編碼者間信度的程度。

如何提升信度　達到完美的信度是非常罕見的情況，但是我們卻可以用以下的四種方法來提升信度：（1）清楚地將構念加以概念化；（2）使用精確的測量尺度；（3）使用多重的指標；以及（4）使用前導測試（pilot tests）。

1. *清楚地將所有的構念加以概念化*　當每個測量方法所要測量的對象只有唯一一個概念時，信度就會增加，而這也意味著我們必須發展出明確的、清楚的理論性定義。我們必須詳細說明每一個構念，以排除來自於其他構念的「雜音」（"noise"）（也就是一些會使人分心的、具有干擾性質的資料）。舉例而言，一個純粹的單一化合物的指標，就會比一個摻雜了其他物質或是灰塵的化合物的指標還要更爲可靠。

讓我們回頭看看之前提過的教師士氣的例子。我應該要把士氣與相關的觀念（例如：心情、個性、精神、工作態度等等）區隔開來。如果我不這麼做的話，我可能就無法確定自己真正在測量的到底是什麼了。例如，我或許會發展出一個關於士氣的指標，但是這個指標同時也可能與個性有關。換句話說，關於個性的構念會污染了關於士氣的構念，並且因而產生了一個不太可信的指標。當研究者使用一個指標來把不同的構念操作化時（例如，以相同的問卷題目來指明士氣以及個性時），就是品質低劣的測量。

2. *提升測量的尺度*　關於測量尺度的問題，將在本章稍後部分進行討論。具有較高、或者是較爲精確測量尺度的指標，會比那些較不精確的測量方法具有較高的信度，這是因爲後者所能獲得的詳細資訊比較少。如果測量了比較具體精確的資訊，那麼除了構念之外，我們不太可能去捕捉到其他的事物。一般的原則是：盡量以可能範圍內最精確的尺度來進行測量。不過，在較高的測量尺度上進行量化的工作會比較困難。舉例而言，在測量士氣時，如果我可以只用低或高兩種選項、或者也可以把士氣從極度低落到極度高昂區分爲十個等級的話，那麼比較好的作法，是以十個等級的方式來測量士氣。

3. *對同一個變項使用多重指標*　提升信度的第三種方法，就是使用多重指標，因爲對相同的構念使用兩個（或者是更多個）指標，會比只使用一個指標還要更好。

[7]圖 7.4 說明了在假設考驗的過程中使用多重指標的情形。一個自變項構念的三個指標被結合成爲一個整體性的量數 A，而依變項的兩個指標則被結合成爲一個單一的量數 B。例如，假定 A 是教師士氣，而我對於教師士氣有三個具體的量數：（a1）教師們對於在問卷中有一道和對學校態度有關的問題的回答；（a2）除了病假以外的缺席次數；以及（a3）其他人聽到教師們抱怨的次數。而對於依變項 B，也就是對學生所付出的額外關注，我也有兩個量數：（b1）教師在放學後與學生進行個別談話的小時數，以及（b2）教師是否經常詢問學生在其他課程方面的進步情況。

有了多重指標，我們就可以進行三角檢證，並且從一個概念性定義內容（也就是概念領域內的樣本）的較爲廣泛的範圍來進行測量。我們可以用一個構念本身的指標來測量該構念的不同層面。除此之外，一個指標或許會有一些缺陷，但是數個不同的測量方法則比較不可能出現相同的錯誤。詹姆斯（James, 1991）曾經把這個原則應用於計算街友的數目上，而他的研究則爲我們提供了一個很好的範例。如果我們只考量人們睡覺休息的場所（例如，計算在街頭或是在公園裡睡覺的人，並且計算在公共收容所裡的街友人數等等），我們就會漏算一些街友，因爲有許多街友可能會有暫時的棲身之所（例如，在某位朋友或者是家人所住的房子裡打地舖）。如果我們使用了政府服務機構的記錄，也會漏算一些人，因爲許多無家可歸的街友會避免和政府以及其他官方機構有所接觸。然而，如果我們結合了官方的記錄以及在各種地方睡覺休息的街友人數，並且也對接受各種政府服務（例如：街頭免費看診服務、免費發放食物服務、臨時收容所等等）的人們進行問卷調查，我們就可以獲得關於街友人數更爲正確的數目。除了能夠捕捉整體的圖像之外，多重指標通常也比單一項目的量數還要更加穩定。

4. *使用前導研究以及複製*　你可以先使用某種測量方法前導版本，以提升信度。你可以發展出一個或者是多個草案、或者是某種測量方法的初步版本，並且先去嘗試各種不同的版本，然後才把最後確定的版本應用於假設考驗的情境中。這個過程需要花費不少時間以及努力。回到先前討論過的例子，在我對於教師士氣所進行的調查中，在確定了最後的問卷版本之前，我已經擬過好幾份不同版本的問卷，

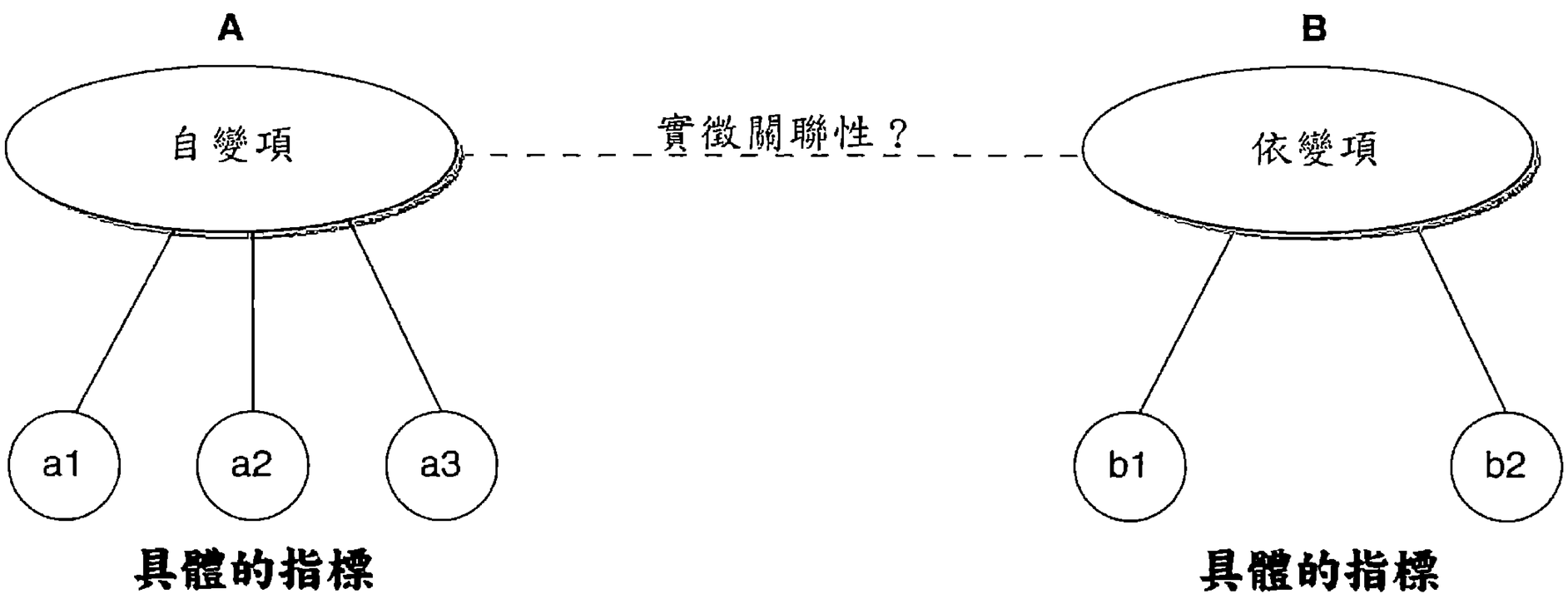


圖 7.4　使用多重指標的測量

測量效度　某個實徵指標與該指標所要測量的構念的概念性定義之間的「契合」程度。

而我測試這些問卷版本的方法，是去向人們詢問在問卷上的問題，並且檢查看看這些題目是否夠清楚。

使用前導測試的原則也可以延伸到複製其他研究者所使用過的測量方法。例如，我搜尋了文獻資料，並且從過去的研究中發現了一些測量士氣的方法。因此，如果其中某一種測量方法是相當良好的，我或許可以利這個測量方法爲基礎來進行研究。當然，我也必須在研究中註明出處。除此之外，我可能也想要增加一些新的指標，並且把這些新指標和之前的測量方法進行比較（請參閱範例 7.1 改善美國宗教歸屬的測量方法）。如此一來，只要使用相同的定義，測量方法的品質就能不斷地增進（關於信度與效度種類的摘要，請參閱表 7.1）。

效度　效度是一個被過度使用的詞彙。有時候，效度被用來指「真實的」或者是「正確的」。事實上，效度有數種一般的類型，而我們在這兒所關心的類型是**測量效度**（measurement validity），而測量效度本身也有好幾種類型。至於非測量型的效度，則會在稍後的部分進行討論。

當我們說某個指標是有效的，是指該指標對於某個特殊的目的以及定義是有效的，如果是針對其他的目的的話，相同的指標可能就不是那麼有效、或者甚至是完全無效的。例如，在前面所討論過的測量士氣的方法（例如，在問卷中詢問教師們對於學校的感覺的問題），在測量教師們的士氣時可能是有效的，但是若用來測量警察的士氣時，就會變成是無效的。[8]

測量效度的核心就是在告訴我們概念性定義與操作性定義彼此之間契合的程度有多高，換句話說，這兩者之間的契合程度越高，測量效度也就越高。要達到效度

範例　7.1

改善美國宗教歸屬的測量方法

量化研究者會去測量個人的宗教信仰（例如，你相信有上帝嗎？你相信有惡魔嗎？你相信有來生嗎？你覺得上帝的模樣像什麼？）、宗教活動（例如，你祈禱的頻率有多高？你參加宗教儀式的頻率有多高？）、以及宗教歸屬（例如，你是否屬於某一個教會或者是宗教團體？如果是的話，是哪一個？）。量化研究者們在把數百種美國的宗教教派進行分類時，若不是分成三個大型教派（新教、天主教、猶太教），就是依照在 1990 時引進美國的分類法來分成三種派別：基本教義派、溫和派、或者是自由派。

史汀斯蘭德以及同事（Steensland and colleagues, 2000）將「歸屬」（affiliation）重新概念化，而且，在檢視過宗教神學以及社會實踐的趨勢之後，他們主張把所有的美國宗教教派分成六個主要類別：主流新教（Mainline Protestant）、福音派新教（Evangelical Protestant）、黑人新教（Black Protestant）、羅馬天主教（Roman Catholic）、猶太教、以及其他教派（其中包括摩門教、耶和華見證人（Jehovah's Witnesses）、回教、印度教、以及一位論教派（Unitarian）等等）。作者們評估了此種把宗教教派分為六種類別的新分類方法，其評估方式是檢視人們的宗教看法與活動、以及他們對於當代社會議題的看法。在美國全國性的樣本當中，他們發現此種新的分類方法比先前的分類方法更能夠區別出不同教派之間的差異之處。

表 7.1　測量信度與效度類型的摘要

信度（可以相信的測量方法）	效度（真實的測量方法）
穩定信度—長期 （利用測試與再測試法來加以驗證）	表面效度—他人在判斷之後認為有意義
代表信度—橫跨各個次群體 （利用次母群分析來加以驗證）	內容效度—捕捉到完整的意義
等值信度—橫跨不同的指標 （利用折半法來加以驗證）	效標效度—與外部的來源意見一致 • 同時效度—同意已經存在的測量方法 • 預測效度—同意未來的行為
	構念效度—擁有一致的多重指標 • 輻合效度—相同構念的指標之間彼此相似 • 區別效度—不同構念的指標之間彼此會有差異

比達到信度還要更加困難。我們無法對於效度擁有絕對的信心，但是有些測量方法會比其他方法*更爲有效*。造成此種狀況的原因，是因爲構念是抽象的想法，而指標則是具體的觀察。存在這兩者之間的差距，就是存在於我們心裡對於世界所抱持的圖像、以及我們在特定的時間與地點所做的事情之間的差距。效度是一個動態過程的一部分，而且這個動態過程會隨著不斷累積證據而成長，而且若是沒有效度的話，所有的測量都會變得毫無意義可言。

有一些研究者會使用對應規則（之前已經討論過）來減少在抽象觀念以及特定指標之間的差距。例如，某一條對應規則如下：若是某位教師認同「過去五年來，狀況越來越糟」以及「沒有進步的希望」這兩種說法的話，這就表示他的士氣低落。有些研究者則會談到*知識的相關*（*epistemic correlation*），這是一種在某個指標以及該指標所要測量的構念之間的假設性相關。我們無法以實徵方式來測量此種相關，但是卻可以對此種相關進行估計。[9]

測量效度的四種類型

1. **表面效度**（face validity）是最爲基本、而且也是最爲容易達到的效度類型。此種效度是由科學社群來判斷某個指標是否真的能夠測量到某個構念。表面效度所處理的是下面這個問題：就表面上來看，人們是否相信定義以及測量方法之間是互相契合的？例如，如果在測量大學生的數學能力時，所使用的測量方式是去問大學生們「二加二等於多少」這類的問題，那麼幾乎沒有人能夠接受此種測量方式。因爲就表面上來看的話，這並不是一種測量大學程度數學能力的有效方法。你可以回想一下之前曾經提過，在科學社群裡有所謂的系統性的懷疑原則，而該原則的意思是指其他人會去仔細地審查研究的各個不同層面。[10]

表面效度　測量效度的一種類型。表面效度所指的是其他人，尤其是科學社群，在進行判斷之後，認為某個指標是測量某個構念的合理方式。

2. **內容效度**（content validity）所處理的問題是：某個測量方法是否能夠呈現某個定義的完整內容？一個概念性定義會包含許多觀念，也就是說，一個概念性定義就好像是一個包含了許多觀念以及概念的「空間」一樣。測量方法應該要能夠抽樣或者是代表在概念空間內的所有觀念或者是區域。內容效度牽涉了三個步驟。首先，詳細說明某個構念的定義裡的內容。接下來，從該定義的所有區域中進行抽樣。最後，發展出一個或者是多個指標，而且這些指標要能夠連結該定義的所有部分。

讓我們來看看一個內容效度的例子。我把*女性主義*（*feminism*）定義爲某人對於一組信念的支持與擁護，而這組信念的內容是要在藝術、追求知識、家庭、工作、政治、以及權力關係等各領域中創造出男女兩性之間完全平等的狀態。我也創造出測量女性主義的方法，在進行測量時，我會詢問兩個調查問題：（1）男性和女性是否應該要同工同酬？以及（2）男性和女性是否應該共同分擔家務？如此一來，我所得到的量數的內容效度就會較低，因爲這兩個問題只問到了薪資以及家務而已，卻忽略了其他的領域（追求知識、政治、權威關係、以及工作與家庭的其他層面）。若要創造一個具有內容效度的量數，我必須把測量的範圍加以擴充、或者是把定義的範圍縮小才行。[11]

3. **效標效度**（criterion validity）是利用某些標準或者是準則來正確地指明某個構念。驗證某個指標的效度的方法，是與相同構念的另外一個測量方法進行比較，而且其他的研究者對於這另外一個測量方法是具有信心的。效標效度的兩種次類型分別爲同時效度以及預測效度。[12]

要具有**同時效度**（concurrent validity），我們必須把一個指標連結到一個既存的、而且我們也已經判定爲有效（也就是具有表面效度）的指標。例如，我們創造出一種新的測驗來測量智力。若要使這項新的測驗具有同時效度，這項測驗就必須與現存的智力測驗具有高度的相關性（假定這兩種測驗所使用的智力定義是完全相同的）。這就意味著大多數在舊的智力測驗中得到高分的人，在新的測驗中應該也會得到高分，反之亦然。這兩種測量方法之間的關聯性或許無法達到完美的地步，但是如果這兩種測量方法所測量的是相同的或者是相似的構念，那麼這兩者應該會得到類似的結果才對。

內容效度　測量效度的一種類型，此種效度要求某種測量方法必須能夠代表某個構念的概念性定義裡的所有面向。

效標效度　一種測量效度，此種效度依賴某種獨立的、外界的驗證方式。

同時效度　一種測量效度，此種效度依賴的是一個既存的、而且已經被接受的測量方法，以驗證某個構念的指標。

預測效度　一種測量效度，此種效度是以發生於未來的、在邏輯上連貫的某個事件或者某種行為來驗證某個構念的指標。

當效標效度可以被某個指標用來預測未來的事件、而且這些事件在邏輯上與某個構念相關時，此種效標效度就被稱作是**預測效度**（predictive validity）。不過，預測效度無法被用在所有的測量方法上。測量方法以及被預測的行動必須與構念有所區別，但是卻要指向相同的構念。預測的測量效度不應該和假設考驗中的預測混爲一談，因爲在假設考驗中，所謂的預測是指某個變項預測了未來一個不同的變項。例如，許多美國高中生參加的學術性

向測驗（Scholastic Assessment Test, SAT）所測量的是學業方面的性向，也就是學生在大學課業方面的表現能力。如果 SAT 具有很高的預測效度，那麼在 SAT 中得到高分的學生之後在大學裡應該也會有很好的表現。然而，如果得到高分的學生在大學時的表現，和得到普通分數或是低分的學生一樣的話，這就代表 SAT 的預測效度很低。

另一個測試預測效度的方法，就是挑選一群具有某些特殊特質的人，並且預測這群人相對於某個構念能夠得到多少分數（非常高或是非常低）。例如，我創造出了一種測量政治保守主義的方法，而且我預測保守團體的成員（例如，約翰柏奇會（John Birch Society）、保守委員會（Conservative Caucus）、美國革命婦女會（Daughters of the American Revolution）、道德多數（Moral Majority）等等）將會在測量中得到高分，而自由主義團體的成員（例如，民主社會黨人（Democratic Socialists）、美國人民之路基金會（People for the American Way）、美國民主行動會（Americans for Democratic Action）等等）的得分則將會很低。我先針對這些不同團體的成員進行了前導測試，以「驗證」我的測量方法是否有效。之後，我所創造出來的這種測量方法，就可以被用來測量一般大眾的政治保守主義。

4. **構念效度**（construct validity）是用於具有多重指標的測量方法上，而構念效度所處理的是下面這個問題：如果該測量方法是有效的，那麼各種不同的指標的運作方式是否一致？構念效度要求一個定義必須要有明確界定的概念範圍。而構念效度的兩種類型分別爲輻合效度以及區別效度。

輻合效度（convergent validity）的應用時機，是當多重指標趨於同樣的結果、或者是彼此互相關聯的時候。換句話說，輻合效度所指的是同一個構念的多重測量方法彼此聯結在一起，或者是以相似的方法運作。例如，我想要測量「教育」這個構念，而我用的方法是去問人們在學校念書幾年、查看學校的記錄、以及要求人們完成一份關於學校知識的測驗。如果這些測量方式的結果不一致（例如，有些人宣稱具有大學的學位、但是卻沒有就讀於大學的記錄，或者是有些具有大學學位的人，在測驗中得到的分數卻沒有高於那些在高中就輟學的人），那麼我的測量方法的輻合效度就很弱，而且我就不應該把所有的三個指標結合成爲單一的量數。

區別效度（discriminant validity）是輻合效度的相反，而此種效度的意思是，雖然同一個構念的不同指標會彼此互相關聯、或者是趨向於同樣的結果，但是這些指標也會與對立的構念具有負面的相關。區別效度告訴我們，如果 A 與 B 這兩個構念彼此之間的差異性很大，那麼 A 和 B 各自的量數就應該不會有相關性。例如，我在一份問卷裡有十道題目是在測量政治保守主義，而且人們也以類似的方式來回答這十道題目。但是，我在同一份問卷裡也加入了五道用來測量政治自由主義的題目。如果這

構念效度　一種測量效度，此種效度使用多重指標，並且有兩種次類型：某個構念的各種指標彼此趨同的程度有多高、或是不同構念的指標間，其差異程度有多大。

輻合效度　一種多重指標的測量效度類型，其基礎來自於下列這個觀念：某一個構念的各種指標將會以類似的方式運作、或者是趨向於同樣的結果。

區別效度　一種多重指標的測量效度類型，其基礎來自於下列這個觀念：不同構念的指標之間會與彼此有所差異。

十道關於保守主義的題目彼此高度相關，並且與另外五道關於自由主義的題目具有負面相關的話，那麼我測量保守主義的方法就具有區別效度。（關於各種測量效度的複習，請參閱圖 7.5）

質化研究中的信度與效度

雖然質化研究接受信度與效度的核心原則，但是我們在質化取向中卻鮮少看見這些詞彙，這是因爲信度與效度與量化測量有著相當密切的關聯性。除此之外，在質化研究中，我們也是以不同的方式來應用信度與效度的原則。

信度 回想一下，前面我們曾經提過，*信度*所指的就是可靠性或者是一致性。在質化研究中，我們會運用各式各樣的技術（例如，訪談、參與、照片、文件研究等等），並且以前後一致的方式來記錄觀察結果。在我們進行觀察的時候，我們所使用的方法必須要有一致性（也就是說，不能搖擺不定或者是反覆無常），而這與穩定信度的觀念相當類似。在質化研究中要達到信度時所會遇到的困難之一，就是我們研究的事物通常是長期不穩定的過程。除此之外，我們也強調在身爲研究者的我們與我們研究的人們之間那種不斷變化以及發展的互動關係的價值。我們相信，研究主題以及我們與研究主題之間的關係是一種不斷演變的過程。我們可以把此種關係比喻成一種不斷演變的關係、或者是一種會隨著時間經過而自然成熟的有機生命體（像是一棵植物）。許多質化研究者認爲量化取向中的信度是冰冷的、僵化的機械式工具，而量化研究者則一再地把此種工具應用於靜態的、無生命的物質之上。

在質化研究中，我們會考量多種資料來源，並且也會採用多重的測量方法。我們不會死守著複製、等值、以及次母群信度等屬於量化－實證主義的觀念。我們認爲，不同的研究者、或者是使用替代測量方法的研究者可能會發現完全不同的結果。之所以會發生這種狀況，是因爲蒐集資料乃是一種互動的過程，在這個過程中，特定的研究者在一個不斷演變的環境中進行研究，此種環境的脈絡使得研究者必須使用一種獨一無二的方式來組合各種測量方法，而且此種組合的方式是無法被重覆的。多樣化的測量方法、以及與不同的研究者之間的互動是有益的，因爲這些測量方法以及互動可以啓發研究主題的不同面向。許多質化研究者質疑量化研究者追求標準的、固定的測量方法的作法，而且也擔心此種測量方法忽視了擁有不同取向的許多研究者所能帶來的益處，也可能忽略了存在於社會世界中的多樣性的關鍵面向。

效度 *效度*一詞所指的是真實性。在質化研究裡，我們比較感到興趣的是達到一種本真性，而不是實現某種單一版本的「真理」（“Truth”）。本真性（authenticity）的意思，是指從每天生活在社會世界中的人們的觀點，以公平的、誠實的、以及平衡的方式去描述社會生活。質化研究者不太關心把一個抽象構念與實徵資料連結在一起，而比較在乎以一種坦率的方式去描繪出社會生活，而且此種描繪的內容必須忠於被研究者的真實生活經驗。在大多數的質化研究裡，我們強調的是去捕捉一種局內人的觀點，並且詳盡地去描述被研究者理解事件的方式（請參閱資訊箱 7.2 效度

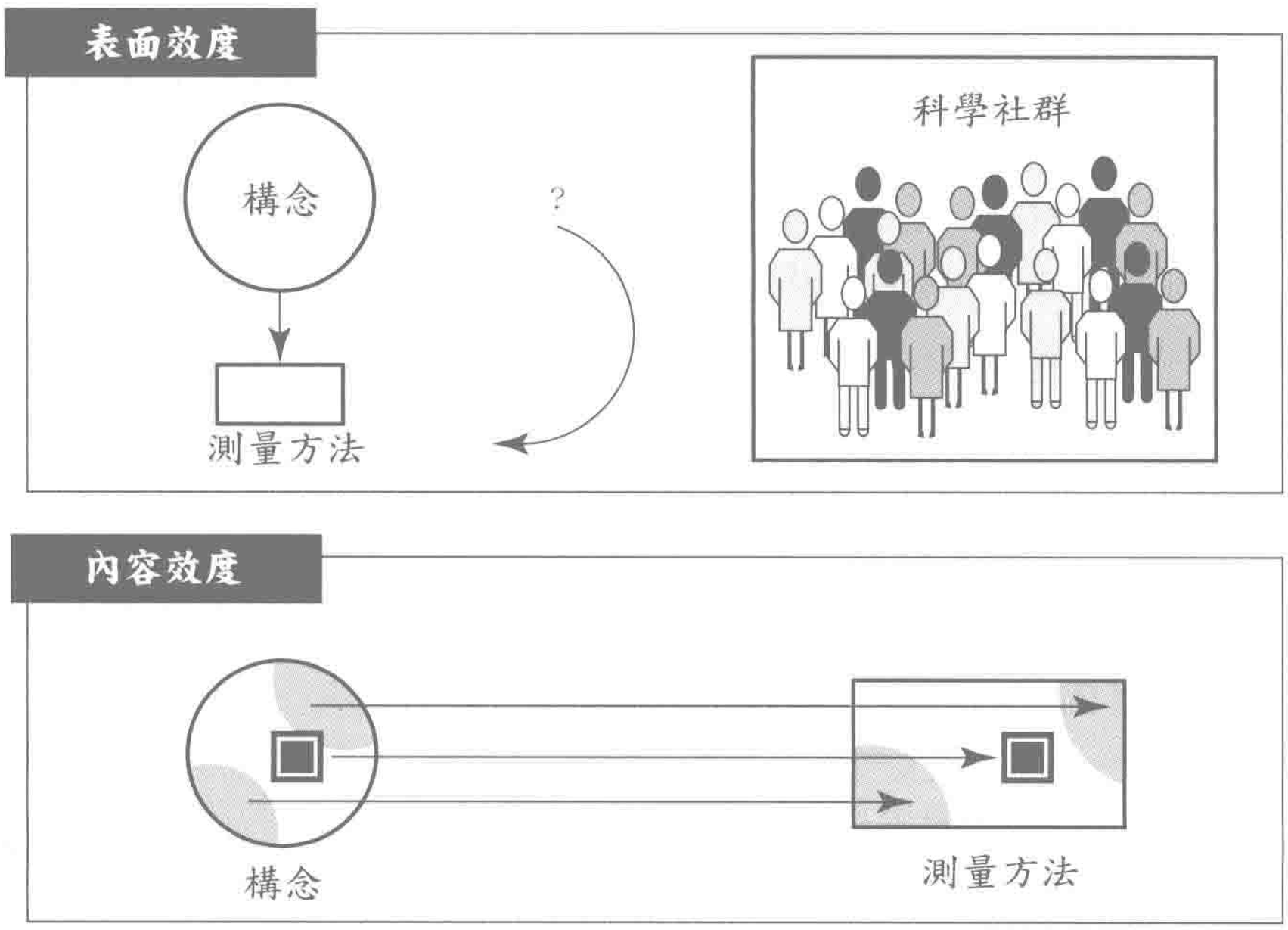

表面效度
構念
?
測量方法
科學社群
內容效度
構念
測量方法
效標效度
預測效度
構念
?
?
?
現存的
測量方法
未來的行為
測量方法
相似的結果
概念效度
輻合效度
區別效度
構念
相同的
構念
多重
測量方法
測量方法
?
相似的
結果
=
=
=
=
=
≠
?
不同的
結果

圖 7.5　效度的類別

在質化研究中的意義）。

在質化研究中已經發展出一些方法來取代量化取向中的效度，像是生態效度（ecological validity）或者是自然歷史方法（natural history methods）等等（請參閱第十三章）。上述這兩者都強調要把局內人的觀點傳達給他人。歷史研究者使用內部以及外部評論（請參閱第十四章）來判斷證據是否爲真。質化研究者堅守效度的核心原則，盡可能維持真誠性（也就是避免錯誤的或者是扭曲的敘述），並且盡力在他們對於社會世界的理解、觀念、與陳述，以及實際發生於社會世界中的事物之間創造出緊密的結合。

信度與效度之間的關係

信度是達到效度的必要條件，而且也比效度更容易達到。雖然在發展出一個概念的有效測量方法時，信度是必要的條件，但是擁有信度並不保證某種測量方法就是有效的。換句話說，信度並不是效度的充分條件。某個測量方法可能會一再地產生同樣的結果（也就是擁有信度），但是該方法所真正測量的事物並不一定會符合某個構念的定義（也就是達到效度）。

例如，我站上一個體重計來測量我的體重。在兩個小時的時間內，當我站上去量體重時，體重計每次都顯示出相同的體重數字。接著，我站到另外一個體重計上面去量體重——這次的體重計是在診所裡的「正式」體重計——而這個體重計上面顯示的體重數字是之前的兩倍。第一個體重計所產生的是可信的（也就是可靠的而且一致的）結果，但是卻不是測量我的體重的有效方法。用示意圖來說明，或許可以幫助你瞭解信度與效度之間的關係。圖 7.6 利用標靶的類比來說明信度與效度這兩個概念之間的關係，而靶心則代表了測量方法與構念的定義之間的契合狀態。

效度與信度通常是互補的概念，但是在某些情況中卻會彼此產生衝突。有時候，隨著效度增加，要達到信度就變得越來越困難，而且反之亦然。當構念極爲抽象而且無法輕易進行觀察、但是又捕捉到某個觀念的「真正本質」（“true essence”）時，就會發生此種狀況。當一個量數是精確的、具體的、而且可以觀察的時候，是最容易達到信度的。例如，*疏離*是一個非常抽象而且主觀的構念。我們也許可以把疏離定義爲一種在內心深處失去個人核心人性的感覺，也就是一種冷漠的、失去目標的感覺，而且會擴散到生活的所有層面（例如，自我意識，以及與其他人、工作、社會、甚至是自然之間的關係）。雖然不太容易，但是我們大多數人都能夠理解疏離的觀念，也就是滲透於個人生存狀態之中的一種沒有方向的失落感。隨著我們越來越深入去探究這個概念的真正意義，要精確地測量這個概念就會變得越來越困難。在問卷上的特定題目也許可以產生比其他方法更爲可信的結果，但是這些問題卻無法捕捉到疏離這個觀念的本質。

資訊補給站 7.2　效度在質化研究中的意義

質化研究中的測量效度，不會要求研究者必須在一個仔細定義的抽象構念、以及一個針對該構念的精確標準化實徵測量方式之間展現出固定的對應關係。在建立質化研究中的效度時，研究測量過程中的其他特徵也都是非常重要的。

首先，若要被認為是有效的，一個研究者的真理宣稱（truth claims）就必須是似乎可靠的，而且，如同范恩（Fine, 1999）所說的，也必須在互為主體的層次上「夠好」（"good enough"）（也就是研究者的真理宣稱能夠讓許多其他的人都能理解）。所謂「似乎可靠的」（plausible），意思是指資料以及關於資料的陳述都不是唯一的，也就是它們並非唯一可能的宣稱，也不是對於世界上唯一真理的精確描述。然而，此種情況不會使得這些資料以及陳述變成僅僅是研究者憑空創造、或者是任意指定的產物。相反地，這些資料以及陳述是強而有力的、具有說服力的描述，能夠揭露出研究者對於實徵資料的真誠經驗。

其次，當一個研究者的實徵宣稱獲得許多各式各樣的實徵資料的支持時，這些實徵宣稱就擁有效度。僅僅只有任何一個特定的實徵細節的話，這個細節可能是平凡的、普通的、或者是「瑣碎的」（"trivial"）。但是把數百個小的、多樣的細節累積起來，這些細節就能一起形成具有重大影響力的證據，而效度就是來自於這些累積的證據。

第三，當研究者持續地在各種不同的資料中搜尋、並且思考這些資料之間的連結時，效度就會因而提高。在自然社會世界中的原始資料並不是整齊地預先包裝好的系統性科學概念，更確切地說，這些原始資料是數量龐大的互異元素，而且這些元素「形成了一個動態的、一致的整體」（Molotch et al., 2000: 816）。當研究者在互異的細節中辨識出緊密的連結性時，效度也會隨之提高。效度的提升不僅僅是來自於互相連結的各種細節的數目，而且也是來自於橫跨多樣不同的領域而創造出的動態連結的網絡。

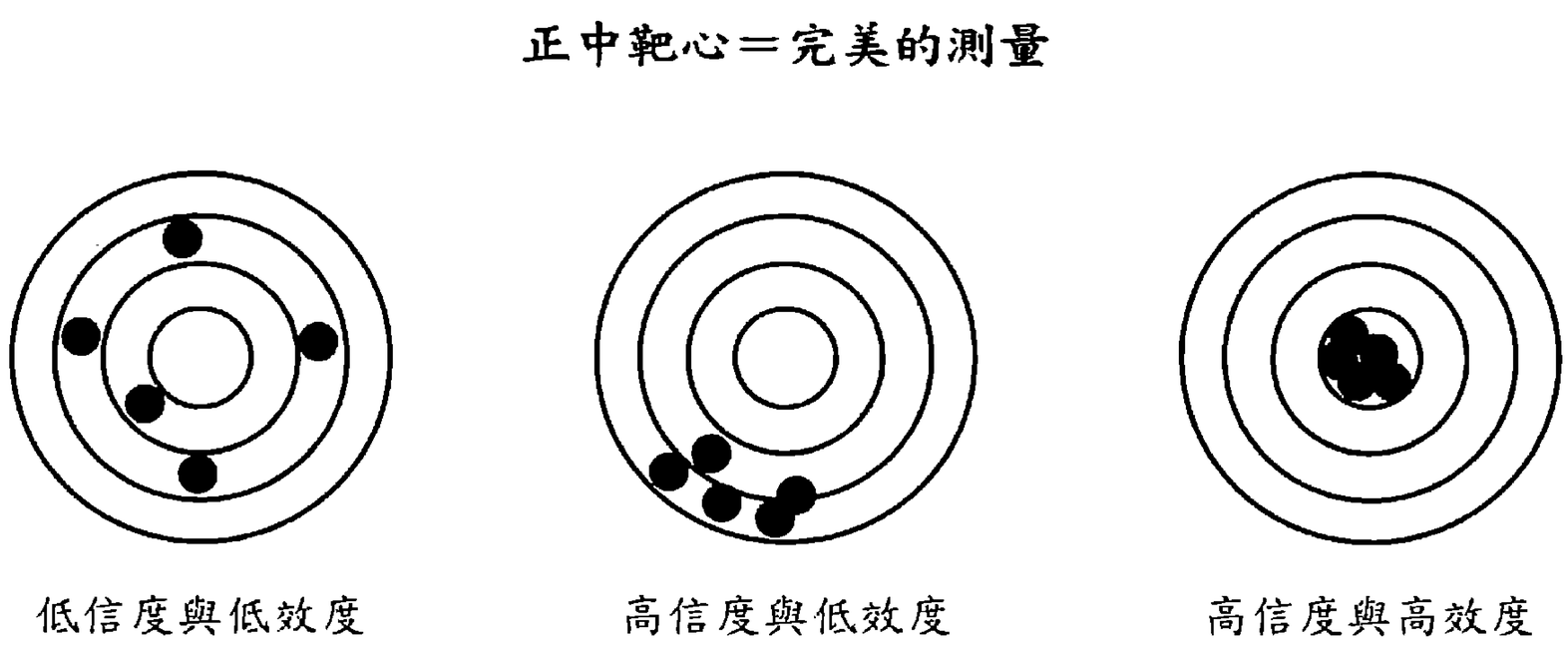


圖 7.6　信度與效度之間關係的說明圖

資料來源：Adapted version of Figure 5-2 An Analogy to Validity and Reliability, page 155 from Babbie, E. R. 1986. *The Practice of Social Research*, Fourth Edition. Belmont, CA: Wadsworth Publishing Company.

「可信」與「有效」兩個詞彙的其他用法

許多字詞都會有多重的定義，因此，在同一個字的各種不同用法之間也會因而造成許多混淆之處。此種情況也同樣會發生在信度與效度這兩個詞彙上。我們會在日常語言當中使用*信度*這個詞。一個*可信*的人就是一個可靠的、穩定的、而且負責任的人，在不同的時間以及狀況裡，這樣的人所作出的反應都是類似的、而且是可以預測的。一輛*可信*的車子是可靠的而且值得信任的，也就是這輛車子在啓動以及行車表現方面都是可以預測的。有時候我們會說一項研究或者是研究結果是可信的，這就意味著其他研究者可以複製這項研究，而且也會得到相似的結果。

內在效度（*internal validity*）是指在一個研究計畫的設計內部沒有一些可能會造成錯誤結論的失誤。[13] 在實驗研究中，我們主要談論的是那些雖然我們試著去進行控制，但是仍然有可能發生的研究結果的替代解釋（請參閱第九章裡的相關討論）。

外在效度（*external validity*）主要也是用於實驗研究裡，而外在效度所指的是，我們是否可以把在一個特定背景中的一小群特定人們身上所發現的結果，向外推論到範圍更廣的其他背景以及許多不同的人們身上。外在效度所處理的是下面這個問題：如果某件事物發生於實驗室裡或者是一群特定的研究參與者（例如，大學生）身上的話，同樣的事物是否也會發生於「真實的」世界（也就是非實驗室的世界）、或者是一般大眾（非大學生）身上（這些問題將於第九章中進行討論）？在評估理論時，外在效度具有相當重要的涵意。如果某個一般性的理論爲真的話，這就暗示著我們可以把該理論的一個測試中所得到的發現，推論到許多其他的情境以及群體上面（請參閱 Lucas, 2003）。

統計效度（statistical validity）所指的是我們針對一個特定的目的使用了適當的統計程序，並且也達到了該程序的數學條件。之所以會有統計效度，是因爲正如同在統計程序的教科書上所討論的一樣，不同的統計考驗或是程序都有各自適合的不同情境。所有的統計程序的基礎，都是來自於所使用數字的數學特性的假定。如果我們把某個統計程序用在不適合的情境中、或者嚴重地違反了該程序的假定的話，那麼，就算在數字的運算過程方面是正確無誤的，最後所產生的也只是無意義的結果而已。這就是爲什麼我們在使用某個統計程序時，必須要先知道該統計程序的設計目的以及假定才行。而這也是爲什麼電腦雖然可以進行正確的運算、但是產生的結果卻可能毫無意義的原因。

量化測量指引

到目前爲止，我們已經討論過測量的原則了。量化研究者已經發展出專門的測量方法，以協助他們爲可信的以及有效的量數創造出操作性定義。本章的這個部分將會簡短地介紹這些觀念以及一些特定的測量方法。

測量尺度

我們可以把可能的量數在一條連續線上依次排列出來。在其中一端是「較高」的量數，而這些量數會包含大量高度具體的資訊，其中有許多精確的、詳細的特性。而在相對的一端則是「較低」的量數，這些量數是粗糙而且比較不精確的，其中只包含了少量的資訊以及一些基本的的特性而已。當我們在測量社會世界的特性時，測量的尺度會影響到我們所能夠得到的資訊數量，而且當我們試著去捕捉一個構念的實徵細節時，測量尺度也會限制了我們能夠使用的指標種類。

測量尺度（level of measurement）的決定方式，是依據在我們對於某個構念的假定中，該構念的精鍊、正確、以及精確的程度有多高。而這也意味著，我們對一個構念加以概念化的方式，會對測量尺度造成很大的影響，因爲這會影響到我們測量該構念的方式，而且在我們蒐集完資料之後，也會限制了我們可以使用的統計程序的範圍。一般而言，我們在測量尺度以及測量的容易度之間必須作出一些權衡取捨。在較低的尺度上進行測量會比在較高的尺度上進行測量更爲簡單也更爲容易。然而，較低的測量尺度只能產生最精糙的資訊，而且在資料分析的過程中也只能使用最少的統計程序。關於這個議題，我們可以從下列的兩個方面來進行討論：（1）連續變項以及間斷變項，以及（2）四種測量尺度。

連續以及間斷變項　變項可以是連續的或者是間斷的。**連續變項**（continuous variables）包含了在一條連續線上移動的大量數值或者是屬性。我們可以把一個連續變項區分成許多較小的部分，而且在數學理論中，這些較小部分的數量是無窮的。連續變項的例子包括溫度、年齡、收入、犯罪率、以及接受教育的程度等等。例如，在測量你所接受的教育程度時，我們可以去測量你完成的學校教育的年數。我們可以把這項資訊再細分爲你花在教室裡聽課、以及在課後的作業或是準備工作上的總時數。我們還可以再更進一步地把這項資訊細分爲你投入於獲得以及處理學校資訊以及知識、或者是學校作業的分鐘數。我們甚至可以再更進一步地把這項資訊分割爲當你在吸收以及處理資訊時、你的大腦從事於特定的認知活動的總秒數。

間斷變項（discrete variables）則是具有一組相對固定的個別數值或者是變項屬性。間斷變項所包含的不是一個由許多數值所組成的平順連續體，而是包含了數量有限的明確類別。間斷變項的例子包括了性別（男性或是女性）、宗教（新教、天主教、猶太教、回教、無神論者）、婚姻狀態（未婚、已婚、離婚或分居、喪偶）或者是學位（高中畢業、社區大學畢業、四年制大學畢業、碩士或是博士學位）。一個變項是連續或者是間斷的，會影響到該變項的測量尺度。

測量尺度　一種把測量變項的資訊組織成四種尺度的系統，其範圍是從名目尺度到比率尺度。

連續變項　在一條連續線上被測量的變項，而且在變項的屬性之間可能會有無限個更小的數值。

間斷變項　在此種變項裡，變項的屬性只能以有限數目的不同的、個別的類別來加以測量。

四種測量尺度 測量尺度是建立在連續變項以及間斷變項之間的差異之上的。較高的測量尺度是連續的，而較低的測量尺度則是間斷的。而這四種不同的測量尺度則會把測量的精確度加以區別成各種不同的類別。[14]

要去決定某個構念的適當測量尺度，並非總是一件容易的事。適當的測量尺度取決於兩件事：我們理解一個構念的方式（該構定的定義以及假定）；以及指標或是測量程序的類型。

我們將一個構念加以概念化的方式，可能會限制了我們在測量該構念時的精確程度。例如，我們或許可以把一些先前列爲連續的變項重新概念化，使其成爲間斷變項。我們可以把溫度當成是一個有數千個較小單位（例如，溫度的度數或者是度數的分數）的連續變項。或者，我們也可以把溫度粗略地分成五個間斷的類別（例如：非常炎熱、炎熱、涼爽、寒冷、非常寒冷）。我們可以把年齡當作是連續的（以年、月、日、時、分、秒來計算），或者是間斷的（嬰兒、兒童、青少年、青年、中年、老年）。

雖然我們可以把連續變項轉變成爲間斷變項，但是我們卻不能反過來把間斷變項轉變成爲連續變項，舉例而言，我們無法把性別、宗教、以及婚姻狀態等等變項轉變爲連續變項。不過，我們卻可以用稍微不同的定義以及假定來處理相關的構念，並使其成爲連續變項（例如：具備男性特質或是女性特質的程度、對宗教虔誠的程度、對於婚姻關係的承諾程度等等）。在較高的測量尺度上進行概念化以及進行測量，有其實際上的理由：我們可以把較高的測量尺度轉化成較低的測量尺度，但是我們卻無法把較低的測量尺度轉化成爲較高的測量尺度。

名目尺度測量 最低的、最不精確的測量尺度，對於此種測量尺度而言，在變項的不同類別之間只有類型方面的差異而已。

順序尺度測量 此種測量尺度可以指明在變項的類別之間的差異，並且允許將不同的類別排出等級或順序。

區間尺度測量 此種測量尺度可以指明在變項屬性之間的差異、將類別排出等級或順序、而且可以測量類別之間的距離，但是沒有真實的零。

比率尺度測量 最高的、最精確的測量尺度，變項的屬性可以依照等級來排序、屬性之間的距離可以被精確地測量出來，而且也有絕對的零。

區分四種尺度 這四種測量尺度的精確程度由最低到最高分別是名目的、順序的、區間的、以及比率的。每一種尺度所提供的是不同類型的資訊（請參閱表 7.2）。**名目尺度測量**（nominal-level measurement）所指明的，是在不同類別之間存在著差異之處（例如，宗教：新教、天主教、猶太教、回教；種族遺傳：非裔、亞裔、白種人、西班牙裔、其他）。**順序尺度測量**（ordinal-level measurement）不但指明了差異，*而且*允許我們把這些類別加以排序或者是排出等級（例如，以字母來代表分數：A、B、C、D、E；測量意見的方法：非常同意、同意、不同意、非常不同意）。**區間尺度測量**（interval-level measurement）不但可以提供上述兩種尺度所能提供的資訊，*而且允許*我們詳細說明在類別之間距離的量（例如，華氏或是攝氏的溫度：5 、45 、90 ；智商分數：95、110、125）。**比率尺度測量**（ratio-level measurement）則是能夠做到其他所有測量尺度能做到的事，*而且*還具有一個真實的零，這個特質使我

表 7.2　四種測量尺度的特質

尺度	不同的類別	等級	測量類別之間的距離	真實的零
名目的	是			
順序的	是	是		
區間的	是	是	是	
比率的	是	是	是	是

們能夠以比例或是比率的方式來說明關係（例如，金錢收入：10 美金、100 美金、500 美金；接受正式學校教育的年數：1、10、13）。在大多數實際的情境中，區間尺度測量以及比率尺度測量之間幾乎沒有什麼太大的差異。

造成混淆的來源之一，是我們有時候會在區間測量尺度中使用任意指定的零，但是這些任意指定的零只是爲了要幫助我們記分數而已，並不是真實的零。例如，當溫度從 30 度上升到 60 度時，雖然表面上的數字似乎變成兩倍，但是溫度並沒有真的變成兩倍。在華氏和攝氏中的零度並不代表完全沒有熱的存在，而只是一種幫助我們比較容易計算溫度的方式而已。例如，水凍結成冰的冰點溫度，在華氏溫標中是 32 度，在攝氏溫標中是 0 度，而在凱氏溫標（Kelvin scale）中則是 273 度。至於水沸騰的沸點溫度，在三種溫標中分別是 212 度、100 度、以及 373.15 度。如果有一個真實零度的話，在溫度數字之間的真正關係應該會是一種比率的關係。例如，從華氏 25 度增加到 50 度就會是「溫暖兩倍」，但是這並不是真實的情況，因爲在沒有一個真實零度的情況下，比率關係是不會存在的。我們可以在水的沸點與冰點溫度的比率上看出這一點。在華氏溫標中，沸點比冰點高 6.625 倍，在攝氏溫標中是 100 倍，而在凱氏溫標中則是 1.366 倍。凱氏溫標具有絕對零度（沒有任何熱的存在），所以凱氏溫標的比率能夠對應於實際的情況。雖然你或許已經相當熟悉這個實體世界裡的例子了，但是當我們以數字來測量態度時，就會出現另外一個任意指定的——不是真正的——零度的例子。我們可能爲一份調查問卷裡的題目指定一個數值（例如：−1 等於「不同意」，0 等於「沒意見」，+1 等於「同意」）。然而，僅僅因爲我們擁有的是數字形式的資料，並不代表我們就可以使用那些要求在數學假定上擁有真實零度的統計程序來處理這些資料。

間斷變項是名目的以及順序的，而我們則可以在間距的或者是比率的尺度上來測量連續變項。在這四種尺度之間，有一種耐人尋味的單一方向關係。我們可以把一個比率尺度的量數轉變成爲區間的、順序的、或者是名目的尺度；我們也可以把一個區間的尺度轉化成順序的、或者是名目的尺度；或者，我們也可以把一個順序的尺度轉化爲名目的尺度。然而，此種轉化的過程卻不能倒過來！之所以會發生此種狀況，是因爲較高的測量尺度比較低的尺度包含了更多精細的資訊。我們可以排除或是忽視較高測量尺度中所包含的精細資訊，但是我們卻無法從一個低尺度的量

數中擠出額外的精細的資訊。

對於順序尺度的測量而言，我們一般會有至少五種順序的類別，並且會試圖爲每個類別都取得許多的觀察結果，這是因爲當我們把一個連續的構念結合成少數幾個順序性的類別時，就會無可避免地造成某種程度的扭曲。而隨著順序類別的數目以及觀察結果的數目增加時，我們就能夠盡量把此種扭曲的狀況減到最低的程度。[15]（請參閱範例 7.2 四種測量尺度的範例）

在我們繼續進行討論之前，你必須把兩件事記在心裡。首先，我們可以測量幾乎任何社會現象。我們可以直接地測量某些構念並且創造出精確的數值（例如，家庭收入），但是有些構念就比較不精確，而且也必須利用一些代理者或是替代物來間接地測量一個變項（例如，犯罪的傾向）。其次，我們可以由其他研究者所創造出來的測量方法中學習到很多知識。很幸運地，我們可以去引用其他研究者的著作。我們不一定每次總是得從頭開始，而可以去使用一個過去的量表或是指數，或者是依據我們的目的來作出修正。測量社會生活的各種層面是一個持續進行的過程，因此我們會不斷地創造出觀念、重新發展並修正理論性定義、而且也會改善舊的或是新的構念的測量方法。

良好測量的原則 在我們考慮是否要使用一個單一指標、量表、或者是指數（接下來會進行討論）來測量一個變項時，必須先知道良好測量的三個特徵：（1）單一變項的屬性或是類別應該要具有相互排除性；（2）這些屬性或者是類別應該也要具有全容性，而且（3）測量應該具有單向度性。

1. **相互排除的特性**（mutually exclusive attribute）是指一個個人或是個案符合、而且也只符合一個變項的類別。例如，我們想要利用四個屬性－基督教、非基督教、猶太教、以及回教－來測量宗教類型這個變項。我們所使用的這四個屬性並不是互相排除的，由於回教與猶太教都是屬於非基督教的宗教信仰，因此一個猶太教徒以及回教徒就同時符合兩個類別：（1）非基督教、以及（2）猶太教或者是回教。另一個缺乏相互排除特性的例子，是利用河港城市、州首都、以及國際機場所在地這三種類別來測量城市的類型。這是因爲一個城市可能同時兼具三種性質（也就是一個州首都同時是河港城市、而且也是國際機場所在地）、這三種類型中任兩者的組合、或者三者都不是。爲了要達到相互排除的特性，我們在創造出類別時，必須注意同一個個案不能同時屬於超過一個以上的類別。

2. **全容的特性**（exhaustive attribute）是指每一個個案都必須屬於或者是符合於至少一個變項的類別。回到剛剛宗教變項的例子，其中共有基督教、非基督教、猶太教、以及回教這四個屬性，而我們現在把非基督教這個類別去掉，以使這些屬性具有相互排除的特性。然而，如此一來，這些屬性就不具有全容的特性，因爲像是佛教徒、印度教徒、無神論者、以及不可知論者就無法歸於任何一個類別

相互排除的特性 一種原則，意指在一個量數中的變項屬性或者是類別的組織方式必須讓回應只能符合或屬於其中一種類別，而且沒有重疊的部分。

全容的特性 一種原則，意指在一個量數中的屬性或是類別應該要包含所有可能的回應。

範例　7.2

四種測量尺度的範例

變項（測量尺度）	測量變項的方法
宗教（名目）	不同的宗教派別（猶太教、天主教、路德教、浸信會）之間沒有等級的分別，只有各自之間的差異（除非有一種信仰被概念化為更接近天堂）。
參與（順序）	「你參加宗教活動的頻率有多高？（0）從來沒有、（1）一年不到一次、（3）一年數次、（4）大約一個月一次、（5）一週二到三次、或者是（8）一週數次。」如果我們問的是一個人參加宗教活動的精確次數的話，那麼這個問題或許可以在比率尺度上進行測量。
智商分數（間距）	大多數的智力測驗都把 100 當作平均值、中間值、或者是正常值。較高或是較低的分數所代表的是與平均值之間的距離。一個智力得到 115 分的人，其得分比參加此項測驗者的平均值稍微高一些，而得到 90 分的人則是比平均值稍微低一些。低於 65 分或者是高於 140 分的人則比較少見。
年齡（比率）	年齡是以年數來進行測量的。此變項有真的零度（出生）。請注意到，一個四十歲的人，所活過的時間是二十歲的人的兩倍。

之內。我們必須創造出能夠涵蓋每一種可能情境的屬性。例如，如果宗教這個變項的屬性包括了基督教、猶太教、回教、或者其他宗教的話，那麼就會具有全容以及相互排除的特性。

3. **單向度性**（unidimensionality）所指的是一個測量方法能夠符合、或者是測量一個單一而且一致的構念。在之前我們討論到構念以及內容效度的時候，就曾經提到過單向度性的概念。單向度性的內容是說，如果我們把數種明確的資訊片段結合成爲單一的分數或者是量數的話，那麼所有的資訊片段所測量的對象，應該要是同樣一件事物。我們有時候會使用一種較爲先進的技術－因素分析－來測試資料的單向度性。

在使用多重指標、量表、或是指數（請參閱下一個部分）來掌握一個複雜構念的不同部分、以及單向度性的標準這兩種觀念之間，我們或許會看到明顯的矛盾情形。不過，此種矛盾情形只是表面上的而已，因爲就理論上而言，構念的抽象層次是有所差異的。我們或許會用多重的次向度來定義一個複雜的、抽象的構念，而且每一個次向度都是這個複雜構念整體內容的一部分。

單向度性　一種原則，意指當研究者使用多重指標來測量一個構念時，所有的指標應該要彼此具有一致性，並且指明一個單一的構念。

對比之下，簡單的、抽象層次較低的構念則較爲具體，而且通常只有一個向度。例如，「女性主義的意識型態」是一個高度抽象而且複雜的構念，這個構念包含了對於社會、經濟、政治、家庭與性關係的特定信念以及態度。這個意識型態的五個信念領域是一個單一的、更爲抽象以及廣泛構念的組成部分。這些部分結合在一起而成爲一個整體，而且會彼此增強並進而形成一組關於女性的尊嚴、力量、以及權力的信念。如果要創造出一個女性主義意識型態的單向度量數，我們必須把它概念化爲一個統一的信念系統，其範圍則包括了從非常反對女性主義到非常支持女性主義的信念。我們可以用連結到這個構念的次部分的多重指標，來測試我們的測量方法的輻合效度。如果某一個信念領域（例如，性關係）在實徵測試中始終與其他的領域格格不入，那麼我們就必須去質疑此領域的單向度性。

我們很容易就會被單向度性給弄糊塗了，這是因爲我們在某個情境中用來測量一個簡單構念的指標，可能在另一個情境中卻指向一個不同的複雜構念的某個部分。我們可以把多個簡單的、具體的構念結合起來，成爲一個複雜的、更爲抽象的構念。測量中的單向度性原則所指的是，當我們要去測量一個構念時，我們必須*針對這個構念的抽象層次*把它概念化爲一個一致的、完整的核心觀念。而此種情況也顯示出概念化過程以及測量過程彼此之間緊密交織在一起的方式。

這兒有一個具體的例子。在某人對於性別平等的態度中，其中關於兩性同工同酬的觀念，是一個比性別意識型態（也就是關於在生活所有領域中的性別關係的一套廣泛的信念）更爲簡單的、更加具體的、而且比較不是那麼抽象的觀念。在測量對於兩性同工同酬的態度時，我們可以把此種態度本身當作是一個單向度的構念，或者是把它當成性別意識型態這個複雜的、較爲廣泛構念的一個比較不是那麼抽象的次部分。這並非意味著性別意識型態就不再具有單向度性。性別意識型態是一個具有數個部分的複雜概念，但是在一個較爲抽象的層次上，性別意識型態也仍然可以具有單向度性。

量表與指數

在這個部分裡，我們所要討論的是量表以及指數。在研究者所創造出來的數百種測量方法中，這兩種是專門化的測量方法。[16] 我們擁有可以測量許多事物的量表以及指數：在官僚組織裡的形式化程度、職業的聲望、人們對於婚姻的調適程度、團體互動的強度、在社區裡的社會活動的水平、某個州的性侵害法律反映女性主義價值的程度、以及某個國家的社會經濟發展程度等等。我們將會檢視測量的原則、考量指數以及量表的建構原則、然後探討一些指數以及量表的主要類別。

由於人們經常交替著使用*指數*（*index*）與*量表*（*scale*）這兩個詞，因此你可能會覺得這兩個詞很容易令人感到困惑。某個研究者所使用的量表對於另一個研究者而言可能是指數，除此之外，量表和指數都可以產生順序尺度或者是間距尺度的量數。更加令人困惑的是，我們可以在單一量數中把量表和指數的技術結合在一起。

然而，量表以及指數都是非常重要的，因爲這兩者都能給予我們更多關於一個變項的資訊，並且相較於使用一個簡單的、單一指標的測量方法，量表與指數更能擴展測量的品質（也就是提升信度與效度）。量表以及指數也可以濃縮與簡化資訊，因此能夠幫助研究者化約資料（請參閱資訊箱 7.3 量表與指數：兩者之間有何不同？）。

指數　把某一個構念或是變項的許多個別量數加總或是結合起來，以創造出一個單一的分數。

指數建構

你隨時都會聽到各種的指數，例如，美國的報紙會報導美國聯邦調查局（Federal Bureau of Investigation, FBI）的犯罪指數，以及消費者物價指數（consumer price index, CPI）。FBI 的指數是警方針對七種所謂指數犯罪（兇殺、重傷害、強暴、搶劫、闖空門、竊盜價值超過 50 美元的財物、以及汽車竊盜）報告的總合，而這個指數一開始時是在 1930 年的統一犯罪報告中的一個部分（請參閱 Rosen, 1995）。至於用來衡量通貨膨脹情況的消費者物價指數，其製作方式則是把購買特定的物品以及服務（例如，食物、租金、以及水電費用等等）的成本加總起來，並且把這個總數與之前時期購買相同物品以及服務的總成本進行比較。自從 1919 年以來，消費者物價指數就一直由美國的勞工統計局所採用，而工資的調升、工會的契約、以及社會保險的給付金等等，都是以消費者物價指數爲基礎來計算的。所謂的**指數**（index），就是把許多項目結合起來，以成爲一個單一的分數數值。一個構念的各種不同成分或是次部分，都會被個別進行測量，然後再結合成爲一個量數。

指數有許多種不同的類型，例如，在一份有二十五道問題的考卷上，答對的題目總數就是一種指數。這種指數是一種綜合的量數，其中每一道問題所測量的是一小部分的知識，而且所有答對以及答錯的題目的數目會被加總起來，以產生一個單

資訊箱 7.3　量表與指數：兩者之間有何不同？

就大部分的目的而言，研究者可以把量表與指數當成是可以互相交換使用的工具，而且社會研究者並未使用統一的術語來區分這兩者之間的差異。

*量表*是一種測量工具，研究者可以藉著這個測量工具而去掌握一個變項的強度、方向、尺度、或者是潛力，並且在一條連續線上面標示出回應或者是觀察結果。一個量表可以使用一個單一的指標或者是多重的指標，而且大多數的量表都是屬於順序的測量尺度。

*指數*也是一種測量工具，在使用此種工具時，研究者會加總或者是結合一個構念的數個不同的指標，以成為一個單一的分數。這一個綜合的分數通常是多重指標的簡單總結，而且被應用在內容效度以及輻合效度當中。指數通常是在間距尺度或者是比率尺度上進行測量。

研究者有時候會在單一量數中結合量表以及指數的特性，當研究者有數個指標都是量表（也就是測量強度或者方向）的時候，此種情況是很常見的。接著，研究者會把這些指標加總以得到一個單一分數，並藉此創造出一個指數。

一的量數。指數可以用來測量人們最想居住的地方（依據失業率、通勤時間、犯罪率、休閒機會、以及天氣等等）、犯罪的程度（依據各種不同的特定犯罪發生情況的總數）、以及某個人的心理健康狀況（依據人們在各種不同的生活領域中的適應情況）等等。

由於指數很容易就可以創造出來，因此我們必須要謹慎小心地去確認在指數中的每個項目都具有表面效度，並且把任何一個沒有表面效度的項目排除在外。針對構念的每一個部分，我們至少都要有一個指標來加以測量。當然，如果能用多重指標來測量構念的各個部分的話，那就更好了。

指數的例子之一，是大學品質指數（請閱範例 7.3 指數的範例）。大學品質指數的理論性定義，是指一所高品質的大學應該具有下列六種卓越的特質：（1）每位教授只需指導少數學生；（2）教授具有高學歷；（3）圖書館的藏書豐富；（4）退學的學生數目極少；（5）許多學生在畢業後繼續深造；（6）教授們曾經出版書籍或者是學術文章。我們依據每個項目分別爲 100 所大學打分數，並且把每一所大學的得分加總起來以製作出大學品質的指數分數，然後我們就可以利用這些指數分數來比較各大學的品質優劣。

我們也可以把指數彼此結合起來。例如，爲了強化我的大學品質指數，我增加了一項教學品質的次指數，而這個指數包含了八個項目：（1）班級平均人數；（2）課堂討論時間的百分比；（3）每位教授教導不同班級的數量；（4）學生在課外時間可以找到教授的頻率；（5）指定閱讀的時間量以及數量；（6）作業促進學習狀況的程度；（7）每位教授對每個學生的瞭解程度；以及（8）學生對於教授講課品質的評分。對於大學品質指數中的其他部分，我們也可以製作出類似的次指數量數，而這些次指數量數也可以結合成爲更全面的大學品質量數，如此一來，就能夠更進一步地闡述「大學品質」這個構念的定義。

接下來，我們要看看當我們在製作指數時所會牽涉到的三個議題：項目的加權、遺失資料、以及使用比率與標準化。

1. 在指數建構中，*加權*（*weighting*）是一個相當重要的議題。除非有另外說明，我們一般都會假定在一個指數中的所有項目都是沒有加權的。同樣地，除非我們有充分的理論性原因來爲每個項目分派不同的權數，否則的話，我們就應該使用相等的權數。一個*未加權的指數*（*unweighted index*）會給予每個項目相等的權數，因此，我們只要把所有項目加總起來就可以了，不必經過任何修正，就好像把每一個項目都乘以 1 一樣（如果是負面的項目，就乘以－1）。而在*加權的指數*（*weighted index*）中，某些項目的價值或者是權數會高於其他的項目。至於權數的大小，可以來自於理論的假定、理論性定義、或者像是因素分析之類的統計技術。

例如，在我們更深入地闡述大學品質指數的理論性定義之後，我們決定學生／教授的比率、以及擁有博士學位的教授數目，其重要性是每個學生可以分配到的圖書館藏書數量或者是追求更高學位的學生數目的兩倍。除此之外，大一新鮮人被退學的百分比、以及每一位教授的著作數量，其重要性則是圖書館藏書或是追求更高學位的學生比例的三倍。當我們用公式來表示時，此種加權的情況會比較容易看得

範例　7.3

指數的範例

以符號來表示的話，則：

Q = 整體大學品質

大學品質指數是根據下列六個項目所製作而成的：

R = 每位教授所指導的學生數目

F = 擁有博士學位的教授比例

B = 每名學生所能分配到的圖書館藏書數量

D = 大一新鮮人退學或是未完成學業的百分比

A = 攻讀更高學位的畢業生百分比

P = 每一個教授的著作數量

未加權的公式：　(−1) *R* + (1) *F* + (1) *B* + (−1) *D* + (1) *A* + (1) *P* = Q
加權的公式：　(−2) *R* + (2) *F* + (1) *B* + (−3) *D* + (1) *A* + (3) *P* = Q

老牌長春藤大學

未加權的公式：　(−1) 13 + (1) 80 + (1) 334 + (−1) 14 + (1) 28 + (1) 4 = 419
加權的公式：　(−2) 13 + (2) 80 + (1) 334 + (−3) 14 + (1) 28 + (3) 4 = 466

地區大學

未加權的公式：　(−1) 20 + (1) 82 + (1) 365 + (−1) 25 + (1) 15 + (1) 2 = 419
加權的公式：　(−2) 20 + (2) 82 + (1) 365 + (−3) 25 + (1) 15 + (3) 2 = 435

大型大學

未加權的公式：　(−1) 38 + (1) 95 + (1) 380 + (−1) 48 + (1) 24 + (1) 6 = 419
加權的公式：　(−2) 38 + (2) 95 + (1) 380 + (−3) 48 + (1) 24 + (3) 6 = 392

出來（請參考範例 7.3）。

每個教授所分配到的學生數量、以及退學的學生比例都是負值，這是因爲當這兩個數字增加時，大學的品質就會隨之下降。加權過以及未加權的指數可能會產生不同的結果。讓我們來看看老牌的長春藤大學、地區大學、以及大型大學的狀況。在未加權之前，這三種大學的指數分數是不分軒輊的，然而，在經過加權之後，這三種大學的品質分數就大不相同了。

在這個例子裡，加權會產生不同的指數分數，但是在大多數的情況中，加權以及未加權的指數所得到的結果是差不多的。研究者所關心的是變項之間的關係，而對於變項之間的關係，加權以及未加權的指數所產生的結果通常是相當類似的。[17]

2. 在建構指數時，*遺失資料*（*missing data*）會是一個嚴重的問題。每當某些個案的資料遺失時，效度與信度就會大打折扣。我們可以用四種方法來試著解決這個問題（請參閱資訊箱 7.4 處理遺失資料的方法），但是沒有任何一種方法能夠完全地解決這個問題。

資訊箱 7.4　處理遺失資料的方法

1. *把所有遺失資訊的個案全部都刪除掉*。如果把芬蘭從研究中刪除掉，對於其他可以獲得資訊的國家而言，這個指數就會是可信的。如果其他的國家也有遺失的資訊，這就會是一個問題了。如此一來，一項針對 50 個國家所進行的研究，就會變成只有針對 20 個國家所進行的研究。除此之外，遺失資訊的個案可能在某些面向有相似性（例如，這些個案都是位於東歐或者是在第三世界），此種情況會限制了研究發現的可推論性。
2. *以平均分數取代沒有資料的個案*。其他國家的平均識字分數可以用來當作替代的分數。此種「解決之道」可以把芬蘭保留在研究裡，但是卻會給它不正確的數值。對於項目數量很少的指數、或者是一個並非「平均」的個案，此種作法會造成嚴重的效度問題。
3. *根據個案的非量化資訊來插入資料*。關於芬蘭的其他資訊（例如，13 歲到 18 歲的青少年就讀中學的比例），可以被用來對其識字率進行有根據的推測。在這個情境中，此種「解決之道」不太能讓人接受，因為在測量芬蘭的識字率方面，此種作法不夠精確，而且這種作法所依賴的是一個未經檢驗的假定——我們可以從其他國家的中學就學率來預測其識字率。
4. *插入一個隨機的數值*。在發展指數的例子中，這是一種不智之舉。但是，如果某個指數擁有非常大量的項目、而且個案的數量也相當龐大時，此種作法或許是可以接受的。然而，如果真的是此種情況的話，把遺失資料的個案刪除掉，或許是一個比較好的「解決之道」，也能夠產生較為可信的量數。

資料來源：Allison (2001).

例如，我針對 50 個國家在 1985 年時的社會發展程度建構了一個指數，而這個指數包含了四個項目：預期壽命、擁有室內水管配線的家戶數比例、識字人口比例、以及每一百個人擁有的電話數目。我找到了一份聯合國的統計數字，以當作我的資訊來源。比利時的分數是 68＋87＋97＋28，而土耳其的分數是 55＋36＋49＋3。然而，在關於芬蘭的資料裡，我發現沒有關於識字率的資料。我檢查了其他的資訊來源，但是沒有任何一個來源有這項資料，因爲當時並未有任何人或者是機構去調查與蒐集芬蘭的這項資料。

3. *比率*（*rate*）以及*標準化*是相關的觀念。你一定曾經聽過犯罪率、人口成長率、或者是失業率等等。有一些指數以及單一指標的量數會以比率來表示。比率牽涉到把一個項目的值加以標準化，如此才能夠進行比較。在一個指數中的項目經常必須要先進行標準化，才能被結合起來。

標準化（standardization）的過程是先選擇一個基數，然後再把原始量數除以這個基數。例如，A 城市在某一年之中總共發生了十起謀殺案，而 B 城市在同一年裡則發生了三十起謀殺案。爲了要比較這兩個城市的謀殺案件，我們必須利用城市的人口數來將謀殺案件的原始數目加以標準化。如果這兩個城市的規模相同的話，那麼 B

標準化　一種程序，此種程序能藉由指定一個共同基礎，並以測量不同單位的方式與統計的方法，調整量數進行真正的比較。

城市就是比較危險的城市。但是，如果 B 城市比 A 城市大很多的話，那麼 B 城市就有可能是比較安全的。例如，假設 A 城市的人口是十萬人，而 B 城市的人口則是六十萬人，那麼 A 城市每十萬人的謀殺率是 10，而 B 城市的謀殺率則只有 5 而已。

標準化讓我們可以在一個共同的基礎上去比較不同的單位。標準化的過程也叫做規範化（norming），此種過程會把相關的、但是特質不同的影響移除，以突顯出重要的差異之處。例如，我們現在有兩班不同的學生，其中一班是美術班，而且班上有十二個會吸煙的學生，另外一班是生物班，班上有二十二個會吸煙的學生。我們可以利用班級人數來將吸煙者的數目加以標準化，就能夠比較吸煙者的比率或者是頻率。假定美術班有 32 名學生，而生物班則有 143 名學生。你早已經知道的標準化方法之一，就是利用百分比率，也就是把所有的量數依照 100 這個共同基數來進行標準化。就百分比率而言，我們可以很容易就看出來，美術班的吸煙者比率（37.5%）超過生物班（15.4%）的兩倍。

在標準化過程中的關鍵問題之一，是決定應該要採用什麼基數。在上述的兩個例子裡，我是如何知道要用城市人數或者是班級人數當作基數的？基數的選擇不一定總是顯而易見的，而是要取決於構念的理論性定義。不同的基數會產生不同的比率。例如，我們可以把失業率定義爲在勞動人口中沒有工作的人數。如此一來，整體失業率的公式就是：

失業率=失業的人數／整體工作者的總人數

我們可以把整體人口分割爲數個次群體，像是白人男性、黑人女性、18 到 28 歲的黑人男性、或者是擁有大學學歷的人，並且爲這些次群體計算其個別的失業率。對於研究問題的理論性定義而言，這些次群體的個別失業率可能具有更大的關聯性。例如，我們認爲失業是一種會影響到全家或是家庭的經驗，因此我們應該是以家戶數爲基數，而不是以個人爲基數。如此一來，失業率的計算方式就會變成：

失業率=家中至少有一人失業的家戶數／家戶總數

不同的概念化過程，會暗示了不同的基數以及不同的標準化方法。當我們把數個項目結合爲一個指數時，我們最好能在一個共同的基數上將所有的項目先進行標準化（請參閱範例 7.4 標準化以及 2000 年奧運的真正贏家）。

量表

當我們想要測量某個人對於某件事物的感覺或者是想法時，我們通常會使用量表，而有些人則會將此稱作是*感覺的硬度或是能力*（*hardness or potency of feelings*）。量表也有助於概念化以及操作化的過程。例如，你相信人們對於特定政策（例如，住宅、教育、外交事務）的判斷，是以一個單一的意識型態面向爲基礎的。利用量表可以幫助你確認，某個單一的構念——例如，「保守的／自由的意識型態」——是否真的是人們對特定政策所採取的立場背後的基礎。

範例　7.4

標準化以及 2000 年奧運的真正贏家

美國的體育迷們對於美國在 2000 奧運比賽中「獲勝」感到非常雀躍歡欣，因為美國在這次的奧運中所獲得的金牌總數在所有國家中是最多的。然而，因為這些體育迷們沒有進行標準化，所以「獲勝」只是一個幻覺而已。當然，在所有參賽的國家中，美國身為全世界最富裕、而且人口第三多的國家，在一對一的比賽中表現其實是可圈可點的。然而，若是要看清楚真相，我們就必須以人口或是財富當作基數來進行標準化。標準化可以藉由調整比賽結果的方式，產生一幅比較正確的圖像，就好像所有的國家都擁有相等的人口以及財富一樣。經過標準化之後的結果顯示出，巴哈馬這人口少於三十萬人的國家（比一個中型的美國城市還要小），就比例上而言是贏得最多金牌的國家。在針對人口規模或是財富進行調整之後，美國甚至都還排不到前幾名。美國似乎只是因為擁有巨大的人口規模以及財富才會成為領袖。美國的體育迷必須刻意對美國擁有的相對優勢視而不見，才能繼續保有美國是世界第一的幻覺。

2000 年雪梨奧運中獲得最多金牌的前十名國家

未經過標準化的排名			*經過標準化的排名**			
排名	**國家**	**金牌總數**	**國家**	**總數**	**人口**	GDP
1	美國	39	巴哈馬	1.4	33.3	20.0
2	俄羅斯	32	斯洛伐尼亞	2	10	10.0
3	中國	28	古巴	11	9.9	50.0
4	澳洲	16	挪威	4	9.1	2.6
5	德國	14	澳洲	16	8.6	4.1
6	法國	13	匈牙利	8	7.9	16.7
7	義大利	13	荷蘭	12	7.6	3.0
8	荷蘭	12	愛沙尼亞	1	7.1	20.0
9	古巴	11	保加利亞	5	6.0	41.7
10	英國	11	立陶宛	2	5.4	18.2
	歐盟 15 國**	80	歐盟 15 國	80	2.1	0.9
			美國	39	1.4	0.4

*「人口」欄的數字所指的是每一千萬人的金牌數，而「GDP」欄的數字所指的是每 100 億美元的金牌數。
**歐盟 15 國指的是把歐盟中的 15 個國家當成一個單一的整體單位。

*資料來源：*Adapted from The Economist, October 7, 2000, p. 52. Copyright 2000 by Economist Newspaper Group. Reproduced with permission of Economist Newspaper in the format Textbook via Copyright Clearance Center.

量表可以測量一個變項的強度、方向、水平、或者是潛力。圖形計分**量表**(scales)是量度(scaling)的一種基本形式。在一條從一個極端到另一個極端的線上，人們可以勾選其中一個點，以表明自己所給予的評分。此種形式的量表相當容易建構與使用，而此種量表所傳達的是一種連續體的觀念，而且指定一些數字可以幫助人們思考數量的問題。量表假定擁有同樣主觀感受的人們，會在圖形量表上標示出相同的位置。圖 7.7 是一個「感覺溫度計」量表的例子，此種量表是被用來測量人們對於社會中的各種不同團體(例如，國家婦女組織、三 K 黨、工會、醫生等等) 的感受。自從 1964 年以來，政治學家就在國家選舉的研究中使用此種量表來測量人們對於候選人、社會團體、以及各種議題的態度。[18]

量表　一種量化資料的量數，通常被用於調查研究中，能夠掌握某個變項構念在一個連續體上的強度、方向、水平、或是潛力。大多數的量表都是屬於順序的測量尺度。

李克特量表　一種經常被用於調查研究中的量表。在此種量表中，人們是以在一條連續線上排列的順序尺度類別，以表示自己的態度或其他的反應。

我們接著要來看的是五種經常被使用的社會科學量表：李克特量表、賽斯通量表、包氏社會距離量表、語意差別量表、以及古特曼量表。每一種量表都代表了在量度邏輯上的些微差異。

1. *李克特量表*　你很可能已經使用過**李克特量表**（Likert scales）了，此種量表被廣泛地應用於調查研究中。這是由李克特（Rensis Likert）在 1930 年代所發展出來的，其目的是要提供關於某人態度的順序尺度量數。[19] 李克特量表也被稱爲*總合評分*（*summated-rating*）或者是*加法量表*（*additive scales*），因爲在計算某人在量表上的分數時，所使用的方法是加總此人在量表上的回應數目。李克特量表通常要求人們表明他們同意或者是不同意某項陳述。我們也可以作出其他修正，像是人們也可能會被問到他們是贊成或是反對、或者他們是否相信某件事「幾乎總是真的」(請參閱範例 7.5 李克特量表類型的範例)。

要製作李克特量表，你至少需要兩種類別，像是「同意」以及「不同意」。然而，只使用兩種類別的話，所產生的是粗糙的量數，而且也強迫把差異性只限制在兩種

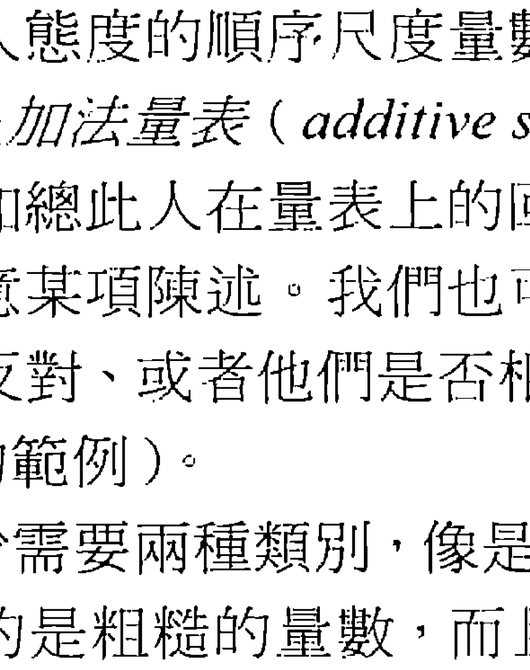


圖 7.7　「感覺溫度計」的圖形評分量表

範例 7.5

李克特量表類型的範例

羅森伯自尊量表（The Rosenberg Self-Esteem Scale）
大體而言，我傾向於覺得自己是個失敗者。

（1）幾乎總是如此
（2）常常如此
（3）有時如此
（4）很少如此
（5）從未如此

學生對教學評分的量表
整體而言，我對這門課程的教學品質所評的分數是：

非常好　　好　　普通　　還不錯　　差勁

漱口水評分的市場研究量表

品牌	*完全不喜歡*	不太喜歡	有點不喜歡	有點喜歡	蠻喜歡的	非常喜歡
X	___	___	___	___	___	___
Y	___	___	___	___	___	___

工作小組監督者量表
我的監督者：

	從來沒有	很少	有時候	經常	總是
讓成員瞭解對他們的期望	1	2	3	4	5
友善而且易於接近	1	2	3	4	5
對所有單位成員一視同仁	1	2	3	4	5

類別之內。一般而言，使用四種到八種類別是比較好的作法。在資料蒐集完畢之後，你可以把不同的類別結合或者是合併，但是如果在蒐集資料時使用的是粗糙的類別，之後就無法使這些類別變得更加精確。你可以在量表的末端增加類別的數目，像是加上「強烈同意」、「有點同意」、「非常強烈同意」等等，以增加類別的數目，不過，你應該要把選項的數目控制在最多八到九個的範圍內。如果選項數目超過這個數字的話，不但會失去意義，而且作答的人也會覺得困惑。除此之外，選項也應該要平均分配（例如，「強烈同意」、「同意」、「強烈不同意」、「不同意」）。南那利（Nunnally, 1978: 521）曾經說過：

> 當量表裡的分階從 2 增加到 20 時，信度的增加速度一開始是非常快速的。但是當分階到了 7 時，信度的增加曲線就開始持平，而在大約 11 階之後，就算再增加分階的數目，信度的增加量也是微乎其微了。

研究者們一直在爭論，是否要在方向性的類別（例如，「不同意」、「同意」）之外提供一個中立的類別（例如，「不知道」、「尚未決定」、「沒意見」）。如果加上一個中立的類別，類別的總數就會變成奇數。

如果有數個李克特量表的項目都是測量同一個構念的話，我們可以把這些項目

結合起來成爲一個綜合指數。讓我們來看看史尼德曼與哈根（Sniderman and Hagen, 1985）所製作的女性平等機會指數（Index of Equal Opportunity for Women）以及自尊指數（Self-Esteem Index）（請參閱範例 7.6 使用李克特量表來製作指數的範例）。在一項大型的調查進行到一半的時候，這兩位研究者詢問受訪者三個關於女性地位的問題。接下來，他們爲受訪者的答案計算分數，並且把這些項目結合爲一個範圍從 3 到 15 的指數。受訪者也回答了數個關於自尊的問題。請注意，在爲這些項目評分時，其中有一個項目（第二個問題）是反向計分的。以此種方式轉換評分方向的理由，是爲了要避免**反應心向**（response set）的問題。反應心向又稱爲*反應風格*（*response style*）以及*反應偏頗*（*response bias*），所指的是有些人因爲懶惰或者是心理傾向，而以同樣的方式（通常是同意）來回答大量的問題。例如，如果題目的敘述方式使得回答「強烈同意」總是意謂著擁有自尊時，我們可能無法知道一個總是回答強烈同意的人，到底是真的擁有高度的自尊心、或者只是單純地傾向於用同意來回答所有的問題。換句話說，這個人在回答「強烈同意」時，可能只是出於習慣、或者是給予正向答案的傾向。因此，研究者會以正向或是反向的方式來交替敘述問題，如此一來，那些總是回答同意的人就會出現答案前後不一致、或者是意見自相矛盾的情況。

我們經常會把許多李克特量表的態度指標結合成爲一個指數。量表以及指數能夠提升信度以及效度。一個指數使用了多重的指標，因而能夠提升信度。而使用多重指標來測量某個構念或是意見的數種不同面向，則能夠提升內容效度。最後，對於個人的意見，指標分數能夠提供更爲精確的量化量數。例如，在測量一個人的意見時，我們可以使用 1 到 4 的數字來進行測量，而非僅僅用四種類別（「強烈同意」、「同意」、「不同意」、「強烈不同意」）而已。

我們也可以使用像是－2、－1、＋1、＋2 之類的分數，而不一定要像是在之前的例子裡一樣以李克特量表上的項目來計分。此種計分方式的優點在於，零分就代表中立狀態或者是完全模糊，而大的負數和大的正數則代表了正反兩方的態度與意見。

我們所指派給每個反應類別的數字是由我們所自行決定的。你要記住的是，使用零這個數值並無法給予量表或是指數一個比率的測量尺度。李克特量表的量數屬於順序的測量尺度，因爲回應所指明的只是一種順序或是等級而已。就算我們不使用 1 到 4 或者是－2 到＋2，而是使用 100、70、50 以及 5 這幾個數字的話，也同樣能夠進行測量。除此之外，我們也不應該誤以爲，只因爲我們指派了數字，所以順序類別之間的距離就是等距的區間。事實上，使用這些數字的最主要原因純粹只是爲了方便起見而已，因此基本的測量尺度仍然只是順序尺度而已。[20]

李克特量長的真正優點，在於其簡單而且易於使用的特性。當我們結合數個排序過的項目時，我們就可以得到一個更爲全面的多重指標測量。李克特量表有兩項限制：以不同的方式來結合數個量表項目，會產生相同的整體分數；而且反應心向也是一個潛在的危險。

反應心向　一種同意一連串問題中的每一道問題、而非仔細思考每一道問題的傾向。

範例 7.6

使用李克特量表來製作指數的範例

史尼德曼以及哈根（1985）製作了指數來測量對於女性機會平等以及自尊的觀念。這兩種指數是把分數直接加總而成的、未經過加權的指數。

女性機會平等指數

問題

1. 女性接受教育以獲得最高職位工作的機會比男性還少。
   強烈同意　　有點同意　　有點不同意　　非常不同意　　不知道

2. 許多有資格的女性無法獲得好工作，而擁有相同技能的男性則比較沒有這方面的問題。
   強烈同意　　有點同意　　有點不同意　　非常不同意　　不知道

3. 我們的社會歧視女性。
   強烈同意　　有點同意　　有點不同意　　非常不同意　　不知道

*分數計算方式：*對於所有的問題，強烈同意 = 1，有點同意 = 2，有點不同意 = 4，非常不同意 = 5，不知道 = 3

最高可能指數分數 = 15，表示受訪者認為女性所擁有的機會是平等的
最低可能指數分數 = 3，表示受訪者認為女性所擁有的機會是不平等的

自尊指數

問題

1. 整體而言，我對自己感到滿意。　　同意　　不同意　　不知道
2. 有時候我覺得自己一點都不優秀。　　同意　　不同意　　不知道
3. 我有時候會覺得其他人沒有把我的意見當一回事。　　同意　　不同意　　不知道

*分數計算方式：*問題 1 與 3：1 = 不同意，2 = 不知道，3 = 同意，問題 2：3 = 不同意，2 = 不知道，1 = 同意

最高可能指數分數= 9，表示具有高自尊
最低可能指數分數= 3，表示具有低自尊

2. *賽斯通量表法*　此種量表的使用情境，是當我們對某個有許多順序面向的事物感到興趣，但是卻想要得到一種測量方式，可以把所有資訊結合成單一區間尺度的連續體。例如，有一家名爲「快潔」（Quick and Clean）的乾洗公司與我們聯絡，這家公司想要知道，與其主要競爭對手「友善清潔」（Friendly Cleaner）公司相比之下，該公司在格林鎮居民心中的形象如何。我們把人們對於乾洗公司的態度加以概念化，並且分成四個面向：對於營業地點、營業時間、服務、以及價格的態度。我們得知人們認爲快潔乾洗公司的營業時間和地點較爲方便，但是價格較高、而且服務態度較差。至於友善清潔公司，人們則認爲該公司的價格較低、服務態度較好，但是營業時間與地點則比較不方便。除非我們知道這四個面向與核心態度——乾洗

公司的形象——之間的關聯方式，否則我們就無法判斷哪一家公司的形象比較好。在 1920 年代晚期時，賽斯通（Louis Thurstone）發展出一些可以在此種情境中指派數值的量度方法，而這些方法現在則被稱作**賽斯通量表法**（Thurstone scaling），或者是*等距區間法*（*method of equal-appearing intervals*）。[21]

賽斯通量表法以是比較判斷法則（law of comparative judgment）爲基礎，來處理當每個人作出獨特判斷時的比較順序態度的議題。也就是說，在每個人作出個別的判斷之後，這個法則會標示或是定出某個人的態度相對於其他人態度的位置。比較判斷法則主張，針對每個受到評判的對象或是概念，我們都可以確認「最通常的反應」。雖然不同的人會作出不同的判斷，個人的判斷還是會群聚圍繞在單獨一個最通常反應的周圍。圍繞著某一個通常反應周圍的個人判斷的分布情況，會依循著一種被稱爲*常態分配*（*normal distribution*）的統計模式。根據比較判斷法則，如果許多人認爲某兩個物體是不一樣的，那麼大多數對於這兩個物體的通常反應就會落在連續線的兩端，而且兩種意見彼此之間有著遙遠的距離。對比之下，如果許多人覺得困惑不解、或者是意見分歧的話，那麼對於這兩個物體的通常反應就會與彼此相當接近。

利用賽斯通量表法，我們就可以發展出許多關於研究對象的陳述（例如，超過 100 句以上），然後再利用許多評判者來協助我們刪除那些模稜兩可的陳述，以把陳述的數目降低（例如，刪到剩下 20 句左右）。每一位評判者在一條基本的連續線（例如，從贊成到不贊成）上對每一句陳述進行評分。我們接著要檢視這些評分的結果，並且依據下列兩個因素來保留某些陳述：（1）評判者之間意見相同，以及（2）該項陳述在可能價值範圍內的位置。最後留下來的陳述組合就是橫跨一定的價值範圍的測量量表。

賽斯通量表法是從大量的陳述開始的，而且這些陳述涵蓋了所有形形色色的意見。每一句陳述都應該要清楚而且精確。「良好」的陳述所指涉的是現在，而且不能被詮釋爲事實。這些陳述不可能被所有人所接受、應該要使用簡單的句子、而且要避免像是*總是*（*always*）以及*絕非*（*never*）之類的字眼。我們可以從回顧文獻中、從大眾媒體中、從個人經驗中、以及去詢問別人的意見中得到撰寫這些陳述的想法。例如，關於乾洗公司的陳述或許可以包括前面列出的四個面向，並且加上下列的這些陳述：

- 我認爲 X 業者在乾洗方面的服務既快速又準時。
- 依我之見，X 業者的店面看起來既乾淨又吸引人。
- 我認爲 X 業者在去除髒污方面做得不好。
- 我相信 X 業者對於清洗外套的收費價格是合理的。
- 我相信 X 業者送回來的衣物是乾乾淨淨的，也熨燙平整。
- 我認爲 X 業者的交貨服務很糟糕。

接下來我們要找到 50 到 300 名評判者，而且這些評判者應該要熟悉陳述中的事物或者是構念。每一位評判者會拿到一組陳述的卡片以及指示，每一張卡片上面者

賽斯通量表法　一種測量方法，在使用此種測量方法時，研究者會給一群評判者許多項目，並且要求他們在一條連續線上把這些項目整理分類成不同的類別，然後研究者再仔細思考分類之後的結果，以挑選出評判者們意見一致的項目。

會有一句陳述，而評判者則把每一張卡片放在數疊卡片的其中一疊裡。卡片的疊數通常是 7、9、11、或者是 13，而這些卡片疊則代表了與被評估的事物或者是概念有關的一定範圍內的值（例如，從喜歡到中立到不喜歡）。每一位評判者都是獨立地把卡片放在評分疊裡面，而且不受其他評判者的影響。

在每一位評判者把所有的卡片都放進卡片疊中之後，我們就可以製作一份把各疊卡片與陳述交叉分類的圖表。例如，100 句陳述與 11 疊卡片可以產生一份 11×100 的圖表，也就是一張有著 11×100 = 1,100 格的表格。分派評分給某句陳述的評判者數目會被寫在每一個格子裡，然後我們可以使用統計量數（不在目前的討論內容之中）來計算每一句陳述的平均評分，以及評判者們同意或者是不同意某句陳述的程度。我們會保留那些評判者之間同意程度最高的陳述，也就是評分者之間的信度，同時也保留那些代表了整個價值範圍的陳述。（請參閱範例 7.7 賽斯通量表法的範例）

利用賽斯通量表法，我們就能夠建構一個態度量表，或者是從大量關於態度的陳述當中挑選出一些陳述。然而，此種方法有下列四種限制：

- 此法測量的是對於某句陳述的態度是同意或是不同意，而不是同意或是不同意的強度。
- 此法假定評分者以及其他人對於陳述在評分系統中出現的位置會有相同的看法。
- 此法相當耗費時間，而且所費不貲。
- 有可能以數種不同的方式得到完全相同的整體分數，因爲就算以不同的方式來組合各種陳述，對這些陳述同意或是不同意的評分仍可能產生相同的平均分數。

3. *包氏社會距離量表*　**包氏社會距離量表**（Bogardus social distance scale）所測量的是將社會團體彼此分隔開來的「社會距離」。我們可以對某個團體使用此種量表，以得知該團體的成員對某個目標或者是「外團體」（"out-group"）所感受到的距離有多遠。包嘉德斯（Emory Bogardus）在 1920 年代發展出此種技術，用來測量不同族群的成員之間與彼此交往的意願。自從那時開始，包式社會距離量表就一直被用來瞭解某個團體裡的成員對於其他團體（例如，宗教的少數團體、或者是偏差團體）的感覺是親近還是疏遠。[22]

此種量表有一個簡單的邏輯。我們要求人們去回應一系列經過排序的陳述，並且把在社會上較爲親密或是親近的情境放在一端，而把比較不具有社會威脅性的情境放在相反的另一端。此種量表的邏輯假定，一個對於另外一個社會團體覺得不自在、而且只能接受一些非威脅性（社會距離較遠）情境的人，在面對比較具有威脅性（社會親密）的情境時，會表現出不舒服或者是拒絕的反應。

包氏社會距離量表　一種測量兩個或者是多個社會團體之間的社會距離的量表，其測量方法是要求某個團體裡的成員指出，對於和其他團體成員之間的各種不同類型的互動或是親近方面，他們感到自在的界限到什麼程度。

我們能夠以數種不同的方式來使用此種量表。例如，我們可以給人們一系列的陳述：來自於 X 團體的人們進入你的國家、待在你的城鎮裡、在你上班的地方工作、住在你的社區裡、成爲你的朋友、並且與你的兄弟或是姐妹結婚。我們可以詢問人們

範例　7.7

賽斯通量表法的範例

測量的變項：關於死刑的意見

步驟一：利用個人經驗、通俗以及專業的文獻、以及他人的看法等等，提出 120 句陳述。

陳述範例

1. 我認為死刑是殘忍的，而且是不必要的刑罰。
2. 沒有死刑的話，將會有更多的暴力犯罪。
3. 我相信死刑應該只用於少數極端的暴力犯罪上。
4. 我不認為任何人會因為害怕死刑而不敢犯下謀殺罪。
5. 我認為只要是犯下謀殺罪的人，就算他們在精神方面有問題，也不能免除死刑。
6. 我相信聖經支持死刑的存在。
7. 對我而言，死刑本身不是問題，但我相信用電椅將人處死是一種殘忍的方式。

步驟二：把每句陳述寫在個別的卡片或是紙片上，並且將這 120 句陳述做成 100 組卡片。

步驟三：找到 100 名願意當評判者的人，然後給每一位評判者一組陳述的卡片，並指示他們把卡片放在 11 疊卡片中的其中一疊，而這些卡片疊的排序是從 1 = 非常不認同的陳述，到 11 = 非常認同的陳述。

步驟四：評判者把每一句陳述放在 11 疊卡片中的某一疊（例如，1 號評判者把 1 號陳述放進 2 號卡片疊中；2 號評判者把相同的陳述卡片放入 1 號卡片疊中；3 號評判者則把該句陳述放入 2 號卡片疊中，而 4 號評判者則把該句陳述放入 3 號卡片疊中等等）。

步驟五：把卡片疊從評判者那兒收回來，並且製作一張表格來總結他們的反應。請參閱下列的表格範例。

每句陳述評分卡片疊的評判者數目

	不支持					*中立*					*支持*	
陳述	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*總數*
1	23	60	12	5	0	0	0	0	0	0	0	100
2	0	0	0	0	2	12	18	41	19	8	0	100
3	2	8	7	13	31	19	12	6	2	0	0	100
4	9	11	62	10	4	4	0	0	0	0	0	100

步驟六：計算評判者們的平均評分以及同意程度。例如，問題 1 的平均分數大約是 2，所以同意度相當高，而問題 3 的平均分數接近 5，因此同意度較低。

步驟七：選出最後 20 句陳述來納入死刑意見量表中。選擇的標準，是那些評判者們表現出意見一致的陳述（也就是大多數人把同一句陳述放在相同或者是相近的卡片疊裡），以及那些能夠反映各種意見的陳述，也就是範圍包括了從支持到中立到不支持的陳述。

步驟八：準備一份有 20 個問題的問卷，並且詢問研究中的受訪者同意或者是不同意這些陳述。

對於陳述中的情境是否覺得自在、或者是否能夠接受此種接觸方式。我們也可以要求人們必須回應所有的陳述，直到他們處於一種覺得不舒服的情境為止。在使用此種方式時，陳述不必要有固定的數目，一般而言，數目大約是五句到九句之間。

我們也可以使用包氏社會距離量表，來看看人們對於某個外團體和另外一個外

團體之間所感覺到的距離有多遠（請參閱範例 7.8 包氏社會距離量表的範例）。我們可以把社會距離的量數當成是一個自變項或者是依變項。例如，我們或許會認爲，對於具有某些特質（例如：教育）的人而言，和一個團體的社會距離是最遙遠的。我們的假設可能如下：白人對於越南人的社會距離感受，和教育之間呈現負相關的關係。也就是說，教育程度最低的白人，其社會距離感是最遠的。在此種情境中，社會距離是依變項，而教育程度就是自變項。

社會距離量表有以下兩個限制。首先，我們必須針對一個特定的外團體以及社會場景來調整陳述的類別。第二，我們不容易去比較某個受訪者對於數個不同團體的感覺，除非該位受訪者同時完成針對所有外團體所設計的類似的社會距離量表。當然，一名受訪者完成量表的方式以及該名受訪者在特定社會情境中的實際行爲之間，也可能會有所差異。

4. *語意差別* **語意差別**（semantic differential）是在 1950 年代所發展出來的，其目的在於爲人們對於某個概念、物體、或者是其他人的感覺提供一個間接的測量方法。語意差別利用許多形容詞來測量主觀的感受，因爲人們通常是透過形容詞來傳達其評估的結果。大多數的形容詞都有兩極的相對詞（例如：光明的／黑暗的、堅硬的／柔軟的、緩慢的／快速的），而語意差別則試圖利用形容詞中的含意來掌握人們所作出的評估結果。藉由此種方式，語意差別就可以利用一種間接的方式來測量一個人的感覺以及評估結果。

在使用語意差別時，我們會爲研究參與者提供一份成對的相反形容詞的列表，而在成對的相反形容詞中間則有一個分成 7 點到 11 點的連續線。我們會要求研究參與者在兩個相反的形容詞中間的連續線上，標示出一個最能夠貼切地表達他們的評估或是感覺的點。這些不同的成對形容詞之間的差異性可以很大，而且應該要充分地混合（例如，正面的形容詞不應該全都放在右側或者是左側）。英文中的形容詞通常會分成三種主要的意義類型：評估（好／壞）、力量（強／弱）、以及活動（主動／被動）。。在這三種類型當中，與評估有關的形容詞通常是最爲顯著而且重要的。

語意差別最困難的部分在於分析結果的過程，因爲我們必須使用高階的統計程序才能對結果進行分析。語意差別的結果可以讓我們知道某個人如何認知不同的概念、或者是人們對於某個概念、物體、或者是某個人的看法。例如，政治分析家可能會發現，年輕的選民認爲他們的候選人是傳統的、軟弱的、遲鈍的、而且介於好與壞的中間。年長的選民則認爲候選人較偏向強勢的、迅速的、好的、而且介於傳統與現代之間。在範例 7.9 語意差別的範例中，某個人爲兩個概念進行評分。對於每一個概念的反應模式，說明了這個人對於這些概念的感覺。這個人對這兩個概念的看法相當不同，而且對於離婚的感覺似乎是相當負面的。

語意差別 一種間接地測量人們的感覺或者是想法的量表，其測量方式是提供人們一個主題或者是物體，以及一張兩極對立的形容詞或是副詞的列表，然後要求人們在兩個形容詞或是副詞中間的數個空格內，標示出能夠代表其感受的空格。

統計的技術可以製作出三度空間的結果示意圖。[23] 而形容詞的三種主要的意義類型則被標示在一個三維的「語意空間」裡。在這個三維的示意圖

範例　7.8

包氏社會距離量表的範例

有一名研究者想要知道大一新鮮人對於來自兩個不同國家（奈及利亞以及德國）的交換學生所感受到的社會距離有多遠。這名研究者想要看看學生對於來自非洲的學生、還是來自歐洲的學生感覺距離比較遠。這名研究者在訪談中使用了下列的一系列問題：

你個人對於來自於（國家名稱）的交換學生是否覺得自在？請用「是」或「否」來表明你的最初反應。

____ 到你的學校裡訪問一個星期
____ 在你的學校裡註冊成為正式的學生
____ 選修了幾門你也有選修的課程
____ 在課堂上坐在你旁邊，並且與你一起準備考試
____ 與你住在同一棟宿舍裡的同一個樓層，而且只相隔幾個房間
____ 與你住在同一間寢室的同性學生
____ 邀你出去約會的異性學生

假設的結果　　*認為感覺自在的新鮮人比例*

	奈及利亞	德國
訪問	100%	100%
註冊	98	100
同班	95	98
一起唸書	82	88
住在同一棟宿舍	71	83
室友	50	76
約會	42	64

結果顯示，大一新鮮人對於奈及利亞的學生感覺比對德國的學生還遠。幾乎所有人對於外國學生來訪、註冊成為正式學生、以及一同修課等等都覺得自在。但是當人際接觸機會增加時，距離的感覺也會隨之增加，尤其是如果接觸的情形牽涉到個人的居住環境、或者是那些並未與教室有直接關聯的活動時。

中，「好」在上面、「壞」在下面，「主動」在左邊、「被動」在右邊，「強」離觀看者較遠、「弱」則離觀看者較近。

5. *古特曼量表法*　**古特曼量表指數**（Guttman scaling index）也被稱作*累積量表*（*cumulative scaling*），此種量表與之前所提過的幾種量表或是指數都不同，因爲我們是在蒐集完資料之後，才使用此種量表來評估資料，而這也意味著我們必須依據古特曼量表法的技術來進行研究設計。古特曼（Louis Guttman）在 1940 年代發展出此種量表，以決定在一組指標當中是否有結構性的關係。他想要知道的是，在關於一個議題的多重指

古特曼量表法　一種研究者在蒐集完資料之後才使用的量數，其目的在於揭露出許多回應之間是否存在著一種等級的模式，也就是那些在「較高等級」作出回應的人，通常也會在「較低等級」作出回應。

範例　7.9

語意差別的範例

請閱讀下列的每對形容詞，然後在最接近你第一印象感覺的空格內畫上記號。所有的答案都沒有對錯之分。

你對於離婚這個想法有什麼感覺？

差	____	x	____	____	____	____	____	____	____	好
深刻	____	____	____	____	____	____	____	x	____	淺
弱	____	____	x	____	____	____	____	____	____	強
公平	____	____	____	____	____	____	____	x	____	不公平
安靜	____	____	____	____	____	____	____	____	x	大聲
現代	x	____	____	____	____	____	____	____	____	傳統
簡單	____	____	____	____	____	x	____	____	____	複雜
快速	____	x	____	____	____	____	____	____	____	緩慢
骯髒	____	x	____	____	____	____	____	____	____	乾淨

你對於婚姻這個想法有什麼樣的感覺？

差	____	____	____	____	____	____	____	____	x	好
深刻	____	x	____	____	____	____	____	____	____	淺
弱	____	____	____	____	____	____	____	x	____	強
公平	____	x	____	____	____	____	____	____	____	不公平
安靜	____	____	x	____	____	____	____	____	____	大聲
現代	____	____	____	____	____	____	____	____	x	傳統
簡單	____	____	____	____	____	x	____	____	____	複雜
快速	____	____	____	____	____	____	____	x	____	緩慢
骯髒	____	____	____	____	____	____	x	____	____	乾淨

標背後，是否有一個潛在的面向或者是累積的強度。[24]

在使用古特曼量表時，我們必須先從測量一組指標或者是項目開始。這些指標或是項目可以是問卷上的題目、投票、或者是觀察到的特徵。我們通常會以一種簡單的「是／否」或者是「出現／未出現」，來測量三到二十個指標。我們會選擇一些我們相信其中具有邏輯關係的項目，然後我們會把結果放入一個古特曼量表中，接著決定在這些項目之中是否有一種等級的模式。

在我們擁有資料之後，我們就可以考慮所有回應的可能結合方式。例如，我們有三個項目：一個兒童是否知道（1）她的年齡，（2）她的電話號碼，以及（3）三個當選的地方政治人物。如此一來，這個小女孩可能只知道她的年齡但是不知道其他的答案、或者是知道所有問題的答案、或者是只有知道她的年齡以及電話號碼等等。這三個項目的答案有八種可能的組合方式、或者是回應的模式，其範圍包括了三者都不知道、到三者通通知道之間的各種組方式。有一種數學方法可以計算組合的數量（例如：23），但是你也可以把三個問題的答案的所有組合方式都寫下來，並且看看這八種可能的組合方式是哪些。

古特曼量表的一種應用被稱爲*量度圖分析*（*scalogram analysis*），此種分析可以

讓我們去測試在資料中是否有一種等級的關係。我們可以把反應的模式區分爲量度的項目以及誤差（或者是非量度的）兩種類別。在兒童知識這個例子中的量度模式可能如下：不知道任何一項、只知道年齡、只知道年齡與電話號碼、以及三者都知道。其他所有的答案組合（例如，知道政治領袖的名字、但是卻不知道自己的年齡）在邏輯上是可能的，但卻是無法量度的。如果我們發現了等級關係，那麼大多數的答案就會符合可以量度的模式。如果存在著一種等級的關係，那麼這些項目就是可以量度的、或者是能夠符合古特曼量表的。同意等級較高項目的人數會比較少，但是所有人都會同意較低等級的項目，不過此種順序不能夠反過來。換句話說，較高等級的項目是建立在中度等級的項目之上，而中度等級的項目則是建立在較低等級的項目之上。

統計程序可以指出各種項目符合預期的等級模式的程度。此種統計程序會產生一個範圍從 0 到 100 的係數。零分表示一種沒有等級結構的隨機模式，而 100 分的分數則表示所有的反應通通都符合等級模式。也有人提出測量可量度性的替代統計技術。[25]（請參閱範例 7.10 古特曼量表的範例）

柯勞格與索耶爾（Clogg and Sawyer, 1981）曾經利用古特曼量表法來研究美國民眾對於墮胎的態度，他們檢視了在哪些情境下人們會認爲墮胎是可以接受的（例如，會危及媽媽健康的情境、因爲強暴而導致懷孕的情境等等），而且他們發現有84.2%的回應符合量表的反應模式。

結論

本章討論了測量的原則以及過程。測量的核心部分是我們進行概念化的方式——或者是我們精鍊或闡明想法以成爲概念性定義的方法，以及把概念性變項操作化爲特定量數的方法——或者是發展出一些程序來連結概念性定義以及實徵的實在。依據一項研究主要是質化或者是量化研究，我們處理這些過程的方式也會有所差異。在量化研究中，我們通常採取比較偏向演繹的路徑，而在質化研究裡，則比較偏向歸納的路徑。儘管如此，這兩種研究取向所追求的是完全相同的目標，也就是在抽象觀念以及實徵資料之間建立起清楚明白的連結。

本章也討論了信度以及效度的原則。*信度*所指的是一個量數的可靠度，而*效度*所指的則是量數的真實性、或者是構念與資料之間的符合程度。在量化以及質化研究中，我們都試著要以一種一致的方式來進行測量，並且也力求緊密地結合抽象的觀念以及實徵的社會世界。除此之外，測量的原則也被應用於量化研究當中，以建立指數以及量表。本章也討論了研究者們目前所使用的一些主要的量表。

除了信度與效度這兩個核心觀念之外，我們現在也已經知道了優質測量的原則：爲概念創造清楚的定義、使用多重的指標、並且在適當的時機將資料進行加權以及標準化。這些原則適用於所有的研究領域（例如：家庭、犯罪學、不平等、種族關係等等），而且也橫跨許多研究技術（例如：實驗、調查）。

範例 7.10

古特曼量表的範例

克羅札特（Crozat, 1998）檢視了大眾對於各種政治抗議形式的反應。他查閱了一份調查資料，該份資料的內容是在 1974 以及 1990 年時，英國、德國、義大利、荷蘭、以及美國等各個國家的民眾對於各種抗議形式的接受程度。他發現大眾的接受程度的模式形成一個古特曼量表。換句話說，那些能夠接受較為激烈的抗議形式（例如：罷工與靜坐）的人，幾乎都能夠接受較為溫和的抗議形式（例如：請願或者是示威），但是接受溫和形式的人並不全然能夠接受較為激烈的形式。除了展現出古特曼量表的用處之外，克羅札特也發現到，不同國家的民眾對抗議的看法相當類似，而且古特曼量表的可量度性程度也會隨著時間而增加。因此，對於抗議活動接受程度的模式在兩個時段裡都可以套用古特曼量表，但是在 1990 年時的模式比在 1974 年時的模式更加接近於古特曼量表的模式。

	抗議的形式				
	請願	*示威*	*抵制*	*罷工*	*靜坐*
古特曼模式					
	否	否	否	否	否
	是	否	否	否	否
	是	是	否	否	否
	是	是	是	否	否
	是	是	是	是	否
	是	是	是	是	是
其他模式（僅作為例子）					
	否	是	否	是	否
	是	否	是	是	否
	是	否	是	否	否
	否	是	是	否	否
	是	否	否	是	是

正如同你很可能已經開始瞭解到的，一個健全的研究計畫意味著在每一個研究階段都必須力求完善。即使研究中其他階段的進行方式完美無缺，但是只要在其中任何一個階段犯下嚴重的錯誤或者是漫不經心，就可能會對研究結果造成無法挽回的傷害。

問題回顧

1. 測量的三個基本部分是什麼？這三者之間如何彼此配合？
2. 信度與效度之間有何差異？兩者之間如何彼此互補？
3. 有哪些可以改善測量信度的方法？
4. 不同的測量尺度之間有何差異？
5. 輻合效度、內容效度、以及同時效度之間有何差異？你是否可以同時達到這三種效度？請解釋你的答案。

6. 爲什麼多重指標通常會比單獨一個指標好？
7. 量表的邏輯以及指數的邏輯之間有何差異？
8. 爲什麼單向度性是量表的一個重要特質？
9. 對指數進行加權的優點以及缺點爲何？
10. 標準化如何幫助研究者更爲容易進行比較？

注解

1. Duncan (1984:220–239) presented cautions from a positivist approach on the issue of measuring anything.
2. The terms *concept*, *construct*, and *idea* are used more or less interchangeably, but their meanings have some differences. An *idea* is any mental image, belief, or impression. It refers to any vague impression, opinion, or thought. A *concept* is a thought, a general notion, or a generalized idea about a class of objects. A *construct* is a thought that is systematically put together, an orderly arrangement of ideas, facts, and impressions. The term *construct* is used here because its emphasis is on taking vague concepts and turning them into systematically organized ideas.
3. See Grinnell (1987:5–18) for further discussion.
4. See Blalock (1982:25–27) and Costner (1985) on the rules of correspondence or the auxiliary theories that connect an abstract concept with empirical indicators. Also see Zeller and Carmines (1980:5) for a diagram that illustrates the place of the rules in the measurement process. In his presidential address to the American Sociological Association in 1979, Hubert Blalock (1979a:882) said, "I believe that the most serious and important problems that require our immediate and concerted attention are those of conceptualization and measurement."
5. See Bailey (1984, 1986) for a discussion of the three levels.
6. See Bohrnstedt (1992a,b) and Carmines and Zeller (1979) for discussions of reliability and various types of it.
7. See Sullivan and Feldman (1979) on multiple indicators. A more technical discussion can be found in Herting (1985), Herting and Costner (1985), and Scott (1968).
8. See Carmines and Zeller (1979:17). For a discussion of the many types of validity, see Brinberg and McGrath (1982).
9. The epistemic correlation is discussed in Costner (1985) and in Zeller and Carmines (1980:50–51, 137–139).
10. Kidder (1982) discussed the issue of disagreements over face validity, such as acceptance of a measure's meaning by the scientific community but not the subjects being studied.
11. This was adapted from Carmines and Zeller (1979:20–21).
12. For a discussion of types of criterion validity, see Carmines and Zeller (1979:17–19) and Fiske (1982) for construct validity.
13. See Cook and Campbell (1979) for elaboration.
14. See Borgatta and Bohrnstedt (1980) and Duncan (1984:119–155) for a discussion and critique of the topic of levels of measurement.
15. Johnson and Creech (1983) examined the measurement errors that occur when variables that are conceptualized as continuous are operationalized in a series of ordinal categories. They argued that errors are not serious if more than four categories and large samples are used.
16. For compilations of indexes and scales used in social research, see Brodsky and Smitherman (1983), Miller (1991), Robinson and colleagues (1972), Robinson and Shaver (1969), and Schuessler (1982).
17. For a discussion of weighted and unweighted index scores, see Nunnally (1978:534).
18. Feeling thermometers are discussed in Wilcox and associates (1989).
19. For more information on Likert scales, see Anderson and associates (1983:252–255), Converse (1987:72–75), McIver and Carmines (1981:22–38), and Spector (1992).
20. Some researchers treat Likert scales as interval-level measures, but there is disagreement on this issue.

Statistically, whether the Likert scale has at least five response categories and an approximately even proportion of people's answer in each category makes little difference.

21. McIver and Carmines (1981:16–21) have an excellent discussion of Thurstone scaling. Also see discussions in Anderson and colleagues (1983:248–252), Converse (1987:66–77), and Edwards (1957). The example used here is partially borrowed from Churchill (1983:249–254), who described the formula for scoring Thurstone scaling.

22. The social distance scale is described in Converse (1987:62–69). The most complete discussion can be found in Bogardus (1959).

23. The semantic differential is discussed in Nunnally (1978:535–543). Also see Heise (1965, 1970) on the analysis of scaled data.

24. See Guttman (1950).

25. See Bailey (1987:349–351) for a discussion of an improved method for determining scalability called *minimal marginal reproducibility*. Guttman scaling can involve more than yes/no choices and a large number of items, but the complexity increases quickly. A more elaborate discussion of Guttman scaling can be found in Anderson and associates (1983:256–260), Converse (1987:189–195), McIver and Carmines (1981:40–71), and Nunnally (1978:63–66). Clogg and Sawyer (1981) presented alternatives to Guttman scaling.

第 8 章

質化與量化抽樣

進行抽樣的理由
抽樣策略
結論

> 對於任何類型的研究而言，抽樣都是一個主要的問題。我們無法去研究我們有興趣的事物的每一個個案，而且我們也不應該這麼做。每個科學研究都試圖藉由研究幾個例子的方式，以找出能夠應用於某種事物中所有個案的研究發現。就如同我們所說的，此種研究的結果是「可以概括推論的」。
>
> —Howard Becker, *Tricks of the Trade*, p. 67

在《我能夠遵守的承諾》（Promises I Can Keep）這項針對低收入母親所進行的深度研究中，艾丁以及柯法拉斯（Edin and Kefalas, 2005）透過廣泛的質化田野工作、以及對於人口普查資料所進行的量化分析，首先在美國賓州的費城地區找到了八個低收入的社區。每個社區都符合了三個篩選的準則：至少有百分之二十的家戶低於貧窮線；在所有的家戶中至少有百分之二十是單親家庭；而且在社區中必須有大量的黑人、白人、以及西班牙裔的居民。在每個社區裡，艾丁以及柯法拉斯在找到願意接受訪談的低收入母親時，有一半是透過當地的專家（教師、社工人員、公共護士、牧師、公司負責人、以及負責公共住宅計畫的官員）進行轉介的方式，而另一半則是藉由在公共電話亭張貼廣告傳單、或者是在街角親自接觸這些母親們的方式。研究中所有的母親在前一年的收入都低於貧窮線的標準。艾丁以及柯法拉斯試著找到各種不同背景的受訪者：50 名白人、50 名非裔美國人、以及 50 名波多黎各人，而且他們也希望在受訪者中有一半的人年紀大於 25 歲、另一半的人則小於 25 歲。最後，他們總共找到 162 名低收入的母親，其中包括了 52 名白人、63 名非裔美國人、以及 47 名波多黎各人。其中只有 40 名母親的年紀超過 25 歲，但是全體的年齡範圍是從 15 歲到 56 歲。作者們認為「雖然研究的樣本不是隨機的、也不具有代表性，但是異質性卻相當高」（p. 238）。

進行抽樣的理由

當我們在進行抽樣時，我們會挑選一些個案並且詳盡地檢視這些個案，然後利用

我們從這些個案所得到的發現，來瞭解一個更大的個案群體。大多數（但不是全部）的實徵研究都會使用抽樣。根據研究的不同，我們用來進行抽樣的方法也會有所差異。大多數討論抽樣的書籍都會強調在量化研究中使用抽樣，並且包含了一些數學以及量化的範例。在量化研究中，抽樣的主要用途是要得到一個具有代表性的樣本（所謂的**樣本**（sample），是指經過挑選而產生的一小群個案或者是單位），而這個具代表性的樣本，能夠以相當接近的方式複製、或者是代表在一個較大的個案集合體中那些我們所感到興趣的特徵。至於這個較大的個案集合體，則被稱作**母體**（population）。

我們會詳細地檢視樣本的資料，而且如果我們的抽樣方式正確的話，就可以把樣本的結果推論到整個母體。在量化研究中，我們必須使用非常精確的抽樣程序，以取得一個具有代表性的樣本。由於這些程序的基礎來自於數學的機率理論，因此也被稱作是*機率抽樣*（*probability sampling*）。

在大多數的量化研究中，我們想要看到母體中有多少個案符合我們所感到興趣的各種不同類別。例如，我們或許想要知道，在以芝加哥地區所有高中學生所形成的母體中，有多少人符合各種不同的類別（例如，高收入的家庭、單親家庭、非法毒品的使用者、因偏差行爲而被逮捕者、以及具有音樂天份的人等等）。我們在量化研究中使用機率樣本的原因，是因爲機率樣本非常有效率。此種樣本不但能夠節省許多的時間以及成本，而且也具有相當高的正確性。相較於蒐集整個母體資訊的過程，一個妥善執行的機率樣本，只需要花費大約千分之一的成本與時間而已，但是卻能產生幾乎完全相同的結果。假定我們想要去蒐集美國境內被診斷出有糖尿病的一千八百萬名病患的資料。只要我們擁有一個設計完善的一千八百名病患的機率樣本，就能把所得到的資料與結果推論到全部的一千八百萬名病患。

機率樣本可能具有高度的正確性。對於大型的母體而言，相較於試著去取得母體中每一個個案的資料，來自於一個設計完善的、仔細執行的機率樣本的資料，就算不是更爲正確、至少也具有同等的正確性，然而，這個事實卻讓許多人覺得困惑。事實上，當美國政府在西元 2000 年規畫進行人口普查時，所有的社會研究者以及受過統計訓練的科學家們都一致同意，比起試著去計算所有人口的傳統方法，機率抽樣可以產生更爲正確的資料。一份仔細規畫的三萬人的機率樣本，其誤差率是極小的、而且是已知的。如果我們試圖去找出美國三億人口中的每一個人，那麼，除非我們付出極大的努力、並且花費龐大的時間和金錢，否則的話，一定會在無意之間產生系統性誤差。順便一提，美國政府實際上正是以傳統的方法來進行人口普查，不過，採取此種方法其實是爲了政治方面、而非科學方面的原因。

在質化研究中會以不同的方法來進行抽樣，而且抽樣的目的也和量化研究不同。事實上，在質化研究中使用抽樣這個詞反而會造成混淆，因爲這個詞已經和量化研究緊密地連結在一起了（請參閱 Luker, 2008: 101）。在質化研究中，當我們在對於母體的類別作出陳述時，我們進行抽樣的目的，並非

樣本　研究者從一個較大的群體中所挑選出來的一小組個案，並且可以推論到整個母體。

母體　一個指由許多個案所形成的大型群體的抽象概念，研究者可以從其中挑出樣本，並且能夠將樣本所得到的結果推論回到整個母體。

是爲了蒐集一小組在數學上能夠精確地複製整個母體的個案。相反地，我們進行抽樣的真正目的，是爲了辨識出在少數個案中發揮作用的相關類別。在量化抽樣中，我們所挑選的是個案或是單位，然後我們把這些個案或是單位當作是社會世界的面向或者是特質的承載體（carrier）。一個由個案或是單位所組成的樣本，「代替」（"stands in"）了較大的母體。換句話說，我們挑選了一些個案或單位來「代替」其他許多的個案或單位。對比之下，質化樣本的邏輯，則是對社會世界的面向或是特質進行抽樣。我們的樣本的面向或是特質突顯了、或者是「照亮了」（"shine light into"）複雜社會生活中的關鍵面向或者是過程。我們挑選一些樣本，以澄清、洞察、或者是理解在社會世界裡的議題或者是關係。在質化抽樣中，我們的目標是去加深對於一個較大的過程、關係、或是社會場景的理解。一個樣本可以提供我們珍貴的資訊或是新的面向，而這些新的面向能夠突顯、強化或者是豐富了關鍵的特質或是情境。我們進行抽樣的目的，在於獲取新的理論洞見，揭露人們或是社會背景的獨特面向，或者是加深對於複雜的情境、事件、或是關係的理解。在質化研究中，「從被研究的人們當中挑選出研究對象的方式，取決於他們與研究主題之間的關聯性、而不是他們所具有的代表性」（Flick, 1998: 41）。

我們不應該過度強調量化與質化之間的區別。在一些情境裡，一項主要爲量化的研究，也有可能會使用質化的抽樣策略，而且反之亦然。儘管如此，大多數的量化研究還是會使用機率樣本或者是類似的樣本，而大多數的質化研究則是使用非機率的方法、以及不具代表性的抽樣策略。

抽樣策略

我們必須避免犯下兩種可能的抽樣錯誤類型。第一種類型的錯誤，是以草率或者是不適當的方式來進行抽樣，而第二種類型的錯誤，則是選擇了與研究目的不相符的樣本類型。第一種錯誤提醒我們，在進行抽樣的時候必須非常小心謹慎而且有條有理。而爲了避免犯下第二種錯誤，我們所使用的抽樣策略，必須符合於我們特定的研究目的以及資料。抽樣策略可以概略地分成兩種廣泛的類型：能夠正確地代表個案母體的樣本，以及其他類型的樣本。我們主要是在量化研究中使用第一種策略，而在質化研究中則使用另外一種策略。

當目標是創造出具有代表性樣本時的抽樣策略

在一個具有代表性的樣本中，我們的目標是要創造出能夠反映或是代表許多其他個案的樣本資料，因爲我們無法直接地去檢視這些個案。我們可以用兩種方法來進行抽樣。第一種方法是比較受到歡迎、而且也被認爲是代表性樣本的「黃金標準」（"gold standard"）的抽樣方法，也就是*機率抽樣*。此種抽樣方法的基礎，建立在超過一世紀之久的仔細推理以及應用數學之上，而且還包括了在自然科學以及量化社會科學中數以千計的研究。在使用機率抽樣策略時，我們試圖去創造一個正確的代

表性樣本，而且這個樣本具有在數學上可以預測的誤差（也就是已知會「偏離目標」的精確機率）。由於此種抽樣策略有數種不同的次類型，因此會變得有點複雜。在我們詳細討論此種抽樣策略之前，讓我們先來看看第二種類型，也就是產生代表性樣本的較爲簡單的方法：使用非機率的抽樣技術。當我們想要得到一個具有代表性的樣本時，這是一種較不正確的替代方法，然而，當機率抽樣無法進行、成本過於高昂、過於耗費時間、或者是不切實際時，非機率抽樣技術就是可以接受的抽樣方式。

非機率抽樣技術　理想上，當我們想要得到一個具代表性的樣本時，我們應該要使用機率樣本。以要求較不嚴格的替代樣本而言，我們有兩種非機率的替代抽樣方法，分別是便利抽樣以及限額抽樣。在**便利抽樣**（convenience sampling）〔又稱爲*偶然性抽樣*（*accidental sampling*）、*可得性抽樣*（*availability sampling*）、以及*隨意抽樣*（*haphazard sampling*）〕中，我們選擇個案的主要標準，是這些個案容易接觸、方便、或者是立即可得的。此種樣本類型可能適合於一些探索性的初步研究、以及一些質化研究，因爲我們的目的並不是要創造出一個具有代表性的樣本。不幸的是，此種方式經常會產生非常不具代表性的樣本，因此，如果你的目的是要創造出一個能夠代表整個母體的正確樣本時，此種方法是*不被推薦*的。

當我們是基於方便性的理由來選擇個案時，我們的樣本可能會嚴重錯表（misrepresent）整個母體的特質。[1] 既然這種方法如此拙劣、而且所得到的樣本也可能非常不具有代表性，那麼，你或許想要知道，爲什麼還會有人使用這種抽樣方法。使用此種抽樣方法的理由其實很單純：因爲便利抽樣的樣本容易取得、便宜、而且還能很快得到。另外一個理由，或許是因爲人們不知道應該如何取得一個良好的代表性樣本。電視節目中所進行的街頭訪談，就是一種便利抽樣。電視節目中的訪談者帶著攝影機和麥克風走到街上，然後就和幾個方便接受訪談的人們談話。白天走路經過攝影棚的人，並不能代表每一個人。相同地，電視節目的訪談者也傾向於挑選那些他們覺得看起來「正常」的人，而且也會避免去訪談那些沒有吸引力的、有殘疾的、窮困的、年邁的、或者是口齒不清的人。隨意抽樣的其他例子，則包括了報紙請讀者從報上剪下問卷並且寄回報社、網站要求使用者點擊一個選項、或者是電視節目要求觀眾打電話到節目中去表達自己的選擇等等。上述的這些例子也許具有娛樂方面的價值，但是卻很容易產生具有高度誤導性的資料，即使有爲數眾多的人們作出回應，此種資料仍然無法代表整個母體。

或許你會覺得很好奇，到底是什麼因素使得上述的樣本不具有代表性。例如，假設你想要瞭解 XYZ 城市裡的每一個人，該市的人口爲一百萬，其中只有一些人會去閱讀報紙、瀏覽某個網站、或者是收看某個電視節目。除此之外，並不是每個會閱讀報紙、瀏覽某個網站、或者是收看某個電視節目的人對於某個議題都同樣地感到興趣。有些人可能會作出回應，而且人數可能很多（例如，五千人），但是這些人都是經過自我選擇的。我們無法把這些自我選擇的人們正確地推論到整個母體。在母體當中有許多人並沒有閱讀報

便利抽樣　一種非隨機的抽樣方式。在使用此種抽樣方式時，研究者所選擇的是他或她碰巧遇上的任何人。

紙、瀏覽特定的網站、或者是收看某些電視節目的習慣，而且就算他們有這樣的習慣，他們也可能沒有興趣或者是動機去參與這些活動。關於代表性樣本的兩個關鍵性概念是：（1）自我選擇會產生一個不具代表性的樣本，以及（2）僅僅因爲某個樣本的數目龐大，並不保證這個樣本就一定會具有代表性。

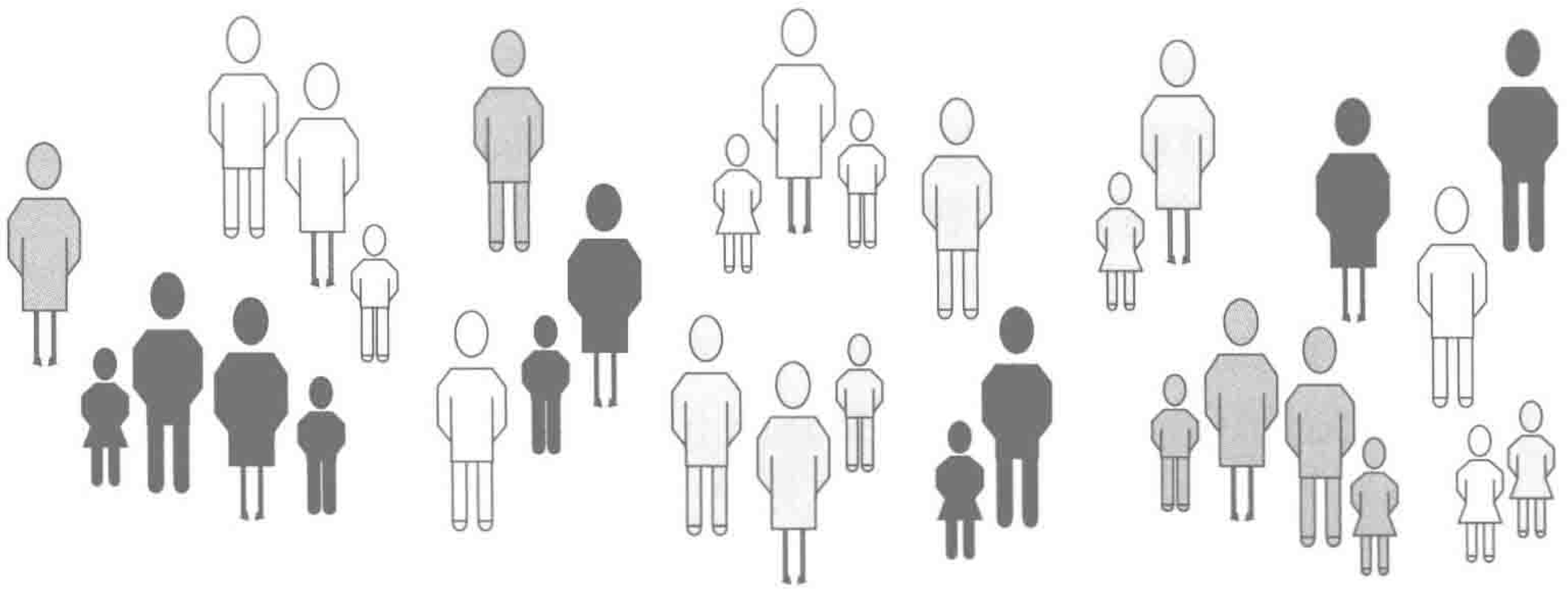
對於許多目的而言，設計完善的**限額抽樣**（quota sampling）是一種可以接受的非機率替代方法，能夠產生半代表性的樣本。[2] 在進行限額抽樣時，我們首先要在進行抽樣的母體中找出相關的類別，以捕捉到在不同單位之間的多樣性（例如，男性與女性；或者是年齡低於 30 歲、介於 30 歲到 60 歲之間、以及超過 60 歲）。接下來，我們還得決定在每個類別裡要有多少個案，這也就是我們所謂的「限額」。因此，我們在一開始的時候，就先把樣本中不同類別的個案數目固定下來了。

讓我們回到剛才從 XYZ 城市的居民裡進行抽樣的例子。在一個 180 人的樣本裡，你選擇了 30 歲以下的男、女居民各 25 人，30 歲到 60 歲的男、女居民各 50 人，而超過 60 歲的男、女居民則是各 15 人。雖然這是作爲母體多樣性的一個起點，但是卻難以正確地代表所有可能的母體特質（請參閱圖 8.1）。話雖如此，限額抽樣仍然確保了一個樣本具有一定程度的多樣性。對比之下，在便利抽樣裡，一個樣本中的每一個人可能都是相同年齡、相同性別、或者是來自於相同背景的人。在本章一開始《我能夠遵守的承諾》研究裡所提到的抽樣方式，就是使用了限額抽樣（同時也請參閱範例 8.1 限額樣本）。

限額抽樣相對較爲容易。我的學生就使用限額抽樣的方式，對於大學部的學生進行了一次意見調查。我們使用了三種限額類別：性別、年級、以及少數／多數族群的身分，以及一種便利的挑選方法（換句話說，一名學生訪談者對任何一名在圖書館裡、教室裡、以及自助餐廳裡的學生進行訪談）。我們事先決定了每個限額類別裡要訪談的學生數目：50%的男性以及 50%的女性；35%的大一新生、25%的大二學生、20%的大三學生、以及 20%的大四學生；10%的少數族群學生、以及 90%的多數族群學生。我們是根據大學官方記錄裡的學生大約比例，來決定每個類別的學生比例。在這項研究裡，一名學生訪談者會先接近某個人，確認這個人是否是一個學生，並且確認這個人的性別、年級、以及少數／多數族群的身分。如果這個人符合某一個尚未額滿的類別（例如，找出五個具有少數族群身分的男性大一新生），這個人就會被包含在樣本裡，而這名訪談者就會繼續接著詢問一些調查的問題。如果這個人不符合限額類別，訪談者就不會詢問任何調查的問題，而是會很快地謝謝這個人，並且繼續去找下一個人。

限額抽樣有三個缺點。首先，此種抽樣法只能捕捉到所有母體多樣性的一些面向而已（例如：性別以及年齡），而且會忽視了其他的面向（例如：種族、在城市中的居住地區、或是收入水準等等）。其次，在每個類別中的固定個案數目，並不一定能夠正確地反映該類別的個案在整個母體中所佔的比例。例如，或許在城市的居民當中，有 20%的人是

限額抽樣　一種非隨機的抽樣方式，在使用此種抽樣方式時，研究者先確定幾個廣泛的類別，以把個案或者是人們放入這些類別，然後研究者會去挑選一些個人，以達到每個類別中預先決定的個案數目。



註：不同的個陰影代表各種不同的膚色。

圖 8.1　限額抽樣

超過 60 歲的，但是在限額中卻只佔了 10%而已。最後，當我們在爲每一個限額類別挑選個案時，所使用的是便利抽樣的挑選方式。舉例而言，我們會把一開始遇到的 25 名年紀在三十歲以下的男性包含在樣本中，即使這 25 個人都是剛剛參加完一場關於財務投資研討會的高收入白人律師。沒有任何事物能夠阻止我們只有去找那些看起來「友善」、而且想要讓我們去挑選他們的人來當作我們的樣本。

機率抽樣技術　在創造出具有代表性的樣本時，機率抽樣是「黃金標準」。機率抽樣具有專門的詞彙，除非你曾學習過這些詞彙，否則將會難以理解這些詞彙的意義。因此，我們接下來將會先回顧一些與機率抽樣有關的詞彙。

機率抽樣的語言　你從一個個案或是單位的較大集合體中抽取出一個樣本，而每一個個案或是單位就是你的**抽樣元素**（sampling element），這就是分析單位、或者是在母體中的一個個案。抽樣元素可以是一個人、一個家庭、一個社區、一個國家、一個組織、一份書面文件、一個象徵性的訊息（電視廣告、展示某個旗幟）、或者是一個社會行動（例如，逮捕、離婚、或是一個吻）。

抽樣元素　用來稱呼被抽樣出來的一個個案或者是一個單位的名稱。

範例　8.1

限額樣本

有兩項研究可以用來說明在量化研究中使用限額抽樣的不同方式。在其中一項研究中，麥可馬宏、麥可艾拉尼、以及艾德加（McMahon, McAlaney, and Edgar, 2007）想要檢視英國民眾對於狂飲（binge drinking）的看法。三位作者注意到，大多數之前的研究都是針對年輕的成年人所進行的研究，而且一些抑止狂飲的運動也都成效不彰。作者們想要得知整個成年人母體對於狂飲的觀感，因此，他們進行了一項調查，其中詢問了人們如何去定義狂飲、他們對於狂飲行為的關切程度、以及造成狂飲的原因和解決之道。作者們結合了限額抽樣以及另一種抽樣方法，對一個城市裡（蘇格蘭的因弗克萊德行政區）（Inverclyde, Scotland）的 586 名成人進行了訪談。在限額抽樣方面，訪談者在一座購物中心的周圍街道上去接近可能的受訪者，並且邀請他們參與這項調查。限額的方式是針對不同性別以及六種年齡類別找到相同的人數。另一種抽樣方法，則是在數個低收入社區裡挨家挨戶地進行訪談。作者們得知了人們對於狂飲行為看法的珍貴資訊，而這些資訊則是來自於整個城市裡橫跨不同年齡團體的男女兩性的成人。作者們發現，人們對於狂飲的定義、以及對於一種「錯誤共識效應」（"false consensus effect"）（此種效應所指的是一小群酗酒程度最高的人認為他們的行為是正常的、而且能夠被社會大眾接受）的支持程度方面有極大的差異。然而，由於研究的樣本不具有代表性，因此，對於大眾之間狂飲行為的普遍程度、以及對於狂飲行為看法的相關研究發現，可能無法反映出這個城市整個母體的行為或者是看法。

另外一個研究，是由畢安、布瑞格、戴維斯、以及格拉斯基維茲（Bian, Breiger, Davis, and Galaskiewicz, 2005）在中國所進行的研究，他們所使用的是一種針對特定目標的限額抽樣方式。作者們想要知道的是在幾個大型的中國城市裡，不同社會階級的人們之間的社會網絡以及社會連結（例如：朋友、家人）的差異。他們在四個中國最大的都會區（上海、深圳、天津、以及武漢）挑選了幾個家庭，在每一個都會區中找到一組的社區，然後在每個都會區中抽樣 100 人來進行研究。他們列出了一張有十三種職業頭銜的清單，以代表中國階級系統中的完整範圍、以及在這四個都市中的百分之八十八的工作人口。他們的限額方式，是在每個城市都找到相同的人數，並且在這十三種職業類別中都找到足夠的家戶數，以進行仔細的分析。因此，雖然大約只有百分之四的人擔任經理級的職位，但是在樣本中卻有將近百分之十的人是經理，而在所有人口中約有百分之四十的人是工廠的工人，但是在樣本中卻只有大約百分之十的人是工廠的工人。這項研究的目標是要去測試一些假設，這些假設是關於一個家庭究竟是與其他相似的、或者是相異的社會階級的人產生社會連結。作者們要求那些參與研究的家庭以書面方式記錄下與他人進行社交拜訪（不論是親自見面或者是透過電話）的日誌，並且記錄訪客的職業。這個過程持續了一年，而且研究者每隔三個月就訪談研究對象一次。研究者在該項研究中最關心的事物，就是要針對各種不同社會階級的社會網絡模式進行比較。例如，經理級的人們是否只與其他經理級的人們互相往來？還是會與來自不同階級的人們互動？工廠裡的工人是否只和其他工廠的工人以及各種較低等職業的人們往來，並且避開那些職業等級較高的人？由於研究目標是在於比較不同社會階級的社會網絡模式，而不是找到一個能夠描述整體中國人的代表性樣本，因此，這項研究在運用限額抽樣方面是高度有效的。

較大的集合體就是母體，但是有時候也會被稱爲*母群*（*universe*）（在第六章中已經加以定義過了）。爲了要定義母體，你必須詳細說明其中的各種元素，並且確認母體在地理上以及時間上的界限，以及任何其他相關的界限。

大多數針對整體美國民眾的大型樣本所進行的機率研究，都會有數個界限。這些研究都把居住在美國本土大陸四十八州裡、已滿十八歲的成年人包括在樣本中，

並且排除被收容在機構裡的人群（例如，那些住在醫院裡、安養之家、軍事基地、監獄與看守所、無家可歸與受害婦女收容所、以及大學宿舍裡的人）。刻意忽略住在阿拉斯加、夏威夷、以及波多黎各的美國公民，並且排除被收容在機構裡的人們，可能會使得統計數據完全走樣——例如，如果把數百萬名被關在監獄裡的人們也加入計算的話，失業率可能會變得更高（請參閱 Western and Pettit, 2005）。許多研究的樣本只包含會說英語的人，但是截至 2007 爲止，有大約百分之五的美國家庭是「在語言方面被隔絕的」（"linguistically isolated"）（也就是家中沒有任何一個 14 歲以上的人可以說非常流利的英語）（U.S. Census Bureau, 2007）。

在取得機率樣本時，我們必須先從母體開始，但是*母體*卻是一個抽象的概念。我們必須把母體加以概念化並且更加精確地加以定義，這個過程類似於在測量過程中的概念化（請參閱第七章），例如，像是住在佛羅里達州坦帕地區（Tampa, Florida）的所有人、或者是在內華達州裡所有的大學生。所謂的**目標母體**（target population）所指的就是我們要進行研究的所有特定個案的集合體（例如，2011 年 5 月 15 日當天在坦帕市裡所有年滿 18 歲、而且未被收容於機構內的合法居民；2010 年 10 月在內華達州內所有合格的兩年或是四年的高等教育機構中註冊的全職學生）。就某些方面而言，目標母體就類似於我們在測量過程中所使用的概念性定義一樣。

由於母體處於不斷變動的情況中，所以我們必須設定時間界限。例如，在一座城市裡，隨時都會有人死去、登上飛機或是下飛機、或者是開車穿越過城市邊界。那麼，我們應該把哪些人算進來呢？當我們把時間固定下來時，是否要把一個剛好出城去度假的長期居民排除在外？母體（例如，2011 年 3 月 1 日凌晨 12 點 01 分時，在威斯康辛州的密爾瓦基市裡所有年滿 18 歲的人）是一個抽象的概念。母體存在於我們的心中，但是卻非常難以具體地加以精確界定（請參閱範例 8.2 母體的範例）。

在我們將母體概念化之後，我們必須針對抽象的母體提出一個操作性定義，而這個過程也同樣類似於測量過程中的操作化（請參閱第七章）。我們把這個抽象想法轉變成爲一個在實徵方面的具體清單，而這份清單則是非常近似於母體中的所有元素。這份清單就是我們的**抽樣架構**（sampling frame）。

抽樣架構有許多種不同的類別，像是電話簿、報稅記錄、以及駕照記錄等等。把某個母體中所有的元素都列出來，聽起來好像是一件很容易的事，但是這個過程通常是非常困難的，這是因爲母體中的元素通常沒有一份正確的、最新的清單。

一個好的抽樣架構對於正確的抽樣是非常重要的。在抽樣架構以及概念上定義的母體之間，若是有任何搭配不當的地方，都有可能會造成誤差。正如同在變項的理論性定義和操作性定義之間的錯誤搭配會減弱測量效度一樣，在抽象母體以及抽樣架構之間的錯誤搭配，也會對我們的抽樣效度造成傷害。在抽樣的歷史上最廣爲人知的一個案例，就和抽樣架構的議題有關。[3]（請參閱資訊箱 8.1 抽樣架

目標母體　由許多個案所組成的、具體詳細說明的大型集合體，研究者從這個集合體中取得樣本，而且可以依據從樣本所得到的結果，來對這個集合體進行推論。

抽樣架構　一份列出母體中所有個案的清單、或是最接近這些個案的清單。

範例　8.2

母體的範例

1. 所有於 2009 年 12 月 2 日居住在澳洲、年滿 16 歲以上的人，且未被監禁於監獄、精神病院、或者是類似的機構中
2. 2005 年 7 月，在加拿大安大略省雇用員工人數超過 100 人的所有商業機構。
3. 在 1988 年 8 月 1 日到 1993 年 7 月 31 日之間，所有住進紐澤西州裡公立或私立醫院的病患。
4. 在 2004 年 11 月 1 日到 11 月 25 日之間，於美國東部標準時間早上 7 點到晚上 11 點在美國三大電視網中所播放的所有商業廣告。
5. 從 1960 年 1 月 1 日到現在現在之間取得醫學學位、並且目前在美國執業的所有醫生。
6. 2004 年間，在溫哥華、英屬哥倫比亞、或是華盛頓州西雅圖地區裡所有男性黑人海洛因毒癮犯。

構以及抽樣的歷史）

假設我們的母體是在 2010 年間所有居住在美國太平洋海岸區域內的所有成年居民。我們聯絡了交通部門，以取得在加州、奧勒岡州、以及華盛頓州內擁有駕照的所有人們的名單。我們知道在母體中有些人是沒有駕照的，而且有些人是無照駕駛或者是根本不開車。即使我們經常去更新這些持有駕照者的清單，但是隨著人們搬進或是搬出某個州，這些清單也會很快就過時了。這個例子顯示出，在我們使用像是駕照之類的官方記錄來當作抽樣架構之前，我們必須先瞭解官員們是如何製作這些記錄的。當奧勒岡州施行了一項新的規定，要求人們必須提供自己的社會安全號碼才能取得駕照時，申請駕照的人數就下降了百分之十（從 2007 年的三個月之內有 105,000 人提出申請，下降到 2008 年同樣的三個月內只有 93,000 人提出申請）。也就是說，有數千個人就這樣憑空從官方記錄中消失了。我們也可以試著去使用報稅記錄，但是，並不是每個人都有繳稅。有些人會刻意逃漏稅、有些人因爲沒有收入所以不必繳稅、有些人是已經過世了或是還沒開始報稅、還有些人是在上次報完稅之後才遷入或是遷出這個地區的。選民投票記錄則會排除掉母體中大約一半的人。在美國只有大約 53%到 77%的合格選民會去登記投票（Table 401, Statistical Abstract of the United States, 2009）。電話簿的情況就更差強人意了。許多人沒有被列在電話簿裡，有些人的電話號碼則沒有出現在電話簿中，而且有些人最近才剛搬家。除了少數的例外情況（例如，在一間大學內註冊入學的所有學生的名單）之外，要得到一個完全正確無誤的抽樣架構是相當困難的。抽樣架構可能會包含了那些在母體之外的個案（例如，電話簿裡可能會列出那些其實已經遷出者的電話號碼）、或者是遺漏了原本應該被包含在母體內的個案（例如，那些沒有電話的人）（請參閱範例 8.3 抽樣架構）。

樣本大小與目標母體數量的比率就叫作**抽樣比**（sampling ratio）。如果目標母體有 5 萬人，而樣本爲 150 人，則抽樣比就是 150/50,000 = 0.003，或者是 0.3%。如果目標母體是 500，而樣本大小爲 100 的話，則抽樣比就是 100/500 = 0.20，或者是 20%。一般而言，我們會使用在抽樣架構裡的抽樣元素數目，來當作我們對於目標母體大小的最佳估計值。

抽樣比　把樣本中的個案數目除以整個母體或者是抽樣架構中的個案總數所得到的值，或者是樣本在母體中所佔的比例。

資訊箱 8.3　抽樣架構以及抽樣的歷史

在抽樣歷史上的一個著名案例，可以用來說明限額抽樣以及抽樣架構的限制。《文學文摘》（*Literary Digest*）是一本美國的知名雜誌，這家雜誌社曾經在 1920、1924、1928、以及 1932 年的美國總統大選之前，寄了明信片給許多美國的民眾。這本雜誌是從汽車登記資料以及電話簿中挑出人們的姓名來當作樣本的，而接到明信片的民眾可以在明信片上註明自己將會投票給哪位候選人，然後再把明信片寄回去。這本雜誌正確地預測了這四次大選的結果。這本雜誌成功預測選舉結果的故事廣為人知，而在 1936 年時，這本雜誌將樣本人數從一百萬人增加到一千萬人。在接獲明信片的民眾當中，有二百四十萬人將明信片寄回雜誌社。而雜誌則根據這些明信片的結果，預測藍登（Alf Landon）將會大勝小羅斯福（Franklin D. Roosevelt）。但是，這一次《文學文摘》的預測失準了，選舉結果是小羅斯福獲得了壓倒性的勝利。另一個由蓋洛普所進行的五萬人隨機樣本則相當正確，其預測結果和真正的選舉結果之間僅有大約百分之一的差距而已。

有數個原因造成此次預測錯誤，但其中最為關鍵的原因則是抽樣的錯誤。雖然這本雜誌挑選了非常大的樣本，但是所採用的抽樣架構並未正確地代表目標母體（也就是所有的選民），因為在其中排除了沒有電話或者是汽車的人，而在 1936 年時，這樣的人在整個母體中佔了相當大的比例。此種抽樣架構排除了母體中高達百分之六十五的人，尤其是在選民中傾向於投票給小羅斯福的人（收入較低者）。這本雜誌在之前的選舉中之所以能夠準確地預測結果，是因為高收入和低收人的人們在投票方式方面並沒有太大的差異。除此之外，在美國經濟大蕭條發生之前的選舉當中，也有較多的低收入者能夠負擔得起電話或是汽車。

《文學文摘》所犯下的錯誤教導我們兩個教訓。第一，正確的抽樣架構是非常重要的。第二，樣本大小的重要性，比不上樣本能夠正確地代表母體的重要性。相較於一個高達一千萬或是五千萬、但是卻不具代表性的樣本，一個就算是僅有五萬人的代表性樣本，就能夠對美國整體民眾作出更為正確的預測。

範例　8.3

抽樣架構

一項由史密斯、密契爾、阿特波、以及李德（Smith, Mitchell, Attebo, and Leeder, 1997）於澳洲所進行的研究，顯示了不同的抽樣架構會對樣本造成什麼樣的影響。該研究的作者們檢視了某個特定的郵遞區號範圍內的 2,557 名年滿 49 歲以上的成人，而這些樣本的取得方式，是由作者們挨家挨戶去進行調查所得到的。在該郵遞區號所包含的全部地址當中，作者們接觸了 80.9%的民眾，而在被接觸到的民眾當中，有 87.9%的人都回答了調查的問題，作者們搜尋了電話簿以及選舉人名冊以找到每一個人。電話簿中列出了該地區 82.2%的人，而選舉人名冊列出了 84.3%的人。然而，年紀較輕的民眾、無殼蝸牛、以及出生地不在澳洲的人們，則非常有可能並未被包括在上述的兩種抽樣架構之內。除此之外，在電話簿中可能不會列出那些職業聲望較高的人們，而選舉人名冊則可能會把未婚人士以及男性排除在名單之外。

除了小型的特殊母體（例如，在某間教室的所有學生）之外，當我們不必去抽樣時，我們會利用一個樣本內的資料來推估較大的母體內的特質。母體的任何一種特質（例如，城市居民中吸煙者的百分比例、所有年滿 21 歲以上的女性的平均身高、相信幽浮的民眾比例等等）就是母體的**母數**（parameter）。母數是母體的真正特性。就大型的母體而言，我們永遠無法非常確定地得知母數爲何，因此我們可以利用樣本資料來推估母數。在樣本中被用來推估母數的資訊，就被稱作**統計值**（statistic）（請參閱圖 8.2）。

母數　由樣本所推估出來的整個母體的特質。

統計值　一個具有多種意義的詞彙，其中包括由樣本所計算得到的關於母數的估計數值。

隨機樣本　利用數學上的隨機方法，像是亂數表或是電腦程式，所產生的樣本，因此在母體中的每一個抽樣元素被挑選為樣本的機率都是完全相等的。

抽樣誤差　樣本偏離母體代表性的程度。

隨機抽樣

在應用數學裡，機率理論所依賴的是隨機過程。*隨機*（*random*）一詞有數種意義。在日常生活裡，這個詞所指的可能是無法預測的、不尋常的、出乎意料的、或者是偶然的。而在數學裡，隨機一詞則有特定的意義：一個沒有任何模式的挑選過程。在數學領域中，隨機過程所指的是每個元素被挑選出來的機率都是完全相同的。只要是真正的隨機過程，我們就可以用數學方法去精確地計算出在許多個案中各種結果出現的機率。

隨機樣本（random samples）能夠產生最有可能真正代表整個母體的樣本，而且隨機樣本也允許我們以統計的方式去計算樣本與母體之間的關係——也就是**抽樣誤差**（sampling error）的大小。而抽樣誤差所指的是在樣本資料以及理想的母體母數之間由於隨機過程而產生的差異。

機率樣本是以隨機挑選過程爲基礎的。隨樣挑選的抽樣過程，會比非隨機挑選的樣本要求更高的精確性、時間、以及努力。正式的數學程序會明確地具體指定要去挑出哪一個人來當作樣本，但是要找到這個特定的人卻有可能是非常困難的事！在抽樣中，隨機並不是指任何人都可以當作樣本，也不是指不必經過思考、或者是

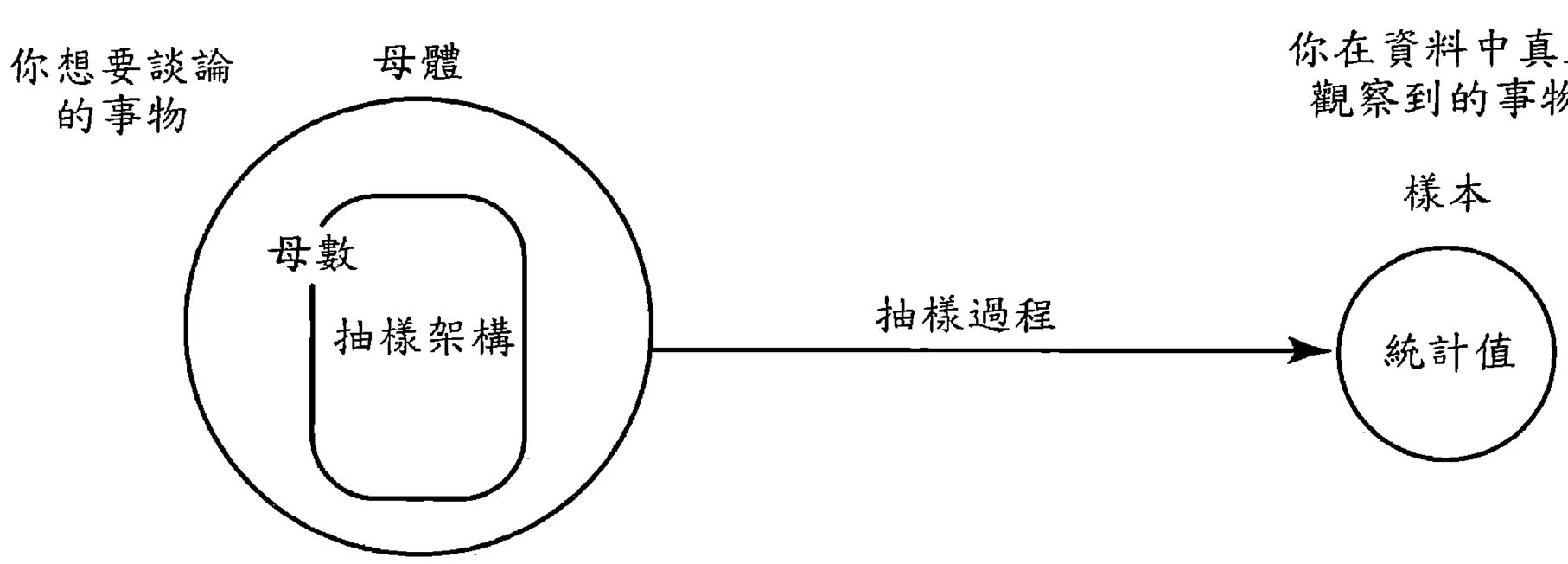


圖 8.2　抽樣邏輯的模型

可以任意指定樣本。舉例而言，如果我們在一項電話調查中使用了真正的隨機抽樣，爲了要聯絡到一個由數學上的隨機過程所指定的特定個人，我們或許必須在一天中的不同時間、以及在不同的日子裡打電話六到七次才能夠找到這個人。[4]

本章並未涵蓋所有關於隨機抽樣的技術性以及統計細節。相反地，我們所討論的是機率抽樣過程運作的基本原則、好的樣本以及壞的樣本之間的差異、如何取得樣本、以及在社會研究中進行抽樣的基本原則。如果你打算以量化研究當作未來的職業，你將會需要更多關於機率的數學以及統計背景。不過，由於篇幅有限，因此本書無法涵蓋這些更爲深入的細節。

隨機抽樣的五種方法

簡單隨機 **簡單隨機樣本**（simple random sample）是所有的機率樣本的基礎模型。在挑選簡單隨機樣本時，我們首先要確認母體以及目標母體，並且確認母體的特定抽樣元素（例如，2011 年 3 月期間在亞歷桑那州普萊斯考特市的所有家庭）。接著，我們提出一個正確的抽樣架構，然後再利用一個真正的隨機過程（稍後會加以討論），從抽樣架構中挑選出元素。除了提出正確的抽樣架構之外，接下來的困難之處，就在於我們必須找到由隨機過程所挑選出來的特定抽樣元素。如果抽樣的元素是一戶家庭，我們或許必須重新拜訪、或者是打五次電話，才能聯絡上這個被挑選出來的特定家庭。

爲了要從抽樣架構中挑選出抽樣元素，我們還必須製作出一份亂數號碼的清單，這份清單將會告訴我們要去挑選哪一個元素來當作樣本。我們所需要的亂數號碼的數目，必須要和抽樣元素的數目相符。亂數號碼的範圍應該從 1（在抽樣架構中的第一個元素）到我們的抽樣架構中的最高數目。如果抽樣架構中列出了 15,000 個家庭，而我們想要從中抽樣 150 個家庭，我們就需要一份有 150 個亂數號碼的清單（也就是從 1 到 15,000 中，利用真正的隨機過程所產生的號碼）。

有兩種主要的方式可以獲得一份亂數表。比較「舊式的」（"old-fashioned"）方法是利用**亂數表**（random-number table）。在大多數關於統計以及研究方法的書裡（包括本書在內，請參閱附錄），你都可以找到亂數表。這些數字都是由真正的隨機過程所產生的，所以任何一個號碼出現在任何一個位置的機率者是完全相等的。目前大多數的人都是利用電腦程式來產生亂數的列表，此種程式不但很容易取得，而且通常也是免費的。

你可能會問，一旦我們從抽樣架構中挑選出了一個元素之後，我們是要把這個元素放回抽樣架構之內、或者是把這個元素與其他元素區隔開來？無限制的隨機抽樣被稱作是「可替代的隨機抽樣」

簡單隨機樣本 一種隨機樣本，在此種樣本中，研究者創造出一個抽樣架構，並且使用一個真正的隨機過程以挑選出個案，如此一來，在母體中的每一個抽樣元素被挑選到的機率都是完全相等的。

亂數表 一張數字的列表，其中數字的排列是沒有任何模式的，而且研究者在挑選個案或者是達到其他的隨機目的時，都可以利用此種列表來進行隨機過程。

（“random sampling with replacement”），換句話說，在抽出一個元素之後，就把這個元素再放回去，這樣一來，這個元素才有機會被再次挑選到。而在沒有替代的簡單隨機抽樣中，我們則會把已經挑選出來成爲樣本的元素「扔掉」（“toss out”）、或者是刻意忽略不管。就社會科學中所有的實用目的而言，幾乎都是使用沒有替代的隨機抽樣過程。

我們可以用一個基本的例子來看出簡單隨機抽樣的邏輯：從一個罐子裡取出彈珠。假設我現在手上有一個裝滿了 5,000 顆彈珠的大罐子，其中有些彈珠是紅色的，而有些則是白色的。這些彈珠就是我的抽樣元素，而這 5,000 顆彈珠就是我的母體（既是目標母體、也是理想母體），而我的樣本數是 100。我不需要抽樣架構，因爲我所處理的只是小型的實體物體而已。而我想要估算的母體母數，則是紅色彈珠在罐子裡所有彈珠中所佔的百分比。

我需要一個隨機過程來挑選出當作樣本的 100 顆彈珠。對於小型的物體而言，這是一件相當容易的事。我只需要閉上我的眼睛、把罐子搖一搖、從罐子裡拿出一顆彈珠，然後把這整個程序重覆一百次就行了。如此一來，我就擁有一個彈珠的隨機樣本。接著，我去計算在樣本中的紅色彈珠的數目，以推估在母體中紅色彈珠相對於白色彈珠所佔的百分比例。此種方式會比直接去計算全部的 5,000 個彈珠容易多了。而計算之後的結果，在我的樣本中有 52 顆白色彈珠以及 48 顆紅色彈珠。

這是否意味著在母體母數裡的紅色彈珠真的剛好就是 48%？可能是、也可能不是。由於隨機抽樣的緣故，我可能沒有抽到特定的樣本。我可以把這一百顆彈珠全部倒回到罐子裡去、把所有的彈珠充分混合、然後再拿出 100 顆彈珠當作第二個隨機樣本，以檢驗我所得到的結果。在第二次的樣本中，共有 49 顆白色彈珠以及 51 顆紅色彈珠。現在問題來了，到底哪一個結果才是正確的？你或許會問，如果從相同的母體中所抽取出來的不同樣本所產生的結果都不同的話，那麼這種隨機抽樣的方式到底有什麼好處呢？我一次又一次地重覆上述的抽樣程序，一直到我抽取了 130 個樣本，而且每個樣本都有 100 顆彈珠爲止（抽樣的結果請參閱表單 8.1）。大多數的人可能會覺得，乾脆把罐子裡的彈珠通通倒出來、然後直接去數這 5,000 顆彈珠，會是比較容易的作法。但是，此處我真正想要瞭解的，其實是抽樣的過程。我所抽出的 130 個不同樣本的結果顯示了一個清楚的模式，也就是最常見的紅色與白色彈珠的混合比例是 50／50。換句話說，接近此種比例的樣本出現的頻率，高於其他混合比例的樣本。因此，母體母數似乎是 50%的白色彈珠以及 50%的紅色彈珠。

數學證明以及實徵測試顯示出，在表單 8.1 中所發現的模式會一直出現。許多不同樣本的組合，就是我的**抽樣分配**（sampling distribution）。抽樣分配所指的是不同樣本的分配情形，可以揭露出來自於許多分次隨機樣本的不同樣本結果出現的頻率。就算樣本數是 1,000，而不是 100；就算彈珠的顏色有 10 種，而不是 2 種；就算母體有 100 顆彈珠或是一千萬顆彈珠，而不是

抽樣分配　由相同的母體所抽取出來的隨機樣本所形成的分配。

5,000 顆彈珠；就算樣本元素是人們、汽車、或者是大學，而不是彈珠，這種模式仍然會繼續出現。事實上，隨著我從母體中抽取越來越多獨立的隨機樣本，這種「鐘形」（“bell-shaped”）的抽樣分配模式會變得越來越明顯。

抽樣分配模式告訴我們，在多次分立的樣本中，真正的母體母數（例如，在前述例子中的 50／50 比例）會比任何其他的結果更常出現。有一些樣本可能會偏離了母體母數，但是此種樣本是比較不常出現的。當我們規畫了像是在圖表 8.1 中的許多隨機樣本時，抽樣分配的形狀看起來就會像是一個常態或者是鐘形曲線。此種曲線在理論上是相當重要的，而且在統計當中幾乎都會使用到這種曲線。在鐘形曲線下方的區域是廣爲人知的，或者，就像在這個例子裡一樣，我們可以很快地計算出我們取得特定數量彈珠的機率。如果真正的母體母數是 50／50 的話，標準的統計圖表可以讓我們知道，在一個隨機樣本中出現 50／50、40／50、或者是其他比例的機率有多少。

數學中的**中央極限定理**（central limit theorem）告訴我們，當抽樣分配中的不同隨機樣本的數目逐漸增加到無窮大時，樣本的模式以及母體母數的模式就會變得更加容易預測。只要隨機樣本的數量夠龐大，抽樣分配就會形成一個常態曲線，而在曲線正中央的地方就是母體母數。

你很可能沒有時間或是精力去抽取許多不同的樣本，而只想取得一個樣本而已。如果你有這種想法的話，你並不是單獨的特例。除非是要去驗證中央極限定理，否則的話，我們幾乎都不會去抽取許多的隨機樣本。我們通常只有抽取一個隨機樣本，但是中央極限定理能夠讓我們從一個樣本去推論到整個母體。雖然中央極限定理是來自於許多樣本的結果，但是卻能夠讓我們去計算某個特定樣本偏離母體母數的機率。不過，在這兒我們將不會去探討計算的方式。

重點在於，隨機抽樣並不保證每個隨機樣本都能完美地代表母體。相反地，隨機抽樣的意思是，大多數的隨機樣本多半會接近母體母數。除此之外，我們也能夠去精確地計算出某一個特定樣本不正確的機率。中央極限定理讓我們能夠去估算某個特定樣本不具代表性的機率、或者是偏離母體母數的程度。中央極限定理也能讓我們去估算出抽樣誤差的大小。當我們在估算抽樣誤差時，是利用從一個樣本所得到的資訊來推估抽樣分配，然後把這個資訊結合中央極限定理以及常態曲線下方區域的相關知識，以建立非常重要的**信賴區間**（confidence intervals）。

信賴區間是一個相當簡單、但是卻非常強而有力的觀念。當電視或者是報紙在報導民意調查結果的時候，你可能會聽到記者說「誤差界限」（“margin of error”）是正負兩個百分點。這就是一種信賴區間，也就是圍繞著一個特定的點的全距（range），而我們則是利用這個特定的點來估算母體母數。

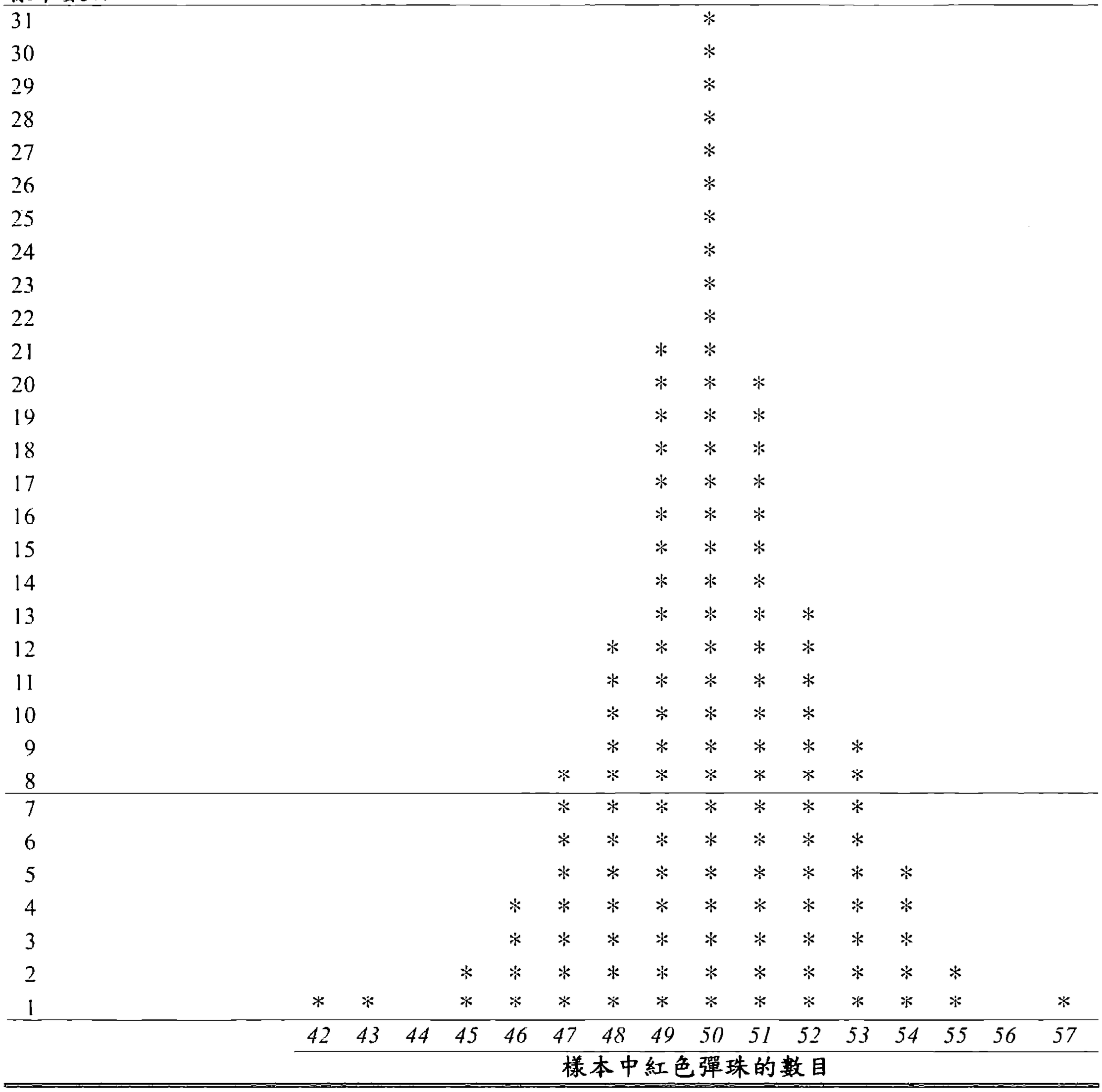
中央極限定理　一種數學的關係，其內容告訴我們，當我們從一個母體中多次抽取隨機樣本時，就會形成常態分配，而一個變項在此種分配曲線的中央部分，就等同於母體母數。

信賴區間　一個數值的範圍，通常會略高以及略低於在一個樣本中所發現的特定數值，而研究者則具有高度的信心，認為母體母數就落在這個數值範圍之內。

圖表 8.1　抽樣分配的範例

紅	白	樣本數目
42	58	1
43	57	1
45	55	2
46	54	4
47	53	8
48	52	12
49	51	21
50	50	31
51	49	20
52	48	13
53	47	9
54	46	5
55	45	2
57	43	1
	Total	130

從一罐有 5,000 顆彈珠的罐子中隨機抽取來的紅色彈珠和白色彈珠的數目，每次抽出 100 個，重覆執行 130 次獨立隨機抽樣。



我們使用全距的原因，是因爲隨機過程的統計是以機率爲基礎的，因而不允許我們去預測一個精確無誤的點，但是卻允許我們以高度的信心水準（例如，95%的信心水準）說，真正的母體母數會落在一個特定的全距之內（也就是信賴區間）。抽樣誤差或者是信賴區間的計算方式，已經超出了此處討論的層次，因此我們將不再進行討論。然而，抽樣分配是一個能夠我們去計算抽樣誤差以及信賴區間的關鍵觀念。因此，我們不能夠說：「這個樣本能夠產生母體母數的完美量數」，但是我們可以說：「我們有 95%的信心，可以確定真正的母體母數與在樣本中所發現的母數差距不到 2%。」（請參閱資訊箱 8.2 信賴區間）

回到之前的彈珠抽樣的例子。我不能說：「根據一個隨機樣本的結果，在罐子裡的紅色彈珠剛剛好正是 2,500 顆。」然而，我卻可以說：「我有 95%的信心，認爲母體母數落在 2,450 到 2,550 之間。」我結合了樣本的各種特質（例如，樣本大小、其中的變異等等）以及中央極限定理，並且有相當的信心程度來預測出圍繞著母體母數的特定全距。

***系統抽樣*　系統抽樣**（systematic sampling）是以一種走捷徑的挑選程序來進行簡單隨機抽樣。系統抽樣的程序大致上與簡單隨機抽樣是完全相同的，唯一的不同之處，就在於進行系統抽樣時不使用亂數表，而是先去計算**抽樣間距**（sampling interval），以創造出一個半隨機的挑選方法。抽樣間距（例如：1／k，其中 k 是某個數目）告訴我們如何從一個抽樣架構中挑選元素，而其方式則是在進行挑選之前，先略過抽樣架構中的一些元素。

例如，我們想要從 900 個名字當中挑出 300 個名字當作樣本。在隨機挑選第一個名字之後，我們在這 900 個名字當中每隔 3 個就抽一個，以得到一個 300 個名字的樣本。而在這個例子裡，我的抽樣間距就是 3。抽樣間距很容易計算，只需要樣本大小以及母體大小（或者是以抽樣架構當作最佳的估計值）。我們可以把抽樣間距想成是抽樣比的倒數。從 900 個名字當中抽出 300 個名字的抽樣比是 300／900 = .333 = 33%，而抽樣間距則是 900／300 = 3。

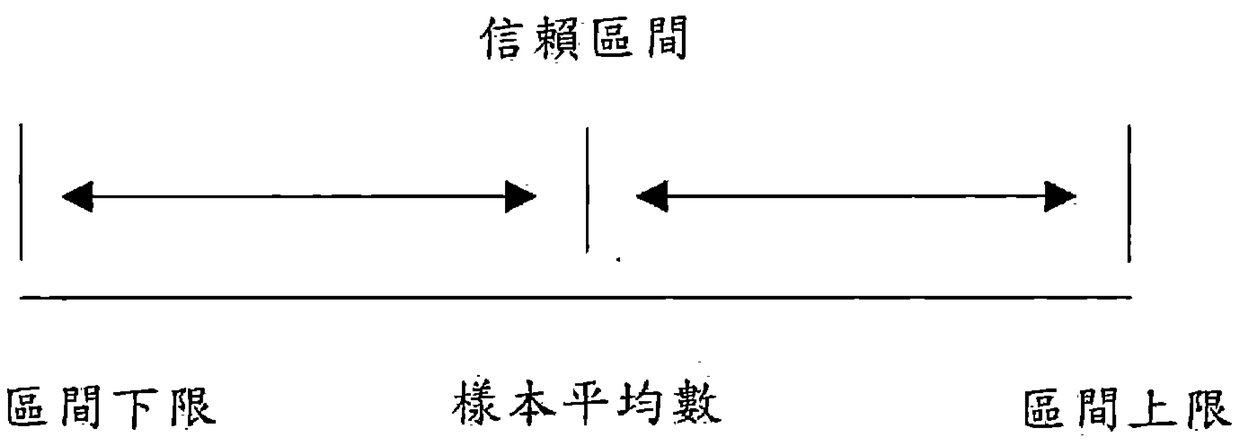
在大部分的情況裡，一個簡單隨機樣本以及一個系統樣本會產生完全相等的結果。但是當樣本中的元素是以某種周期或是模式組織而成的時候，就不能用系統性抽樣來取代隨機抽樣。例如，在我們的抽樣架構中，其組織的方式是在一份已婚夫妻的清單上先列出先生的名字，然後再列出太太的名字（請參閱表 8.1）。像是這樣的模式，如果我們使用了系統抽樣的話，那麼我們的樣本就不具有任何的代表性。由於個案的組織方式，我們所得到的系統樣本可能不具有代表性，而且只包含了太太而已。當我們的樣本架構是由成對的夫妻所組成的時候，偶數的抽樣間距就會得到都是先生或者都是太太的樣本而已。

系統抽樣　一種隨機樣本，在此種樣本中，研究者利用抽樣間距而在抽樣架構中每隔 k 個（例如，3 個或是 12 個）個案就挑選出一個個案來當作樣本。

抽樣間距　抽樣比的倒數，當研究者以系統抽樣的方式來挑選個案時，就會使用到抽樣間距。

資訊箱 8.2　信賴區間

信賴區間是一個相當簡單、但是卻強而有力的觀念，其背後具有極佳的數學邏輯以及一些很棒的公式。如果你擁有良好的數學背景，這個概念對你會很有幫助。如果你對於有許多希臘符號的複雜數學公式覺得緊張的話，這裡有一個簡單的例子，其中只有一個簡單的公式而已（已經將希臘符號減到最少的程度了）。區間所指的是一個範圍，其中包括了比母體的某個特質（也就是母體母數，像是平均值或者是統計平均數）的估計值略高以及略低的數值。區間具有上限以及下限，而下列的例子所示範的是以一種簡化過的方式來計算信賴區間，並且也顯示出樣本大小以及樣本同質性會如何影響信賴區間。



假設你抽取了一個由 9 個 12 歲的兒童所組成的樣本。你為這些兒童量體重，並且計算出他們的平均體重，也就是平均數是 90 磅，而標準差是 36 磅（平均數以及標準差都是統計量數，我們將在第十二章裡再進行討論）。你可以用希臘字母 μ 來代表母體母數。

接下來是你如何以一個簡單隨機樣本來計算出母體平均數的信賴區間的方式。然後，你可以利用把樣本平均數（也就是你對於 μ 的最佳估計值）加上以及減去特定數值範圍的方式，以估算出 μ 的信賴區間。

為了要計算出樣本平均數的信賴區間，你首先必須計算出平均數的標準誤（standard error of the mean）。為了方便起見，我們以簡化的方式稱為標準誤就可以了。而標準誤所指的就是你對於抽樣分配的變異性所估計的值。你可以利用另外一個希臘字母 σ 來代表標準差，並且在這個字母的下方標上 m 這個字母，表示這是你對於抽樣分配中的標準差的估計值。因此，標準誤就是來自於母體所有可能隨機樣本的抽樣分配中的標準差。

在估計抽樣分配的標準差時，方式是先得到樣本的標準差、並且稍微地加以調整。為了簡化這個例子，你可以略過調整的過程，並且假定抽樣分配的標準差就等同於樣本的標準差。而為了要得到標準誤，你必須用你的樣本大小（以字母 N 來表示）來進行調整。計算公式如下：

$$\sigma_m = \frac{\sigma}{\sqrt{N}}$$

讓我們把這個例子變得更為具體一些。例如，讓我們看一下這 9 名 12 歲兒童的體重。為了要得到平均數的抽樣分配，你必須使用平均數 90 磅，以及標準差 36/3 = 12 （請注意 9 的平方根是 3）。信賴區間具有上限以及下限，而下面則是計算上限以及下限的公式：

下限　$M - Z_{.95}\sigma_m$

上限　$M + Z_{.95}\sigma_m$

除了原本的 σm 之外，現在又多了兩個其他的符號：

在公式裡的 M 代表了你的樣本平均數。$Z_{.95}$ 則代表了在 95%信心水準（最典型的水準）的信賴區間的鐘形或是常態曲線下的 Z 分數（將在第十二章中進行討論）。對於常態曲線而言，Z 分數是一個標準的數字（也就是說，對於 95%的信心水準而言，這個分數都是一樣的，而且剛好就是 1.96）。雖然我們可以使用除了 95%以外的信心水準，但是這卻是最常被使用的典型信心水準。

現在，你已經擁有在計算信賴區間的上限以及下限時所需要的所有資訊了。在進行計算時，你必須把平均數 90 加上以及減去 1.96 個標準差，計算過程如下：

下限　90 – (1.96)(12) = 66. 48

上限　90 + (1.96)(12) = 113.52

這意味著你有 95%的信心，認為真正的母體母數會落在 66.48 磅到 113.52 磅之間。在決定信賴區間的上限以及下限時，你必須以樣本平均數（在我們這個例子中，就是 90 磅）加上以及減去一定的數量。我們使用 1.96 這個數字的原因，是因為當你想要達到 95%的信心水準時，Z 分數的值就是 1.96。而你則是根據樣本大小以及樣本的標準差，計算出平均數的標準誤為 12。

在看到從 66 磅到 113 磅之間的廣泛範圍之後，你或許會覺得範圍很大，而且你可能也會想問，為什麼樣本這麼小，只有僅僅 9 名兒童而已？

接下來將要說明增加樣本大小會對信賴區間造成什麼樣的影響。讓我們假定現在的樣本不是 9 名兒童，而是 900 名 12 歲的兒童（幸運的是，900 的平方根很容易就可以計算出來：30）。如果所有的條件都仍然完全相同的話，以 900 人為樣本的 σm 就是 36/30 = 1.2。如此一來，你的信賴區間如下：

下限　90 – (1.96)(1.2) = 87.765

上限　90 + (1.96)(1.2) = 92.352

利用這個大了許多的樣本，你可以有 95%的信心，認為平均體重的母體母數會落在 87.765 磅到 92.352 磅之間。

接下來所要說明的，是一個同質性非常高的樣本會對信賴區間造成什麼影響。讓我們假設你的標準差不是 36 磅，而是 3.6 磅。如果其他所有條件都仍然完全相同的話，在標準差為 3.6 的情況下，你的 σm 就是 3.6/3 = 1.2。

現在，你的信賴區間如下：

下限　90 – (1.96)(1.2) = 87.765

上限　90 + (1.96)(1.2) = 92.352

在樣本同質性很高的情況下，你可以有 95%的信心，認為平均體重的母體母數會落在 87.765 磅到 92.352 磅之間。

讓我們回顧一下，當樣本大小以及標準差改變時，信賴區間的變化情形：

樣本大小 = 9，標準差 = 36，信賴區間為 66 磅到 113 磅。

樣本大小 = 900，標準差 = 26，信賴區間為 87.765 磅到 92.352 磅。

樣本大小 = 9，標準差 = 3.6，信賴區間為 87.765 磅到 92.352 磅。

圖 8.3 舉例說明簡單隨機抽樣以及系統抽樣，請注意到在每個樣本中所抽到的名字是不一樣的。例如，亞當絲（H. Adams）在兩個樣本中都有出現，但是德羅拉德（C. Droullard）只有出現在簡單隨機樣本之中，這是因爲任何兩個隨機樣本完全雷同的機會可以說是微乎其微。

抽樣架構中總共包含了二十位男性以及二十位女性（性別標示在每個名字後面的括弧裡）。簡單隨機抽樣的結果得到三名男性以及七名女性，而系統抽樣則得到了五名男性以及五名女性。這是否意味著系統抽樣比較正確呢？其實不然。爲了確認這一點，我們利用不同的隨機號碼來抽出一個新的樣本，也就是從最後一個亂數開始，並且取前兩個數字（例如，由 11921 中選取 11，然後從 43232 中選取 43）。除此之外，我們也以一個不同的隨機號碼開始，重新抽出一個新的系統樣本。上一次開始的隨機號碼是 18，而我們這次則試試從 11 這個亂數號碼開始。我們會發現什麼？每一種性別各有多少人？[5]

分層抽樣　當我們使用**分層抽樣**（stratified sampling）時，我們首先依據補充的資訊，把母體分成幾個次母體〔也就是層（strata）〕。[6] 在把母體分爲幾層之後，我們接著會從每一個次母體中抽出一個隨機樣本。在分層抽樣裡，我們可以控制每一層的相對大小，而不是讓隨機過程去控制每一層的相對大小。此種作法能夠確保樣本的代表性、或者是將樣本中不同分層的比例固定下來。當然，關於分層的必要資訊並非總是隨手可得的。

一般而言，如果分層的資訊正確的話，分層抽樣所產生的樣本會比簡單隨機抽樣的樣本更加能夠代表母體。我們可以用一個簡單的例子來說明其中的原因。想像一個母體中有 51%的女性以及 49%的男性，則母體母數的性別比是 51 比 49。透過分層抽樣，我們分別在女性以及男性當中進行抽樣，如此一來，樣本中的性別比就可以維持在 51 比 49。如果我們使用的是簡單隨機抽樣的話，我們所得到的隨機樣本就有可能會偏離母體中真正的性別比例。因此，利用分層抽樣的方式，我們在代表母體時的錯誤就會比較少，而且抽樣誤差也會比較小。

當我們感到興趣的分層在母體中所佔的比例很小、而且隨機過程可能會偶然漏失這一個分層時，我們就必須使用分層抽樣。舉例而言，我們利用來自於大學的註冊組所得到的資訊，從 2 萬名大學生當中抽取了 200 人來當作樣本。在資料中顯示，在這 2 萬名學生中，有 2%、或者是 400 名學生是離婚的婦女，並且還養育了至少一名 5 歲以下的兒童。就我們的研究而言，這個群體相當重要，必須納入樣本當中。在一個具有代表性的樣本中，將會有 4 個這樣的學生（200 名學生的 2%），但是如果我們所抽取的是簡單隨機樣本的話，就有可能會偶然地漏掉這些學生。利用分層抽樣，我們從註冊組那兒取得這 400 名學生的名單，並且以隨機的方式從中選出 4 名學生。就重要的分層而言，此種方式可以確保我們的樣本能夠真正地

分層抽樣　一種隨機樣本，在進行此種抽樣方式時，研究者首先找出一組相互排除的、以及全容的類別，然後依據這些類別來劃分抽樣架構，最後再從每一個類別裡利用隨機挑選的過程來挑選出個案。

表 8.1　對周期性資料進行系統抽樣時的問題

個案	
1	先生
2[a]	太太
3	先生
4	太太
5	先生
6[a]	太太
7	先生
8	太太
9	先生
10[a]	太太
11	先生
12	太太

隨機開始的號碼 ＝2；抽樣間距 ＝4。

[a]是指被選入樣本的個案。

代表母體（請參閱範例 8.4 分層抽樣的範例說明）。

在一些特殊的情況裡，我們可能想要讓樣本中分層的比例與在母體中的真正比例不同。例如，在母體中的阿留申人（Aleuts）只有 0.5%，但是我們想要特別去研究阿留申人。於是，我們就盡量多抽樣，好讓阿留申人在樣本中的比例高達 10%。在使用此種非比例的分層樣本時，我們不能夠在未經過特殊調整的情況下，直接從樣本去推論母體。

在某些情況下，我們會讓某個分層或者是次群體的比例與在母體中的真正比例不同。例如，戴維斯以及史密斯（Davis and Smith, 1992）指出，在 1987 年的綜合社會調查（General Social Survey）（將在第十一章裡加以解釋）中，非裔美國人被過度抽樣了。一個對美國整體人口所進行的隨機樣本中應該只有包含 191 名黑人。戴維斯以及史密斯另外抽取出了一個非裔美國人的樣本，把人數增加到 544 人。191 名黑人受訪者在隨機樣本中所佔的比例約爲 13%，與美國人口中的黑人比例約略相等。而 544 名黑人在非比例的樣本中佔了約 30%。若是研究者想要使用這整個樣本去推論美國人口的話，就必須先進行調整、減少被抽樣的非裔美國人數目才行。如果研究者想要將焦點集中在那些與某個次群體最爲相關的議題上面時，非比例的抽樣方式就能有所幫助。在上述的案例中，利用 544 名受訪者、而非只有 191 名受訪者的樣本，研究者就能更加正確地對非裔美國人進行推論。這是因爲較大的樣本比較有可能反映出非裔美國人這個次母體的完整多樣性。

圖 8.3　如何抽取簡單隨機樣本以及系統樣本

1. 依序將抽樣架構中的每個個案進行編號。這份 40 個名字的名單是以字母順序來排列的，編號是 1 號到 40 號。
2. 決定樣本大小。我們將要抽取兩個 25%（10 個名字）的樣本。
3. 若是要抽取*簡單隨機樣本*的話，先找出一份亂數表（請參閱本圖中的摘要，完整版本的亂數表請參閱附錄 B）。在使用亂數表之前，先計算樣本所需要的最大位數（例如，在 40 個名字的情況下，必須要有兩位數；若是 100 到 999 個名字的話，就必須有三位數；若是 1,000 到 9,999 個名字的話，就必須有四位數）。從亂數表中任何一個地方開始（我們從左上角開始），然後取出一組數字（我們所選的是最後兩個數字）。把抽樣架構中與選出的亂數相對應的號碼標記起來，表示這個個案已經被選入樣本中了。如果數字太大（超過 40），就忽略這個數字。如果某個數字出現的次數超過一次的話（在這個例子裡，10 和 21 都出現了兩次），就忽略第二次。繼續這個程序，一直到樣本中的個案數目（在我們這個例子裡是 10 個）達成為止。
4. 如果所抽取的是系統樣本的話，先從隨機挑選一個號碼開始。最簡單的方式，就是閉上眼睛、在亂數表上用手隨便一指，然後挑出在抽樣架構中最接近的號碼。在這個例子，我們所挑選的數字是 18。從這個隨機號碼開始，依照抽樣間距（在我們的例子中，抽樣間距是 4）去計數，以找出第一個號碼。把這個號碼標記起來，然後再依據抽樣間距，以找出下一個號碼，一直到名單的末端為止。然後繼續依抽樣間距去計數，就好像名單的開頭部分就接在名單的末端一樣（就像是一個圓圈一樣）。繼續計數，直到接近一開始時所挑選的號碼為止，或者，如果抽樣間距能夠剛好平分抽樣架構的整體數量，就一直計數到回到一開始時所挑選的號碼為止。

編號	*姓名（性別）*	*簡單隨機*	*系統*	*編號*	*姓名（性別）*	*簡單隨機*	*系統*
01	Abrams, J. (男)			21	Hjelmhaug, N. (男)	Yes*	
02	Adams, H. (女)	Yes	Yes (6)	22	Huang, J. (女)	Yes	Yes(1)
03	Anderson, H. (男)			23	Ivono, V. (女)		
04	Arminond, L. (男)			24	Jaquees, J. (男)		
05	Boorstein, A. (男)			25	Johnson, A. (女)		
06	Breitsprecher, P. (男)	Yes	Yes (7)	26	Kennedy, M. (女)		Yes(2)
07	Brown, D. (女)			27	Koschoreck, L. (女)		
08	Cattelino, J. (女)			28	Koykkar, J. (男)		
09	Cidoni, S. (男)			29	Kozlowski, C. (女)	Yes	
10	Davis, L. (女)	Yes*	Yes (8)	30	Laurent, J. (男)		Yes(3)
11	Droullard, C. (男)	Yes		31	Lee, R. (女)		
12	Durette, R. (女)			32	Ling, C. (男)		
13	Elsnau, K (女)	Yes		33	McKinnon, K. (女)		

14	Falconer, T. (男)	Yes (9)	34	Min, H. (女)	Yes	Yes(4)
15	Fuerstenberg, J. (男)		35	Moini, A. (女)		
16	Fulton, P. (女)		36	Navarre, H. (男)		
17	Gnewuch, S. (女)		37	O'Sullivan, C. (男)		
18	Green, C. (男)	START,	38	Oh, J. (男)		Yes(5)
		Yes (10)	39	Olson, J. (男)		
19	Goodwanda, T. (女)	Yes	40	Ortiz y Garcia, L. (女)		
20	Harris, B. (男)					

亂數表摘要（供簡單隨機樣本使用）

15010	18590	00102	42210	94174	22099
90122	38221	21529	00013	04734	60457
67256	13887	94119	11077	01061	27779
13761	23390	12947	21280	44506	36457
81994	66611	16597	44457	07621	51949
79180	25992	46178	23992	62108	43232
07984	47169	88094	82752	15318	11921

*在選擇的亂數中出現兩次的號碼。

叢集抽樣　我們使用**叢集抽樣**（cluster sampling）來處理下列的兩個問題：對於一個分散的母體，研究者缺乏一個良好的抽樣架構；以及接觸抽樣元素的成本過高。[7]例如，目前沒有任何一份單一的名單列出北美的所有汽車技工。即使我們擁有一個正確的抽樣架構，要去接觸這些分散在北美各地的汽車技工，成本也會過於高昂。因此，我們只能使用一種包含了多重階段以及叢集的抽樣設計，以取代使用單一的抽樣架構。

叢集（*cluster*）是一個包含了最終抽樣元素的單位，但是可以暫時被當作是一個抽樣元素來看待。我們首先對叢集進行抽樣，然後我們再從第一階段裡所挑選出來的那些叢集當中，去抽取出第二個樣本。我們以隨機的方式對叢集進行抽樣，然後再從被挑選出的叢集當中，隨機地挑選出元素來納入樣本。此種作法的實際用處相當大，因爲即使在我們不可能爲抽樣元素製作一個抽樣架構時，我們也能夠針對叢集製作出一個良好抽樣架構。一旦我們有了一個叢集的樣本之後，針對每個叢集中的元素製作一個抽樣架構就變得容易多了。而對於在地理位置上較爲分散的母體而言，第二個好處是每個叢集中的元素在實際上的距離是比較接近的，因此在尋找或是接觸每個元素時，就會省時省力多了。

叢集抽樣　一種隨機樣本，此種樣本使用了多重階段，並且經常被用來涵括廣泛的地理區域。在此種抽樣方式裡，集合的單位會被隨機挑選出來，然後從這些被挑選出來的集合單位或者是叢集裡去抽取樣本。

範例　8.4

分層抽樣的範例說明

某家綜合醫院 100 名人員的樣本，依職位進行分層

職位	母體		簡單隨機樣本	分層樣本	與母體相較
	母體數量(N)	*百分比(%)*	*數量(n)*	*數量(n)*	的誤差
行政人員	15	2.88	1	3	−2
專任醫師	25	4.81	2	5	−3
實習醫師	25	4.81	6	5	+1
正式護士	100	19.23	22	19	+3
助理護士	100	19.23	21	19	+2
醫療技術人員	75	14.42	19	14	+5
雜工	50	9.62	8	10	−2
助理	75	14.42	5	14	+1
維修人員	30	5.77	3	6	−3
清潔人員	25	4.81	3	5	−2
總數	520	100.00	100	100	

隨機在 15 名行政人員中挑選 3 名，在 25 名專任醫師中挑選 5 名，以此類推。

註：傳統上，*N* 代表在母體中的數量，*n* 則代表了在樣本中的數量。

在簡單隨機樣本中，正式護士、助理護士、以及醫療技術人員被抽樣的數量過多，而行政人員、專任醫師、維修人員、以及清潔人員被抽樣的數量則是太少。分層抽樣則確保每個職位都具有正確的代表比例。

在叢集抽樣的各個階段中，我們都會抽取出數個樣本。在三階段的抽樣中，第一階段是針對大型的叢集進行隨機抽樣；第二階段是針對每一個被挑選出來的大型叢集裡的小型叢集進行隨機抽樣；而在最後一個階段裡，則是針對每一個被挑選出來的小型叢集裡的元素進行隨機抽樣。例如，我們想要從楓樹市（Mapleville）取得一份個人的樣本。首先，我們對該市的所有街區進行隨機抽樣，然後針對街區裡的家戶進行隨機抽樣，最後則是針對家戶裡的個人進行隨機抽樣（請參閱圖表 8.2）。雖然我們沒有一份楓樹市所有居民的正確名單，卻有一份該市所有街區的正確清單。在挑選出街區的隨機樣本之後，我們先去計算每個被挑選出來的街區裡的所有家戶數，以製作出每一個街區的抽樣架構。然後，在對家戶進行抽樣的階段裡，我們利用家戶的清單來抽取出一個隨機樣本。最後，我們則是在每個被抽樣的家戶中抽取出一個特定的個人來納入樣本。

叢集抽樣的花費通常不會像簡單隨機抽樣那麼高昂，但是卻比較不是那麼精確。在叢集抽樣的每一個階段都會引進抽樣誤差，因此，一個多重階段的叢集樣本，其抽樣誤差會多於一個單一階段的隨機樣本。[8]

圖表 8.2　叢集抽樣的範例

目標：在楓樹市隨機抽取 240 人作為樣本。

步驟 1：楓樹市總共有 55 個行政區，隨機挑選出 6 個行政區。

1　2　3*　4　5　6　7　8　9　10　11　12　13　14　15*　16　17　18　19　20　21
22　23　24　25　26　27*　28　29　30　31*　32　33　34　35　36　37　38　39　40*
41　42　43　44　45　46　47　48　49　50　51　52　53　54*　55

* = 隨機選取的行政區。

步驟 2：把挑選出來的行政區再劃分為街區。每個行政區裡有 20 個街區。從每個行政區裡隨機挑選出 4 個街區。

第三行政區的範例（在步驟 1 中所挑選出來的）：

1　2　3　4*　5　6　7　8　9　10*　11　12　13*　14　15　16　17*　18　19　20

* = 隨機選取的街區。

步驟 3：把街區再劃分為家戶，並隨機挑選家戶。

第三個行政區裡的第四個街區的範例（在步驟 2 中所挑選出來的）：

第四個街區裡包含了單親家庭、雙拼式房屋、以及四間住戶的公寓建築。這個街區被橡樹街、河濱路、南方大道、以及綠景路所包圍。在這個街區裡總共有 45 個家戶。從這 45 個家戶裡隨機挑選出 10 個家戶。

1	橡樹街 1 號	16	"	31*	"
2	橡樹街 3 號	17*	河濱路 154 號	32*	"
3*	橡樹街 5 號	18	河濱路 156 號	33	"
4	"	19*	河濱路 158 號	34	綠景路 156 號
5	"	20*	"	35*	"
6	"	21	南方大道 13 號	36	"
7	橡樹街 7 號	22	"	37	"
8	"	23	南方大道 11 號	38	"
9*	河濱路 150 號	24	南方大道 9 號	39	綠景路 158 號
10*	"	25	南方大道 7 號	40	"
11	"	26	南方大道 5 號	41	"
12	"	27	南方大道 3 號	42	"
13	河濱路 152 號	28	南方大道 1 號	43	綠景路 160 號
14	"	29*	"	44	"
15	"	30	綠景路 152 號	45	"

* = 隨機選取的家戶。

步驟 4：在每一個家戶內選擇一個受訪者。

叢集抽樣摘要：

每個家戶中隨機挑選 1 個人
每個街區裡隨機挑選 10 個家戶
每個行政區裡隨機挑選 4 個街區
在城市中隨機挑選 6 個行政區
1 × 10 × 4 ×6 = 樣本中有 240 人

當我們在使用叢集抽樣時，我們必須決定叢集的數目以及每個叢集裡的元素數目。例如，如果我們要使用兩階段的叢集抽樣從楓樹市裡挑選出一個 240 人的樣本的話，我們可以先隨機挑選出 120 個叢集，然後從每個叢集中隨機挑選出 2 個人，或者，我們也可以先隨機挑選出 2 個叢集，然後再從每個叢集裡隨機挑選出 120 個人。哪一種方式比較好呢？一個擁有較多叢集數目的設計是比較好的，因爲在每個叢集裡的元素（例如，住在同一個街區裡的人們）通常會比較相似（例如，住在同一街區裡的人們，其相似度通常會比住在不同街區裡的人們還要高）。如果只有挑選了少數的叢集，就會挑選到許多類似的元素，如此一來，樣本就可能比較無法代表整個母體。例如，我們可能挑到 2 個居民相對較爲富裕的街區，並且從這二個街區中再各自抽出 120 個人。相較於另一個從 120 個街區裡各自再挑出 2 個人的樣本而言，這樣的樣本就比較不具有代表性。

當我們要在一個廣大的地理區域中進行抽樣，並且必須找到每個元素時，使用叢集抽樣可以大幅度地減低旅行成本。一如往常地，我們必須在正確性以及成本之間作出權衡取捨。例如，艾倫、李卡度、以及芭芭拉分別親自訪談了一個有 1,500 名學生的樣本，而這些學生則代表了北美地區所有大學生所組成的母體。艾倫取得了一份所有學生的正確抽樣架構，並且使用了簡單隨機抽樣。他走遍了 1,000 個不同的地點，並在每個地點訪談一名或是二名學生。李卡度則是從一份列出全部 3,000 所大學的名單中隨機抽取了三所大學來當作樣本，然後親自造訪這三所大學，並從每一所大學中挑選出 500 名學生進行訪談。芭芭拉則是抽取了一份由 300 所大學所組成的隨機樣本，她親自走訪這 300 所大學，並且在每一所大學裡都挑選了 5 名學生來進行訪談。如果走訪每一個地點的平均成本是 250 美元的話，艾倫的整體旅行費用爲 25 萬美元、李卡度的旅行費用是 750 美元、而芭芭拉的旅行費用則是 7 萬 5 千美元。艾倫的樣本非常精確，而芭芭拉的樣本雖然精確度稍微差了一點，但是旅行的成本卻只有艾倫的三分之一而已。至於李卡度的樣本雖然是最便宜的，但是卻不具有任何的代表性。

家戶內抽樣（Within-Household Sampling）　當我們在叢集抽樣中抽樣出一個家戶或者是相似的單位（例如，家庭或者是居住單位）之後，下一個問題就是我們應該要選擇誰來當作樣本。如果我們以第一個接電話、應門、或者是收信的人當作樣本時，就可能會引入潛在的偏誤。只有在第一個回應的人是由一個真正隨機的過程所產生的結果時，我們才能挑選這個人來納入樣本中。然而，此種情況其實是很罕見的。某些人很少在家，而在某些家戶裡，則可能有某個特定的人（例如，丈夫）比家中其他人更可能去接電話或者是應門。因此，研究者必須使用家戶內抽樣，以確保在隨機挑選出一個家戶之後，在這個家戶中當作樣本的個人也是由隨機的方式所挑選出來的。

我們可以用數種方法在一個家戶中隨機挑選出一個人。[9] 最常見的方法，就是在得知家戶的人數以及組成之後（請參閱表 8.2），用一份選擇表來詳細指定你應該要挑選哪一個人（例如，最年長的男性成員、最年輕的女性成員等等）。此種作法可以排除任何選擇第一個應門、接電話，或者是訪談者選擇看起來最友善的人所可能產生的偏誤。

比例機率抽樣法　我們有兩種方法抽取叢集樣本。剛剛提過的方法是等比例的、或未加權的叢集抽樣。此種抽樣方法之所以是等比例的，是因爲每一個叢集的大小（或是每一個階段的元素數目）都是相同的。然而，比較常見的情況卻是，每個叢集團體的大小數量都不一樣。出現此種情況時，我們就必須在抽樣的各個階段調整比例。

先前所提及關於艾倫、芭芭拉、以及李卡度的例子，正好可用以說明未加權的叢集抽樣問題。芭芭拉自一份列出全部 3,000 所大學的名單上，抽取了一個由 300 所學校所組成的簡單隨機樣本。但是，除非每一所大學的學生數目完全相同，否則她的抽樣方式便有問題。在她的抽樣方法中，每一所大學被挑選的機率是完全相等的，也就是 300／3,000、或 10%的機會。但每一所大學裡的學生數目並不相同，因此，每個學生被選入她最後樣本中的機會其實並不相同。

芭芭拉列出了每一所大學，並且從這份名單當中進行抽樣。一所擁有 4 萬名學生的大型大學、以及一所只有 400 名學生的小型大學，被挑選到的機會是完全相同的。但是如果她選擇了那所大型的大學，在該大學中某一名學生被挑中的機會是 4 萬人中選 5 個（5／40,000 = 0.0125%），而在那所小型大學中，某個學生被挑選到的機會是 400 人中選 5 個（5／400 = 1.25%）。兩相比較之下，在小型大學中的學生被選入樣本的機會多出 100 倍。在大型的大學中，學生被選中的總體機率是 0.125%（10 × 0.0125），而在小型的大學中，學生被選中的總體機率則是 12.5%（10 × 1.25）。因此，芭芭拉違反了一條隨機抽樣的原則：每一個元素被選入樣本當中的機會應該要是完全相同的。

如果芭芭拉使用了**比例機率**（probability proportionate to size, PPS）抽樣法（又可譯爲機率隨樣本大小成比例的抽樣法、依樣本大小成比例的機率抽樣法），而且正確地進行抽樣的話，那麼每個最終的抽樣元素或者是學生被挑選到的機率就會完全相等。如果她想使用比例機率抽樣法的話，她就必須在抽樣的第一個階段裡，對每一所大學被挑到的機會進行調整。她必須給予擁有較多學生的大型大學更多被挑中的機會，而給予小型的大學較少被選中的機會。她必須依據所有入學學生在母體中的比例，去調整每所大學被選中的機率。因此，一所擁有 4 萬名學生的大學被選中的機率，就會比一所只有 400 名學生的大學高出 100 倍。（請參閱範例 8.5 比例機率（PPS）抽樣法）

隨機數碼撥號　**隨機數碼撥號**（random-digit dialing, RDD）是一種抽樣的技術，主要是應用於以電話對大眾進行訪談的研究計劃當中。[10] 這項技術並不是以一般出版的電話簿來當作抽樣架構。在使用電話簿來當作抽樣架構時，研究者可能會漏掉三種人：沒有電話的人、最近剛剛搬家的人、以及電話號碼未登錄在電話簿上面的人。那些沒有電話的人（例如，貧窮的人、未受教育的人、以及短期居住者等等）在任何的電話訪談當中都會被遺漏掉，但是在先進的工業化國家中，有高達 95%的人都擁有電話。下列數種類型的人，其電話號碼通常都沒有被登錄在電話簿

比例機率（PPS）　一種在叢集抽樣中所進行的調整，使用的時機是當每一個叢集中所包含的抽樣元素的數目並不相同時。

隨機數碼撥號（RDD）　一種隨機挑選個案來進行電話訪談的方法，此種方法是把所有可能的電話號碼當作是抽樣架構。

表 8.2　家戶內抽樣

在被抽為樣本的家戶內選擇個人。被選中的號碼是在圖表 8.2 中被挑選出來的家戶。

號碼	*姓氏*	*成人（18 歲以上）*	*被挑選的受訪者*
3	艾伯	1 男 1 女	女性
9	巴拉達瓦吉	2 名女性	最年輕的女性
10	狄皮亞札	1 男 2 女	最年長的女性
17	伍西維	2 男 1 女	最年輕的男性
19	塞瑞	2 名女性	最年輕的女性
20	泰勒	1 男 3 女	次年長的女性
29	威魯	2 男 2 女	最年長的男性
31	王	1 男 1 女	女性
32	葛瑞	1 男	男性
35	莫克林	1 男 2 女	最年長的女性

選擇表的範例（只計算成人）

男性	*女性*	*選擇對象*	*男性*	*女性*	*選擇對象*
1	0	男性	2	2	最年長的男性
2	0	最年長的男性	2	3	最年輕的女性
3	0	最年輕的男性	3	2	次年長的男性
4+	0	次年長的男性	3	3	次年長的女性
0	1	女性	3	4	第三年長的女性
0	2	最年輕的女性	4	3	次年長的男性
0	3	次年長的女性	4	4	第三年長的男性
0	4+	最年長的女性	4	5+	最年輕的女性
1	1	女性	5+	4	次年長的男性
1	2	最年長的女性	5+	5+	第四年長的女性
1	3	次年長的女性			
2	1	最年輕的男性			
3	1	次年長的男性			

＋ = 或者更多

範例　8.5

比例機率（PPS）抽樣法

亨利想要對住在河谷市的居民進行一個小時的親身訪談。由於河谷市的幅員相當廣大，為了要減少他的旅行時間以及開銷，因此亨利使用了叢集抽樣設計。最近一次的人口普查資料指出，該市總共有大約 49 萬人左右，而亨利只能訪談大約 220 人，也就是整體人口的 0.05%。他首先從河谷市的稅捐處以及消防局取得全市的地圖，並且取得每一個城市街區的人口普查資訊。他發現河谷市裡總共有 2,182 個街區。首先，他認為他可以隨機挑選出 10%的街區（也就是 218 個街區），親自走訪這些街區並且計算住宅單位，然後在每一個住宅單位（例如，像是房子、公寓等等）中找一個人來進行訪談。但是，每一個街道的地理範圍以及人口數量都不相同。他研究了各個街區的人口密度，並且估算出每一個街區的人口數，然後根據街道平均大小，他提出了關於人口密度的五種分類方式。他的分類方式如

下表所示：

街區類型	*叢集的數目*	*每個街區的平均人數*
超高密度	20	2,000
高密度	200	800
中密度	800	300
低密度	1,000	50
半鄉村區	162	10

亨利明白，若是沒有經過調整的話，隨機選擇街區的方式，無法讓每個人都有同等被選中的機會。例如，1 個擁有超高人口密度的街區，其人口數量等同於 40 個低人口密度的街區。於是，亨利依據街區的大小調整比例，而最簡單的方法，就是根據最小的叢集、或者是半鄉村的街區，把所有的街區轉換成同等大小的單位。例如，在一個超高密度的街區裡，其人口數量是半鄉村街區的 2,000／10、或者是 200 倍之多，因此，亨利增加了此種超高密度街區被選中的機率，使其被選中的機率是一個半鄉村街區的 200 倍。基本上，亨利製作了依每 10 人調整的叢集單位（因為這正是半鄉村街區的平均人數），並且在第一個抽樣階段裡以這些叢集單位取代不同的街區。原本的 162 個半鄉村街區並沒有任何變化，但是對於超高密度的街區而言，就等同於有 20 × 200 = 4,000 個單位，而對於高密度的街區而言，就等同於有 200 × 80 = 16,000 個單位，以此類推，到了最後總共有 49,162 個單位。亨利接著利用調整過的叢集單位，把所有的單位加以編號，而其中許多街區則有多重的數字號碼。例如，在第一個超高密度的街區裡，他所指定的單位編號是從 1 到 200，以此類推。他的編號情況如下：

1　第 1 個超高密度街區
2　第 1 個超高密度街區
3　第 1 個超高密度街區

……依此往下繼續編號

3,999　第 20 個超高密度街區
4,000　第 20 個超高密度街區
4,001　第 1 個高密度街區
4,002　第 1 個高密度街區

……依此往下繼續編號
49,160　第 160 個半鄉村街區
49,161　第 161 個半鄉村街區
49,162　第 162 個半鄉村街區

亨利仍然想要訪談大約 220 個人，並且想要從每一個調整過的叢集單位中挑選出一個人來。他使用簡單隨機抽樣的方式，從 49,162 個調整過的叢集單位中挑選出 220 個單位。然後，他可以把這些叢集單位轉換回原來的街區。例如，假設亨利隨機挑到了 25 以及 184 這兩個數字，由於這兩個數字都是屬於第 1 個超高密度街區裡的編號，因此這就代表他必須從這一個街區裡挑選出兩個人。如果他隨機挑選到的號碼是 49,161，則代表他必須在編號第 161 號的半鄉村街區裡挑選出一個人來訪談。亨利接著可以去到每一個被挑選到的街區，確認在這些街區裡的住宅單位，並且在這些住宅單位中隨機挑選樣本。當然，亨利或許也可以在選擇出一個住宅單位之後，使用家戶內抽樣的方法來找到受訪的個人。

上面：想要避開討債機構的人；非常富有的人；想要擁有隱私的人，以及想要避免接到騷擾電話、推銷員的電話、以及惡作劇電話的人。在美國的一些都會地區裡，未登錄的電話號碼高達 50%。除此之外，由於人們會搬家或是更換住處，因此，每年出版一次的電話簿裡都會列出一些已經搬走的人們的電話號碼、而且也來不及列出那些最近才剛剛搬進某個區域的人們的電話號碼。

如果我們使用隨機數碼撥號的話，我們會以隨機的方式來選擇電話號碼，因此就能避開電話簿的問題。母體是電話號碼、而不是有電話的人。隨機數碼撥號並不困難，但是卻較花時間，而且可能會讓負責打電話的人感到相當受挫。

以下是隨機數碼撥號在美國的運作方式。美國國內的電話號碼有三個部分：一個三位數的區碼、一個三位數的交換碼或者是中央碼、以及一個四位數的號碼。例如，威斯康辛州的麥迪遜地區的區碼是 608，而在這個區碼之內還有許多交換碼（例如，221、993、767、455 等等），但是並非全部的 999 個可能的三位數交換碼（從 001 到 999）都是有效的。同樣地，在一個交換碼中，並不是所有的四位數號碼（從 0000 到 9999）都在使用中。有些號碼被保留作爲將來的擴充之用、有些可能被斷線、有些則可能是在某人搬家之後暫時停用。因此，一個可能接通的美國電話號碼包括了一組有效的區碼、一組有效的交換碼、以及在交換碼當中的一組四位數的數字。

在進行隨機數碼撥號時，研究者要先確認有效的區碼以及交換碼，然後再隨機選擇四位數的號碼。這裡有一個問題，也就是研究者可能選到一個交換碼中的任何一個號碼，而這意味著有一些被選中的號碼可能是沒有提供服務的號碼、被斷線的號碼、收費的公共電話號碼、或者是公司企業的電話號碼。只有一些號碼是研究者真正想要的號碼：有效的住宅電話號碼。一直到研究者真正地打了電話之前，都不可能知道某個號碼是不是一個有效的住宅電話號碼，而這也意味著研究者可能會花費很多時間撥打那些已經被斷線的號碼、或者是公司企業的號碼等等。許多研究機構通常會利用電話來選出隨機的號碼，並且自動撥打這些號碼。此種作法可以加速整個訪談的過程，但是仍然必須要有人在一旁聽電話，並且確認某個號碼是否爲有效的住宅電話號碼（請參閱資訊箱 8.3 隨機數碼撥號）。

在隨機數碼撥號中的抽樣元素是電話號碼，而不是個人或者是家戶。數個家庭或是個人可能會共用相同的電話號碼，而在其他的情況中，每個人則可能各自有專用的一個電話號碼。而這意味著在接通了一個有效的住宅電話號碼之後，我們還必須進行第二階段的抽樣，也就是家戶內抽樣，以挑選出要進行訪談的個人。

範例 8.6（以 2006 年的綜合社會調查爲例）以實際例子說明了，許多抽樣的術語以及觀念可以如何被一起使用於一個具體的真實生活情境當中。

與樣本大小有關的決定

新的社會研究者經常會問：「我的樣本應該要有多大？」而這個問題的最佳答案就是：「視情況而定。」樣本大小取決於母體特質、研究者採用的資料分析類型、以及爲了符合研究目的所需的樣本精確性的信賴程度。正如同前面已經提過的，一個

資訊箱 8.3 隨機數碼撥號（RDD）

在過去的十年裡，隨機數碼撥號調查的參與情況已經逐漸下滑了。此種情況肇因於數個因素，像是新的來電過濾技術、因為過多的電話行銷手法而使民眾日漸關心隱私的問題、非家戶電話號碼大量出現、以及行動電話使用者的數量暴增（大多數隨機數碼撥號的樣本只有包含傳統的室內有線電話）等等。當林克等人（Link et al., 2008）以美國成年人口來針對一種新的技術——以地址為基礎的抽樣法（address-based sampling, ABS）——與隨機數碼撥號進行比較時，他們估計隨機數碼撥號的抽樣架構可能會漏失掉 15%到 19%的成年人口。雖然以地址為基礎的抽樣法在某些方面優於隨機數碼撥號，卻仍然會有一些其他的限制，其中包括與隨機數碼撥號相較之下，在樣本中有過多會說英語的非西班牙裔人口、以及較多受過教育的人。隨機數碼撥號的議題之一，牽涉到以電話來聯絡某人的作法。研究者可能撥打同一個電話號碼十幾次，但是卻都沒有人接電話。此種沒人接電話的情況，究竟只是因為某一個合乎樣本資格的人沒有時間或機會接電話、還是因為這個號碼根本就不是真正連結到某個人的有效號碼？一項針對這個議題所進行的研究（Kennedy, Keeter, and Dimock, 2008）曾經估計過，在研究者回撥了六次仍未有人接聽的電話號碼中，有大約一半的（47%）號碼是屬於合乎樣本資格、但是未被聯絡到的個人。

大型的樣本並不能保證就是一個具有代表性的樣本。一個沒有進行隨樣抽樣、或者是抽樣架構不佳的大型樣本，其樣本代表性還不如一個經過仔細隨機抽樣、以及擁有優良抽樣架構的小型樣本。

我們可以從兩方面來討論樣本大小的問題。方法之一是對母體提出假定，並且使用關於隨機抽樣過程的統計公式。利用此種方法來計算樣本大小時，所需要的統計討論已經超出了本書的範圍。[11] 我們必須對於可以接受的信賴程度（或者是誤差的數目）、以及母體變異的程度提出假定。一般而言，當母體的變異程度越大時，統計分析就必須要越精確、而且就必須有更多同時被檢驗的變項，而當樣本精確度所要求的信賴程度越高時（例如，在重要的健康結果、巨大的財務損失、或者是無辜的人可以獲得自由或被監禁時，一點點差異都可能造成極大的影響），樣本也就必須要越大。相反地，當樣本來自於同質性很高的母體、只需要對少數幾個變項進行簡單的資料分析、而且是用於低風險的決定時，即使是很小的樣本，也可能就會和大型樣本具有完全相同的效果。

第二種決定樣本大小的方法，則是一種經驗法則，也就是傳統的、或者是被廣泛接受的數量。我們使用經驗法則的原因，是因爲我們鮮少擁有統計估計方法所要求的那些資訊。除此之外，經驗法則所給予的樣本大小，也相當接近於統計方法所要求的樣本大小。經驗法則的基礎，是來自於過去那些符合統計方法要求的樣本所得到的經驗。

樣本大小的主要原則之一，就是母體越小，抽樣比就得越大，這樣一來，才能取得一個具有極高機率可以產生與整個母體相同結果的樣本。對於較大的母體而言，即使抽樣比較小，也可以得到同樣良好的樣本，這是因爲隨著母體變大，樣本

範例　8.6

樣本範例：2006 年的綜合社會調查

抽樣有許多用來描述各種不同類型樣本的術語。以下所提到的一個複雜的樣本，能夠說明研究者使用這些術語的方式。我們可以看看在社會學中眾所周知的美國全國性調查在 2006 年的樣本，而這項調查就是綜合社會調查（General Social Survey, GSS）（將在第十一章裡進行討論），而且從 1972 年以來就一直在持續進行當中。經過這許多年，這項調查的抽樣已經依據最複雜的社會科學抽樣技術而更新調整過好幾次，以求能夠在實際的成本限制之內得到一個具有代表性的母體。這項調查中的母體定義是指在美國境內全體民眾當中，定居於美國的成年人（年滿 18 歲或以上）。目標母體則包括了所有能夠說英語或是西班牙語、心智健全、而且住在家戶裡的成年人，但是排除那些住在機構環境中（例如監獄以及療養院等等場所）的人。在地區性的街區或是行政區域的層次上，研究者使用一種複雜的多階段地區機率樣本。在街區的層次上，研究者依據性別、年齡、以及職業狀況而使用限額抽樣的技術。研究者挑選出數量相等的男性與女性受訪者，而年齡在 35 歲以上和以下的受訪者數量也是相等的。

樣本設計則是結合了叢集樣本以及分層樣本。美國的領土被區分為標準都會統計區域（standard metropolitan statistical areas, SMSAs）（這是一種美國人口普查局的分類方式）以及非都會區的郡。而在進行挑選之前，研究者先依據地區、年齡、以及種族將標準都會統計區域以及各郡加以分層。接著，研究者則是根據在每個郡或是標準都會統計區域裡的家戶單位數量，利用比例機率（PPS）來對各個叢集進行調整。

至於抽樣設計則有三個基本階段。階段一：從已經分層的「主要抽樣單位」（"primary sampling unit"）當中，隨機挑選一個「主要抽樣單位」（像是一個美國人口普查區域、標準都會統計區域的一部分、或者是一個郡）。研究者也會依據在某個地理區域內或者是其他區域內是否有穩定的郵寄地址等等標準，將這些抽樣單位進行分類。階段二：隨機挑選出較小的地理單位（例如，一個地區性人口普查的區域、一個郡裡的某些部分），以及階段三：在街區裡或者是類似的地理單位裡，隨機挑選出家戶。在這個最後的階段裡，研究者把家戶當作是抽樣元素，並且從街區的地址中隨機挑選出家戶。在挑選出一個地址之後，就會有一個訪談者去與這個家戶接觸，並且從這個家戶裡選擇一個符合資格的受訪者。訪談者會從一個限額選擇表（quota selection table）中挑選可能的受訪者，並且依據這個選擇表去訪談某一種類型的受訪者（例如，家中的次年長者）。訪談者會使用電腦輔助面訪系統（computer-assisted personal interviewing, CAPI）（請參閱第十章的內容）。

在 2006 年的樣本當中，研究者首先找出了 9,535 個可能的家戶地址或者是地點。然而，在研究者排除了沒人居住的地址、以及家中沒人會說英語或是西班牙語的家戶地址之後，這個數字只剩下 7,987。在考量了拒絕參加調查的人、生重病的人、不符合資格的人、或者是沒有完成訪談的人（23.3%）之後，最後的樣本為 4,510 個人。（若要進一步瞭解細節，請上網連結到 http://publicdata.norc.org:41000/gss/Documents/Codebook/A.pdf）。

大小在精確度方面的增加程度也會跟著遞減。

在實務層面上，這代表了對於小型的母體而言（數量在 500 以下），我們需要較大的抽樣比（大約 30%）或者是約 150 人的樣本，而對於大型的母體而言（超過 150,000 以上），在所有的條件都相同的情況下，我們可以用較小的抽樣比（約 1%）就能得到相當良好的正確性，而且樣本大小只要大約 1,500 左右，就能達到相同的正確性。你可能會注意到，雖然當母體數量爲 150,000 時，其數量增加了 30 倍，但

是樣本大小卻只增加了 10 倍而已。而面對著非常大型的母體時（超過 1 千萬），我們只要用非常小的抽樣比（約 0.025%）、或者是大約 2,500 的樣本大小，就能夠達到相當的正確性。當抽樣比非常小的時候，母體的大小就不再有任何關聯性，對於數量大約 2 億的母體、或者是約 1 千萬的母體而言，大約 2,500 的樣本大小所具有的正確性都是差不多的。這些都只是近似的數量大小，而實際方面的限制（例如，成本）也會對研究者最後所決定的樣本大小造成一些影響。

有一條相關的原則是，對於小型的樣本而言，只要樣本大小稍微增加一些，就能夠大幅地增加正確性。當樣本大小的增加幅度相同時，小型樣本的正確性的增加幅度，將會高於大型樣本的正確性增加幅度。例如，當樣本大小從 50 增加到 100 時，能夠把誤差從 7.1%降低到 2.1%，但是當樣本大小從 1,000 增加爲 2,000 時，卻只能夠把誤差從 1.6%降低到 1.1%而已。[12]（請參閱表 8.3）

注意到我們對於資料分析的計畫會影響到需要的樣本大小。如果我們想要分析在母體裡的許多小型次群體的話，我們就需要較大的樣本。假設我們想要看看住在城市裡的年長黑人女性與其他的次群體（年長的男性、其他年紀以及種族的女性等等）之間的比較情形。由於這個次群體在整個樣本中所佔的比例較小（例如，大約 10%），因此我們需要一個大型的樣本。根據經驗法則，針對每一個我們想要進行分析的次群體，我們大概需要 50 個個案。如果我們想要去分析一個在樣本中只有約 10%的群體，爲了要進行這項次群體分析，我們的樣本大小就應該要是 10 乘以 50（也就是大約 500 個個案）。你或許會想要問，在你蒐集到樣本的資料之前，你要如何才能知道想要研究的群體在樣本中只佔了 10%？這個問題問得很有道理。我們通常必須使用各種不同的其他資訊來源（例如，過去的研究、關於某個區域內的官方人口統計資料等等），接著作出估計，然後再依據這個估計值去規畫我們所需要的樣本大小。

表 8.3　99%的信賴程度時，不同母體的隨機樣本大小

母體大小	樣本大小	樣本相對於母體的抽樣比(%)
200	171	85.5%
500	352	70.4%
1,000	543	54.3%
2,000	745	37.2%
5,000	960	19.2%
10,000	1,061	10.6%
20,000	1,121	5.6%
50,000	1,160	2.3%
100,000	1,173	1.2%

推論　我們抽取機率樣本的理由，就是要從樣本去對母體作出推論。事實上，有一個統計資料分析的次領域就叫作*推論統計*（*inferential statistics*）（請參閱第十二章）。雖然我們所直接觀察的是在樣本裡的資料，但是我們所感到興趣的對象不僅僅是樣本而已。如果我們從某個大學校園裡的 10,000 名學生中，取得一份由 300 名學生所組成的樣本，那麼，我們真正的興趣其實並不在於這 300 名學生身上，而是想要利用從他們身上得到的資訊，來對整個 10,000 名學生的母體作出推論。因此，在我們所具體擁有的事物（在樣本資料中所測量的各種變項）、以及我們真正感興趣的事物（母體母數）之間，其實是存在著一些差距的（請參閱圖 8.4）。

在第七章裡，你已經看到我們可以利用在抽象構念與具體指標之間的差距來表達測量的邏輯。具體的、可觀察的資料的量數，是抽象構念的近似值，而我們則利用這些近似值來對我們真正感到興趣的事物（也就是構念以及因果法則）作出估計。正如同抽樣架構、抽樣過程、以及推論能夠縮減抽樣的差距一樣，概念化以及操作化也可以減少測量的差距。

我們可以藉由直接觀察構念的量數以及樣本中的實徵關係的方式，把抽樣的邏輯以及測量的邏輯整合在一起（請參閱圖 8.4）。我們利用在樣本中以實徵方式觀察到的事物，來推論出抽象的因果法則以及母體中的母數。同樣地，在抽樣的邏輯以及效度的測量邏輯之間，也有類似的關係。在測量中，我們想要得到構念的有效指標，也就是說，我們所想要得到的是具體而能夠觀察的指標，而且這些指標能夠代表那些無法看見的抽象構念。而在抽樣中，我們所想要得到的是幾乎沒有抽樣誤差的樣本，換句話說，我們想要得到的是個案的具體集合，而且這些集合能夠正確地代表無法看見的、抽象的母體。一個有效的量數會相當接近於它所代表的構念。一個良好的樣本則是幾乎沒有抽樣誤差，而且也能夠讓我們得到相當接近於母體母數的估計值。

我們會盡可能地去減少抽樣誤差。對於同樣良好的抽樣架構以及精確的隨機挑選過程而言，抽樣誤差主要來自於兩個因素：樣本大小以及母體的變異性。在其他所有條件都相同的情況下，樣本越大，則抽樣誤差就會越小。同樣地，當母體的同質性越高時，抽樣誤差就會越小。我們可以用下列這種方式來思考：如果我們可以選擇從一個有 1,000 顆紅色以及白色彈珠的罐子中抽樣／挑選 10 顆或是 50 顆彈珠出來，以決定紅色彈珠的數量時，比較好的作法是挑選 50 顆彈珠出來。同理可推，比起罐子裡僅有兩種顏色的彈珠時的情況，如果在罐子裡有十種不同顏色的彈珠時，我們就比較無法正確地預測紅色彈珠的數量。

抽樣誤差也和信賴區間有關。如果有兩個幾乎完全相同的樣本，差別只在於其中一個樣本較大的話，那麼這個較大的樣本就會有比較小的抽樣誤差，以及比較狹窄的信賴區間。同樣地，如果有兩個幾乎完全相同的樣本，差別只在於其中一個樣本裡的個案同質性較高的話，那麼這個同質性較高的樣本就會有比較小的抽樣誤差，以及比較狹窄的信賴區間。狹窄的信賴區間所代表的是，對於一個既定的信賴程度而言，我們能夠對母體母數作出更爲正確的估計。

這裡有一個例子來說明上述的觀念。假設你想要去估計泥水匠的年收入數目，而且你手上有兩個樣本。樣本一是在 80%的信心水準的情況下，母體母數的估計值

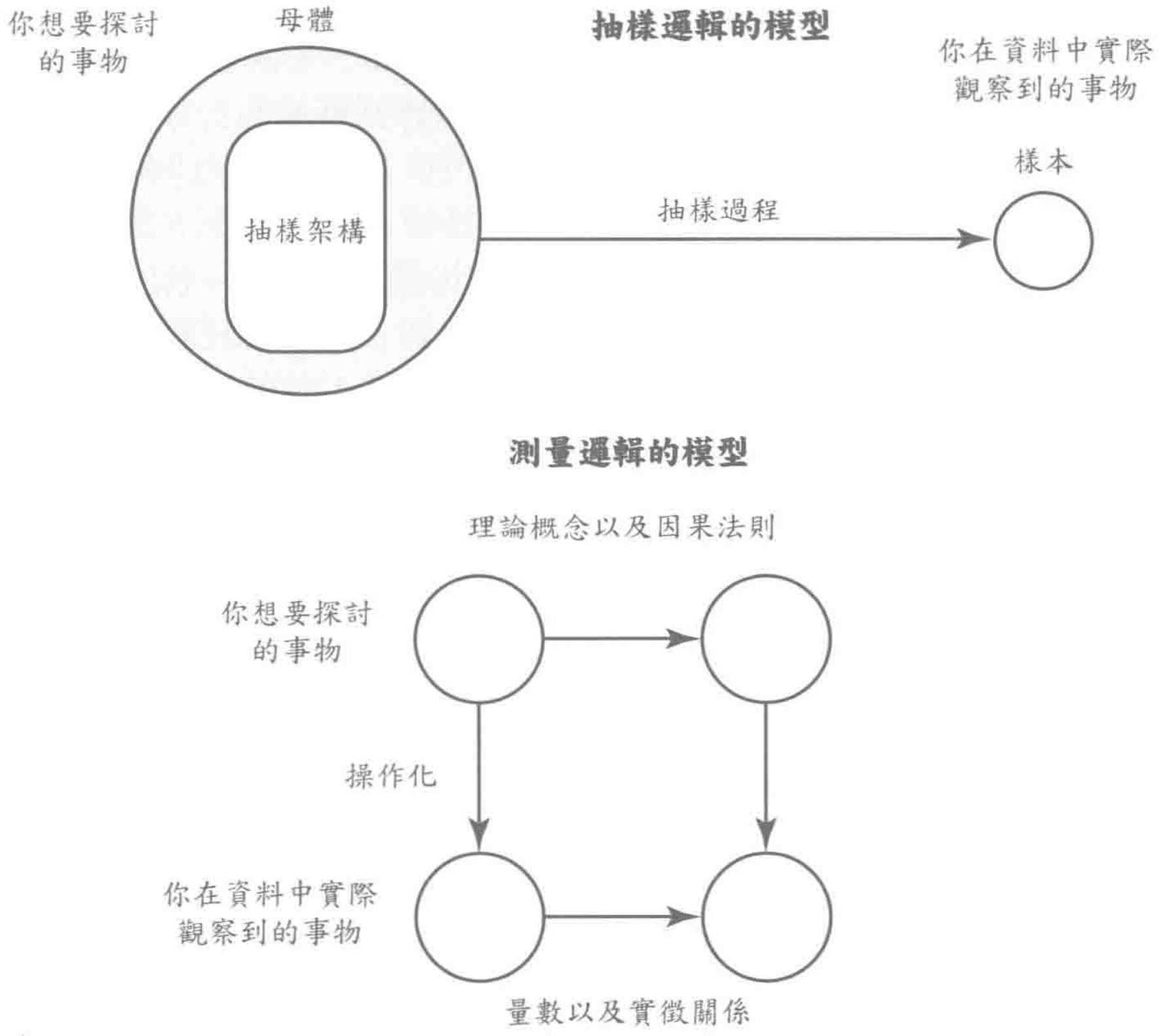


圖 8.4　抽樣邏輯以及測量邏輯的模型

爲 33,000 美元，而信賴區間則爲 30,000 美元到 36,000 美元。不過，你真正想要得到的，是 95%的信賴區間，如此一來，信賴區間的範圍就會變成從 25,000 美元到 45,000 美元。一個抽樣誤差較小的樣本（因爲樣本較大），在信心水準爲 95%的情況下，其信賴區間則可能爲 30,000 美元到 36,000 美元。

當目標並非獲得代表性樣本時的抽樣策略

在質化研究裡，研究的目的或許並不會要求研究者必須從數量龐大的個案中取得一個具有代表性的樣本。相反地，非機率樣本通常比較能夠符合質化研究的目的。在非機率樣本中，你不必事先決定樣本大小，而且對於從中抽取樣本的較大群體或者是母體，你也只具備有限的知識而已。機率樣本要求以數學理論爲基礎，並且事先規畫好研究的取向，而非機率樣本則與此不同。非機率抽樣通常會逐漸選擇個案，並且以個案的具體內涵來決定該個案是否要被選爲樣本。表 8.4 列出了各種非機率抽樣的技巧。

立意或判斷抽樣

立意抽樣（purposive sampling）〔也被稱爲*判斷抽樣*（*judgmental sampling*）〕是在特殊情況下非常有用的抽樣方式。此種抽樣方式主要是應用於探索性研究、或者是田野研究當中。[12] 此種抽樣方式是利用專家的判斷來挑選個案，或者是在挑選個案時，心中有著一個特定的目的。如果目標是要取得一個具有代表性的樣本、或者是去挑選出「平均的」或是「典型的」個案時，此種抽樣方法就不適用。在立意抽樣裡，所挑選出來的個案幾乎都無法代表整個母體。

立意抽樣　一種非隨機的樣本，在此種樣本中，研究者會廣泛地使用各種不同的方法，以找出在一個極為特殊的、難以接觸的母體當中的所有可能個案。

表 8.4　非機率樣本的類型

樣本類型	原理
隨意	抽取任何方便取得的個案。
限額	利用隨意的方式，在數個事先決定的類別中取得事先設定的個案數目，而個案的類別將會反映出母體的多樣性。
立意	利用各種不同的方法，取得所有符合特定標準的個案。
滾雪球	利用一個或是少數個案的轉介以取得其他個案，然後再利用這些個案的轉介，以取得更多的個案，以此類推。
異例	取得一些與主流模式大相逕庭的個案（一種特殊類型的立意樣本）。
逐次	一直抽取樣本，直到沒有額外的資訊或者是新的特質為止（通常會和其他抽樣方法一起使用）。
理論	所取得的個案有助於揭露某個特定背景／主題在理論上的重要特質。
適應性	依據多重階段來取得個案，像是先使用滾雪球抽樣，然後再使用立意抽樣等等。此種樣本主要是應用於隱藏性母體。

在挑選出一些特別能夠提供豐富資訊的特殊個案時，立意抽樣是最適合的方法。例如，假設我們想要使用內容分析來研究各種雜誌，以找出其中的文化主題。我們可以使用三份特別受到歡迎的女性雜誌來進行研究，因爲這些雜誌都能夠引領流行。而在本章開頭部分的《我能夠遵守的承諾》這項研究中，研究者則是利用立意抽樣挑選了八個社區。我們通常會使用立意抽樣來挑選出一些難以接觸的、特殊的母體（例如：妓女）中的成員。我們根本就不可能列出所有妓女的名單，然後再從名單中隨機挑選出樣本。相反地，爲了要找出一些妓女，研究者可能要運用地區性的知識（例如：妓女經常拉客的地點、與妓女有接觸來往的社會團體等等）以及地區的專家（例如：在刑事單位工作的警察、其他的妓女等等），以找到或許可以納入研究計畫中的妓女。由於研究目標是要盡可能找出最多的個案，因此研究者可能必須運用許多不同的方法來找到個案。

我們也會在深度調查當中使用立意抽樣來找出特殊類型的個案，以對不同類型的個案獲得更加深入的瞭解（請參閱範例 8.7 立意抽樣）。

範例　8.7

立意抽樣

在她的《深入組織化的種族主義》（*Inside Organized Racism*）這項研究中，布莉（Kathleen Blee, 2002）運用了立意抽樣來研究一些加入了種族仇視組織的婦女。她進行這項研究的目的，是要去瞭解女性為什麼、以及如何變得積極參與種族仇視組織（例如：新納粹黨、三 K 黨）之中。她想要「取得一個具有廣泛基礎的、全國的種族主義團體中女性成員的樣本」（p. 198）。由於沒有任何一份列出所有此類組織的名單，而且此類組織對於成員資料都是相當保密的，所以根本就無法進行機率抽樣。

布莉避免使用滾雪球抽樣，因為她想要訪談的是彼此之間沒有任何關聯的女性。為了取得研究的樣本，她一開始先去研究「在美國境內一年的時間之中，由每一個自稱是種族主義者、反猶太人者、白人至上主義者、基督徒認同者（Christian Identity）、新納粹黨、白人勢力光頭黨、以及白人分離主義組織等等所散佈的」（p. 198）各種溝通資訊與媒介（影片、書籍、通訊記錄、雜誌、傳單、網站等等）。她也從那些監視種族主義團體的反種族主義組織中取得一些名單，並且也在杜蘭大學（Tulane University）以及堪薩斯州立大學的圖書館中，檢視了有關右翼極端主義的檔案資料。她找出了超過一百個活躍的組織，而從這些組織當中，她發現了那些擁有女性成員或是女性激進份子的組織，並且把名單縮小到三十個種族主義的組織。然後，她試圖找出屬於這些組織的女性成員。她所要接觸的女性成員分散於美國四大主要區域中的十五個州，而且彼此之間的意識型態重點都不一樣。

接著，她採取了某種型式的叢集抽樣。她首先找出組織，然後找出在這些組織中活躍的女性成員。為了找到進行訪談的女性，她運用了親自接觸以及由熟知內情者轉介的方式，而這些熟知內情者包括了「執行假釋的官員、矯正罪犯的官員、報紙的記者、其他的種族主義激進份子以及之前的激進份子、取締幫派的聯邦以及各州的專案小組、律師、以及其他的研究者等等」（p. 200）。她最後找到了三十四名年齡介於 16 到 90 歲之間的女性，並且對每一個人進行了兩次長達六小時的生活史訪談（life history interviews）（請參閱第十三章）。

*資料來源：*Excerpt from page 198 of Inside Organized Racism: Women in the Hate Movement, by Kathleen M. Blee. © 2002 by the Regents of the University of California. Published by the University of California Press.

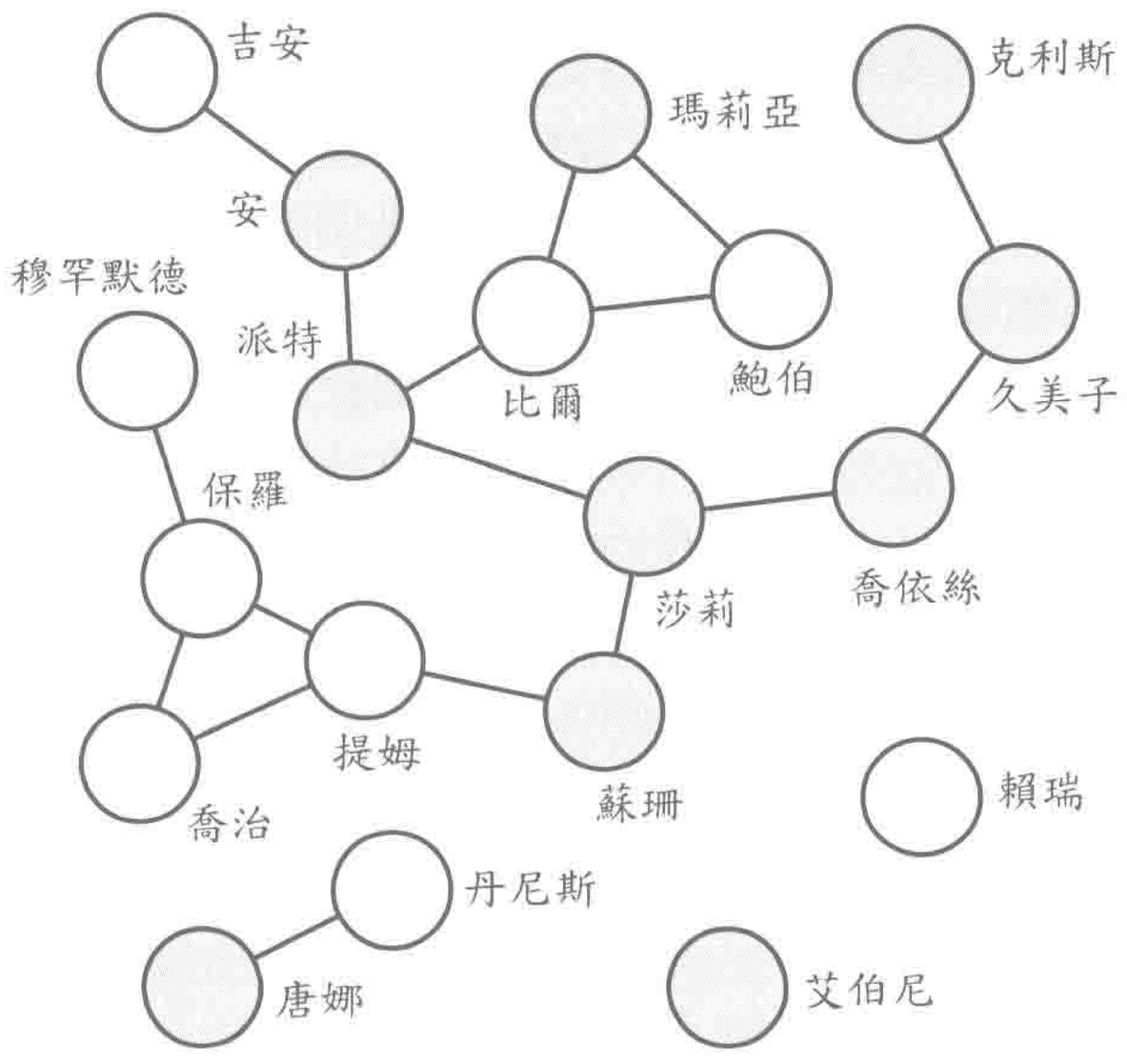
滾雪球抽樣

我們經常會對由人們或是組織所形成的相互關聯的網絡感到興趣。[13] 這種網絡可能是在全世界探索同一個問題的所有科學家、在某個中型城市裡的菁英人士、某個組織化的犯罪家庭裡的成員、在大型銀行以及企業裡的董事會成員、或者是在大學校園裡與彼此有過性經驗的人士等等。其中最重要的特質，就是每一個人或是單位之間，是透過直接或是間接的聯繫而與彼此連結起來的。但是這並非代表每個人與在網絡中的所有其他人都直接認識、彼此進行互動、或者是受到彼此的影響。更確切地來說，從整體的角度來看的話，透過直接與間接的聯繫，大部分的人都處於一個相互聯結的網絡之內。

例如，莎莉以及提姆彼此並不認識對方，但是卻有一個共同的好朋友蘇珊，因此，他們之間就有了間接的連結。這三個人都是同一個友誼網絡中的一部分。研究者會畫出一種*社會關係圖*（*sociogram*）來代表此種網絡，在社會關係圖中，各個圓圈會以直線與彼此連結在一起。這些圓圈就代表了每個個人或者是個案，而其中的直線則代表了友誼或者是其他的聯繫（請參閱圖 8.5）。

滾雪球抽樣（snowball sampling）〔又稱爲*網絡抽樣*（*network sampling*）、*連鎖轉介抽樣*（*chain referral sampling*）、*聲望抽樣*（*reputational sampling*）、以及*受訪者驅動抽樣*（*respondent-driven sampling*）〕是一種在網絡中辨識並且抽樣個案的方法。此種方法以雪球作爲比喻，雖然雪球在一開始時很小，但是當我們在濕淋淋的雪上滾動雪球、並且黏上更多雪時，雪球就會變得越來越大。滾雪球抽樣是一種多重階段的技術，從一個或者是少數的個人或是個案開始，然後根據最初個案的連結而擴散開來。

例如，我們想要研究在我們社區中青少年之間的友誼網絡。我們或許會先從三個彼此互不相識的青少年開始，並且請這三個人都各自指名四個親近的好友。接著我們去找每一組的四個朋友，並且請他們各自再指名四個親近的好友。這個過程繼續進行到下一輪的四個人，並且再次重覆。很快地，我們就可以找到許許多多的人。在樣本中的每一個人都與最初的三名青少年



註：圖圈中不同的陰影表示各種不同的膚色。

圖 8.5　友誼關係的社會關係圖

滾雪球抽樣　一種非隨機的樣本，在抽取此種樣本時，研究者先由一個個案開始，然後根據由該個案所得到的關於相互關係的資訊，去辨識出其他的個案，並且一再地重覆這個過程。

有直接或者是間接的關聯性，而且其中有一些人可能會提到相同的人。當沒有人再提到新的名字，也就是意味著已經出現一個封閉的網絡時；或者是這個網絡太過於龐大，已經超過我們所能研究的範圍時，整個過程就會停止。而在樣本中所包括的人，則是那些至少由一個在網絡中的人所指名爲好朋友的人。

異例抽樣

當我們所感到興趣的個案，是與主要的模式、主流事物、或者是其他個案的普遍特質大不相同的個案時，我們就必須使用**異例抽樣**（deviant case sampling）〔也稱爲*極端個案抽樣*（*extreme case sampling*）〕。與立意抽樣類似的是，我們會使用許多不同的技術，來找出一些具有特殊特質的個案。異例抽樣的目標是要找出一群異常的、不同的、或者是特殊的個案，而且這些個案並不能代表整個母體。我們之所以挑選出這些個案，正是因爲這些個案是不尋常的。藉由考量某些不屬於一般模式的個案、或者是包含一些在主流事件之外的個案，我們有時候可以得到更多關於社會生活的知識。

例如，假定我們想要研究高中的退學生，而且先前的研究指出，大多數的退學生都是來自於低收入的、單親的家庭，而且通常是少數民族。除此之外，其家庭背景通常是家長以及／或者兄弟姐妹本身的教育程度不高、或者自己本身也是退學生。除此之外，許多退學生也會參與非法活動。在使用異例抽樣時，我們或許會去尋找一些屬於多數民族團體、沒有任何犯罪記錄、以及來自於穩定的雙親家庭以及中高收入家庭的退學生。藉由檢視這些非典型的退學生，我們或許可以更加瞭解學生退學的真正原因。

逐次抽樣

逐次抽樣（sequential sampling）也和立意抽樣相當類似。我們使用立意抽樣的目的是要試圖去盡可能找出最多相關的個案。逐次抽樣則與此不同，因爲我們會一直持續蒐集個案，一直到新資訊的數量終止了，或者是個案的差異性達到了某種程度爲止。逐次抽樣的原則就是一直蒐集個案，直到我們到達一個飽和點爲止。用經濟學的術語來說，就是指我們一直蒐集資訊，直到邊際效益或者是額外個案所增加的利益呈現平穩狀態或者是明顯下降爲止。逐次抽樣要求我們持續地評估所有已經蒐集到的個案。例如，我們找到了六十位年齡超過 70 歲、而且已經寡居十年以上的寡婦，並且計畫對她們進行深度訪談。視我們的目的而定，我們可能不需要再去另外找出二十位在生活經驗、社會背景、以及世界觀方面與之前的六十位受訪者大同小異的寡婦。

異例抽樣　一種非隨機的樣本，特別是由質化研究者所使用。在抽取此種樣本時，研究者會刻意挑選不尋常的、或者是不符合一般模式的個案，以提供關於社會過程或者是某種背景的深刻洞見。

逐次抽樣　一種非隨機的樣本，在此種樣本中，研究者試圖去盡可能找出最多的相關個案，直到時間、財務資源、或者是研究者的體力耗盡為止，或者是直到個案裡已經沒有新的資訊或是多樣性為止。

理論抽樣

在**理論抽樣**（theoretical sampling）中，我們所抽取的樣本（例如，人們、情境、事件、時間區段等等）是來自於紮根理論。逐漸成長的理論興趣會引導了樣本個案的挑選方式。研究者依據樣本所能提供的新洞見來挑選個案。例如，一位田野研究者可能在週一到週五的時間內去觀察一個地點以及一群人。理論上，研究者或許會質疑，在其他的時間裡、或者是當該地點的某些面向改變的時候，這群人是否還會表現出相同的行爲。然後，研究者可以挑選其他的時段（例如，夜晚以及週末），以獲得一幅更加完整的圖像，並且瞭解重要的情境是否一如往常。

適應性抽樣以及隱藏性母體

相對於對一般的母體或者是可見的、可以接觸的人們進行抽樣，對**隱藏性母體**（hidden populations）（也就是那些從事偷偷摸摸的或者是隱匿活動的人們）進行抽樣，是在對於偏差行爲或者是污名化行爲（像是性暴力事件的受害者、非法藥物的使用者等等）的研究中一再出現的議題。此種方法說明了以創意的方式應用抽樣的原則、混合了質化以及量化的研究風格、以及結合了機率以及非機率的技術。

適應性抽樣（adaptive sampling）是一種依據初期觀察結果來進行調整的設計。[14] 例如，就像在進行滾雪球抽樣一般，我們要求非法藥物的使用者轉介其他的藥物使用者。然而，依據我們的研究主題，我們對於透過網絡來追蹤個案的方式進行調整。我們可能辨識出一個地理區域，隨機地將此區域劃分爲數個部分，然後透過像是隨機數碼撥號之類的策略、或者是藉由張貼招募廣告傳單的方式，在那個區域內挑選出參與者。一旦我們找到了在目標的隱藏性母體裡的成員，我們可以在滾雪球抽樣技術中利用他們來找到其他的成員。研究愛滋病的研究者、或者是曾經針對隱藏性母體進行抽樣的非法藥物使用者的研究，通常都會依賴修改過的滾雪球抽樣技術，而且也都能夠爲我們提供寶貴的資訊。（請參閱範例 8.8 隱藏性母體）

理論抽樣　一種非隨機的樣本，在此種樣本中，研究者挑選特定的時間、地點、或者是事件來進行觀察，以發展出一個社會理論、或者是評估理論性觀念。

隱藏性母體　一種由人們所組成的母體，其中的成員所從事的是偷偷摸摸的、不被社會認可的、或者是隱匿的活動，而且很難去找到以及研究這些成員。

適應性抽樣　一種應用於隱藏性母體的非機率抽樣技術，在使用此種技術時，研究者可能會使用數種不同的方法來找出以及召募樣本，包括滾雪球或者是轉介的方法。

結論

本章討論了機率以及非機率抽樣方法（請參閱重點摘要 8.1 樣本的類型）。關鍵的重點之一，是抽樣的策略應該要符合特定研究的目的。一般而言，機率抽樣比較常被使用於代表性樣本之中，而且也允許研究者使用統計考驗的技術來進行資料分析。除了簡單隨機抽樣之外，本章也提到了其他的機率

抽樣方法，其中包括了系統抽樣、分層抽樣、隨機數碼撥號、以及叢集抽樣等等。雖然本書並未涵蓋作爲隨機抽樣基礎的數學原理或者是統計理論，但是關於抽樣誤差、中央極限定理、以及樣本大小的許多討論內容，都指出當我們的目標是要得到一個具有代表性的樣本時，機率抽樣能夠產生最爲正確的樣本。

範例　8.8

隱藏性母體

以下有三個關於隱藏性母體的研究，可以說明對此種母體進行抽樣的困難之處。馬丁以及迪恩（Martin and Dean, 1993）從紐約市民當中抽取男同志來當作研究樣本。當作樣本的男同志必須是住在紐約市內、年滿 18 歲以上、未被診斷出患有愛滋病、而且會與其他男同志從事性行為的人。作者們先採用立意抽樣，利用五個不同的來源找到了 291 位受訪者。他們首先接觸了紐約市裡 150 個其成員主要是同性戀者或者是雙性戀者的組織。他們接著對這些組織進行過濾並且留下 90 個組織，因為在這些組織裡都有一些符合研究目的的男同志成員。而從這 90 個組織裡，研究者依據每個組織的成員人數，抽取出一個分層隨機樣本，該樣本中包含了 52 個組織。最後，他們從每一個組織中隨機挑選出五個成員來進行研究。馬丁和迪恩的研究的相關報導後來出現在當地的新聞媒體上，在報導出現之後，有另外的 41 位自願者主動要求加入這項研究。作者們也找到了由已經參加小型先驅研究的受訪者所轉介的 32 名男同志、來自於一年一度的紐約市同志遊行的 72 名男同志、以及他們在一家紐約市診所裡接觸到並且邀請參加研究的 15 名合格的受訪者。作者們接著使用滾雪球抽樣，方法是要求這 291 位男同志中的每一個人都要提供其他 3 名可能願意參與研究的男同志朋友。每一個同意參與研究的朋友也被要求提供其他 3 名男同志朋友。整個過程一直持續到已經從原本的 291 位男同志向外擴展了五個層次為止。研究者們把他們的樣本拿來和其他的研究進行比較，其中包括了在舊金山的兩個男同志的隨機樣本、一個包括 500 名男同志的隨機數碼撥號的樣本、以及利用舊金山人口普查區域所取得的 823 名男同志的叢集樣本。他們的樣本在種族、年齡、以及「出櫃」的比例方面，都與舊金山地區的樣本相當類似。

黑卡索恩（Heckathorn, 1997, 2002）研究了在康乃迪克州的兩個小型城市以及周遭地區裡活躍的毒品注射者。截至 1996 年 7 月為止，醫護人員已經在這些地區診斷出 390 個愛滋病例，而且其中有一半的個案都和注射毒品有關。由於每一個被抽樣的元素都必須符合特定的標準，因此可以算是立意抽樣。黑卡索恩也使用了一種修改過的、有著「雙重獎勵系統」的滾雪球抽樣方法。對於每一個完成訪談的人，他都會給予金錢的獎勵，若是受訪者能夠找到另一名新的受訪者，則可以再獲得一次金錢的獎勵。第一個受訪者會被要求不要把新受訪者的身分透露給研究者知道，此種作法有時候被稱作是掩護（masking）（也就是保護朋友）。此種作法可以避免「告密」（"snitching"）的議題以及「對毒品宣戰」的污名，這兩種情況在美國特別嚴重。此種修改過的滾雪球抽樣也類似於逐次抽樣，因為在一段時間之後，能夠找到的受訪者會越來越少，直到研究者到達一個飽和點或者是均衡點為止。

王等人（Wang et al., 2006）利用一種受訪者驅動的抽樣方法，在三個俄亥俄州的鄉下地區找到了 249 名違法藥物的使用者，以檢視藥物濫用的情形以及健康照護的需求。參與者必須要年滿 18 歲以上、未接受藥物濫用的療程、而且在前一個月裡並未使用過古柯鹼或者是甲基安非他命等毒品，才算是符合成為樣本的資格。在找到了一名符合資格的參與者之後，研究者會給他或她 50 美元當作參與研究的獎勵。而當參與者召募了其他符合資格的同儕時，就可以再額外得到 10 美元的獎勵。在一個滾雪球的過程裡，每一個後續的參與者也都被要求必須轉介其他符合資格的人。作者一開始先找到了 19 個人，其中只有比一半多一些的人（19 人中只有 11 人）轉介了符合資格而且願意參加研究的同儕。在大約 18 個月的時間裡，研究者總共找到了 249 名參與者來進行研究。他們把研究的樣本，和非法藥物使用者母體估計值的特質進行比較，並且發現到，最初找到的參與者的種族身分（白人）造成了在該種族類別裡的樣本過多的現象。否則的話，他們的抽樣方法似乎能夠抽取出一個該隱藏性母體的合理樣本。

重點摘要　8.1

樣本的類型

非機率樣本的八種類型

樣本類型	*原理*
適應性	運用關於一個隱藏性母體可能地點的知識以取得一些個案，利用隨機技術或者是召募參與者，然後利用滾雪球抽樣技術從少數的個案進行擴展。
便利	以任何便利的方式來取得任何個案。
異例	取得與主流模式差異極大的個案（一種特殊類型的立意抽樣）。
立意	利用各種不同的方法，取得所有符合特定標準的可能個案。
限額	利用隨意的方式，在數個事先決定的類別中取得事先設定的個案數目，而個案的類別將會反映出母體的多樣性。
逐次	一直抽取個案，直到沒有額外的資訊或者是新的特質為止（通常與其他的抽樣方法一起使用）。
滾雪球	利用一個或是少數個案的轉介以取得其他個案，然後再利用這些個案的轉介，以取得更多的個案，以此類推。
理論	所取得的個案有助於揭露某個特定背景／主題在理論上的重要特質。

機率樣本的四種類型

樣本類型	*技術*
叢集	針對大型的叢集單位製作一個抽樣架構，從叢集單位中抽取一個隨機樣本，為每一個挑選出來的叢集單位內的個案製作一個抽樣架構，然後抽取一個個案的隨機樣本，以此類推。
簡單隨機	為所有的個案製作一個抽樣架構，然後利用純粹隨機過程（例如，使用亂數表或者是電腦程式）挑選出個案。
分層	針對數個個案的類別的每一個製作抽樣架構，從每一個類別中抽取出一個隨機樣本，然後把數個隨機樣本結合起來。
系統	製作一個抽樣架構，計算出抽樣間距 1／k，隨機挑選一個起點，然後取出每隔 1／k 的個案。

本章也討論了數種非機率樣本的類型：便利抽樣、異例抽樣、限額抽樣、逐次抽樣、滾雪球抽樣、以及理論抽樣。除了便利抽樣之外，這些類型的抽樣方法所最適合的研究，是那些研究目的不在於得到能夠高度代表母體的樣本的研究。

在我們進入下一章之前，先重新陳述一條所有社會研究的基本原則，或許會對你有所幫助：不要把研究過程的步驟切割開來，相反地，應該要學習去瞭解各個步驟之間的相互關聯性。研究設計、測量、抽樣、以及特定的研究技術等等，這些步驟彼此之間都是唇齒相依的。不幸的是，由於在教科書中呈現資訊時會受到一些限

制，因此只能依次個別說明每一個部分。在實務上，當我們在設計研究以及發展測量方法時，也必須同時思考資料蒐集的問題。同樣地，抽樣的議題也會影響到研究設計、變項的測量、以及資料蒐集的策略。正如同你會在將來各章中所看到的，良好的社會研究取決於在數個不同的步驟裡同時控制品質，其中包括了研究設計、概念化、測量、抽樣、以及資料的蒐集與處理。在任何一個步驟裡犯下嚴重的錯誤，都可能會使整個研究計畫失去價值。

問題回顧

1. 何時要使用立意抽樣？
2. 何時適合使用滾雪球抽樣的技術？
3. 何謂抽樣架構？抽樣架構爲什麼很重要？
4. 當母體中有數個次群體、而且研究者想要確保每一個群體都能納入樣本中時，何種抽樣方法是最佳的方法？
5. 研究者如何從抽樣比來決定抽樣間距？
6. 研究者何時應該考慮使用機率比例抽樣法？
7. 在隨機數碼撥號中的母體爲何？此種抽樣法是否能夠避免抽樣架構的問題？試解釋之。
8. 研究者如何決定應該使用多大的樣本？
9. 抽樣的邏輯與測量的邏輯之間有何種關聯性？
10. 何時使用隨機數碼撥號？此種方法的優點與缺點爲何？

注解

1. See Stern (1979:77–81) and Beck (1983) on biased samples.
2. Babbie (1998:196), Kalton (1983:91–93), and Sud-man (1976a:191–200) discuss quota sampling.
3. For a discussion of the *Liter Digest* sampling error, see Babbie (1998:192-194), Dillman (1978:9–10), Frey (1983:18–19), and Singleton and colleagues (1988:132–133).
4. See Traugott (1987) on the importance of persistence in reaching sampled respondents for a representative sample. Also see Kalton (1983:63–69)on the importance of nonresponse.
5. Only one name appears in both. The stratified sample has six males and four females; the simple random sample has five males and five females. (Complete the lower block of numbers and then bgin at the far right of the top block.)
6. Stratified sampling techniques are discussed in more detail in Frankel (1983:37–46), Kalton (1983:19–28), Mendenhall and associates (1971:53–88), Sudman (1976a:107–130), and Williams (1978:162–175) .
7. Cluster sampling is discussed in Frankel (1983:47–57), Kalton (1983:28–38), Kish (1965), Mendenhall and associates (1971:121–141, 171–183), Sudman (1976a: 69–84), and Williams (1978:144–161).
8. For a discussion, see Frankel (1983:57–62), Kalton (1983:38–47), Sudman (1976a:131–170), and Williams (1978:239–241).
9. Czaja and associates (1982) and Groves and Kahn (1979:32–36) discuss within-household sampling.

10.For more on random-digit dialing issues, see Dill-man (1978:238–242), Frey (1983:69–77), Glasser and Metzger (1972), Groves and Kahn (1979:20–21, 45–63), Kalton (1983:86–90), and Waksberg (1978). Kviz (1984) reported that telephone directories can produce relatively accurate sampling frams in rural areas, at least for mail questionnaire surveys. Also see Keeter (1995).

11.See Grosof and Sardy (1985:181–185), Kalton (1985:82–90), Kraemer and Thiemann (1987), Sudman (1976a:85–105), and Williams (1978: 211–227) for a technical discussion of selecting a sample size.

12.For further discussion on purposive sampling, see Babbie (1998:195), Grosof and Sardy (1985:172–173), and Singleton and associates (1988:153–154, 306). Bai-ley (1987:94–95) describes "dimensional" sampling, which is a variation of purposive sampling.

13.Snowball sampling is discussed in Babbie (1998:194–196), Baily (1987:97), and Sudman (1976a:210–211). For discussion of sociometry and sociograms, also see Bailey (1987:366–367), Dooley (1984:86–87), Kidder and judd (1986:240–241), Lindzey and Byrne (1968:452–525), and Singleton and associates (1988:372–373) . Network sampling issues are discussed in Galaskiewicz (1985), Granovetter (1976), and Hoffmann-Lange(1987).

14.On adaptive sampling, see Martsolf et al. (2006), Thompson and Geber (1996), Thompson (2002), and Thompson and Collins (2002).

第 9 章

實驗研究

適當的技術
實驗簡史
隨機分派
實驗設計的邏輯
內在與外在效度
實際的考量
實驗研究的結果：進行比較
與倫理有關的議題
結論

實驗的與眾不同之處，就在於研究者的活動，因為研究者決定了要在何種條件下進行調查或研究。不論是在整體或者是部分，研究者……創造、建立、或者是控制了研究的背景環境。

——Willer and Walker, *Building Experiments, Testing Social Theory*, p. 2.

佩吉爾（Pager, 2007）想要知道，當受刑人在服刑期滿並且離開監獄之後，曾經入監服刑的記錄，會對受刑人在找工作方面造成什麼樣的影響。除此之外，他對於種族的差異是否會造成影響也感到相當好奇。他設計了一個田野實驗，在這個實驗裡，他雇用了大約 18 歲到 22 歲左右的男性「試驗者」（"testers"）。在這些試驗者當中，有一半是白人，而另外一半則是黑人。在 2001 年時，這些試驗者應徵了在密爾瓦基都會地區的報紙上所刊登的一些基層的工作。這些工作不要求任何的工作經驗，只要求應徵者要具備高中畢業證書而已。佩吉爾依據試驗者的種族、身體外表、以及表現的風格等等，將他們一一配對。接著，他訓練這些試驗者、確認他們在面試時的技巧、並且為每一名試驗者製作了一份假的履歷表。針對每一個種族中一半的試驗者，他製作了一份有犯罪記錄的履歷表，其中顯示出該名試驗者曾經因為持有毒品而被判刑確定，並且曾經入監服刑 18 個月。至於另外一半的試驗者，其履歷表則大致相同，只是上面沒有任何的犯罪記錄。佩吉爾隨機分派這些試驗者去應徵廣告上的工作。在這項研究中，自變項是試驗者的種族以及犯罪記錄，而依變項則是某個雇主是否回電並且同意錄用某個試驗者。佩吉爾發現到，在履歷表上有犯罪記錄、以及種族為黑人的試驗者，得到工作的機會明顯少了很多。當他同時檢視兩個自變項時，他發現一個有犯罪記錄的白人，比起一個擁有同樣資格、但是沒有犯罪記錄的黑人，更有可能得到一份工作。在威斯康辛州以及其許多其他的州裡，法律已經明文規定，當應徵者的犯罪記錄與所應徵的工作無關時，雇主不得因為種族或是犯罪記錄的因素而對應徵者作出任何歧視的行為。佩吉爾也檢視了一些相關資料，這些資料顯示出，他在密爾瓦基地區發現的明顯種族效果，在其他主要的都會地區可能會造成更大的影響。

在先前的章節裡，你已經學到了研究設計的基礎知識，而在本章以及接下來的三章中，我們的焦點則是在於能夠產生量化資料的研究技術上面。此處我們將先從實驗研究開始。

實驗研究是建立在實證主義原則的基礎上的。[1] 自然科學家（例如：化學家或者是生物學家）、以及在相關應用領域內的研究者們（例如：農業、工程、以及醫藥等等），都會進行實驗。我們也會在教育、犯罪司法、新聞學、市場行銷、護理、政治學、心理學、社會工作、以及社會學裡運用實驗去檢視許多的社會議題以及理論。在本章一開始的部分，曾經提到佩吉爾（Pager, 2007）針對種族以及犯罪記錄在找工作方面的影響進行過一項實驗，而正如同該項實驗所顯示出的，實驗可以爲我們提供強而有力的證據，以證明一個或者是兩個變項是以何種方式去影響一個依變項的。

在常識的語言裡，*進行實驗*的意思，是指在一個情境中調整某個事物，然後把所得到的結果和未經過修改時的狀況進行比較。例如，我試著發動我的車子，然而，讓我驚訝的是，車子竟然無法發動。我把電池的接頭清理一下以「進行實驗」，因爲我心中有一個簡單的假設，也就是我認爲造成車子無法發動的原因，應該是電池接頭沒有清理乾淨的關係。然後，我試著再發動車子一次，這一次，我已經調整了某個事物（清理了電池的接頭），並且把結果（車子是否能夠發動）與之前的情況（車子無法發動）進行比較。每一項實驗都是以一個「關於原因的假設」爲起點。而我的假設是在電池接頭上累積的污垢阻礙了電流、並且使得車子無法發動，因此，一旦我清理了這些污垢之後，車子應該就可以順利發動了。這一個常識的實驗相當簡單，但是卻清楚地說明了在一項實驗中的三個重要步驟：（1）從一個因果假設開始，（2）調整在情況中某一個和原因有密切關聯的特定面向，以及（3）對調整之後以及未調整之前的結果進行比較。

在本章的開頭部分所提到的實驗研究裡，佩吉爾（Pager, 2007）的假設是，種族身分以及犯罪記錄會影響到一位合格的應徵者是否能夠得到工作。他以種族身分來挑選出試驗者，並且製作了假的履歷表，如此一來，他就能以和種族身分以及犯罪記錄有關的方式來調整找工作的情境。然後，他就能夠以種族背景以及犯罪記錄來比較應徵工作的結果。

與其他的社會研究技術相較之下，實驗研究在測試因果關係方面能夠提供最爲強力的證據，這是因爲我們是以有意識的方式來設計一項實驗，以滿足因果關係的三項條件（也就是自變項在時間順序上必須早於依變項、關聯性的證據、以及排除其他可能的替代性原因）。

適當的技術

剛剛接觸社會研究的人，或許會對何種研究技術最適合某個特定的研究問題感到相當苦惱。由於在研究技術以及研究問題之間並沒有一個現成的、固定的搭配方式，因此，要選擇何種研究技術，就可能會是一個相當困難的決定。在作出決定時，

我們必須要作出「充分知情後的判斷」（“informed judgment”）。你可以藉由下列的一些方法，來發展出判斷的技巧，像是去瞭解各種不同研究技術的優點與缺點、閱讀許多已出版研究中的方法論的部分、協助一名經驗豐富的社會研究者進行研究、以及自己進行研究以獲得實務的經驗等等。

實驗能夠強而有力地去測試以及聚焦在關於因果關係的證據上。與其他的研究技術相較之下，實驗研究具有其優點以及限制，而這些優點與限制可以幫助我們瞭解在什麼情況下最適合使用實驗研究。

實驗通常是人為操控的，而且是對於複雜的社會世界以有目的的方式去加以簡化。我們一般傾向於認為「人為操控的」（“artificial”）就意謂著負面的，但是韋伯斯特以及賽爾（Webster and Sell, 2007: 11）則主張：

> 實驗最大的優勢，就在於實驗是人為操控的這個事實。換句話說，實驗允許我們在一個特定的情境之下進行觀察，而這個情境是由研究者所設計以及創造的，而不是在自然界中所發生的。

*人為操控*意謂著實驗者有意識地控制研究情境，並且有目的地整合在理論上相關的變項，同時也排除在一個假設中沒有因果重要性的變項。*人為操控*也意謂著一個具體的焦點以及目標明確的效果，而這些效果是我們在自然界中可能無法輕易遇上的。我們在實驗中包含了自變項以及依變項，但是卻會排除不相關的變項、或者是**混淆變項**（confounding variables）（也就是指那些並非我們假設考驗中的一部分的變項）。我們可以把此種情況類比於一個在自然界中發現純鈉（pure sodium）的化學家。在一個受到控制的實驗室情境中，化學家可以精確地把純鈉和其他的純化學物質混合在一起，以研究其效果。此種受到控制的、乾淨的實驗室是人為操控的、純鈉也是人為操控的、而化學家用來混合純鈉的化學物質也是人為操控的，然而，其結果卻可以產生新的知識以及化合物，而且這些新知識與化合物可能對於實際的世界具有極大的用處。

社會科學的實驗具有非常強而有力的邏輯，然而，我們仍然會面臨許多實際上以及倫理上的限制。在實驗中，我們會操弄真實世界的某些面向，然後檢查結果，然而，我們卻不能夠為了得到科學知識，而去操弄人類生活中的許多領域。在進行實驗時，我們所能處理的研究問題，是那些具有我們能夠操弄的特定情境的問題，以及那些清楚地合乎人類研究的倫理標準的問題。因此，一項實驗無法直接去回答像是下列的這些問題：完成大學教育的人，是否能夠比那些未讀過大學的人增加更多的年收入？有弟弟或妹妹的人，其領導技巧是否比獨生子或是獨生女更好？加入較多不同組織的人，是否比較願意在選舉中去投票？我們不能刻意讓某些人去上大學、並阻止其他人去上大學，以確定在往後的人生裡，哪些人賺的錢比較多。我們也不能強迫人們去加入或者是退出某些組織、或者是禁止人們加入任何組織，然後去看看他們是否會去投票。雖然我們不能去操弄許多

混淆變項　在實驗研究中，並非是正在接受檢驗的假設中的一部分的那些因素，但是卻可能對研究者感興趣的變項產生效果，並且威脅到內在效度。

我們覺得有興趣的情境或者是變項，我們卻能夠在模擬此種干預或是情境時，充分發揮我們的創意。

實驗研究的技術通常最適合用在那些範圍較爲狹窄的議題上。我們通常可以在短時間內利用有限的資源去設計並且進行眾多的實驗，並且仍然可以去測量在理論上具有重要性的假設。例如，我們可以在一個月的時間之內、以及在只花費極少成本的情況之下，去複製一個像是尼凡（Niven）所進行的研究（請參閱範例 9.1 關於死刑意見的新聞報導）。

一般而言，實驗比較適合於微觀層次（例如，個人或者是小型團體的心理現象）的理論議題，而比較不適合鉅觀層次的理論議題。而這也正是爲什麼社會心理學家以及政治心理學家會去進行實驗的原因。當某個研究問題需要去考量到一些在整個社會中、或者是長時間運作的條件時，實驗就無法輕易地去處理這樣的問題。

實驗鼓勵我們去把少數的幾個因果變項隔離出來、並且當作操弄的目標。雖然實驗能夠展現出一個或者是兩個變項的因果效果，但是如果我們想要同時考量數十個變項時，實驗就無法派上用場了。如果某個研究問題要求我們同時檢視許多變項

範例　9.1

關於死刑意見的新聞報導

尼凡（Niven, 2002）注意到，近幾十年來，美國人民對於死刑的民意調查表現出壓倒性的支持態度（大約 75%到 80%）。然而，針對犯下謀殺罪的犯人，如果人們可以在判死刑或不得假釋的無期徒刑（life imprisonment without parole, LIWP）之中作出選擇的話，對於死刑的支持度就會下降大約一半。尼凡發現，在關於死刑意見的媒體報導中，有超過九成的報導都提到民眾以壓倒性的比例支持死刑，但是卻很少有報導提到，許多人寧可把不得假釋的無期徒刑當作是同樣罪行的替代懲罰。尼凡提出一個假設，也就是如果人們大量接觸到那些關於民眾高度支持把不得假釋的無期徒刑當作替代懲罰的媒體報導，則人們對於死刑的支持態度就會有所改變。為了測試他的假設，他在超過兩個禮拜的時間之內，到邁阿密國際機場的等候區找了 564 名參與者加入他的研究。他把這些參與者隨機分派，閱讀三篇報紙文章中的一篇，而這也正是他的自變項。其中一篇報紙文章的內容，是關於民眾高度支持死刑的報導；另外一篇內容，是關於民眾高度支持不得假釋的無期徒刑報導；而第三篇文章的內容，則與死刑議題無關，是關於機場的擴建計畫。他告訴這些受訪者一個封面故事：該項研究的主題是關於報紙文章的寫作風格。接著，為了掩飾這個實驗的目的，尼凡請參與者完成了一份關於報紙文章的清晰性以及組織性的問卷。在問卷中，他也以想瞭解具備不同政治信念的人會如何對同一篇文章作出反應為前提，加入了一個關於政治信念的部分。這個部分包含了他的依變項，也就是三道問題，其內容分別為受訪者支持或反對將犯下謀殺罪的犯人判處死刑、受訪者偏好死刑或不得假釋的無期徒刑、以及請受訪者估計，將來採行死刑的州會變多還是變少。結果顯示，在閱讀了民眾一致支持死刑的文章的參與者、及閱讀了機場擴建計畫文章的控制組之間，對於死刑的態度沒有任何差異。在這兩組中都有超過 80%的參與者支持死刑，比一半多一點點的人偏好死刑勝於不得假釋的無期徒刑，而大多數人都認為將來會有更多的州採行死刑。那些閱讀了關於不得假釋的無期徒刑文章的人，對於死刑的支持度明顯低了許多（約為 62%），偏好不得假釋的無期徒刑勝於死刑（其差距約為 57%比 43%），並預測將來採行死刑的州會變少。因此，尼凡的假設獲得支持，也就是那些提到大眾廣泛支持死刑的媒體報導，只會強化民眾支持死刑勝過不得假釋的無期徒刑的態度。

所造成的影響、或者是要求我們去評估橫跨許多複雜情境或是社會團體的條件時，就不適合利用實驗來處理這樣的問題。

利用每一個實驗只考慮在一個特定情境中的一或兩個變項的方式，實驗研究能夠提供聚焦過的假設考驗。在匯集、比較、以及整合來自於許多個別實驗的發現之後，知識就能緩慢地向前進展。此種建立知識的方式，與其他研究技術建立知識的方式有所差異，因爲在其他的研究技術中，一項研究可能會在許多社會情境的複雜範圍中，同時檢視十五到二十個變項。

傳統也會影響到最適合使用實驗技術的研究問題。利用實驗研究的技術，研究者們已經針對許多主題建立了龐大的文獻資料，此種情況讓研究者能夠以快速、順暢的方式來溝通這些主題。除此之外，這些文獻資料也讓研究者可以藉由作出一些修改的方式，來複製過去的實驗，並且精確地分離出特定變項的效果。然而，在實驗方面的專門知識也可能會是一項限制，因爲專精於此種主題的研究者，通常傾向於期待每個人都使用實驗的方法。這些研究者會以一項良好實驗的標準來評估新的研究，而且在接受以及吸收來自於非實驗研究的新知識時，其步調可能會比較緩慢一些。

我們也可以在同一項研究中混合使用實驗以及非實驗的方法，以擴展我們的知識。例如，我們想要研究人們對於坐在輪椅上的人會抱持著什麼樣的態度。我們可以去調查一千位民眾，以瞭解他們對於坐輪椅者的看法。我們也可以進行一項田野研究，自己坐在輪椅上與其他人接觸，並且親自去觀察人們在真實生活的情境中對我們的反應爲何。我們也可以設計一項實驗，在實驗中，當我們在與其他人進行互動時，我們有時候是坐在輪椅上、有時候是站著、有時候則是不使用輪椅而自行走路，然後去注意人們對於每一種不同情境的反應爲何。爲了能夠充分地測試理論、並且發展出較爲完整的瞭解，我們也可以去結合來自於所有種類研究的知識（請參閱範例 9.2 利用實驗以及調查方法來測試並且應用社會認同理論）。

實驗簡史

社會科學是從心理學開始的，並且從自然科學借用了實驗的方法。一直到二十世紀之後，心理學才完全地接受實驗法。[2] 德國的心理學家以及生理學家馮德（Wilhelm M. Wundt, 1832-1920）將實驗的方法引進了心理學。在十九世紀晚期時，德國是研究教育的中心，吸引了世界各地的社會科學家來到德國進行研究。馮德爲心理學的實驗創設了一間實驗室，這間實驗室後來成爲社會研究的模範。到了 1900 年，美國以及其他地區的大學也設立了心理學實驗室來進行實驗研究。然而，當時一位著名的哲學家以及心理學家詹姆士（William James, 1842-1910）並未使用或是接受實驗法。實驗取代了在心理學中較爲哲學性的、內省的、整體的取向，而此種取向比較接近於我們在第四章中曾經討論過的詮釋取向的社會科學。

從 1900 年到 1950 年之間，社會研究者不斷地詳盡發展實驗法，使得實驗法在一些領域中變得根深蒂固。實驗吸引研究者之處，就在於它能提供客觀的、不帶偏

範例 9.2

利用實驗以及調查方法來測試並且應用社會認同理論

特蘭蘇（Transue, 2007）在一項研究中結合了實驗的邏輯以及調查研究方法，並且利用把一個抽象的社會科學理論應用於一個真實的公共政策議題的方式，對這個理論加以測試。他這項研究對於一個逐漸增加的文獻種類作出了貢獻，而此種文獻資料顯示出，由於美國人們彼此之間會以一種微妙的方式強調種族之間的差異，因而通常使得不同種族在公共議題方面的意見分歧情況變得更為明顯。

根據社會認同理論（social identity theory），我們會自動地把其他人分為內團體（in-groups）（也就是我們所歸屬的團體）以及外團體（out-groups）（也就是我們沒有參與的團體）。這些團體形成了社會界線的基礎、以及對於其他人的社會距離或是親密感覺的基礎。除此之外，我們也會有多重的認同。自我歸類理論（self-categorization）是在這個較大理論中的一個分支，該理論主張我們會依據我們許多認同當中較為活躍的那一個，來把其他人重新歸類為內團體以及外團體的成員。社會界線以及社會距離的感覺都是由最為顯著的內團體來決定的。我們對於內團體的成員會覺得比較親近，而對於顯著的外團體中的成員，則會覺得比較疏遠。當某個事物碰巧啟動了某個特定的認同時，就會引起促發（*priming*）的過程。一旦某個認同被啟動了之後，這個認同通常會對後續的行為或是思考產生較大的影響力。一旦某個認同被喚醒了之後（也就是該認同被促發了），就會變成我們在思考自己時的主要認同，並且因而影響了我們的行為。

在大多數之前對於社會認同理論所進行的研究中，研究者都是以學生作為便利樣本，在實驗室裡進行實驗，並且測試一個暫時的、人為創造的認同在一個人為議題上所造成的效果。特蘭蘇（Transue, 2007）則是試圖想要提升此種研究的外在效度。為了提升外在效度，他所使用的是大型的成人隨機樣本、一個實際的社會認同、以及一個真實的公共政策議題。他的研究所使用的是電話調查方法，利用隨機數碼撥號（請參閱第八章的內容，第十章將會討論電話調查），在 1998 年夏天時針對在明尼阿波里斯市的白人美國公民進行抽樣，並且得到一份 405 人的隨機樣本。特蘭蘇考量了兩個實際的認同：種族以及國家。之前的相關研究顯示出，在種族方面帶有偏見的白人，一旦被促發、或是被提醒了他們的種族之後，在投票時會比較可能以種族主義者的方式來進行思考。而特蘭蘇的研究則是建立在此種發現之上。另外，他所檢視的真實政策議題，則是受訪者是否支持繳稅來補助公立學校的經費。

對於社會認同這個自變項而言，特蘭蘇先把調查的受訪者隨機分派為幾個次群組，並且問受訪者下列兩個問題之中的一個問題：「對於自己的種族團體，你覺得有多親近？」或者是「對於其他的美國人，你覺得有多親近？」這個問題促發或者是提高了認同的意識。之後，在調查時，他則向這些隨機分派的受訪者問了兩個與繳稅來補助公立學校有關的問題，他問受訪者覺得這麼做的原因是「為了改善公立學校的教育品質」還是「為了提高少數民族的受教機會」，而這就是主要的依變項。特蘭蘇所提出的假設是，被促發了種族認同的白人、會比被促發了美國國家認同的白人，更傾向於反對繳稅去幫助少數民族這件事。特蘭蘇也認為，一般而言，被促發了美國國家認同的白人、會比被促發了種族認同的白人，更加強烈地支持繳稅來補助公立學校經費的公共政策。

特蘭蘇發現，那些被促發了種族認同、而且被問到關於幫助少數民族問題的白人，對於繳稅來補助公立學校經費的支持度最低。支持度最高的，則是被促發了美國國家認同、以及被問到廣泛地幫助公立學校的白人。特蘭蘇也檢視了那些對自己的種族團體認同感較為強烈的白人，並且把他們和那些種族認同較為薄弱、或者是沒有種族認同的白人進行比較。與社會認同理論一致的是，他發現那些種族認同最為強烈的白人，對於繳稅來提高少數民族受教機會這件事的抗拒程度也是最高的。在這項研究中，一個被促發的種族自我認同會增加了個人的種族內團體的意識，並且提高了與種族類別有關的社會界線。對於個人的種族內團體的強烈認同，會增加對於種族外團體成員的社會距離感，並且降低提供協助的意願。

見的、以及科學的方法，讓研究者得以研究人們的心理以及社會生活，而在當時，對於社會生活所進行的科學研究才剛剛開始獲得大眾的廣泛接受。而以下的四個趨勢，則加速了實驗社會研究的發展與擴張：行爲主義的興起、量化觀念的散佈、研究參與者的各種改變、以及實驗法在實際方面的應用。接下來，讓我們簡短地回顧一下每一個趨勢。

1. *行爲主義*（*behaviorism*）是心理學中的一門學派，原本是由美國的華生（James B. Watson, 1878-1958）所創立的，後來則是由史金納（B. F. Skinner, 1904-1990）加以發揚光大。行爲主義強調去精確地測量可以觀察的行爲、或者是內在心理生活在外表所顯示出的種種跡象，並且倡導使用實驗以進行嚴謹的實徵假設考驗。

2. *量化*（*quantification*）、也就是以數字來測量社會現象的作法，在 1900 到 1950 年之間出現了大幅成長擴張的情況。研究者將社會構念加以概念化成爲量化的量數，並且把其他無法量化的構念（例如，精神、意識、意志等等）摒除於實徵研究之外。此種作法的例子之一，就是利用智力測驗來測量心智能力。智力測驗一開始是由法國人比奈（Alfred Binet, 1857-1911）所發展出來的，在 1916 年時，有一些研究者把這個測驗翻譯成英文，並且作了一些修正。智力測驗很快地就被廣泛地使用，並且也吸引了許多人，因爲智力測驗可以用一個單一的數字來代表像是心智能力之類的主觀事物，並且也能夠以一種客觀、科學的方法來將人們排出等第。在 1921 年到 1936 年之間，光是討論智力測驗的出版文章就超過 5,000 篇之多。[3] 許多量表以及指數的技術也在這個時期被發展出來，而且社會研究者也開始使用應用統計。

3. 隨著時間經過，被當作研究參與者的人們也有所變化。早期的社會研究報告都會包括參與研究的特定個人的姓名，而且其中許多人都是專業的研究者。之後的研究則是隱匿了研究參與者的姓名，並且只報告參與者們的行動結果。隨著時間經過，研究者開始轉移方向，使用大學生或者是學童來當作研究參與者。研究者與被研究的人們之間的距離變得更爲疏遠，而此種距離疏遠的情況，反映出一種要求實驗者對接受研究的人們表現出更爲超然、疏遠、以及客觀態度的趨勢。研究者把在研究中減少與研究參與者的情感連結的作法，視爲是在實證主義中變得更爲中立的、或者是價值中立以及真正「科學的」作法。

4. 隨著研究者意識到實驗在實際層面的應用，有越來越多的企業、政府、健康照護機構、以及學校開始將實驗法用於應用的目的上面。例如，美國軍方在第一次世界大戰期間就採用了智力測驗，將數千名士兵分配到不同的軍事職位上。在工廠中推行「科學管理」運動的領導者泰勒（Frederick W. Taylor, 1856-1915），也提倡在工廠中使用實驗法。他與工廠中的管理階層合作，調整了工廠的環境，以當作增加工人生產力的方法。而在 1920 年代時，教育研究者則是進行了許多教學法的實驗、以及教室中的學生人數對於學習效果所造成的影響的實驗。

到了 1950 年代以及 1960 年代時，研究者變得更爲關心那些或許會不知不覺溜進實驗設計中的替代解釋或是混淆變項的可能來源。研究者在設計實驗時會盡量減少此種潛在的錯誤，並且逐漸地在分析資料時使用統計程序。在社會科學實驗日漸

嚴謹的設計中，由坎培爾以及史丹利（Campbell and Stanley, 1963）兩人所合著的書是一個重要的轉捩點，因爲這兩位作者定義了在實驗法當中的基本設計以及議題。

1970 年代時，有越來越多的研究者會評估研究方法論的嚴謹程度。另一個相關的趨勢，則是欺瞞的情況日漸增多，並因而引起大眾越來越關心研究的論理議題。例如，對參與者進行簡報的作法，雖然現在已相當普遍，但是此種作法其實是一直到 1960 年代才開始被研究者所採用。[4] 在過去三十年裡，整體趨勢則是傾向於使用較爲複雜精巧的實驗設計、以及在資料分析中使用統計的技術。

實驗與理論

我們所進行的社會科學實驗有兩種類型：以實徵爲基礎的、以及理論導向的（請參閱 Willer and Walker, 2007a, 2007b）。雖然進行實驗的實際過程大同小異，但是每一種類型都具有不同的目的。大多數的研究所進行的，都是以實徵爲基礎的實驗。

在以實徵爲基礎的實驗中，我們的目標是要去決定某個自變項是否會對一個特定的依變項造成明顯的效果，換句話說，我們想要去詳細記錄並且描述一個效果（也就是該效果的大小、方向、或是形式）。我們通常會在一個受到控制的環境中，以實徵的方式展現某個效果，並且將此種情況推論到「真實生活」的情境中（請參閱本章稍後部分針對外在效度所進行的討論內容）。換句話說，我們會把我們的發現推論到自然的、或者是「真實世界」的環境中。例如，在社會心理學家艾許（Solomon Asch, 1955）的著名實驗中，研究者利用要求八名學生看著三條直線的方式，說明了遵從於團體壓力的效果。在艾許說明了團體從眾（group conformity）的力量之後，我們就可以把這個效果從他要求八名學生看著三條直線的特定研究，推論到那些由各種類型的人們所組成的各種大小不同的團體上，而在其中的人們所從事的，則是真實生活中的工作或是任務。在本章開頭部分曾經提到佩吉爾（Pager, 2007）所進行的一項研究，而該項研究也是一項以實徵爲基礎的研究。該項研究揭露出種族以及犯罪記錄在求職時所造成的效果；同樣地，尼凡（Niven, 2002）針對新聞報導以及死刑意見所進行的研究（請參閱範例 9.1），也是一項以實徵爲基礎的研究，而他的研究則顯示出閱讀新聞報導對於死刑意見所會造成的效果。

在理論導向的實驗中，我們會以演繹的方式來進行實驗，其方法是把一個我們相信這個世界運作的抽象模式（也就是一個理論），轉變成爲一個具有特定量數的特定研究。實驗是理論模型的複製品。當我們從一個理論導向的實驗去進行推論時，我們就是把這個理論當作是世界運作方式的一種模式來進行推論。我們主要的任務就是去考驗這個理論，並且去瞭解是否有實徵證據可以支持這個理論。我們所關心的並不是去找到自變項的大型效果，更確切地說，我們所關心的，是要去確認一個理論的具體預測是否能夠密切地符合實徵發現。我們不會擔心對理論所進行的實驗測試是否受到高度的人爲操控或者是無法切合於自然世界。我們最關心的，是實徵結果是否能夠符合我們的理論。我們會進行許多的複製實驗，以一再地顯示出證據符合理論、或者是理論能夠經得起無數次的檢驗。的確，正如同韋伯斯特以及賽爾（Webster and Sell, 2007:

21）所主張的，「除了對理論所具有的影響之外，實驗結果本身其實一點也不有趣。」

我們經常會在實驗中使用統計技術，以瞭解由理論所預測的結果發生的機率有多高。如果理論預測某個結果的發生機率很低，但是該結果卻經常出現的話，我們對於該項理論正確性的信心就會減少。以下有一個簡單的例子可以說明此種情況。我的朋友相信他自己可以辨別出五種不同品牌的可樂之間的差異之處。我在四天的時間內，請他喝了二十杯這五種品牌的可樂。在這二十杯可樂中，每一種品牌的可樂都各佔了五分之一、也就是四杯，但是其排列順序是混合的、不規則的。如果他在喝這二十杯可樂時，每次都能答對可樂的品牌，我就能夠有信心地認爲他確實能夠分辨出這些可樂之間的差異。就機率而言的話，他猜對的機會應該只有 20%而已。如果一個像是我朋友能夠分辨可樂差異的理論，每一次都是百分之百正確的話，我們對於該理論的信心就會增加，但是百分之百正確的情況其實是相當罕見的。然而，如果我的朋友答對的正確率有 90%的話，我就會認爲他的判斷力非常好，但是仍未到達完美的地步。如果他的正確率僅僅只有 30%的話，那麼，他的正確率其實只有比單純的機率高一些而已，而我對他的判斷力就不會具有太大的信心了。在測試一個理論時，該理論的預測是否能夠遠遠超過我們單純以機率而言所期待出現的情況、或者是該理論是否能夠經得起重覆測試的考驗，都會影響到我們對於一個解釋的所具有的信心程度。

在範例 9.2 裡曾經提到一項由特蘭蘇所進行的研究，而該項研究就具有理論導向實驗的特徵。針對自我分類理論這一個之前已經通過許多實驗考驗的理論，他試圖去複製對於這個理論的考驗方式。他應用了在選擇內團體認同時會啓動的自我分類的促發效果，並且提出了支持該項理論的證據。而他這項研究與先前研究的不同之處，就在於他的研究結合了調查方法以及一個真實的政策議題。另外一個在本章稍後部分會提到的針對接觸假說所進行的研究（請參閱範例 9.7 大學室友的田野實驗），雖然是一項應用於真實生活情境中的研究，但也是一個理論導向的實驗。雖然我們通常一開始時會在高度控制的人爲情境中進行理論導向的實驗，但是我們可能會將這些實驗加以延伸、並且在自然的情境中複製這些實驗。

隨機分派

身爲研究者，我們總是在進行比較。俗話說：「要拿蘋果和蘋果作比較，不要拿蘋果和橘子作比較」，這句話真正在談論的主題其實並不是水果，而是進行比較的方式。換句話說，這句話真正的意思，是指我們應該要拿本質上相同的東西來進行比較，才能算是有效的比較。

進行比較的方法有許多種。[5] 我們可以去比較同一個人在不同時間的變化（例如，在完成一項訓練課程之前以及之後），而這就是一種受試者內實驗（within-subject experiment）。然而，通常我們比較感興趣的，並不是一個處遇或是自變項是否造成某個人的改變，而是該處遇或自變項是否能夠造成某種結果。我們可以在兩個時間

點去比較一群人（例如，一群由三十個人所組成的群組在完成一項訓練課程之前以及之後的平均值）。我們也可以去比較同一群由三十個人所組成的群組在接受一系列的處遇之後的狀況（例如，依序接受三個訓練課程），以瞭解我們是否每次都能看出效果，而此種方式就被稱作是組內實驗（within-group experiment）。除此之外，我們也可以去比較兩組各由 15 名參與者所形成的群組，其中一組接受特定的實驗處遇，而另外一組則沒有接受實驗處遇（例如：訓練課程），此種方式就稱作是組間實驗（between-group experiment）。

隨機分派可以創造出類似的群組，以方便我們進行組間的比較。爲了要達到比較的目的，對於一些可能會造成因果關係的替代解釋的變項，我們希望這些群組在這些變項方面不要有差異。例如，我們想要比較兩個群組，以確認完成一項消防訓練課程對於每個人對火災反應能力的影響效果。因此，除了修習這門課程以外，我們會盡量使這兩個群組在所有方面都很類似。如果這兩個群組除了修習課程與否之外，其他條件都是相同的話，我們就可以有信心地對這兩個群組的結果進行比較，並且知道我們所發現的任何差異都是由該項課程所造成的。如果這兩個群組之間存在著差異的話（例如，其中一組裡有一個經驗豐富的消防隊員、或者是其中一組的成員年紀比較輕而且體能狀況也比較好），當我們在對這兩個群組進行比較時，我們就無法確定訓練課程是否爲造成我們觀察到的差異的唯一原因。

進行隨機分派的原因

隨機分派（random assignment）是一種把個案（例如，個人或者是組織等等）分派到各群組以進行比較的方法。這是一種把一群研究參與者畫分爲兩組或是更多組的方法，以增加研究者的信心，確定這些群組之間不會有系統性的差異。隨機分派是一種純粹機械式的方法，而且分派的過程是自動的。你不能依據你自己或者是研究參與者的偏好、或者是參與者的個人特質（例如，你認爲某個人看起來很友善、某個人想要和朋友分在同一組、把所有晚到的人都分到同一組等等）來進行分派。

隨機分派中的隨機是指在統計或者是數學意義上的隨機，而不是在日常意義上的隨機。在日常生活中，當我們說到*隨機*時，我們可能是指沒有計畫的、隨意的、或者是偶然的。而在機率理論裡，*隨機*所指的是一種過程，在此種過程中，所有的個案被挑選到的機會都是完全相同的。利用隨機挑選，你就能夠以數學方法計算出某個特定個案出現在某一個群組、而非其他群組中的機率。例如，你現在有五十位研究參與者，並且利用一個隨機過程（像是拋擲一枚重量均衡的硬幣）來把某些人分到某個組別（硬幣有人像的那一面朝上）或者是另外一個組別（硬幣的反面朝上）。利用這種方法，所有的研究參與者最後被分到某一組的機會都是完全相等的。

隨機分派 在實驗一開始時，利用一個隨機過程將研究參與者畫分為不同群組的作法，如此一來，實驗者就可以把所有的群組當作是相等的。

隨機過程的美妙之處，就在於經過許多次個別的隨機事件之後，就會出現我們可以相當正確地預測可能會發生的情況。雖然整個過程是完全取決於

機率的、而且也不可能去預測某個特定時間點所會出現的特定結果，但是當我們在檢視過許多情況之後，仍然可以作出極爲正確的預測。

隨機分派是*沒有偏誤的*（*unbiased*），這是因爲我們想要確定一個假設的想法、或者是一名研究參與者的個人興趣，並不會介入於挑選的過程之中。*沒有偏誤*並非意味著在每個特定的隨機分派挑選過程中的所有群組都會完全相同，而是指我們可以得到接近於此種理想狀況的分派結果：我們可以利用數學方法來決定選擇某個個案的機率，而且從長遠來看的話，在經過了許多次個別獨立的挑選過程之後，所有群組的平均值將會是完全相同的。

隨機抽樣以及隨機分派都是把個案選入研究中的過程。當我們在進行隨機分派時，我們會使用一個隨機過程來把一群個案分成兩組或者是更多組。而當我們進行隨機抽樣時，我們則是先從一個較大的個案集合體中挑選出一組較小的個案組合（請參閱圖 9.1）。我們可以先抽樣、然後再進行隨機分派。例如，我們可以先進行抽樣，以取得一個較小的個案組合（例如，從 20,000 人當中挑選出 150 人），然後利用隨機分派的方式，將這個較小的個案組合再分成數個不同的組別（例如，把這 150 人分成 3 個各有 50 人的小組）。

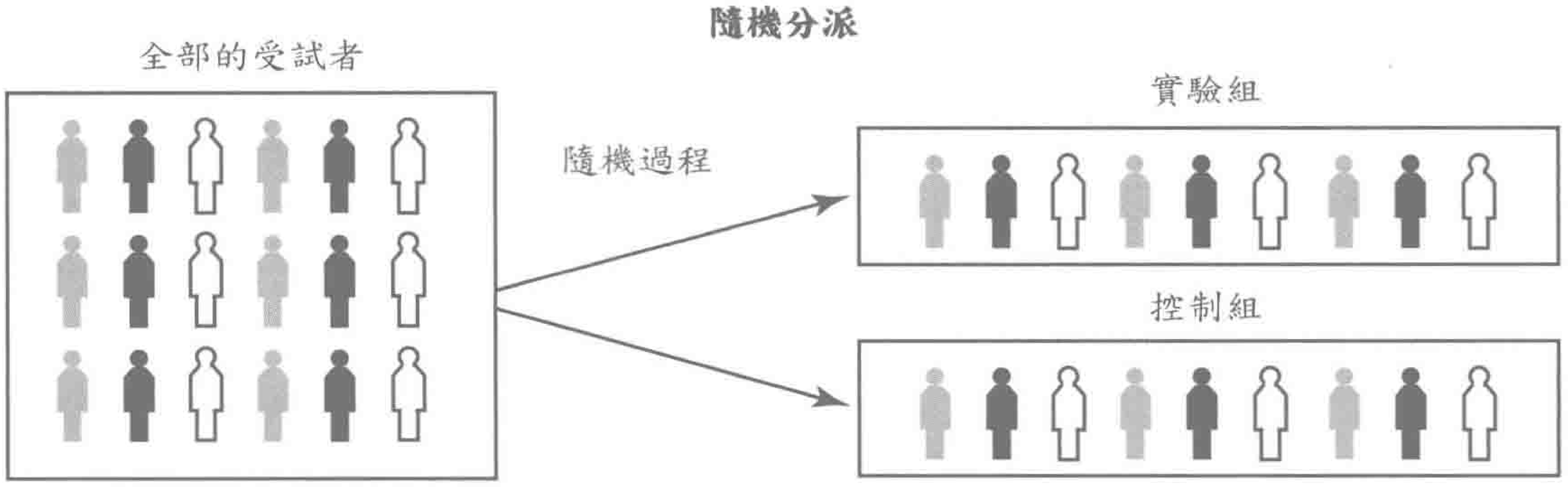


圖 9.1　隨機分派以及隨機抽樣

如何進行隨機分派

隨機分派在實際上是相當簡單的。我們從一群個案（也就是個人、隊伍、公司、或者是任何的分析單位）開始，然後利用一個隨機過程把這一群個案分成兩個或者是更多個群組，像是要求人們報數、拋擲硬幣、或者是擲骰子等等。例如，我們想把 32 個人分成兩個 16 人的群組。我們可以要求每個人在一張標準大小的紙條上寫下自己的姓名，把所有的紙條放入一頂帽子裡，由研究者閉上眼睛把所有的紙條充分混合，然後由研究者在閉上眼睛的情況下去抽出第一組的十六個人的名字，最後再把剩下的十六個人分到第二組裡。真正具體的情況可能會比較不一樣，而且所分出來的群組也可能有些差異。例如，雖然出現的機率很低，但是具有同一特質的所有個案仍然有可能都被分到同一組裡。例如，我們有 32 名研究參與者，其中 16 人是男性，而另外 16 人則是女性，但是在經過隨機分派之後，最後所有的男性都被分到同一組，而所有的女性則全部被分到另外一組。就隨機的機率而言，此種情況的確是可能出現的，但是卻極爲罕見（請參閱圖 9.2 如何進行隨機分派）。

配對與隨機分派

如果隨機分派的目的是要得到兩個（或是更多個）相等的群組，你或許會想要問，若是把每一個群組內的個案特質加以配對的話，不是會比較簡單嗎？有一些研究者會把不同群組內的個案依據某些特質來加以配對，像是年齡以及性別等等。配

圖 9.2　如何進行隨機分派

對是隨機分派的替代方案，但卻是一個不常被使用的作法。

在進行配對時，會產生一個問題：要進行配對的相關特質是什麼？研究者能夠找到正確的配對嗎？每個不同的個案之間有數以千計的差異之處，而且我們無法得知哪些特質才是具有相關性的。例如，我們比較了兩個各有 15 名學生的群組。由於第一組裡有八名男學生，因此我們在第二組裡也要有八名男學生。在第一組裡有二名男學生是獨生子，其中一名是來自於父母離異的家庭、而另外一名則是來自於完整的家庭，此外，其中一名又高又瘦、而且是一名猶太教徒，而另外一名則既矮又胖、而且是一名天主教徒。爲了要將這兩個群組進行配對，我們是否需要去找到一個來自於父母離異的家庭、又高又瘦、而且是獨生子的猶太教徒，以及一個來自於完整的家庭、既矮又胖、而且也是獨生子的天主教徒？這名又高又瘦的猶太教男學生只有 22 歲，而且是一名醫學院的預科生。至於另外一名又矮又胖的天主教男學生則是 20 歲，而且主修會計。我們是否也要針對這兩名男學生的年齡以及事業抱負進行配對？如此一來，真正的配對很快地就會變成一項不可能的任務。

實驗設計的邏輯

實驗的語言

在實驗研究裡，許多研究都把參與者稱爲**受試者**（subjects），不過，在最近幾年裡，*研究參與者*（*research participants*）這個稱呼則是越來越常被研究者所使用。

實驗的各個部分　實驗總共可以分成七個部分，不過，並不是所有的實驗都具備了這些部分，而有些實驗則是除了這七個部分之外，還增加了其他的部分。

1. 處遇或者是自變項
2. 依變項
3. 前測
4. 後測
5. 實驗組
6. 控制組
7. 隨機分派

在大多數的實驗裡，我們會創造出一種情境、或者是進入一個正在進行中的情況並且去加以調整。我們所進行的調整，就叫做**處遇**（treatment）（或者也被稱爲刺激或操弄）。這個詞彙來自於醫學，原本是指醫生對病人所施予的治療或是處理方式，也就是醫生利用一種生理的、或是心理的治療方式或是處遇來介入並且改變病人的狀況。處遇就是自變項、或者是一

受試者　一種對於實驗研究中參與者的傳統稱呼。

處遇　在實驗研究中的自變項。

群自變項的組合。在本章一開頭的部分所提到的研究裡，佩吉爾（Pager, 2007）就擁有兩個自變項：一個是固定的特質（試驗者的種族）、而另外一個則是受到操弄的特質（在僞造的履歷表上的犯罪記錄）。而在尼凡（Niven, 2002）的研究中（範例 9.1），實驗裡的處遇是參與者在機場等候區裡所讀到的三篇新聞報導中的其中一篇。而在特蘭蘇（Transue, 2007）的研究裡（範例 9.2），實驗中的處遇則是參與者在電話調查中被問到的兩個關於認同的問題中的其中一個問題。

有時候我們會費很大的工夫來創造出處遇。有些人可能會使用像是閱讀不同的僞造記錄的方式、閱讀不同的新聞報導的方式、詢問不同的調查問題的方式、或者是觀看不同影片的方式（請參閱範例 9.4）等等較爲簡單的處遇，但是其他的處遇則可能相當複雜，例如像是讓研究參與者置身於有許多精密設備的情境中、經過布置的實體場景中、或者是設計過的社會情境中等等。在第四章中，你已經讀過著名的米爾葛蘭以及辛巴度的實驗，而在這些實驗中的處遇，就是屬於複雜的處遇（請同時參閱本章稍後部分的範例 9.6）。我們希望所施予的處遇能夠造成影響以及產生特定的反應、感覺、或者是行爲（請參閱本章稍後討論實驗真實性的部分）。

依變項、也就是實驗研究中的結果，就是因爲處遇而發生改變的具體情況、社會行爲、態度、感覺、或者是參與者的信念等等。我們可以使用紙筆指標、觀察、訪談、或者是生理反應（例如，心跳或者手心冒汗等等）來測量依變項。

在一般情況中，我們在實驗過程中去測量依變項的次數會超過一次以上。所謂的**前測**（pretest），就是在引入處遇之前先對依變項進行測量。至於**後測**（posttest），就是在已經把處遇引進實驗之後，再去測量依變項。

爲了要達到比較的目的，我們通常會把參與者分成兩組或是更多組。在簡單的實驗裡通常只有兩個組別，而且只有其中一個組會被施予處遇。**實驗組**（experimental group）就是有接受處遇、或者是有出現處遇的群組。而沒有接受處遇的群組就是**控制組**（control group）。當自變項有許多不同的值時，實驗中就可能會使用超過一個以上的實驗組。

進行實驗的步驟 依循著研究過程的基本步驟，我們會先決定一個主題、將主題濃縮爲一個可以測試的研究問題、然後再發展出一個含有變項的假設。一個相當關鍵的早期步驟，就是要規畫一個具體的實驗設計（稍後將進行討論）。當我們在進行規畫的時候，我們必須決定要使用的群組數量、如何以及何時創造出處遇的情境、測量依變項的次數、以及參與者群組在研究中從頭到尾所要經歷的狀況。我們通常會針對實驗進行*前導測試*（*pilot test*）〔也就是對實驗進行一次「模擬演練」（"dry run"）〕。

在我們找到了自願加入的參與者、並且把他們隨機分派到不同的組別之後，實驗就開始進行了。我們會給予參與者明確的、事先規畫好的指示。接著，我們可能會在引入處遇之前，先在前測之中去

前測 在施予處遇之前，先去測量實驗中的依變項。

後測 在施予處遇之後，才去測量實驗中的依變項。

實驗組 在實驗研究中接受處遇的參與者。

控制組 在實驗研究中未接受處遇的參與者。

測量依變項。然後，我們會僅僅讓其中一組接受處遇（或者是較高程度的處遇）。最後，我們會在後測中去測量依變項。我們也會在參與者離開之前，訪談他們對於實驗的反應。我們會把依變項的測量結果記錄下來，並且檢視每一組的結果，以確認假設是否得到支持。

實驗中的控制　在實驗當中，控制是非常重要的。[6]我們想要控制實驗情境的所有面向，以把處遇的效果孤立出來。藉由控制混淆變項的方式，我們就能夠排除掉替代的解釋，而這些替代的解釋可能會破壞我們建立因果關係的企圖。

我們有時候會使用欺瞞的手段來控制實驗情境（請參閱本章後面部分的與倫理有關的議題）。所謂**欺瞞**（deception），就是指我們刻意地利用書面或者是口頭的指示、其他人的行動、或者是實驗情境中的某些面向來誤導研究參與者的作法。當我們在使用欺瞞時，可能也會牽涉到使用**同謀**（confederate），也就是某個假裝是另一個研究參與者或者是旁觀者的人，但是其實這個人是爲研究者工作的，而且會刻意地誤導其他的參與者。在米爾葛蘭的實驗中，就使用了同謀，而在本章稍後的範例 9.6 中所描述的關於身心障礙同事的研究中，一樣也使用了同謀。

使用欺瞞的目的，是要控制參與者所看見或是聽到的事物、以及他們認爲正在發生的事，而這通常意味著我們必須創造出一個**封面故事**（cover story），也就是一個關於研究目的的錯誤解釋，我們會把這些錯誤解釋告訴參與者們，以誤導他們並且掩飾真正的研究目的。封面故事有助於滿足參與者的好奇心，但是會減少要求特徵（demand characteristics）（請參閱本章稍後的部分）。許多研究都使用了封面故事（請參閱在範例 9.1、9.4、9.6、以及 9.7 當中的研究）。

設計的類型

我們把實驗的各個部分（例如，前測、控制組等等）結合起來，成爲一個**實驗設計**（experimental design）。有些實驗中沒有包括前測、有些實驗沒有包含控制組、而有些實驗中則有許多的實驗組。一些被廣泛使用的標準設計已經被賦予專門的名稱了。學習標準設計是相當重要的，而其理由有以下兩個原因。首先，在閱讀研究報告時，研究者可能只有提到某個標準設計的名稱、而不是去詳細地說明此種標準設計的內容。第二，標準設計舉例說明了結合設計的各個部分的常見方式。我們可以在我們所進行的實驗中使用這些標準設計，或者是創造出一些經過變化的設計。

接下來，我們以一個簡單的例子來說明各種不同的設計。假設你想要瞭解，如果服務人員（包括男服務生和女服務生）先介紹自己的名字、並且在

欺瞞　實驗者對參與者所說的一種關於實驗本質的謊言，或者是透過實驗者的行動或是實驗情境來創造出一種錯誤的印象。

同謀　一個為實驗者工作的人，此人會裝作也是研究參與者、或者是在其他參與者面前使用實驗的封面故事來欺瞞這些參與者。

封面故事　欺瞞的一種類型，在使用封面故事時，實驗者會告訴參與者一個錯誤的故事，好讓參與者依據實驗者所希望的方式去行動，並且不知道真正的假設為何。

實驗設計　對實驗的各個部分所進行的規劃以及安排。

送上餐點後的八到十分鐘內回頭詢問顧客「請問您對餐點覺得滿意嗎？」的話，他們是否會得到更多的小費。在這個例子中，依變項就是服務人員得到的小費數量。而你進行研究的地點，則是在一個鎮上的兩家完全相同的餐廳，這兩家餐廳位於小鎮的不同兩邊，但是顧客的類型以及小費的平均數量都是相同的。

古典實驗設計 所有的設計都是**古典實驗設計**（classical experimental design）的變種，而古典實驗設計也是我們到目前為止所一直在討論的設計類型，其中包含了隨機分派、前測與後測、實驗組、以及控制組。

範例 你讓四十名新進的服務人員接受了完全相同的二個小時的訓練課程，並且指示這些服務人員遵循一份腳本，在這份腳本中，他們不必向顧客介紹自己的名字、而且在顧客用餐期間也不必回頭去確認顧客的反應。接著，你把這四十名服務人員隨機分為兩個各二十人的群組，並且把這兩組服務人員派到兩家餐廳的其中一家去開始上班。你記錄了所有參與者在一個月內所得到的小費數量（前測分數）。接下來，你「重新訓練」在第一家餐廳裡的二十名參與者（實驗組）。你教導他們從今以後要向顧客介紹自己的名字，並且在送上餐點之後的八到十分鐘之內，要回頭去詢問顧客「請問您對餐點覺得滿意嗎？」（實驗處遇）。你提醒在第二家餐廳裡的那一組（控制組）繼續依照之前的方式工作，也就是不必介紹自己的名字，也不用回頭去確認顧客對於餐點的反應。在第二個月過完之後，你記錄了兩個群組所得到的小費數量（後測分數）。

前實驗設計 有一些設計中缺乏隨機分派，因而是一種妥協折衷或者是抄捷徑的方案。我們會在一些難以使用古典設計的情況中使用此種**前實驗設計**（preexperimental design）。由於此種設計具有一些缺陷，因而在推論因果關係時會造成一些困難之處。

古典實驗設計 一種實驗設計的類型，其中包括了隨機分派、控制組、實驗組、並且會針對每一個組別進行前測以及後測。

前實驗設計 此種設計所指的是缺乏隨機分派、或者是抄捷徑的實驗計畫，因此此種設計的證明效果比古典實驗設計薄弱很多。當實驗者無法使用古典實驗設計中的所有面向時，可以利用此種實驗設計作為替代方案，但是此種設計的內在效度比較差。

單次個案研究設計 一種只有一個實驗組、一次後測、但是沒有前測的實驗計畫。

單次個案研究設計 **單次個案研究設計**（one-shot case-study design）也被稱作是*單組唯後測設計*（*one-group posttest-only design*），在此種設計中，只有一個群組、一次處遇、以及一次後測。由於只有一個組別而已，因此也沒有隨機分派。

範例 你找到一個由四十個新進服務人員所組成的群組，並且讓每個人都接受了一個二小時的訓練課程，在訓練課程中，你指示這些服務人員先介紹自己的名字、並且在送上餐點後的八到十分鐘內回頭詢問顧客「請問您對餐點覺得滿意嗎？」（實驗處遇）。之後，這些服務人員開始正式上班，而你則記錄下所有人在一個月期間所得到的小費數量（後測分數）。

單組前測後測設計（one-group pretest-posttest design）　此種設計中包含了一個群組、一次前測、以及一次後測，但是缺乏控制組以及隨機分派。

範例　你讓四十名新進的服務人員接受了完全相同的二個小時的訓練課程，並且指示這些服務人員遵循一份腳本，在這份腳本中，他們不必向顧客介紹自己的名字、而且在顧客用餐期間也不必回頭確認顧客的反應。接著，你把這四十名服務人員隨機分爲兩個各二十人的群組，並且把這兩組服務人員派到兩家餐廳的其中一家去開始上班。你記錄了所有參與者在一個月內所得到的小費數量（前測分數）。接下來，你「重新訓練」全部的四十名服務人員，並且教導他們從今以後要向顧客介紹自己的名字，並且在送上餐點之後的八到十分鐘之內，要回頭去詢問顧客「請問您對餐點覺得滿意嗎？」（實驗處遇）。在第二個月過完之後，你記錄了兩個群組所得到的小費數量（後測分數）。

此種設計是單次個案研究設計的改良版本，因爲你分別在施予實驗處遇之前以及之後去測量依變項。但是由於在此種設計中缺乏控制組，因此我們並不知道在前測以及後測之間，是不是有除了處遇之外的事情發生並且影響了後來的結果。

靜態群組比較　**靜態群組比較**（static group comparison）又稱爲*唯後測非同等群組設計（posttest-only nonequivalent group design）*，在此種實驗設計中有兩個群組、一次後測、以及一次處遇。此種設計缺乏隨機分派以及前測，而此種設計的缺點之一，就是兩個群組之間在任何後測結果方面的差異，有可能是在實驗之前各組之間的差異所導致的，而不是由處遇所造成的。

範例　你讓四十名新進的服務人員接受了完全相同的二個小時的訓練課程，並且指示這些服務人員遵循一份腳本，在這份腳本中，他們不必向顧客介紹自己的名字、而且在顧客用餐期間也不必回頭去確認顧客的反應。這些新進的服務人員可以自由選擇在兩家餐廳的其中一家去上班工作，但是兩家餐廳最後都一定要各有二十個人。在一個月之後，你「重新訓練」在第一家餐廳裡的二十名參與者（實驗組）。你教導他們從今以後要向顧客介紹自己的名字，並且在送上餐點之後的八到十分鐘之內，要回頭去詢問顧客「請問您對餐點覺得滿意嗎？」（實驗處遇）。在第二家餐廳裡的那一組（控制組）則被「重新訓練」繼續依照之前的方式工作，
也就是不必介紹自己的名字，也不用回頭去確認顧客對於餐點的反應。在第二個月過完之後，你記錄了兩個群組所得到的小費數量（後測分數）。

準實驗以及特殊設計　就像古典實驗設計一樣，這些設計能夠比前實驗設計更加確定地去鑑別出因果關係。在一些難以使用古典實驗設計、或者是不適合使用古典實驗設計的情境中，**準實驗設計**（quasi-experimental designs）能夠幫助我們測試因果關係。我們把此種設計稱爲*準*實驗設

靜態群組比較設計　一種有兩個群組、沒有隨機分派、而且只有一次後測的實驗計畫。

準實驗設計　一種證明效果比前實驗設計強大的實驗計畫，在此種設計中會使用古典實驗設計的變化形態，而其使用時機則是在特殊的情境中、或者是當實驗者對於自變項的控制相當有限時。

計的原因，是因爲這些設計其實是古典實驗設計的變化形態。其中有一些有隨機化但是缺乏前測、有一些使用的群組超過兩組、而有一些則以針對某一個群組的長時間觀察來取代控制組。一般而言，比起在古典實驗設計中的情況，研究者在準實驗設計中對於自變項的控制力量是比較弱的（請參閱表 9.1）。

雙組唯後測設計（*two-group posttest-only design*） 此種設計基本上和靜態群組比較幾乎完全相同，只有一個例外之處：群組是隨機分派的。除此之外，此種設計除了沒有前測之外，其餘部分都和古典實驗設計完全相同。隨機分派減少了在施予處遇之前的群組差異，但是在沒有前測的情況下，你無法確定實驗開始時所有的群組在依變項方面都是同等的水準。

林德以及史托梅茲（Rind and Strohmetz, 1999）在一份使用了有隨機分派的雙組唯後測設計的研究中，檢視了餐廳中的小費數量。在該項研究中的處遇，是在顧客帳單背面寫上一些關於即將推出的特別活動的訊息。研究中的參與者是在一家紐澤西州的高檔餐廳中用餐的八十一群顧客。研究中的處遇是一名女性服務人員是否在帳單的背後寫上一個關於餐廳即將推出的特別活動的訊息，而依變項則是小費的數量大小。研究者們拿一疊隨機洗牌過的卡片給一名有兩年工作經驗的服務人員。在這些卡片當中，有一半寫著「不必留下任何訊息」、而另一半則是寫著「留下訊息」。在這名女性服務人員把帳單拿給顧客時，她先隨機從自己的口袋中抽出一張卡片，如果卡片上寫著「留下訊息」，她就把關於即將推出的特別活動的訊息寫在帳單背面。如果卡片上寫著「不必留下任何訊息」，那麼她就不寫任何東西在帳單背面。實驗者記錄了小費的數目以及在某一桌用餐的顧客群的數目，而且實驗者也指示該名服務人員以同樣的態度來對待所有的顧客。實驗結果顯示出，收到即將推出特別活動訊息的顧客們，在給小費時會比較大方一些。

間斷性時間序列 在**間斷性時間序列設計**（interrupted time-series design）中，你會在一段時間內針對一個群組的依變項，在施予實驗處遇之前（前測）以及之後（後測）多次地去測量依變項的量數。

同等性時間序列 **同等性時間序列**（equivalent time-series design）是一種類似於間斷性時間序列設計的單一群組設計。此種設計是一種長時間的實驗設計，但是，在此種設計中所施予的處遇不只一次，而是會多次施予相同的處遇。就像間斷性時間序列設計一樣，我們會在施予處遇之前以及之後去測量依變項數次。有一項針對酒類販售以及自殺率所進行的研究（請參閱範例 9.3 間斷性時間序列、酒類販售權、以及自殺率），可以作爲說明同等性時間序列設計的例子。

拉丁方塊設計 有時候，我們所感到興趣的，是數個不同的自變項如何在不同的順序或者是時間次

間斷性時間序列設計 一種實驗計畫，在此種計畫中，依變項會在許多時間點上被週期性地加以測量，而且實驗者會在這些測量的時間點之間施予實驗處遇，通常只有一次處遇而已。

同等性時間序列設計 一種實驗計畫，在此種計畫中會長期針對一個群組進行數次重覆的前測、後測、以及處遇。

表 9.1　古典實驗設計的比較

實驗設計	隨機分派	前測	後測	控制組	實驗組
古典	是	是	是	是	是
單次個案研究	否	否	是	否	是
單組前測／後測	否	是	是	否	是
靜態群組比較	否	否	是	是	是
雙組唯後測	是	否	是	是	是
時間序列設計	否	是	是	否	是

序內去影響了一個依變項，而**拉丁方塊設計**（Latin square design）則使我們能夠去檢視此種情況。例如，有一名地理老師想要教導學生三個不同的單元：閱讀地圖、使用羅盤、以及經緯度（longitude/latitude, LL）系統。這三個單元可以用不同的次序來教導，但是這名老師想要知道哪一種次序對學生的學習效果幫助最大。在某一個班級裡，學生先學習如何閱讀地圖、然後是如何使用羅盤、最後才學習經緯度系統。在另一個班級裡，學生先學習如何使用羅盤、然後學習如何閱讀地圖、最後才學習經緯度系統。而在第三個班級裡，這名老師先教導學生學習經緯度系統，然後是羅盤的使用方法、最後才學習如何閱讀地圖。在每一個單元結束後，這名老師都會讓每一班考一次該單元的小考，而在學期結束時，則是讓學生們接受一次綜合考試。由於學生們都是被隨機分派到各班去的，所以老師就能夠看出以某種次序來教導這些單元是否能夠提升學習的效果。

所羅門四組設計　我們相信前測的量數可能會對處遇或者是依變項造成影響。前測有時候可能會使得參與者對於處遇變得較爲敏感、或者是改善了參與者在後測的表現（請參閱接下來對於測驗效應的討論）。因此，所羅門（Richard L. Solomon）發展出了**所羅門四組設計**（Solomon four-group design）來處理前測效應的問題。此種設計結合了古典實驗設計以及雙組唯後測設計，並且將參與者隨機分派到四個群組中的某一組。例如，一名精神治療師想要知道一種新的訓練方法是否能夠改善病患的應變技巧。這名治療師先進行一個 20 分鐘的測試，以測量病患的應變技巧，而測試內容則是對於壓力事件的反應。由於病患可能會從這個測試當中學到一些應變技巧，因此這名精神治療師決定使用所羅門四組設計。這名精神治療師先把病患隨機分成四個群組。其中有兩組接受了前測，然後這兩組中的一組接受了新的訓練方法，另一組接受舊的訓練方法。至於另外兩組則是沒有接受前測，而在這兩組中的其中一組接受了新的訓練方法，另外一組則接受了舊的

拉丁方塊設計　一種實驗計畫，其目的在於檢視參與者接受不同處遇的順序或是時間次序是否會造成效果。

所羅門四組設計　一種實驗計畫，在此種計畫中，參與者會被隨機分派到兩個控制組以及兩個實驗組中。其中只有一個控制組和一個實驗組會接受前測，但是四個群組全部都會接受後測。

範例　9.3

間斷性時間序列、酒類販售權、以及自殺率

許多經濟方面的利益團體對政府施加強大的壓力，要求政府允許他們從酒類的銷售當中收取更多的利潤。在加拿大西部的大多數地區，幾乎整個二十世紀當中都是由一個公立的專賣公司掌握了酒類的銷售以及分配。私有化的支持者指出了將酒類販售私有化所可能帶來的經濟利益，其中包括販售先前由政府擁有的零售管道、以及出售販賣酒類的許可執照以促進酒類的銷售量等等。其他人則指出了私有化對於酒精的飲用量以及健康所可能帶來的衝擊。針對將酒精飲料的販售權利私有化所進行的研究指出，私有化將會大幅地增加人們取得酒類飲料的方便性以及飲用量。

亞伯達省以三個階段的方式來達成酒類販售的私有化：在 1985 年開放私人經營的酒類商店；在 1989 年到 1990 年之間開放私人經營的冰啤酒商店、以及允許位於郊區的飯店銷售烈酒以及葡萄酒；最後在 1994 年全面開放所有的酒類商店的私有化。在私有化之後，亞伯達省內販售酒類的商店數目急速增加，而且烈酒的消費量也大幅度地成長，而在同一時期內，加拿大其他省分的酒類消費量卻是下滑的。亞伯達省的酒類販售私有化一直被人們認為和犯罪案件的增加有所關聯，像是非法侵入酒類販售店的案件、以及在禁止未成年人購買酒類方面執法較為寬鬆等等。亞伯達省內的酒醉駕車死亡率，在加拿大全國當中也是最高的。許多過去的研究也顯示出，在自殺率以及飲用酒精飲料之間有著強烈的關聯性。

札爾克曼以及曼恩（Zalcman and Mann, 2007）運用了一個三階段的間斷性時間序列設計，來檢視從 1976 年到 1999 年之間，亞伯達省的酒類販售私有化作法對於自殺率所造成的影響。他們所考量的，是在經過每一個私有化的階段之後，自殺率是否有所改變。除此之外，兩位研究者也針對亞伯達省的自殺率、以及相同年份內在安大略省內的自殺率進行比較，因為在安大略省內，政府一直擁有販售酒類的專賣權。

這兩位研究者發現，在 1985 年酒類零售商的私有化之後，亞伯達省內的男性與女性自殺率都上升了，其中男性的自殺率上升了 51%，而女性的自殺率則上升了 35%。在 1989 到 1990 對烈酒以及葡萄酒進行私有化之後，男性以及女性的自殺率也再度明顯地上升，據估計男性以及女性的自殺率分別增加了 17%以及 51%。而在 1994 年的最後一波私有化之後，男性的自殺死亡率再次明顯地增加，據估計大約增加了 19%，但是女性的自殺率則沒有變化。此種自殺率增加的情況，部分是短期間內自殺案例突然暴增所造成的，但是長期的自殺率仍然呈現出一種逐年上升的趨勢。藉由長時間追蹤自殺率的變化、以及與「控制組」（也就是安大略省）的自殺率進行比較的方式，兩位作者為酒類販售私有化所帶來的後果提出了證據。

訓練方法。然後這四個群組全部接受相同的後測，而後測的結果則被拿來進行比較。如果兩個接受處遇（新訓練方法）的群組擁有類似的結果、而且兩個控制組（舊訓練方法）的結果也很相近的話，那麼這名精神治療師就可以確定前測學習並未造成問題。然而，如果兩個接受過前測的群組（一個是處遇組、另一個是控制組）所得到的結果和另外兩個沒有接受前測的群組不同的話，那麼這名精神治療師就能斷定，前測本身可能會對依變項造成影響。

多因子設計　有時候我們會對兩個或者是多個自變項同時產生的效果感到好奇。**多因子設計**（factorial design）將兩個或者是多個自變項結合在一起，並且加以運用。我們所觀察的是變項中各個類別〔有時候也稱爲*因子*或是*因素*

多因子設計　一種實驗計畫，此種計畫考量了數個自變項同時發揮作用時所造成的影響。

（*factors*）〕的每一種組合。當每一個變項包含了數種不同的類別時，各種組合的數量就會增加得很快。在此種設計類型中，處遇並不是指每一個自變項，反而是指變項類別的每一種組合方式。研究者會以一種簡略的方式來討論多因子設計，像是一個「二乘三的多因子設計」就被會寫成 2 × 3，而這代表有兩個處遇（自變項），其中一個處遇有兩種類別，而另外一個則有三種類別。而一個 2 × 3 × 3 的設計，則意味著總共有三個自變項，其中一個自變項有二種類別，而另外兩個自變項則各自有三種類別。

例如，克莉桑以及同事（Krysan and associates, 2009）想要研究人們在選擇社區時的偏好，但是要同時檢視一個社區裡的種族以及社會階級特質，是一件困難的事，所以他們就使用了一個多因子設計（請參閱範例 9.4 社區偏好的多因子實驗）。在他們研究中的三個自變項分別是參與者的種族（共有二種類別：黑人或者是白人）、社區的組成方式（共有三種類別：全部白人、全部黑人、黑人與白人混合在一起）、以及社會階級（共有五個層次）。依變項則是對於某個社區的偏好程度，評分的分數是從 1 到 7。在這項研究裡的設計是 2 × 3 × 5 的多因子設計（作者們也詢問了參與者對於自己的種族團體的認同程度有多高）。

在多因子設計當中，處遇對於依變項可能會產生兩種效果：主要效果以及交互作用效果。只有*主要效果*（*main effects*）才會出現在單因子或者是單一處遇的設計當中。換句話說，我們只是單純地去檢視處遇對於依變項所造成的影響而已。而在多因子設計當中，自變項類別的特定組合，可能會產生單一因子無法造成的效果。我們把此種效果稱爲**交互作用效果**（interaction effects），因爲這是在某種組合方式中的各個類別在進行交互作用之後所產生的效果，而此種效果超越了由每一個變項單獨作用所產生的效果。交互作用效果特別受到研究者的關注，因爲此種效果顯示出，不只有單一自變項可以造成影響，不同自變項的特定組合也能產生獨特的效果、或者是某些變項只有在特定的條件下才會產生影響。

穆埃勒－強森以及戴米（Mueller-Johnson and Dhami, 2010）（請參閱範例 9.5 模擬審判以及由年齡與罪行所產生的交互作用效果）以一個類似審判的情境、以及讓參與者充當陪審團的方式，進行了一場模擬審判的研究。研究者呈現了違法者各種特質的不同組合，以瞭解這些組合對於判決結果的影響（請參閱圖 9.3）。作者們利用年齡、健康、罪行的嚴重程度、以及某個違法者先前被判決有罪的記錄，創造了一個 2 × 2 × 2 × 2 的多因子設計。他們發現罪行的嚴重程度、年齡、以及先前的判刑記錄都具有主要效果。換句話說，那些犯下嚴重罪行的人、較年輕的違法者、以及那些之前曾經被判刑過的人，會比那些犯下輕微罪行的人、較年長的違法者、以及之前沒有任何判刑記錄的人，被判處更長的刑期。研究者們也發現了一些交互作用效果，其中之一是對於那些先前有過判刑記錄的人，年齡以及罪行嚴重程度會產生交互作用效果。

交互作用效果　兩個自變項結合起來並且同時運作後，對於依變項所造成的效果。此種效果大於每個自變項分開運作時的效果總和。

範例　9.4

社區偏好的多因子實驗

克莉桑以及同事（Krysan and associates, 2009）進行了一個實驗，以研究美國的黑人以及白人對於社區的偏好情形。雖然過去也曾經有一些研究檢視了這個議題，然而，同時檢視種族以及社會階級因素是非常困難的，而且也幾乎不可能去判斷人們選擇了某個社區的原因到底是為了該社區的社會階級特質、還是為了種族方面的特質。作者們提到：「以下兩個問題是我們研究分析的核心：（1）人們對於社區的偏好究竟是與膚色無關的、還是具有種族意識的？（2）如果對於社區的偏好是與種族意識有關的，那麼此種偏好所反映出的是人們想要與自己同類的人住在一起的渴望、或者是想要避免與其他種族團體住在同一社區的欲望？」（p. 529）。在 2004 年到 2005 年之間，作者們在底特律地區挑選了超過 700 名參與者、而在芝加哥的都會地區也挑選了將近 800 名參與者。為了要區隔出在社區裡的階級效果以及種族效果，作者們讓參與者看了一些影片，影片內容是由不同的社會階級以及種族團體所居住的社區。作者們總共製作了十三部影片，而在影片中的社區居民的組合方面，社會階級有五個層次、種族混合狀況則有三個層次。

> *我們挑選出不同的社區以傳達不同的社會階級層次，我們所依據的假定是受訪者會根據一些特質來推論出社會階級，像是房子以及住宅的大小、房屋的維護狀況、以及其他由觀察所得到的線索等等。接著，在每一個不同的社區裡，其中的居民所顯示出的種族則有三種不同的變化：（1）所有的居民都是白人；（2）所有的居民都是黑人；（3）有三個居民是白人、而另外二個居民則是黑人。（p. 537）*

其中一部影片當作控制組，裡面沒有顯示出居民。而在其他的每一部影片裡，都有五個人（作者所找來的演員）出現在影片中，充當社區的居民並且從事一些日常活動。作者們提到：

> *在每一個社區裡，我們都會安排一個場景，在這個場景中，會有三個人一起在車道上、在房子前面的庭院裡、在郵箱旁、或者在一部正在維修的車子周圍一起聊天。這些居民們穿著短袖的襯衫、但是沒戴帽子，如此一來，就能夠讓受訪者更加容易看出這些居民的種族身份。在每一個社區裡，屬於相同社會階級裡的居民都會以相近的年齡、性別、以及服裝樣式來進行配對。*

為了要進行操弄確認，作者們在實際進行研究之前，先讓一小群其他的參與者看了這些影片，以確認人們會如同研究者們所預想地那樣，能夠看出影片中的社區裡的社會階級以及種族的組成方式。在看完影片之後，作者們要求參與者在一份 7 點的李克特量表上對每一個社區進行評分，分數是從非常喜歡到非常不喜歡。作者們說（p. 539）：「我們的依變項，就是對於四個社區的喜歡程度評分，因此，我們的分析單位就是影片。由於每一個受訪者都看過也都評分過相同的基準影片——一個沒有居民的上層勞動階級社區，所以我們把對於這個社區的評分也包括進來，當作是一個受訪者層次的控制。」作者們使用了多因子設計，其中有三個自變項：研究參與者的種族、社區的社會階級、以及社區中的種族組成方式。作者們隨機分派參與者去觀看相同社區裡的不同種族組成方式的影片。在他們所得到的許多發現中，作者們提到（p. 538）：「我們最重要而基本的結論就是，就其本身而言，種族形塑了白人看待居住空間的方式，但對於黑人的影響則比較小一些。對於居住地區的偏好，並非單純只是對於某個社區裡的社會階級特質所作出的反應，而是受到住在該社區裡的人們的種族所影響的。」

設計記號　一種被用來標示實驗中各個部分、以及畫出各個部分示意圖的符號系統。

設計記號

我們可以用許多種方法來設計實驗，而**設計記號**（design notation）是一種以各種符號來代表實驗

範例　9.5

模擬審判以及由年齡與罪行所產生的交互作用效果

穆埃勒－強森以及戴米（Mueller-Johnson and Dhami, 2010）創造出了一個模擬審判的情境。他們創造了一個類似於審判的情境，並且請參與者擔任陪審團的團員。作者們呈現了違法者各種特質的不同組合，以瞭解這些組合對於陪審團判決結果的影響，而判決的結果是指在監獄中服刑的刑期長度。實驗中的陪審團團員是由來自於英國一所大學裡的 47 名學生（36 名女性以及 11 名男性）所擔任。作者們利用年齡、健康、罪行的嚴重程度、以及某個違法者先前被判決有罪的記錄，創造了一個 2×2×2×2 的多因子設計。在過去的實驗中，他們發現了健康、先前的判刑記錄、以及罪行的嚴重程度都具有主要效果。健康情況不佳的人，被判的刑期比較短；而不論之前被判有罪的次數多寡，年紀較大的人（66 歲到 72 歲）被判處的刑期，都會比年紀較小的違法者（21 歲到 26 歲）所被判的刑期來得短。先前曾有被判刑記錄、以及犯下的罪行較為嚴重的年輕違法者，則會被判處較長的刑期。在當前的研究中，他們所調查的是兒童性侵害案件的犯法者。在先前被判決有罪記錄的這個變項中，其類別分別是沒有任何記錄、或者是曾與四歲以上的兒童有過性行為；而在罪行嚴重程度這個變項中，其類別分別是隔著衣服去觸摸一名七歲女童的生殖器官一次、或者是在一年的期間內觸摸裸體七歲女童的生殖器官十次。參與者通常在 15 分鐘之內就能決定判決的結果。作者們發現在年齡、罪行的嚴重程度、以及先前的判刑記錄之間出現了有趣的交互作用效果。對於曾經有過被判刑記錄的違法者而言，年紀較大的人所被判處的刑期，會比一個罪行較輕的年輕人還要長，但是會比一個罪行較重的年輕人還要短。也就是說，對於年紀較大的違法者而言，當先前的判刑記錄以及較輕的罪行結合在一起時，會導致被判的刑期較長。而此種情形也符合了「下流的老男人」的刻板印象。

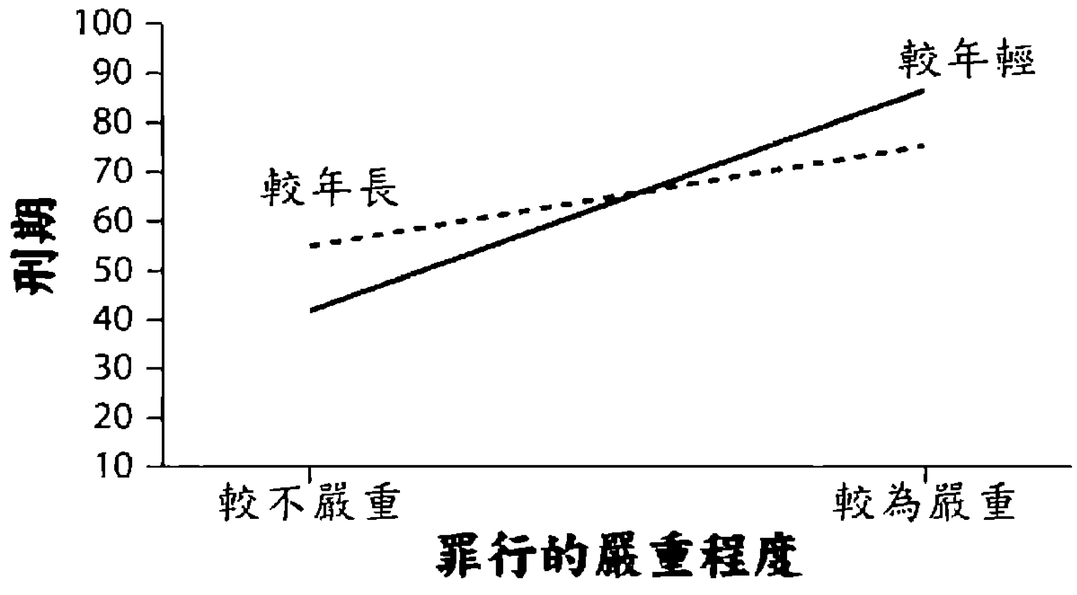


圖 9.3　在模擬審判中，對於曾有犯罪記錄的性侵害犯人所判處的刑期

設計中的各個部分的速記系統。[7] 利用排成兩列的五到六個符號，此種系統就可以表達出原本要使用複雜的、整段式的敘述才能說明的實驗中的各個部分。一旦你學會了設計記號，你會發現在思考以及比較不同的設計時會比較容易一些。設計記號使用了下列的這些符號：O = 對依變項的觀察；X = 處遇、自變項；R = 隨機分派。各個不同的 O 依照時間順序，利用下方小字標號從左到右來表示其順序。前測就是 O_1，而後測就是 O_2。當自變項有超過兩個以上的層次時，各個 X 就會用小字標出號碼來作出區隔。符號都是依照時間順序由左到右排列。R 是第一個，其後則是前測、處遇、然後是後測。我們以列的方式來排列符號，而每一列則代表了一群參與者。例如，一個有三個群組的實驗的表示方式是在 R（如果有使用隨機分派的話）後面跟著三列的 O 以及 X。

各列之間是上下對齊排好的，這是因爲在每一個群組中的前測、處遇、以及後測是在大約同一個時間內進行的。表 9.2 展示了許多標準實驗設計的記號。

內在與外在效度

內在效度的邏輯

當除了自變項之外沒有其他事物對依變項造成影響時，就是具有**內在效度**（internal validity）。除了自變項之外，任何會影響依變項的事物，都會對內在效度造成威脅。這些事物就是混淆變項，因爲它們混淆了實驗的邏輯。而實驗的邏輯就是要把除了變項之間關係以外的所有其他事物，通通從你的假設當中排除。混淆變項使你無法確定地認爲，處遇就是造成依變項產生變化的真正因果因素。你或許曾經聽過有人把混淆變項稱爲**人工製品**（artifacts），這是因爲這些令人討厭的、或者是造成混淆的變項並非來自於你所檢視的自然關係，而是來自於特定的實驗安排。

內在效度　實驗者強化一項因果解釋的邏輯嚴謹性的能力，其方法是透過實驗設計來排除在處遇與依變項之間的關聯性的潛在替代解釋。

人工製品　實驗研究中的一種事物，所指的是一種混淆變項的類型，此種混淆變項並非是假設的一部分，但是卻會影響到實驗的操作或是結果。在田野研究中，人工製品所指的是由人類所創造出來的、具有文化意義的實體物體。更具體地說，人工製品所指的是某個文化中的成員們所使用的、或者是賦予意義的物體，而我們則會去研究此種被賦予的意義，以更加瞭解一個文化的背景或者是其中的成員。

選擇偏誤　一種會威脅到內在效度的偏誤，發生此種問題的原因，是在實驗一開始時，對於依變項而言，所有的群組並不是相等的。

人工製品之所以會偶然地出現，是因爲在研究的準備過程當中，你可能在無意之間引入了某個改變了事物的東西。例如，在實驗的參與者抵達之前，你先清理了一個房間，而這個實驗的主題是一直持續不睡覺的狀態對於情緒所造成的效果。但是你用來擦拭桌椅的清潔劑會對許多人造成過敏的現象。於是，你的實驗結果顯示出，睡眠不足的人會出現較爲嚴重的過敏現象。然而，此種過敏的現象並不是來自於睡眠不足所造成的，而是你使用的清潔劑所造成的非預期之中的副作用。你應該要排除人工製品以及混淆變項——也就是除了處遇之外，所有可能會影響依變項的事物。而要排除人工製品或者是混淆變項的方法，就是去控制實驗情境以及利用實驗設計。接下來，我們將要檢視幾個對於內在效度的主要威脅。

對於內在效度的威脅

以下是對於內在效度的 12 種威脅。[8]

1. *選擇偏誤*　當一項實驗裡有超過一組以上的參與者時，就有可能會產生**選擇偏誤**（selection bias）。你想要對所有的群組進行比較，但是群組之

表 9.2　實驗設計記號摘要

設計名稱	設計記號
古典實驗設計	O　X　O R O　O
前實驗設計	
單次個案研究	X　O
單組前測／後測	O　X　O
靜態群組比較設計	X　O O
準實驗設計	
雙組唯後測	R → X　O O
間斷性時間序列	O　O　O　O　X　O　O　O
同等性時間序列	O　X　O　X　O　X　O　X　O
拉丁方塊設計	R O　X_a　O　X_b　O　X_c　O O　X_b　O　X_a　O　X_c　O O　X_c　O　X_b　O　X_a　O O　X_a　O　X_c　O　X_b　O O　X_b　O　X_c　O　X_a　O O　X_c　O　X_a　O　X_b　O
所羅門四組設計	R O　X　O O　O X　O O
多因子設計	R X_1　Z_1　O X_1　Z_2　O X_2　Z_1　O X_2　Z_2　O

註：字母的下標符號表示不同的處遇變項。數字的下標符號表示同一個處遇變項的不同類別，例如，在性別變項裡的男性或是女性。

間卻有一些差異、或者是並未形成相同的群組。此種問題會出現在沒有隨機分派的設計中。例如，你設計了一個研究侵略性的雙組實驗。如果你沒有使用隨機化、或者是隨機化並未發揮效果，處遇組就可能偶然地出現差異。你可能找到六十個在各種校園

歷史效應 一種對於內在效度的威脅，此種威脅發生的原因，是因為在實驗過程中發生了某些事，並且影響了依變項。歷史效應不在實驗者的規畫之中，而且也在實驗者能夠控制的範圍之外。

成熟效應 一種在實驗中對於內在效度的威脅，發生此種威脅的原因，是因為在實驗過程中出現了成長、厭倦等等的自然過程，並且影響了依變項。

測驗效應 一種對於內在效度的威脅，產生此種威脅的原因，是因為在前測中所進行的測量過程，可能會對依變項造成影響。

活動中都相當主動積極的研究參與者。在偶然的情況下，在你的實驗組中的許多自願者可能已經參加過足球、橄欖球、曲棍球、以及摔角等活動，而在你的控制組當中的自願者，則可能是音樂家、象棋社的成員、芭蕾舞者、以及畫家等等。另外一個與選擇偏誤有關的例子，則是一個針對人們躲避交通壅塞的能力所進行的實驗。如果被分到某一組的人都是來自於交通狀況較爲單純的郊區、而被分到另一組的人則是成長於交通繁忙的大都市裡的話，就可能會發生選擇偏誤的問題。通常，你可以藉由比較前測分數的方式，來偵測出選擇偏誤。如果在前測的分數當中，你沒有看到群組之間的分數有任何差異時，那麼很可能就沒有選擇偏誤的問題。

2. *歷史* **歷史效應**（history effect）所指的是在實驗過程中發生了與處遇無關的事件，並且影響了依變項。歷史效應比較有可能出現在長期的實驗當中。例如，有一項爲期兩個星期、主題是評估對於寵物狗的感覺的實驗，在實驗進行到一半時，附近一家犬舍突然發生大火，有許多狗被燒死或者是受傷，而在新聞報導的畫面中，則顯示出許多受傷的動物、以及許多因爲這場意外而痛哭的地方人士。

3. *成熟* **成熟效應**（maturation effect）這個威脅是指在實驗過程當中，參與者發生一些除了處遇以外的生理、心理、或者是情緒的過程，並且影響了依變項。根據依變項以及研究設計的不同，成熟效應發生的時間可能是數小時、數個月、或者是數年。例如，在一個一整天八個小時的、針對推理能力所進行的實驗裡，參與者可能會覺得厭煩或者是昏昏欲睡，結果使得他們的分數變低。另外一個例子，則是對於一到六年級的兒童遊戲方式所進行的實驗。兒童們在生理的、情緒的、以及成熟度方面的改變，會影響到他們的遊戲方式，而這些改變之所以會發生，可能是兒童們逐漸長大所致，不一定是因爲處遇所造成的影響。包含了前測以及控制組的設計，有助於決定在實驗當中是否出現了成熟效應或者是歷史效應，這是因爲隨著時間的流逝，實驗組以及控制組都會表現出類似的改變。

4. *測驗* 有時候前測量數本身也會影響到實驗。此種**測驗效應**（testing effect）會威脅到內在效度，因爲除了處遇之外，還有其他的事物也影響了依變項。所羅門四組設計可以幫助我們偵測出測驗效應。例如，你進行了一項前測，以確定參與者對於地質學和地理學的知識有多少。而你的處遇是在兩天的時間之內，讓參與者觀看一系列關於地質學和地理學的影片。如果參與者記得前測的題目，而且此種情況影響了他們所學到（也就是注意到）的知識、或者是他們在後測中回答題目的方式，就是出現了測驗效應。如果發生了測驗效應，你就不能說只有處遇影響了依變項，因爲對於前測的記憶以及處遇兩者都對依變項造成了影響。

5. *工具*　此種威脅與穩定信度有關。當*工具*（*instrument*）或者是依變項的測量方式在實驗過程中改變了，就會產生此種威脅。例如，在一個減重的實驗裡，體重計的彈簧在實驗過程中鬆掉了，因此在後測中量出的體重數字就變低了。另外一個例子，則是在實驗當中的處遇是讓參與者去觀看一部影片，但是有一些播放設備壞掉了，因此造成部分的參與者無法觀看這部影片。

6. *實驗對象流失*　當一部分的研究參與者無法持續參與整個實驗過程時，就會產生**實驗對象流失**（experimental mortality）或者是減損效應（attrition）的問題。雖然*流失*（*mortality*）這個字的本意是指死亡，但是這不必然是指參與者已經過世了。如果許多參與者在實驗過程中的某個階段就離開的話，我們就無法知道如果他們留下來的話，最後的結果是否會有所不同。例如，你一開始的時候找到了六十個人來進行一項減重的實驗。而到了實驗結束時，只有四十個人留下來，而且各自在沒有副作用的情況下減少了 5 磅。在實驗中途就離開的那二十個人的情況可能與留下來的這四十個人不同，並且因而改變了整個實驗結果。也許這個實驗對於那些中途離開的人也很有效，而且這些人是在減了 25 磅之後才離開的。或許，這個實驗可能讓他們覺得不舒服並且使得他們不得不提早退出，或者是他們覺得沒有效果，所以決定提早退出。我們必須注意並且報告在實驗所有階段中的參與者人數，以偵測出這個對於內在效度的威脅。

7. *統計回歸效應*　統計回歸效應是一個不容易從直覺上理解的效應，這是一種極端值的問題、或者是隨機誤差把群組結果移向平均值的趨勢。統計回歸效應發生的方式有以下兩種。

發生**統計回歸效應**（statistical regression effect）的情況之一，是當參與者在依變項方面是屬於異常情況的時候。由於此種參與者較爲特殊或是極端，因此就不太可能在某一個方向作出更進一步的反應。例如，你想要知道玩暴力的電動遊戲是否會使得人們變得更具攻擊性。你的參與者是一群從高度戒備的監獄中所挑選出來的暴力罪犯。你對他們進行前測、讓他們去玩極端暴力的電玩遊戲 60 個小時、然後再進行後測。有可能出現的情況是，這些罪犯一開始就具有極端的攻擊性，因此，你的處遇就無法使得他們更具有攻擊性了。僅僅受到隨機的機會所影響，有些人在後測接受測量時，甚至可能會出現攻擊性降低的情況。[9]

第二種會出現統計回歸效應的情況，則牽涉到測量工具的問題。如果在你進行測量時，大多數的人在某個變項上的得分都非常高（到了最高分的程度）或是非常低（到了最低分的程度），那麼，僅僅由於隨機的機會，就可能在前測以及後測之間產生變化。例如，你讓八十位參與者接受了一次簡單的數學測驗，而其中有七十七位參與者得到滿分。你對這些參與者施予處遇，以提高他們的數學分數。然而，由於有這麼多的參與者已經得到滿分，隨機誤差就有可能會降低群組平均分數，這是因爲那七十七位已經得到滿分的參與者只能往一個方向移動，也就是答錯一題，而

實驗對象流失　由於某些參與者無法全程參與整個實驗，因而對內在效度造成威脅。

統計回歸效應　一種對於內在效度的威脅，此種威脅來自於測量工具產生了極端的數值，以及一種隨機誤差把極端結果移往平均值的趨勢。

且只有三個人的分數有可能會提高。如此一來，僅僅由於機率因素使然，在後測中的群組平均分數可能會顯得比較低。你必須監視分數的分佈範圍，才能夠偵測出統計回歸效應。

8. *處遇擴散或污染* **處遇擴散**（diffusion of treatment）的威脅，是指在不同組別中的研究參與者彼此聯絡交流，並且得知了其他人所接受的處遇。避免此種威脅的方法，就是把不同的群組隔離開來，或者是要求研究參與者承諾不會把任何事情透露給其他的參與者知道。例如，在一個持續進行一整天的實驗中，你找了八十名參與者來實驗記憶單字的一些方法。處遇組被教導了一個簡單的方法，而控制組則被告知去使用任何一種他們想要使用的技巧。在中途的休息時間裡，在處遇組裡的成員把關於這個新方法的事情告訴了在控制組裡的參與者。在休息時間結束之後，控制組也開始使用這個新方法。你或許可以在實驗結束之後對參與者進行訪談時，詢問他們有關可能發生的擴散情況，以減少這項威脅。

9. *補償行為* 在某些實驗中，當實驗者提供了某種有價值的物品給某一組的參與者、但是卻沒有給另一組的參與者，而且此種差別待遇被參與者發現了之後，就可能會發生**補償行為**（compensatory behavior）。此種在群組之間的不平等狀況可能會造成一種想要減少差異的欲望、群組之間的競爭對立、或者是引起了怨恨性的怠工（resentful demoralization）。除了處遇之外，此種行為也可能會影響到依變項。例如，某一所學校裡的學生接受了一項處遇，也就是午餐休息時間變長，以提升在學習方面的成效，但是在另外一所學校裡，學生的午餐休息時間仍然維持正常。一旦此種不平等的狀況被學生知道的話，在控制組（也就是午餐休息時間維持不變的那一所學校）裡的學生可能會特別努力地學習，以克服此種不平等的狀況。另外一種可能發生的狀況，則是在控制組裡的學生因為此種不平等的處遇而發生了怨恨性怠工的情況，並且在學習方面所付出的努力也跟著變少。除非你能夠取得外界資訊（請參閱之前關於處遇擴散的討論內容），否則很難去偵測到這項威脅。

處遇擴散 一種會對內在效度造成威脅的擴散情況。當處遇從實驗組中「溢出」（"spills over"），而且控制組中的參與者因為得知了處遇而修正了他們的行為時，就會發生此種威脅。

補償行為 一種會對內在效度造成威脅的行為。當控制組的參與者修正了他們的行為，以補償自己沒有獲得處遇的情況時，就會產生此種威脅。

實驗者期望 一種實驗反應的類型，發生此種狀況的原因，是因為實驗者間接地使得參與者意識到實驗的假設或者是預期得到的結果。

10. *實驗者期望* 如果實驗者的行為間接地傳達了預期的結果時，實驗者的行為就有可能威脅到內在效度，[10] 而此種情況則被稱作是**實驗者期望**（experimenter expectancy）。由於對假設抱持著強烈的信心，即使是最誠實的實驗者也可能會在無意之間傳達出想要得到的發現。例如，你想要研究參與者對於身心障礙人士的反應。你深深地相信，女性對身心障礙者會比男性更加敏感。透過眼神接觸、聲音的音調、暫停、以及其他的非言語溝通方式，你可能在無意中鼓勵了女性的研究參與者回報她們對於身心障礙人士擁有正向的感受，而你對於男性參與者的非言語行為則剛好相反。

雙盲實驗　一種實驗研究的類型，在此種實驗中，參與者以及代表實驗者去直接與他們接觸的人，都不知道實驗的細節為何。

要求特徵　一種實驗反應的類型，當實驗研究中的參與者得知了關於假設的線索，並且因此而改變自己的行為時，就會產生此種對內在效度的威脅。

雙盲實驗（double-blind experiment）正是一種用來控制實驗者期望的設計。在此種實驗中，那些唯一能與參與者直接進行接觸的人，並不知道研究假設以及處遇的細節。此種設計之所以被稱爲是*雙盲*的，是因爲參與者以及和他們接觸的人全都不知道實驗的細節爲何（請參閱圖 9.4）。例如，你想要知道某種新的藥物是否具有效果。你在研究中使用了綠色、黃色、以及粉紅色的藥丸。你把新的藥物放在黃色的藥丸中、把舊的藥物放在粉紅色的藥丸中、而綠色藥丸則是*安慰劑*（*placebo*）（也就是不具有任何效果的處遇）。幫你分送藥丸以及記錄各種顏色藥丸效果的助手並不知道哪一種顏色的藥丸中含有新的藥物。只有你自己才知道哪一種顏色的藥丸中含有新的藥物並且檢查結果，但是你與研究參與者之間沒有任何接觸。在醫學研究當中幾乎都會強迫使用雙盲設計，因爲在此種研究中很有可能會出現實驗者期望效應。

11. *要求特徵*　另一個對內在效度的威脅則是與實驗反應（reactivity）（將在本章的下一個部分進行討論）有關，而此種威脅則被稱爲**要求特徵**（demand

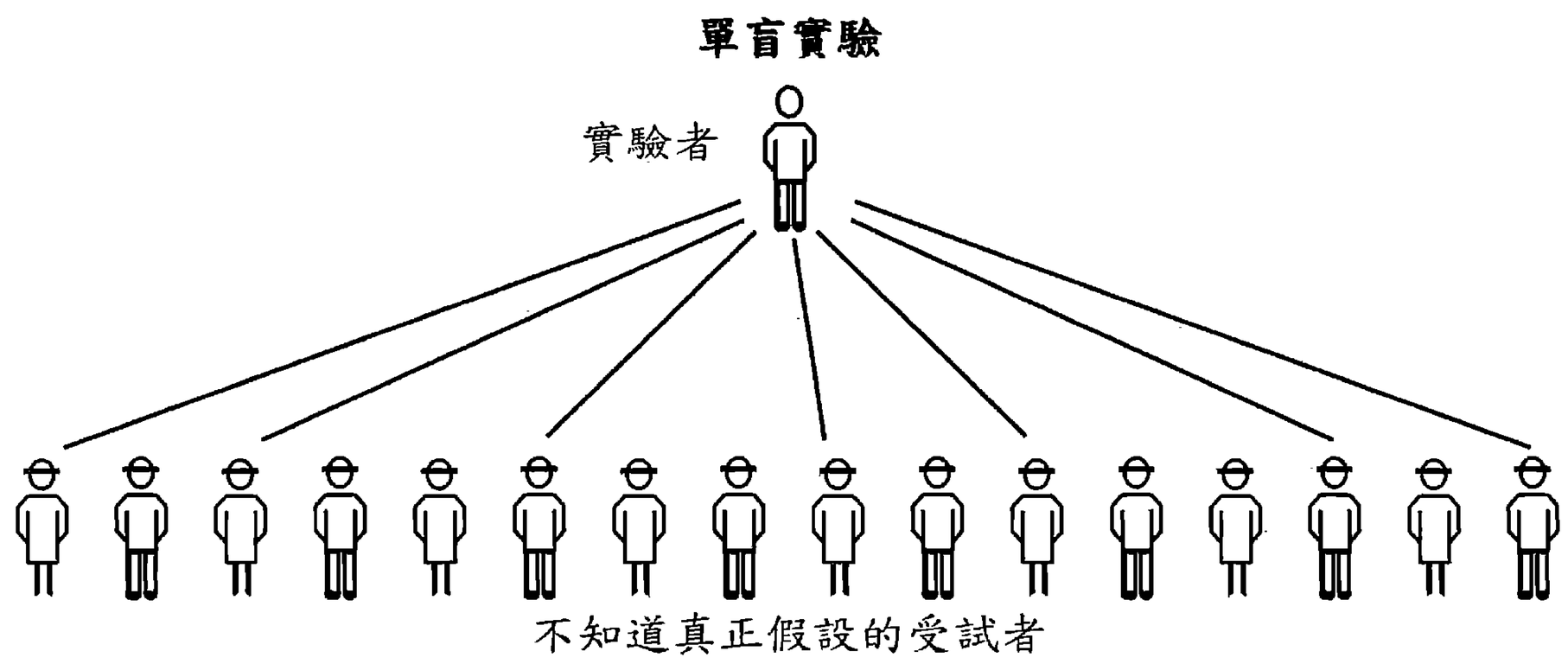


圖 9.4　雙盲實驗：單盲（普通）實驗、以及雙盲實驗的說明圖例

characteristic）。當研究參與者無意間得知關於假設或者是實驗目標的線索、並且據此修正自己的行爲以符合他們認爲研究對他們的要求（也就是支持假設）時，就會產生此種威脅。參與者經常會作出上述的事情來取悅研究者，而這也正是我們爲什麼通常會以封面故事的形式來使用一些溫和的欺瞞策略。

12. *安慰劑效應*　對內在效度造成威脅的最後一種類型，就是**安慰劑效應**（placebo effect）。所謂的安慰劑，就是一種空的、或者是沒有作用的處遇，像是在醫學研究中所使用的含糖藥丸。當你給一些參與者安慰劑、但是他們的反應卻好像他們接受了真正的處遇時，就會發生安慰劑效應。例如，你進行了一項幫助老煙槍戒煙的實驗。你給一些參與者一種藥丸，裡面含有幫助他們減少尼古丁依賴性的抑制尼古丁藥物，而其他的參與者所得到的則是安慰劑（空的藥丸）。如果接受安慰劑的參與者也停止吸煙了，那麼，單純參加這項實驗、並且服用某種他們認爲可以幫助自己戒煙的藥物，就對這些參與者產生了效果。僅僅因爲相信安慰劑具有效果，就可能會影響到依變項（關於對內在效度的威脅的摘要，請參閱表 9.3）。

實驗者通常會進行操弄確認以增加內在效度。**操弄確認**（manipulation check）所指的是一個確認在理論上顯著的變項（例如，在假設中的自變項、依變項、以及中介變項）的過程，其目的在於確認測量效度（例如，變項是否能夠真的測量到理論的概念）、實驗的情境是否具有預期達到的效果、或者是實驗真實性（實驗真實性的問題將在本章稍後部分進行討論）的程度。我們利用操弄確認來確定在我們實驗中的變項以及條件如同我們所預期般地正常運作，並且幫助我們排除對於內在效度的可能威脅。

我們會利用前測、前導測試、以及實驗的簡報來確認我們的「操弄」（也就是我們在一個實驗情境中的測量方式以及干預）。我們可能會針對某些實驗條件創造出一次前測。例如，你要求一位實驗中的同謀假裝他或她是一名身心障礙人士，並且要求初步的研究參與者去觀察這名同謀。爲了進行確認，你去詢問這些參與者是否相信該名同謀是真的有身心障礙的問題、或者只是在演戲而已。在社區偏好的研究中（請參閱範例 9.3），研究者在把一些關於不同社區的影片實際運用於研究當中之前，就先讓一小群參與者觀看過這些影片。而這麼做的目的，是爲了確認人們會像研究者所預期般地能夠辨認出影片中的居民種族差異、以及社區的社會階級。如果你計畫在實驗中提供參與者一些書面或是口頭上的指示，你可能會先找一些初步的參與者來對這些指示進行前測。你可以詢問這些參與者你的指示是否夠清楚、以及他們是否如你所預期般地瞭解這些指示。

實驗程序進行一次「模擬演練」或是前導測試，也算是一種操弄確認。在前導測試的過程當中以及之後，你可以去尋找潛在的缺陷、錯誤、或者是容易造成誤解的地方。你可以看看實驗情境的每一個部分是否都能夠順利進行，並且也對參與者產生了你所預期的效果。除此之外，如果你使用了欺瞞的

安慰劑效應　當參與者並未接受真正的處遇、只是接受了一個沒有作用的處遇，但是其反應卻好像他們接受了真正的處遇時，就是發生了安慰劑效應。

操弄確認　一種單獨去測量自變項或者是依變項的作法，以確認這些變項的測量效度以及／或是實驗的真實性。

表 9.3　內在效度與外在效度的議題

內在效度	外在效度
選擇偏誤	母體推論性
歷史效應	自然推論性
測驗效應	理論推論性
成熟效應	生活真實性
工具	實驗真實性
實驗受試者流失	霍桑效應
統計回歸效應	
處遇擴散	
補償行為	
實驗者期望	
要求特徵	
安慰劑效應	

話，你也可以確認一下參與者是否注意並且接受你所說的「封面故事」。

在前導測試或者是實際的實驗結束之後所進行的實驗簡報（experimental debriefing）也可能是一種操弄確認。在進行實驗的簡報時（與倫理的簡報不同，因爲倫理的簡報所強調的是移除謊言或是欺瞞），你會針對實驗的細節去訪談參與者。你想要瞭解的，是他們認爲發生了什麼事、他們是否覺得全心投入並且認真地看待實驗情境、以及他們是否感受到任何的困惑、焦慮、或者是不舒服。你可能在此種訪談中討論補償行爲以及要求特徵、或者是處遇擴散的問題。有時候，如果實驗者發現某位參與者誤解了實驗的一個關鍵面向、看穿了欺瞞的封面故事、或者是因爲要求特徵而改變了其行爲（請同時參閱本章稍後部分對於實驗反應的討論）的話，實驗者可能會把該名參與者從研究資料當中排除。例如，如果有一名參與者透露說他或她並不接受研究者所說的封面故事，但是相信該項研究是關於對身心障礙人士的反應（這確實是真正的研究目的）、並且依據此種想法而作出反應時（也就是展現出要求特徵），實驗者或許就必須把這名參與者的資料移除（請參閱範例 9.6 誰願意幫助一名身心障礙的同事？）。

外在效度與田野／實地實驗

即使我們排除了所有與內在效度有關的問題，外在效度仍然會是一個潛在的議題。**外在效度**（external validity）是指把實驗發現向外推論的能力與有效性。如果一項研究缺乏外在效度，該研究的發現可能只有適用於一個特定的實驗而已。由於我們在基礎研究中所尋找的是一般的理論知識、而在應用研究中所尋找的是能夠連結到真實生活問題

外在效度　將研究發現推論到特定的研究之外的能力。

範例 9.6

誰願意幫助一名身心障礙的同事？

米勒以及韋納（Miller and Werner, 2007）利用兩種處遇條件以及一個控制組，針對助人行為進行了一個實驗室的實驗。作者們想要知道，哪一種類型的人比較有可能去幫助一名有身心障礙的同事。過去的研究已經發現，在人格特質以及對於身心障礙人士的態度之間存在著正向的關係。米勒以及韋納測量了三種人格特質：對於公平的偏好、女性特質、以及印象管理。公平偏好（equity preference）所指的是認為每個人都必須完成一定份量的工作才能得到報酬。有些人是比較仁慈的（也就是說，只要是努力工作的人，就算生產力比較差，也應該得到同等的報酬），而有些人則比較注重權利（也就是說，如果有人做的工作比較少的話，就絕對不能得到比較多的報酬）。傳統的女性特質（feminine traits）所指的是和善、樂於助人、以及同情理解的特質。印象管理（impression management）所指的是有意識地把自己的某一方面表現給別人看到。在印象管理這個項目得到高分的人，在公共場合中會有意識地表現出符合他人預期的形象。作者們在大學部的企業管理課程當中找到了超過 500 名學生，並且要求這些學生完成了一份測量人格特質的調查問卷。從這些學生裡，作者們挑選出了 133 名自願者來進行實驗。他們操弄了三種層次的身心障礙情況，也就是他們主要的自變項：沒有障礙、心理障礙、以及肢體障礙。他們也進行了一次操弄確認，方法是去要求另外一個有 84 名參與者的群組去閱讀關於各種不同人們的描述，並且去評定這些描述中所提到的人們是肢體障礙的、心理障礙的、或者是沒有任何障礙的。作者們提到（p. 2668）

> *為了減少混淆變項，在實驗中的每一個階段都使用了完全相同的同謀。這名同謀是一名男性的研究生，而且就讀於該所大學裡與商業課程無關的博士班。之所以使用相同的同謀，是因為如此一來，在種族、外表吸引力、個性、以及其他特質方面就不會有任何變化，也不會因此而造成參與者在回應方面的差異。這名同謀是一名體型瘦削的 25 歲白人學生。*

在一開始的時候，每一位參與者以及這名同謀都一起準備並且閱讀一份自傳。在肢體障礙的情境中，這名同謀坐在輪椅上，而其自傳中則提到他之前曾經發生過一場車禍意外，因而使得他從此之後都必須坐在輪椅上。而在心理障礙的情境中，這名同謀則表現出無法勝任撰寫自傳的工作，並且告訴其他人說他曾經發生過一場車禍意外，因而使得他腦部受傷，並且有一些短期記憶方面的障礙。接著，參與者以及這名同謀被要求去完成一項複雜的工作，工作內容是把紙摺好並且放入信封內。這項工作需要用到一些肢體動作以及心算的能力。每個人都被告知，完成這項工作後會得到報酬，而且必須在剛剛好五分鐘之內完成工作。要完成這項工作，動作必須要相當迅速，但是要在指定的時間之內完成這項工作，其實是相當容易的。在肢體障礙的情境中，坐在輪椅上的同謀無法移動身體來完成這項工作。而在心理障礙的情境中，對於在時間限制內完成這項工作所必須擁有的心算能力，這名同謀則表示出有極大困難的樣子。而在沒有任何障礙的情境中，這名同謀則僅僅是動作比較緩慢而已。在三種情境當中，參與者們都可以清楚地看出來這名同謀無法在時間內完成工作。而依變項則為是否有任何參與者去幫助這名同謀。研究者們將整個實驗過程錄影下來，並且找了一名受過訓練的、獨立的觀察者，針對參與者對於這名同謀所提供的協助數量多寡來加以評分。研究結果顯示出，公平偏好以及印象管理會對依變項造成影響，但是女性特質則沒有影響。在仁慈的公平偏好以及印象管理方面得到高分的人，比較願意去幫助別人。在肢體障礙的情境中，伸出援手的人比在心理障礙的情境中還要多；但是在這兩種障礙的情境中，伸出援手的人都比在沒有任何障礙的情境中來得多。在實驗之後所進行的簡報訪談階段中，研究者告訴參與者們關於研究的真正目的，並且詢問他們認為研究的主題為何。研究者後來摒棄了五名參與者的資料，「因為在他們所提出的評論中，透露出了他們的評分可能是有偏差的。像是此類評論的例子包括：『我覺得那一個身心障礙的學生只是一個誘餌』、『我認為你們想要看看我們對於一個身心障礙的人會有什麼反應』等等……」（p. 2671）

的研究發現，因此，缺乏外在效度的研究發現就會變得幾乎毫無用處可言。然而，在〈捍衛外在無效度〉（*In Defense of External Invalidity*）這篇被廣泛引用的文章裡，穆克（Mook, 1983）認爲，從實驗向外推論到自然的、真實生活的環境，並不是許多實驗的目標。相反地，我們或許還有其他的理論目的（請參閱稍後關於理論推論性的部分）。

外在效度的議題可能會相當複雜。的確，就如同泰伊（Thye, 2007: 81）所說的：「在與實驗相關的議題中，或許受到最多誤解的，就是外在效度了。」而造成此種情況的原因，就在於外在效度會牽涉到數種不同形式的推論性。[11] 外在效度所處理的是以下三個關於推論性的問題：我們可以從實驗中一組特定的參與者來推論到整個母體嗎？我們可以從一個受到高度控制以及人爲的實驗情境推論到較爲自然的、「真實世界」的情境嗎？我們可以從一個特定實驗的實徵證據，推論到在變項之間關係的抽象理論模型嗎？爲了處理這些問題，我們可以把外在效度想成涉及到以下三種形式的推論性，而這三種形式的推論性不一定總是會互相重疊：母體的、自然的、以及理論的（請參閱圖 9.5）。

母體推論性　此種形式的外在效度的關鍵問題，就是我們是否能夠正確地把我們在一項研究中從一組特定的人們身上所得到的發現，推論到所有人們或是個案所形成的母體或者是母群。爲了要將研究發現加以推論，我們應該要詳細說明我們打算要推論的母體爲何，並且思考提出支持此種推論的證據。例如，我們以來自於某一所大學的某個課程中的 100 名大學生自願者，進行了一個實驗。那麼，我們可以把這個實驗的發現推論到什麼人的身上呢？是在同一年裡同一所大學所有課程中的大學生？還是在相同年代裡同一個國家中的所有大學生？或者是所有時代裡的所有人群？爲了提升一個實驗中母體推論性形式的外在效度，我們會從一個母體中抽取隨機樣本，並且以這些被抽樣的參與者來進行實驗。

當大多數的人們聽到*外在效度*這個詞時，首先會想到的就是**自然推論性**（naturalistic generalization）。自然推論性的關鍵問題，就是我們是否能夠正確地把我們在一個人爲創造的、受到控制的、像是實驗室一般的背景情境中所得到的發現，推論到「真實生活」的自然背景情境。在進行自然推論時，我們必須考量兩個議題：生活真實性以及實驗反應。

生活真實性（mundane realism）所處理的問題，是一個實驗或情境是否像是真實的世界。例如，你進行了一項關於學習的研究，在研究中，你要求參與者記住四個字母的無意義音節。如果你讓參與者們學習真實生活的事實資訊、而不是去學習那些爲了實驗而發明的無意義音節的話，這項研究的生活真實性就會比較強烈一些。[12]

自然推論性　把從一個人為創造的、受到控制的、像實驗室一般的實驗背景情境中所得到的發現，推論到「真實生活」的自然背景情境的能力。

生活真實性　一種外在效度的類別，也就是指實驗的情境看起來相當真實、而且非常類似於實驗室外的真實生活情境。

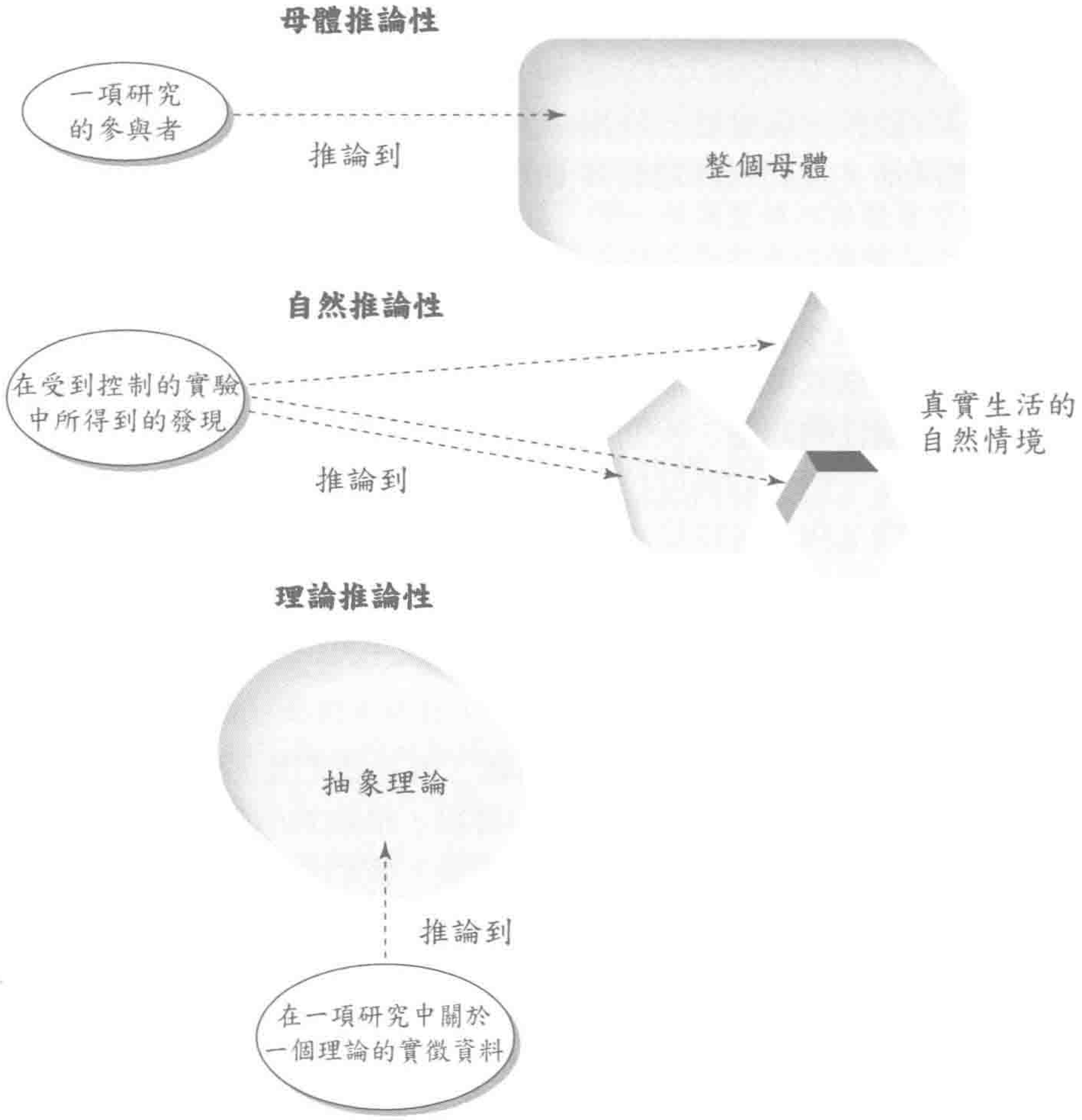


圖 9.5　外在效度推論性的三種類型

實驗反應（reactivity）是指人們因爲意識到自己正在一項研究中，因而刻意做出某種反應。當人們知道某個人正在研究他們時，就有可能在實驗中做出和在真實生活中不同的反應。**霍桑效應**（Hawthorne effect）就是一種特殊的實驗反應。[13] 這個名稱源自於梅堯（Elton Mayo）於 1920 年代以及 1930 年代於伊利諾州西屋電力公司霍桑廠內所進行的一連串實驗。研究者調整許多工作條件的層面（例如，照明、休息時間等等），並且測量生產力。他們發現，在每次進行調整之後，不論所作的調整是什麼，生產力都會提高。這個令人好奇的結果之所以發生，是因爲工人們所回應的對象並不是處遇，而是他們身爲實驗的一部分所得到的額外關注、以及知道他們正在被他人觀

實驗反應　一種對於外在效度的一般性威脅所造成的結果，產生此種結果的原因，是因為參與者意識到他們正在參與一個實驗、並且正在被研究當中。

霍桑效應　一種實驗反應結果，其名稱源自於一個知名的案例，在此案例中，參與者是針對自己正在一個實驗中的這個事實而作出反應，並非是針對處遇而作出反應。

察的事實。之後的研究則質疑在此研究中的工人所作出的回應是否真的曾經發生過，但是這個名稱仍然被廣泛地使用，以指稱由於研究者的關注所造成的結果。

對於外在效度而言，實驗反應的問題，在於我們是否能夠正確地將發生於一個在人們意識到自己正在被研究的情境中的活動，推論到自然的情境中。實驗反應最有可能發生在一個受到高度控制的實驗中，在此種實驗中的研究參與者知道某個實驗者創造出這些特定的條件、並且也正在觀察他們的行爲或者是反應。

讓我們假設你在一間大學教室或者是實驗室裡進行了一個實驗，而且參與者們也都知道他們正在參與一項研究。你要求參與者們去從事一些人爲創造的任務（例如，拼出一個拼圖）、或者是利用欺瞞而創造出一個特定的人爲狀況（例如，告訴參與者們，一個爲你工作的同謀擁有像天才般的智力）。在完成任務之後，你要求參與者完成一份問卷，在問卷中你列出了一些問題，這些問題是關於參與者們對於擁有高智商的人的感覺如何。那麼，你可以把你的研究發現推論到日常生活中的什麼背景情境呢？是推論到所有真實生活中有各種不同智力程度的人們的工作情境？還是所有類型的工作任務以及所有的社會地位？或者是關於其他人的所有態度，而這些態度是在日常生活中自然形成，並且保留在平常的想法、行爲、或者對話之中？爲了要提升實驗中自然推論性形式的外在效度，你可能必須進行田野實驗。

理論推論性（theoretical generalization）所處理的問題，是我們是否能夠正確地從我們想要測量的一個抽象理論裡的概念以及關係，推論到在一個特定實驗中的一組量數以及活動的安排。這很有可能是三種推論性當中最困難的一種類型，因爲其中包含了數個其他的觀念：實驗真實性、測量效度、以及控制混淆變項（高內在效度）。**實驗真實性**（experimental realism）所指的是實驗處遇或者是背景情境對於人們所造成的衝擊與影響。換句話說，當人們參與了一個實驗，並且真正地被實驗所影響時，就是具有實驗真實性。如果人們不受影響、而且實驗對於他們沒有造成任何衝擊時，實驗的真實性就比較薄弱。

田野／實地實驗　我們會在實驗室裡受到控制的條件下進行實驗，也會在真實生活或者是田野情境中進行實驗，不過在後者的情況中，我們對於實驗條件會比較沒有太多的控制力。控制的程度會在一條連續線上變化。在其中一端是受到高度控制的**實驗室實驗**（laboratory experiment），此種實驗的進行地點是在一個專門化的背景或者是實驗室之中；而在另外一端的則是**田野實驗**（field experiment），此種實驗的進行地點是在「田野」中，也就是像一節地鐵車廂、一家賣烈酒的商店、或者是公開的人行道之類的自然背景環境。在田野實驗中的參與者通常不會察覺到他們正在參與一個實驗，而且會以自然的方式來作出反應。例如，研究者要求一名同謀在一節地鐵車廂裡假裝心臟病發作，然後看看周圍的旁觀者會有什麼樣的反應。[14]

實驗真實性　一種外在效度，所指的是實驗情境讓人覺得合乎實際情況，因此實驗的事件會對參與者造成真正的衝擊。

實驗室實驗　一種實驗研究，在此種研究中，實驗者對於人為的背景情境有極大的控制力。

田野實驗　一種發生於自然背景環境中的研究。

有一些田野實驗，像是特蘭蘇針對種族認同以及繳稅支持公立學校所進行的實驗、或者是克莉桑以及同事針對社區偏好所進行的實驗（請參閱範例 9.2 以及 9.3）等等，會牽涉到聚集一群參與者、並且給他們一些實際的選擇。其他的則是「自然實驗」（“natural experiments”），在此種實驗中，就像是在亞伯達省的酒類販售權私有化的研究（請參閱範例 9.3）一樣，會出現類似實驗的情境，而且研究者對此種情境無法完全掌控。在一種田野自然實驗的相關類型當中，研究者能夠針對一個關鍵變項進行隨機分派，正如同在大學室友的混合種族個案中一樣（請參閱範例 9.7 大學室友的田野實驗）。

實驗者控制的程度與內在、外在效度有關。實驗室實驗通常會有較高的內在效度、但是外在效度則會比較低。此種實驗在邏輯上較爲緊密、而且受到較好的控制，但是比較向外推論的能力就會比較弱一些。田野研究通常會有較高的外在效度、但是內在效度就會比較低。此種實驗向外推論的能力比較強大、但是控制的程度卻比較少。準實驗設計是比較常見的。例如，在牽涉到大學室友的實驗中，室友的情境非常真實、而且持續了數個月之久。比起把人們放在一個實驗室的背景環境、並且詢問他們在假定的情況下將會有什麼反應，上述的實驗會有比較高的外在效度。

實際的考量

每一種研究技術都有其「箇中訣竅」（“tricks of the trade”），也就是從經驗中所學習到的實用策略。這些訣竅可以解釋在經驗豐富的研究者所進行的成功研究、以及新手研究者所面臨的困境之間的差異。以下將討論其中的三種差異。

規劃以及前導測試

所有的社會研究都必須先進行規畫。在規畫的階段中，我們會去考量可能的替代解釋或是對內在效度的威脅、發展出一個井然有序的系統來記錄資料、並且對我們將會使用到的任何設備或儀器（例如，電腦、攝影機、錄音機等等）進行前導測試。在經過前導測試之後，我們會去訪談參與者，以找出實驗中需要改善的一些層面。

給參與者的指示

大部分的實驗都會涉及到給予參與者一些指示，以爲實驗的進行「作好準備」（“set the stage”）。我們必須很小心地用言語給予指示，並且依循著一個預先準備好的劇本，如此一來，所有的參與者才能聽到同樣的指示，而這樣才能確保信度。當實驗者利用欺瞞來創造出一個具有真實性的封面故事時，這些指示也是很重要的。亞倫森以及卡爾史密斯（Aronson and Carlsmith, 1968: 46）指出：「新手實驗者最常犯下的錯誤之一，就是關於指示的說明過於簡短。」

範例　9.7

大學室友的田野／實地實驗

接觸假說（contact hypothesis）提到，在與外團體進行親密、長期的接觸之後，能夠減少對於外團體所抱持的偏見。修克以及法吉歐（Shook and Fazio, 2008）想要「評估不同種族之間的關係，並且測試團體之間進行接觸的效果」（p. 719）。然而，當我們以自我回報的態度量表來測量偏見程度時，人們通常會控制自己的偏見反應，好讓自己在表面上看起來似乎沒有偏見，就算他們可能具有偏見的態度時亦然。一種測量隱藏的或是「習慣性的」（"automatic"）種族偏見的間接技術，就是以秒數來測量當人們看見不同種族的人們的視覺影像、並且搭配著各種不同的形容詞時，他們作出回應的時間長度（請參閱 Fazio et al., 1995）。回應的速度之所以能夠間接地測量種族偏見，是因為當我們試圖去隱藏真正的態度時，我們作出回應的速度就會變得比較緩慢一些。為了創造出一個長期的田野實驗，兩位作者利用隨機分派的方式，將參與者分派到不同的大學宿舍裡去，並且以宿舍房間不足的理由，阻止參與者們任意交換室友。這項研究的參與者總共有 136 名白人以及 126 名黑人大學新鮮人。在經過隨機分派之後，有些人的室友是和自己相同種族的人，有些人的室友則是和自己不同種族的人。室友的種族是實驗中的自變項。研究者要求這些學生在學期一開始的最初兩週去參加一次聚會，而在學期最後兩週時，也要再參加一次聚會。在聚會當中，研究者們會去詢問學生們一些議題，包括對室友的滿意度、與室友之間的互動、以及社會網絡等等。這些學生們也完成了一份關於種族態度以及團體之間焦慮感的問卷。除此之外，作者們也創造了一系列的任務，要求學生對在電腦螢幕上出現的各種不同影像作出回應。在利用數個這種電腦任務以創造出一個「封面故事」之後，最後的任務是要去針對一些臉部影像以及搭配的形容詞作出回應。其中一半的臉部影像是黑人、另一半則是白人，而這就是測量種族偏見的間接方式。因此，在種族態度以及不同種族之間的社會互動方面，作者們就擁有許多前測以及後測的測量結果。如同在過往的室友研究當中一樣，他們的結果顯示出，與相同種族的室友相較之下，在不同種族的室友之間的社會互動較少，而且對於室友的滿意度也較低。而在經過一個學期的相處之後，對於相同種族室友的滿意度稍微下降一些，但是對於不同種族室友的滿意度則稍微上升一些。對於不同種族的室友而言，在一學期總共三個月的相處時間裡，種族之間的焦慮感下降了，而室友之間的社會互動則增加了。但是對於相同種族的室友而言，不論是種族偏見的直接或者是間接測量結果都維持不變。然而，在擁有不同種族室友的學生之間，偏見的程度則是明顯地下降許多，而此種情形正好符合接觸假設所作出的預測。

實驗後的訪談

在實驗結束時，我們應該要去訪談參與者，而這麼做的原因有以下三個。首先，如果我們使用了欺瞞，我們在倫理上必須向研究參與者進行**簡報**（debrief）（也就是解釋實驗的真正目的，並且回答參與者的任何問題）。其次，我們可以得知參與者的想法，以及他們對於情境的定義如何影響了他們的行為。最後，我們可以解釋不把實驗真正的本質透露給其他潛在參與者知道的重要性。

簡報　在實驗結束之後，藉由和參與者談話的方式以蒐集資訊。如果在實驗中使用了欺瞞的話，可以向參與者解釋實驗的真正目的，此外，實驗者也可以藉此機會得知參與者們的感受為何。

實驗研究的結果：進行比較

對於所有的研究而言，比較都是極為重要的。

在仔細地檢視實驗研究的結果之後，我們就能夠得知對於內在效度的可能威脅，以及處遇對於依變項所造成的影響或是效果。在本章中所討論過的每一項研究中，研究者都小心謹慎地分析量化資料以檢視自變項的效果，並且考量了與內在效度有關的潛在問題。

在這兒有一個說明如何進行比較的範例（請參閱圖 9.6），這個範例是五個使用古典實驗設計的一系列減重實驗。在這個範例中，在安立克瘦身診所的實驗組裡的三十名研究參與者，平均減掉了大約 50 磅，但是在控制組裡的另外三十名參與者則完全沒有減掉任何體重。其中只有一個人在中途退出實驗。而蘇珊科學節食計畫也有同樣引人注目的結果，但是在她的實驗組裡總共有十一個人在中途退出實驗，而這種情況則是顯示出了實驗對象流失的問題。在卡爾的卡路里計算計畫中，實驗組裡的參與者們減去了 11 磅的體重，而在控制組裡的參與者們則是只有減去了 2 磅的體重。但是，在一開始的時候，控制組與實驗組之間的體重差異平均高達 31 磅之多，而這種情況則顯示出存在著選擇偏誤方面的問題。娜塔莉營養中心沒有任何實驗對象流失或者是選擇偏誤的問題，但是實驗組所減掉的體重卻沒有比控制組還要多，而這則顯示出實驗的處遇沒有效果。寶琳減磅計畫也避開了選擇偏誤以及實驗對象流失的問題。在她的實驗組裡的人們減了 32 磅，但是控制組裡的人們也減掉了同樣的體重。此種情況顯示她的實驗可能出現了成熟效應、歷史效應、或者是處遇擴散的問題。因此，在安立克瘦身診所中的處遇，似乎是最有效的處遇。

與倫理有關的議題

在大部分的實驗中，倫理的考量都是一個相當重要的議題，因爲實驗研究通常會具有侵擾性（也就是說會干涉到日常的活動）。實驗的處遇可能會牽涉到把人們放置於設計好的社會背景情境中、要求人們從事一些特定的活動、或者是去操弄人們的感覺或是行爲等等。當我們在進行實驗時，我們會去傾聽人們所說的話、觀察他們的行動、並且記錄他們的反應。倫理方面的要求，限制了可以被允許的侵擾數量以及類型。我們絕對不可以讓研究參與者處於對身體有任何危害的情境中，而且當我們讓研究參與者處於令人困窘或者是會引起焦慮的情境時，我們必須先小心地作好預防措施。除此之外，我們也必須持續地監測以及控制實驗事件，以確保我們的研究是安全而且合乎倫理的。[15]

有些時候，我們或許會在社會實驗當中使用欺瞞，以暫時地誤導參與者。雖然此種不誠實的作法或許是可以被接受的，但是只有在沒有其他的方法能夠達到一個特定的研究目標時，才能這麼做。即使是爲了一個具有極高價值的目標，我們也只能有限地使用欺瞞。欺瞞的程度以及類型不能夠超過特定目標所要求的最低限度。除此之外，我們必須永遠在最短的時間之內對研究參與者進行簡報，讓他們知道曾經被暫時地欺瞞過，並且向他們解釋真正的狀況爲何。

圖 9.6　古典實驗設計的減重實驗之結果比較

安立克瘦身診所

	前測	*後測*
實驗組	190(30)	140(29)
控制組	189(30)	189(30)

娜塔莉營養中心

	前測	*後測*
實驗組	190(30)	188(29)
控制組	192(29)	189(28)

蘇珊科學節食計畫

	前測	*後測*
實驗組	190(30)	141(19)
控制組	189(30)	189(28)

寶琳減磅計畫

	前測	*後測*
實驗組	190(30)	158(30)
控制組	191(29)	159(28)

卡爾的卡路里計算計畫

	前測	*後測*
實驗組	160(30)	152(29)
控制組	191(29)	189(29)

進行比較時所用的符號

	前測	*後測*
實驗組	A (A)	C (C)
控制組	B (B)	D (D)

比較

	A-B	*C-D*	*A-C*	*B-D*	*(A)-(C)*	*(B)-(D)*
安立克	1	49	−50	0	−1	0
蘇珊	1	48	−49	0	−11	0
卡爾	31	37	−8	−2	−1	0
娜塔莉	2	1	−2	−3	−1	−1
寶琳	1	1	−32	−32	0	−1

A－B	這兩個群組一開始的體重是一樣的嗎？如果不一樣的話，就很可能是發生了選擇偏誤。
C－D	這兩個群組最後的結果是一樣的嗎？如果不一樣的話，可能是處遇沒有效果，或者可能發生了強烈的歷史、成熟、或是處遇擴散的效應。
A－C	實驗組是否有所改變？如果沒有任何變化，處遇可能是無效的。
(A)－(C) 與 (B)－(D)	在實驗組以及控制組裡的參與者人數是否有變化？如果很多人中途退出的話，實驗對象流失的問題就有可能會對內在效度造成威脅。

詮釋

安立克：沒有明顯威脅到內在效度的事物，處遇有發揮效果。
蘇珊：實驗對象流失的威脅可能會造成問題。
卡爾：可能有選擇偏誤的問題。
娜塔莉：沒有明顯威脅到內在效度的事物，但是處遇也沒有效果。
寶琳：可能有歷史效應、成熟效應、以及處遇擴散的問題。

註：數字是平均磅數。在括弧內的數字是每個群組的參與者人數。隨機分派用於實驗組或控制組。

結論

本章所討論的內容是實驗研究，而在大多數的實驗設計中，我們都會使用隨機分派來創造出兩個（或是更多個）群組，我們可以把這些群組視爲相等、並且加以比較。實驗研究爲因果關係提供了精確而且相當明確清楚的證據。實驗研究嚴謹地依循了社會科學中的實證主義取向的原則，並且產生了我們可以用統計來加以分析的量化結果。

本章也檢視了實驗的各個部分如何結合起來，以產生不同的實驗設計。除了古典實驗設計之外，我們也討論了前實驗設計與準實驗設計、以及設計記號。

本章也指出了各種對於內在效度的威脅，這些威脅都可能產生除了處遇之外的替代解釋，同樣地，我們也討論了關於外在效度的問題，以及田野研究將外在效度中的自然推論性加以最大化的各種方法。

實驗研究真正的優點，就是在建立因果關係的證據方面所有具有控制力以及邏輯的嚴謹度。一般而言，與其他的研究技術相較之下，實驗通常較爲容易複製、花費較少、耗費的時間也比較少。然而，實驗研究也有一些限制。首先，有些問題無法利用實驗法來處理，因爲控制以及實驗操弄完全無法應用於這些問題之上。另外一個限制，則是實驗通常一次只能測試一個或者是少數的假設。此種情況會使得知識出現片段化的情況，因此研究者就必須把許多研究報告裡的結果結合起來。外在效度也是一個潛在的問題，因爲許多實驗所依賴的通常是大學生的小型非隨機樣本。[16]

本章也說明了，對實驗結果進行仔細的檢驗以及比較，可以讓我們察覺到研究設計中的潛在問題。最後，本章也提出了一些在實驗中關於實際以及倫理方面的考量。

在接下來的幾個章節裡，你將會看到其他的研究技術。非實驗法的邏輯通常會與實驗的邏輯大相逕庭。實驗者只把焦點集中在少數假設上、把對於一個依變項造成影響的一個或是兩個自變項當作研究的對象、而且通常只檢驗少數的參與者。對比之下，在使用其他的研究技術時，我們通常會同時測試許多假設、測量許多自變項以及依變項、並且使用一個大型的隨機抽樣的研究參與者樣本。

問題回顧

1. 實驗的七個要素或部分是什麼？
2. 前實驗設計與古典實驗設計的差異之處爲何？
3. 何種實驗設計允許依照不同順序對數種處遇進行測試？
4. 有一名研究者說：「這是一個三乘二的設計，其中的自變項是恐懼的程度（低、中、高）以及逃離的容易程度（容易／困難），而依變項則是焦慮。」這段話代表了什麼意思？假定只使用後測的隨機分派，那麼其設計記號是什麼？
5. 間斷性時間序列設計以及同等性時間序列設計的差異之處爲何？
6. 內在效度的邏輯是什麼？使用控制組的作法如何契合於內在效度的邏輯？

7. 所羅門四組設計如何展示出測試的效果？
8. 何謂雙盲實驗？爲什麼要使用雙盲實驗？
9. 在田野實驗與實驗室實驗中，何者具有較高的內在效度？何者具有較高的外在效度？請解釋原因。
10. 實驗真實性與生活真實性之間的差異是什麼？

注解

1. Cook and Campbell (1979:9–36, 91–94) argued for a modification of a more rigid positivist approach to causality for experimental research. They suggested a "critical-realist" approach, which shares some features of the critical approach outlined in Chapter 4.
2. For discussions of the history of the experiment, see Danziger (1988), Gillespie (1988), Hornstein (1988), O'Donnell (1985), Scheibe (1988), and Webster and Sell (2007:6–9).
3. See Hornstein (1988:11).
4. For events after World War II, see Harriis (1988), and Suls and Rosnow (1988). For a discussion of the in-creased of deception, see Reynolds (1979:60).
5. See Field and Hole (2003) for a review of different comparisons.
6. Cook and Campbell (1979:7–9) and Spector (1981:15–16) discuss control in experiments.
7. The notation for research design is discussed in Cook and Campbell (1979:95–96), Dooley (1984:132–137), and Spector (1981:27–28).
8. For additional discussions of threats to internal va-lidity, see Cook and Campbell (1979:51–68), Kercher (1992), Spector (1981:24–27), Smith and Glass (1987) , and Suls and Rosnow (1988).
9. The example is borrowed from Mitchell and Jolley (1988:97).
10. Experimenter expectancy is discussed in Aronson and Carlsmith (1968:66–70), Dooley (1984:151–153), and Mitchell and Jolley (1988:327–329).
11. For discussions of external validity, see Aronson and Carlsmith (1968:22–25), Cook and Campbell (1979:70–80), Lucas (2003), and Zelditch (2007).
12. For a discussion of external validity, see Lucas (2003), Mook (1983), Willer and Walker (2007b), and Vissersi et al. (2001).
13. The Hawthorne effect is described in Roethlisberger and Dickenson (1939), Franke and Kaul (1978), and Lang (1992). Also see the discussion in Cook and Camp-bell (1979:123–125) and Dooley (1984:155–156). Gil-lespie (1988, 1991) discussed the political context of the experiments and how it shaped them.
14. See Piliavin and associates (1969).
15. See Hegtvedt (2007) for a recent review of ethical is-sues in experiments.
16. See Graham (1992).

第 10 章

調查研究

調查研究的歷史
調查研究的邏輯
建構問卷
調查的類型：優點與缺點
調查訪談
合乎倫理的調查
結論

每一種資料蒐集的方法，包括調查在內，都只是近似於知識而已。每一種方法都提供了對知識不同的一瞥，而且在單獨使用時也都各有限制。在進行調查之前，研究者應該要先問問自己，對於手邊的問題而言，這是不是最適當的、最為有效的方法。在研究像是公共意見的問題方面，調查是具有高度價值的，但是對於其他的問題，調查可能就毫無價值可言了。

——Donald P. Warwick and Charles A. Lininger, *The Sample Survey*, pp. 5-6

在公共民意調查當中，大多數的美國人都說他們願意投票給一位符合資格的女性總統候選人。對於符合資格的女性候選人的支持度，已經從 1937 年的 33%，穩定地上升到 2005 年的 92%。然而，當調查研究者在詢問人們一些具有爭議性的議題時，研究者們知道可能會發生社會期望效應（也就是人們會給予不誠實的意見，好讓他們自己能夠符合於一般的社會規範）。史崔伯等人（Streb et al., 2008）提出了一個假設：在接受調查時，許多美國人在這個議題上面都沒有說實話。要測試這樣的假設，必須要發揮創造力才能辦得到。作者們製作了一份關於四個議題的清單（例如，汽油價格上漲、被要求必須繫上安全帶等等），並且詢問其中有多少個議題「使你覺得生氣或者是心煩」。他們以相同的問題製作了第二份完全相同的清單，但是其中加入了第五個議題：「由一位女性來擔任總統」。他們針對每份清單各自隨機挑選了 1,000 個人，並且進行了電話訪談。作者們發現到，當女性擔任總統這個議題出現在清單上面時，使人們覺得生氣或者是心煩的議題數目會提高 26%，而此種情況則顯示出，在接受民意調查時，每四個人當中就有一個人所給的是不誠實的、刻意去符合社會期望的答案，因為他們實際上是反對女性去競選總統的。

調查是最被廣泛使用的社會科學資料蒐集技術。調查有許多種不同的使用方式，而且也有許多種不同的形式，像是電話訪談、網路民意調查、以及各種不同種類的問卷等等。所有這些關於調查的使用方式以及不同形式，都是依據專業社會研

究調查的原則而來的。許多人會說他們準備要進行一項調查來取得資訊，事實上，他們應該要說的是他們需要以最適當的方法來獲得良好的資料才對。調查能夠提供我們正確的、可靠的、以及有效的資料，但是若想要達到此項目標，則必須經過努力以及認真的思考才行。一般大眾對於調查技術已經相當熟悉，而且進行一項調查也相當容易，然而，此種情況卻可能會是一個缺點。雖然調查已經被廣泛地使用、而且也相當受到歡迎，但是，在不謹慎的情況下，調查也很容易產生誤導的結果。如同在本章一開始的部分所描述的社會期望偏見（將在本章稍後部分進行討論）所顯示的，在使用調查方法時必須要謹慎小心才行。在本章中，你將會學習到關於調查研究的知識以及此種技術的限制。

調查研究是在社會科學的實證主義取向中逐漸成長的。[1] 正如同葛洛夫斯（Groves）所說的：「調查所產生的資訊在本質上是屬於統計的。調查是量化的巨獸」（1996:389）。大多數的調查是去詢問大量的人們〔通常被稱爲*受訪者*（*respondents*）〕關於他們的信念、意見、特質、以及過去或是現在的行爲（請參閱資訊箱 10.1 在調查中所詢問的資訊）。正因爲如此，當我們想要得知人們自我回報的信念或者是行爲時，使用調查就是相當適當的。大多數的調查會一次詢問許多問題，因此同時測量許多變項。此種情況允許我們在單獨一次的調查中去蒐集描述性的資訊，並且測試許多假設。

我們可以把調查運用於探索性的、描述性的、或者是解釋性的研究當中。然而，當我們向受訪者提出「爲什麼」的問題時（例如，你覺得爲什麼會發生犯罪案件？），我們應該要十分地謹慎小心。[2] 此類的問題或許會讓我們得知人們的信念以及主觀的理解，但是人們通常會抱持著一些不完整的、錯誤的、或者是扭曲的看法。我們不能把人們提到或者是認爲某些事物發生的原因，和在社會世界裡真正的因果關係給混淆在一起。

調查研究的歷史

現代的調查可以追溯到早期的人口普查形式。[3] 人口普查（census）所指的是政府針對某個區域內整體人口的某些特質所蒐集的資訊。例如，《末日審判書》（*Domesday Book*）就是由英國國王威廉一世在 1085 年到 1086 年之間對於英格蘭地區所進行的人口普查。早期的人口普查是針對需要課稅的財產以及需要服兵役的年輕男性進行調查。在代議民主制度發展之後，官員們則是利用人口普查來得知一個地區的人口數量，並且運用這項資訊來指派選舉代表、以及分配公共建設的資金。

社會研究中的調查是從美國以及英國在 19 世紀時的社會改革運動開始的。調查幫助人們詳細記錄了由早期的工業化所造成的都市情況以及貧窮的問題。早期的調查都是描述性的，而且也沒有使用科學的抽樣方法或者是統計分析。例如，在 1851 年到 1864 年之間，梅修（Henry Mayhew）根據他在倫敦街頭與人們所進行的對話、以及他對於日常生活的觀察，出版了四冊的《倫敦的勞工以及倫敦的貧民》（*London Labour and the London Poor*）（1889-1902）。之後的一些研究，像是布斯（Charles

資訊箱 10.1　在調查中所詢問的資訊

雖然類別會有重疊的部分，但是以下的資訊都可以在調查去詢問受訪者：

1. 行為　你多久刷一次牙？你在最近一次市長選舉中是否有去投票？你最近一次拜訪一名親近的親戚是什麼時候？
2. 態度／信念／意見　你認為市長的工作內容是什麼？你會不會覺得當你不在場的時候，別人會說你的壞話？你認為我們國家最近所面臨的最大問題是什麼？
3. 特質　你是已婚、從未結婚、單身、離婚、分居、或者是喪偶？你是否加入某個工會？你的年齡多大？
4. 期望　你在未來的 12 個月之內是否計畫要買一輛新車？你認為你的孩子將會接受多少的學校教育？你認為這個鎮的人口將會增加、減少、或者是維持不變？
5. 自我歸類　你認為你自己是自由派、溫和派、還是保守派？你認為你的家庭屬於哪個社會階級？你會自稱為是一個極度虔誠的人、或者是一個沒有宗教信仰的人？
6. 知識　在上一次的選舉中，哪位候選人當選了？在這個城市的人口中，非白人市民約佔多少百分比？在這個國家裡私自持有一份馬克斯的《共產黨宣言》是合法的嗎？

Booth）的十七冊《倫敦人民的勞動與生活》（*Labour and Life of the People of London*）、以及榮恩崔（B. Seebohm Rowntree）的《貧窮》（*Poverty*）（1906）等著作，都詳細記錄了在倫敦市區裡的貧困情況。至於《1895 年霍爾館地圖與報告》（*Hull House Maps and Papers of 1895*）以及杜波伊斯（W. E. B. DuBois）的《費城黑人》（*Philadelphia Negro*）（1899）則是詳細記錄了美國的都市情況。

二十世紀早期，在加拿大、英國、以及美國等地所進行的社會調查運動（Social Survey Movement），則是把調查方法當作是質化社區田野研究的一部分來加以運用。社會調查運動是一種行動導向的社區研究計畫，該計畫會去訪談人們並且記錄各種社會狀況，以取得人們對於社會政治改革的支持。到了 1940 年代時，實證主義的、量化的調查已經幾乎取代了此種早期形式的調查研究。

早期的社會調查針對特定區域提供了一幅詳細的實徵圖像，並且結合了量化以及質化的資料來源，而其目標則是在於讓社會大眾瞭解快速工業化所帶來的問題。社會調查的早期領導者——像是在霍爾館以及睦鄰組織運動中的凱莉（Florence Kelly）和亞當絲（Jane Addams），以及非裔美國人杜波伊斯等等——都是處於主流的學術生涯之外的。由於在那個時代的種族歧視以及性別歧視的緣故，凱莉、亞當絲、以及杜波伊斯都無法獲得穩定的學術職位。早期的社會調查提供了二十世紀初期的日常社區生活的圖像，而且令人印象深刻。例如，在 1914 年出版的六冊《匹茲堡調查》（*Pittsburgh Survey*）當中，就包括了來自於面對面訪談的資料，關於健康、犯罪、以及工業傷害的統計資訊，以及直接觀察的結果。

到了 1920 年代以及 1930 年代時，研究開始使用統計抽樣技術，尤其是在發生了《文學文摘》（*Literary Digest*）的事件之後（已經在第八章中討論過了）。研究者們製作了態度量表以及指數（在第七章中已經討論過），以一種更為正確的、量化的

方式來測量民眾的意見以及主觀的信念。而在應用領域（例如，農業、教育、健康照護、新聞業、市場行銷、公共服務、以及慈善事業）中的專業人士則調整了調查技術，以測量消費者行爲、輿論、以及地區性的需求。

到了 1930 年代時，信奉實證主義取向的專業研究者，迅速地取代了那些曾經運用調查來記錄地區性社會問題的社會改革者。這些專業研究者融合了來自於自然科學中的原理，並且努力試圖使調查法變得更爲客觀、量化、以及非政治性。在進步時代（Progressive Era, 1895-1915）結束後，許多學術研究者試圖讓他們自己遠離社會改革政治。當時，在研究者以及大學之間爲了地位、聲望、以及資金而與彼此互相競爭，而此種競爭情況則加速了調查法的重新導向、或者是實證主義的「現代化」。有數個調查研究中心也在這個時期裡成立，其中包括了普林斯頓大學裡的「輿論研究辦公室」、在李克特（Rensis Likert）（你在第七章裡已經讀過關於他的資料了）主導之下所成立的「美國農業部計畫調查處」、以及哥倫比立大學裡的「廣播研究辦公室」等等。除此之外，《民意季刊》（*Public Opinion Quarterly*）這本專門提倡以及促進調查研究的刊物，也於 1937 年開始發行。數個大型的私人基金會（卡內基、洛克斐勒、以及聖賢基金會）則爲量化的、實證主義導向的社會研究提供了加速擴張所需要的資金。[4]

調查研究在第二次世界大戰期間急遽地擴張，尤其是在美國。學術界的社會研究者以及來自於產業界的實務工作者齊聚在美國的華盛頓特區，一起爲了戰爭而努力工作。調查研究者獲得大量的資金贊助以及政府的支持，以研究百姓與士兵的士氣、消費者需求、生產能力、敵軍的宣傳手法、以及轟炸的效果等等。戰爭時期的合作過程，幫助研究者以及應用的實務工作者從彼此身上學習許多知識，並且也獲得了進行許多大模規調查的寶貴經驗。學術研究者幫助實務工作者瞭解並且重視精確的測量、抽樣方法、以及統計分析。實務工作者則幫助學術界人士學習在組織以及進行調查時的一些實務層面。政府官員以及企業主管們發現，使用來自於大模規調查的資訊，可以帶來許多實際的益處。學術的社會科學家也瞭解到，他們可以利用調查資料來促進對於社會事件的理解、以及測試各種理論。

在第二次世界大戰結束之後，政府官員們很快地結束了大型的政府調查機構。這麼做的原因是爲了要削減支出，而且也是因爲政治保守人士擔心改革者可能會利用調查法來記錄社會問題。他們擔心此種關於受到不平等待遇以及貧困狀況的資訊，可能會被用來支持一些保守人士們所極力反對的政策，像是幫助失業的勞工、或者是在實施種族隔離的南方各州中去提倡對於黑人的種族平等運動等等。

在二次大戰結束之後，許多研究者重新回到大學裡，並且成立了新的社會研究組織，像是 1947 年在芝加哥大學所成立的「全國意見研究中心」（National Opinion Research Center）等。李克特則離開了美國農業部，並且在 1949 於密西根大學創立了「調查研究機構」（Institute for Survey Research）。

一開始，大學對於接受這些新的調查研究中心顯得有些猶豫不決，因爲這些研究中心所費不貲，而且必須雇用許多人力。傳統的社會研究者也對量化研究心存疑慮，並且對此種把在私人企業中廣受歡迎的技術帶進大學的作法深表懷疑。應用研

究以及企業導向的民意調查者的文化，與基礎研究者的學術文化格格不入。然而，在美國以及其他先進國家中，使用調查法的情況卻快速地增加。到了 1948 年時，法國、挪威、德國、義大利、荷蘭、捷克斯拉夫、以及英國都已經各自成立了全國性的調查研究機構。[5]

從 1950 年代到 1960 年代之間，收錄調查研究的各種刊物也快速地成長。例如，在 1939 到 1950 年之間，社會學的期刊當中有大約 18%的文章使用了調查法；而到了 1964 年至 1965 年時，這個比例已經上升到了 55%。到了 1960 年代時，高等教育以及社會科學迅速地擴張，同時也刺激了調查研究的發展。而在 1970 年代時，研究者們開始在研究中運用電腦，而電腦則提供了針對大規模量化資料進行統計分析的能力，而且也有數以百計的研究生學習調查研究的技術。[6]

自從 1970 年代之後，量化調查研究就被大量地運用於私人企業、政府、以及許多學術領域中（例如，溝通、教育、經濟學、政治學、公共衛生、社會心理學、以及社會學等等）。僅僅在美國境內，專業的研究產業就雇用了超過六萬人之多。其中大多數人是兼職的工人、助手、或者是半專業人士，而在各個大學裡，則總共約有六千名全職的專業調查研究者在從事設計以及分析調查的工作。[7] 魏斯柏格（Weissberg, 2005: 11）認爲，調查研究已經從許多使用此種技術的領域（例如，社會學、政治學、行銷學）當中區隔出來，成爲一門獨立的學科領域。

在教育、健康照護、管理、行銷、政策研究、以及新聞學等各領域當中的專業人士都會使用到調查研究。而全世界各地區以及各國的政府在進行政策方面的決策時，也者會提供經費來進行調查，以取得相關的資訊。私人部門的調查產業包括了民意調查（例如，蓋洛普、哈里斯、以及羅普民意調查等等）、行銷（例如，尼爾森、市場現況、市場研究公司等等）、以及非營利研究組織（例如，數學政策研究機構、蘭德研究機構等等）。[8] 除此之外，調查研究也有數個專業的組織。[9]

在過去的二十年之間，研究者不斷地去探討調查過程本身，並且發展出關於調查訪談中的溝通—互動過程的各種理論。研究者們現在已經能夠確切地指出在問卷設計中的視覺以及其他線索的效果、辨識出問題措辭方式以及排列方式所可能造成的影響、針對社會期望進行調整、融合與電腦相關的科技、並且對於調查受訪者願意合作或者是拒絕受訪的情況提出理論來加以說明。[10]

調查研究的邏輯

如同在第九章中所討論過的一樣，在進行實驗研究時，我們會把一小群人分成相同的群組、測試一個或是兩個假設、操弄情境以使某些參與者接受實驗處遇、並且控制背景情境以減少對於內在效度的威脅（也就是混淆變項）。在一項實驗結束時，我們會擁有一些量化資料，並且會去比較參與者在依變項上面的反應。不過，調查研究所依循的是一個截然不同的邏輯。在進行調查研究時，我們通常會找許多受訪者來當作樣本，並且詢問所有人完全相同的問題。我們以這些問題來測量許多變項，並且同

時測試許多假設。我們會從關於過去行爲、經驗、或者是特質的問題中去推論出時間順序。例如，就時間順序而言，一個人接受學校教育的年數或者是種族，其發生時間會早於這個人現在的態度。我們以統計的方式來分析變項之間的關聯性，以找出因果關係。我們也預期會有一些可能的替代解釋，並且用其他的調查問題（也就是控制變項）來測量這些替代解釋。之後，我們會以統計方式來檢視這些變項的效果，以排除替代解釋。調查有時候會被稱作是關聯性（correlational）的研究，因爲研究者不像在實驗中一樣地去控制以及操弄情境或是條件。在調查研究中，我們會使用控制變項，以利用統計的方式去接近實驗者對於混淆變項所進行的控制。

進行調查的步驟

在進行一項調查時，研究者會先從一個理論的、或者是應用的研究問題開始。我們可以把調查研究的過程區分爲數個步驟，就如同圖 10.1 所顯示的一樣。第一個階段就是先發展出一項研究工具——一份調查問卷或者是訪談大綱。受訪者可以自行閱讀在問卷上面所列出的問題，並且自己圈選出答案。所謂的*訪談大綱*（*interview schedule*），就是一組由訪談者向受訪者提出的問題，而且也是由訪談者去記錄下受訪者的反應與答案。爲了簡化相關的討論，接下來我只使用*問卷*這個詞彙。

調查研究是以演繹方式來進行的。首先，我們把變項加以概念化，並且針對每一個變項進行操作化，以成爲一個或者是多個調查問題。而這意味著，爲了要達到清晰性以及完整性，我們可能必須重覆地改寫、修正調查問題。一旦我們提出了一組調查問題之後，我們必須在問卷上組織安排這些問題，並且把這些問題分成不同的群組以及排列出問題的順序。我們應該要以我們的研究問題、受訪者的類型、以及調查的類型（請參閱本章稍後部分關於調查類型的討論）來引導我們進行這個程序。

讓我們假設你正要進行一項調查。在你準備製作一份問卷時，先去想想你要如何記錄並且組織你的資料。你也應該利用一小組的受訪者來針對問卷進行前導測試，而且這些受訪者要與最後進行調查時的受訪者相似。如果你要使用到訪談者的話，你必須以問卷來訓練這些訪談者。在前導測試以及受訪者訓練中，你必須去詢問受訪者以及訪談者，看看他們認爲問卷上的問題是否清楚，而且你也必須去探索他們的詮釋，以瞭解你預期的意義是否夠清楚（請參閱本章稍後部分關於前導測試以及認知訪談的討論內容）。[11]

這也是你可能要抽取受試者樣本的階段。在計畫階段結束之後，你已經準備好要去蒐集資料了。你必須以親自拜訪、撥打電話、寄送信件、或者是利用網路等方式去找到被抽樣出來的受訪者。在找到受訪者之後，你必須向受訪者說明完成問卷或是訪談的指示。而你所提出的問題通常是依循著一種簡單的「刺激／反應」或者是「問題／答案」的形式。你必須立刻正確地記錄下受訪者的答案或是回應。在所有的受訪者都完成了問卷、而且你也感謝他們的參與之後，你就可以把這些資料組織起來，並且準備好對這些資料進行統計分析了。

進行調查必須要有良好的組織能力。一項大型的調查可能會相當複雜而且昂貴，

此種調查會牽涉到協調其他的人、經過許多不同的步驟、以及正確地記錄資料。[12] 你必須仔細記錄受訪者對每一份問卷上的每一道問題的答案。爲了幫助你自己處理這項工作，你應該指派一組識別號碼給每一位被抽樣的受訪者，並且把這組號碼附在問卷上面。

在蒐集完所有的資料之後，你可能會想要檢閱個別問卷上的答案、儲存原始的問卷、並且把問卷的資訊轉換成電腦能夠讀取的格式以進行統計分析。一絲不苟的記錄工作以及製作標籤是必要的，如果你過於草率或者是漫不經心的話，你可能會遺失資料，或者是到最後只得到毫無價值的、不正確的資料。

在調查研究中，有許多種犯下錯誤或是誤差的方式（請參閱資訊箱 10.2 調查研究中的誤差來源）。誤差可能發生在抽樣以及挑選受訪者的過程中（已經在第八章當中討論過）、在製作問卷或者是訪談的過程中、以及在處理或是分析資料的過程中。接下來，我們要看看當你在爲一份調查研究的問卷撰寫問題時，所應該避免的一些誤差。

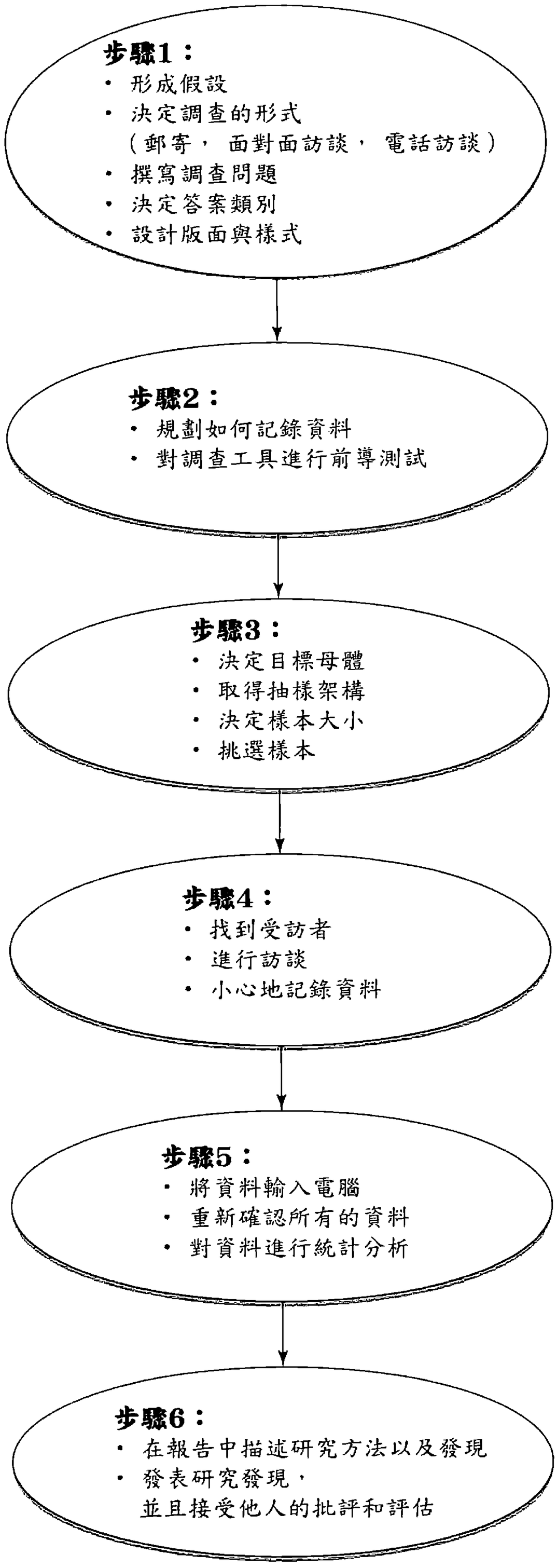


圖 10.1　調查研究過程的步驟

建構問卷

撰寫良好問題的原則

針對如何撰寫調查問卷這個問題，目前在市面上已經有數十本相關的書籍了，因此，在這兒只有回顧一下基本原則而已。撰寫良好的調查問題，是一份結合了藝術與科學的工作。最好的作法，是把整份問卷當作是一個完整的整體，在一些介紹的話語以及告訴受訪者如何輕鬆並且清楚地作答的指示之後，其中的各個

資訊箱 10.2 調查研究中的誤差來源

誤差（error）所指的是在所得到的值以及「真正的值」（"true values"）之間的差異。當調查資料（得到的值）無法正確地反映出在一個母體中的受訪者的真正行為、信念、以及理解時（真正的值），就是發生了誤差。而這些真正的行為、信念、以及理解，才是研究者所試圖要去瞭解的資訊。

1. 選擇受訪者時的誤差
   a. 抽樣誤差（sampling errors）（例如，使用了非機率的抽樣方法）
   b. 涵蓋誤差（coverage errors）（例如，使用了一個較差的抽樣架構，而遺漏了某些群體的人）
   c. 在抽樣單位的層次上所發生的無回應誤差（nonresponse errors）（例如，一名受訪者拒絕回答）
2. 回答調查問題時的誤差
   a. 針對特定調查問題的無回應誤差（例如，受訪者跳過或是忽略某些問題）
   b. 由受訪者所引起的測量誤差（例如，受訪者沒有用心在聽指示）
   c. 由訪談者所引起的測量誤差（例如，訪談者在念問題或是記錄答案時過於草率）
3. 調查執行的誤差
   a. 調查後誤差（postsurvey errors）（例如，在整理資料時、或者把資料轉換為電子格式時犯下一些錯誤）
   b. 模式效應（mode effects）（例如，由於調查方法所產生的差異，像是透過信函、親自訪談、或者是透過網路等等不同的方法）
   c. 可比較性誤差（comparability errors）（例如，在針對同一群受訪者調查相同的議題時，不同的調查組織、國家、或者是調查時間所得到的資料卻不相同）

*請參閱：*Weisberg （2005: 10-28） 以及 Willis （2005: 13-17）

問題能夠流暢地從一個問題銜接到另一個問題。

在試圖撰寫良好的調查問題時，有兩個關鍵原則可以爲我們指引方向：避免可能的混淆狀況，並且將受訪者的觀點謹記在心。避免可能的混淆狀況說起來簡單，但是做起來卻不容易。你應該要讓調查問題所提供的量數是有效的而且是可信的。量數要既有效而且可信，就意味著受訪者應該要能夠如你所預期地去快速掌握每個問題的意義、完整而且誠實地回答問題、並且相信他們的答案是有意義的。

你不會希望你的調查問題讓受訪者感到困惑或是受挫，而這意味著如果受訪者的異質性很高、來自於你不熟悉的生活情境、或者是擁有和你不同的事物優先順序的話，你就必須更爲謹慎小心才行。如果受訪者所使用的詞彙與你不同、或者是思考的方式與你不同時，你就必須要格外注意。

你希望調查問題對於所有的受訪者都同樣地清楚、具有關聯性、而且有意義，但是你會面臨一個兩難的困境。如果受訪者的背景以及參考架構南轅北轍，那麼即使是

相同的問題措辭方式，對於每個人也不見得會產生相同的意義，然而，由於你之後要把所有的答案結果轉換成爲數值資料以進行分析，所以你必須讓每個人都聽到相同的問題。如果每一個問題都是針對每一個受訪者所量身打造的話，你就無法知道，在資料當中的變異性究竟是來自於問題的措辭方式、抑或是受訪者之間的真正差異。

撰寫調查問題需要技巧、練習、耐心、以及創意。當你在撰寫調查問題時，有十個你應該要避免的錯誤。藉由瞭解這十個錯誤，你就能夠瞭解撰寫問題的原則。不過，以下的這份清單只有包含一些研究者經常遇到的潛在問題。[13]

1. *避免使用術語、俚語、以及縮寫*　術語以及專業詞彙可能會以許多不同的形式出現。像是鉛管工人會提到「*小蛇*」（*snakes*，即鉛管通道），律師會談到一份「*最大誠意*」（*uberrima fides*）的合約，而心理學家則會提到「*伊底帕斯情結*」（*Oedipus complex*，即戀母情結）等等。俚語則是在一個次文化中的一種術語。例如，街友們會提到「*雪鳥*」（*snowbird*，即有海洛因毒癮的人），而滑雪者則會談到「*右腳在前*」（*goofy foot*）（譯註：此爲滑雪和衝浪的共同術詞，指在滑雪或衝浪時把右腳放在前面的人）等等。在某個專業領域裡的人、或者是某個特殊的次文化中的成員，或許會對這些術語或者是俚語感到熟悉自在，但是這些詞彙只會讓圈外人覺得困惑不解而已。除此之外，你也要避免使用縮寫或者是首字母縮寫語，這是因爲相同的縮寫詞彙可能會有許多不同的意義。例如，我接到一封來自於「中西部社會學協會」（Midwest Sociological Society, MSS）的信函。但是，如果你在字典裡去查 MSS 這個首字母縮寫語的話，你會發現 MSS 這個詞可以用來指稱數十種不同的意思，像是「製造業標準化協會」（Manufacturers Standardization Society）、「海洋系統模擬程式」（Marine Systems Simulator）、「醫學院學生協會」（Medical Student Society）、以及「明尼蘇達洞穴學協會」（Minnesota Speleological Society）等等。我屬於「亞洲研究學會」（Association for Asian Studies, AAS）這一個專業的學會，但是有六個其他的學術組織也使用了相同的縮寫字母：「美國天文學協會」（American Astronomical Society）、「美國自殺學學會」（American Association of Suicidology）、「美國聽力學協會」（American Audiology Society）、「美國航天協會」（American Astronautical Society）、「美國古物研究學會」（American Antiquarian Society）、以及「亞述學術協會」（Assyrian Academic Society）。

當你在對一般社會大眾進行調查時，你應使用大眾文化的語言（也就是說，那些在電視上或是在地區性報紙上的、大約是八年級學童能夠瞭解的詞彙）。調查研究者已經發現，受訪者經常誤解了基本的詞彙，並且對許多字彙的意義感到困惑。舉例而言，有一項調查詢問受訪者他們是否認爲電視新聞是公正的。研究者後來察覺到，有大量的受訪者忽略了「公正的」（impartial）這個詞彙，而研究者原本認爲這是一個每個人應該都知道的詞彙。只有不到半數的受訪者如研究者預期一般地以適當的意義來詮釋這個詞彙。有超過四分之一的受訪者完全不知道這個詞彙的意義、其他人則賦予這個詞彙與一般狀況不同的意義、另外還有十分之一的受訪者所認定的意義與原本真正的意義剛好完全相反。在另外一個案例中，有大約四分之一學歷

在高中以下的受訪者（在美國成年人口中大約佔了 20%）不知道*陰道交媾*（*vaginal intercourse*）是什麼意思。[14]

2. *避免語意曖昧、混淆、以及含糊不清*　語意曖昧以及含糊不清是大多數撰寫問題的研究者覺得最苦惱的問題。研究者很容易就會做出一些使受訪者覺得困惑的隱性假設。例如，像是「你的收入多少？」這個問題，可能是指週薪、月薪、或者是年薪；家庭或者是個人收入；稅前或者是稅後收入；今年的收入或者是去年的收入；來自於薪資的收入或者是來自於所有來源的收入等等。此種混淆的狀況可能會使得受訪者對於問題的答案出現前後不一致的情況。如果你詢問的是去年的稅前家庭年收入，你就應該清楚明確地向受訪者詢問這個問題。許多受訪者可能不知道去年的稅前家庭年收入到底是多少，但是他們可能會告訴你他們每週拿回家裡的薪資有多少（關於超出受訪者回答能力的問題，請參閱接下來的第 6 項）。[15] 不明確的字眼或者是回答的類別，也是造成曖昧不清的來源。例如，對於「你是否有規律地去慢跑的習慣？是___否___」這個問題的回答，取決於「*規律地*」（*regularly*）這個詞的意義。有些受訪者可能把「規律地」定義爲每天，其他人則可能認爲是一週一次。爲了減少混淆的情況以及獲得更多資訊，你應該要更爲清楚明確一些：不要去問某個人是否規律地去慢跑，而應該去問他是否「大約一天慢跑一次」、「一週慢跑數次」、「一週慢跑一次」等等。（請參閱資訊箱 10.3 改善不清楚的問題）

3. *避免情緒性的語言以及聲望的偏誤*　詞彙除了明顯的字面意義之外，也有暗示性的隱含意義。相同地，社會上的頭銜以及職位（例如，總統、專家等等）也代表了聲望以及地位。帶有強烈情緒暗示的字眼、以及與高地位的人們有所關聯的議題，可能會影響了受訪者回答調查問題的方式。研究者最好能夠使用中性的語言並且避免帶有情緒「包袱」的字眼，因爲受訪者可能會對這些情緒性的字眼作出回應，而不是對於實質的議題作出回應。例如，像是「有一些殘忍的恐怖份子威脅著要從愛好和平的人們手中奪走自由，對於付錢給這些恐怖份子的作法，你有什麼想法？」這個問題，就充滿了情緒性的字眼：*殘忍的*、*自由*、*奪走*、以及*和平*。

當調查問題包含了一些關於一個極有聲望的個人、團體、或者是機構的詞彙，而且受訪者對於這個極有聲望的個人或團體的感覺影響了他或她回答一個問題的方式時，就會發生**聲望的偏誤**（prestige bias）。如此一來，你就不會知道你所測量的到底是受訪者對於某位知名人士的感覺、還是他們對於某個議題的真正想法。例如，你問了一個問題：「大多數醫生認爲那些在吸煙者身邊的人會因爲吸入二手煙而導致肺部病變。你是否同意這樣的說法？」對於那些認爲同意醫生的意見比較好的人而言，即使他們個人不同意這樣的說法，仍然可能會回答同意。

聲望的偏誤　在撰寫調查研究問題時的一種難題，當某個受到高度尊敬的團體或者是個人會與某個答案選項之間產生關聯性的時候，就會發生此種難題。

模稜兩可的問題　一種包含了超過一個議題以上的調查問題，並且可能造成受訪者的困惑或是得到模糊不清的答案。

4. *避免模稜兩可的問題*　這也是一種避免語意含糊不清的作法。你的每一個問題都應該與一個、而且只有一個主題有關。而**模稜兩可的問題**

資訊箱 10.3　改善不清楚的問題

以下是由經驗豐富的專業研究者所撰寫的三個調查問題。由於在進行過前導測試之後，有 15%的受訪者要求研究者對這些問題作出進一步的說明、或者是給了不適當的答案（例如：不知道），因此這些研究者就修改了原始的措辭方式。如同你可以看到的，問題的措辭方式是一種藝術，而且會隨著練習、耐心、以及前導測試而改善。

原始的問句	問題	修正後的問句
你是否規律地運動？	怎樣才算是運動？	你是否規律地作任何包括走路在內的運動、嗜好、或是身體活動等等？
你在一週中平均有幾天會吃奶油？	乳瑪琳（人造奶油）算不算是奶油？	接下來這個問題只與奶油有關，不包括乳瑪琳在內。你在一週當中有幾天會吃奶油？
（在有關蛋的問題之後）你在正常的一天之內有幾餐是有蛋的？	一餐中有多少顆蛋？何謂正常的一天？	在你有吃蛋的那些日子裡，你通常會吃幾顆蛋？

	對問題的回答比例		要求澄清的比例	
	原始	*修正*	*原始*	*修正*
運動的問題（回答「是」）	48%	60%	5%	0%
奶油的問題（回答「沒有」）	33%	55%	18%	13%
蛋的問題（回答「一個」）	80%	33%	33%	0%

資料來源：調查問題改編自 Fowler, Survey Research Methods, Sage Publications. 1992.

（double-barreled question）則是同時包含了兩個或者是更多個混在一起的問題。例如，你詢問了以下的這個問題：「你的雇主是否提供了年金以及健康保險？」如果某個受訪者的公司只提供健康保險、但是沒有年金的話，那麼這名受訪者既不能回答「是」、也不能回答「否」。一位聽到「以及」這個字眼、但是認爲這個字眼所指的是「以及／或者」的受訪者，就會回答「是」。而另一位聽到「以及」這個字眼、並且認爲這個字眼所指的是「兩者都有」或是「而且還有」的受訪者，就會回答「否」。若是使用了模稜兩可的問題，你就無法確定受訪者的意圖爲何。如果你想詢問的是這兩樣事物是否發生的話，你應該提出兩個個別的問題，而且每一個問題都只和一個議題有關。在資料分析時，你就可以去看看那些對某一道問題回答「是」的人，是否也對另一道問題回答「是」。

5. *避免引導式的問題*　你總是想要讓受訪者相信所有的回答選項都是同樣合理的，而且也絕對不想讓他們察覺到你所期望或是想要得到的答案。而*引導式問題*（*leading question*）所指的是那些因爲其措辭的緣故而引導受訪者去作出某種回答

的問題。引導式問題有許多不同的種類，例如，像是「你不抽煙，是不是？」之類的問題，就可能會引導受訪者去選擇不抽煙的答案。

引導式問題可能會將受訪者引導到正面或者是負面的答案。例如，「市長是否應該花更多稅收去使本市極佳的道路維持在超級好的狀況？」這個問題，就會引導受訪者選擇「否」這個答案。而另外一個問題：「本市一些街道上有大型的坑洞，這些坑洞不但相當危險，而且也迫使駕駛人支出高昂的修車費用。你認爲市長是否應該分配經費來修理這些街道呢？」則會引導受訪者回答「是」。

6. *避免超出受訪者回答能力的問題*　問一些受訪者不知道的問題，只會徒增困擾、挫折感、以及獲得不正確的答案。受訪者無法總是能夠回想起過去的細節，而且可能不知道特定的資訊。例如，去詢問一名 40 歲的受訪者「當你六歲的時候，你對你哥哥有什麼樣的感覺？」這種問題，就很有可能是毫無價值的，同樣地，詢問受訪者一些他們一無所知的議題（例如，關於外交事務的一個專門性議題、或者某個組織的內部政策等等），也是毫無意義的。受訪者或許會給你一個答案，但是這個答案卻是不可信的、而且毫無意義的。當許多受訪者可能都不知道某項議題時，你就必須使用特殊的問題格式（我們將在本章稍後部分進行討論）。

你應該要試著用受訪者在思考時所使用的詞彙來撰寫調查問題。例如，幾乎沒有任何受訪者能夠回答下列這個問題：「你在去年爲自己的車子加了多少加侖的汽油？」但是受訪者或許能夠回答他們通常每一週的加油量，如此一來，你就可以自己去進行計算而估計出一年的加油總量。[16]

清楚的、有意義的問題可以增加正確性並且減少誤差。清楚的問題包含了內在的線索，並且明顯地表達出對比性。例如，與其去問「你是否會拿錢給上一段婚姻中所生的子女？」，比較好的問法是「你是否會支付子女的撫養費？」而對於那些回答「是」的受訪者，接下來的問題可能是「除了子女的撫養費之外，你是否還會支付贍養費？」以及「你是否還有其他財務上的義務，像是支付健康保險、學費、或者是房貸或房租的費用呢？」[17]

7. *避免錯誤的前提*　如果你在開始一個問題時所提出的前提是受訪者所不同意的、而且還以這個前提來提供答案選項的話，受訪者可能會覺得相當受挫而且不知道該如何回答這個問題。在大約兩年之前，我曾經經歷過一個具有錯誤前提的問題，不過我並不是在一項調查當中遇到這個問題的。就在某家航空公司停止供應免費的機上餐點之後不久，我剛好搭乘了這家航空公司的班機。一位空服員交給我一份可以自由選擇的餐點，並且問我：「您要付現還是刷卡？」我遲疑了一下子，然後赫然發現這其實是一個圈套，目的是要迫使我購買現在可以自由選擇、但是我並不想要的餐點。我回答「都不要」，並且很快地把這份餐點退回去。在剛剛這個情境下的錯誤前提，就是我想要買這份餐點，而我則對這個錯誤的前提覺得有點不高興。顯而易見的是，這個錯誤的前提也已經惹得其他乘客很不高興，因爲在六個月之後，空服員就再也不敢用這種花招來誘騙乘客額外付費來購買餐點了。

8. *避免詢問過於久遠的未來意圖*　避免去問人們他們在一些假設的情況之下會怎麼做。像是「假設在這條路上開了一家新的雜貨店，請問你會不會去這家店裡

購物？」之類的問題，通常只是在浪費時間而已。最好的作法，是詢問關於現在或者是最近的態度以及行爲。受訪者對於明確的、具體的、有意義的問題所回答的答案，會比對一些與切身經驗相差過遠的問題所回答的答案更爲可靠。

9. *避免雙重否定*　日常語言當中的雙重否定，在文法上是不正確的、而且容易造成混淆。例如，「我不是沒有工作」這句話在文法上以及邏輯上的意思是指我現在有一份工作。有些人會利用第二個否定詞來強調語氣。雖然此種明顯的錯誤相當罕見，但是雙重否定的微妙形式也會相當令人困惑。此種情況可能會發生在當我們去詢問受訪者同意或是不同意某項敘述的時候，例如，你提出了一個問題：「請問你同意或是不同意學生應該不要被要求去參加一個全面性的考試才能畢業？」這個問題會讓人覺得很困惑。如果不同意的話，就變成是雙重否定，也就是「不同意」「不要」做某件事。你應該總是使問題維持既簡單又直接的狀況。

10. *避免重覆或者是偏重某種答案類別*　你應該使你的答案類別或者是選項彼此互斥、全容、並且平衡（你已經在第七章裡面讀過關於彼此互斥以及全容的屬性了）。*彼此互斥*（*mutually exclusive*）的意思是指答案的類別之間沒有重疊的情況。如果重疊的類別是數值範圍的話，這是很容易修正的（例如，把 5-10、10-20、20-30 改爲 5-9、10-19、20-29）。而語意模糊不清的選項則有可能會造成重疊的答案類別，例如：「你對於自己的工作滿意嗎？或者是其中有你不喜歡的地方？」假定我整體而言對我的工作很滿意，但是其中有一些特定的事物是我真的很不喜歡的，那麼我要怎麼回答呢？*全容*（*exhaustive*）的意思是指每一位受訪者都可以找到適當的選項。例如，如果你去問受訪者「你目前正在工作當中、還是失業中？」這個問題的話，就會排除掉那些既不是正在工作當中、也不算是失業的受訪者，像是全職的家庭主婦、度假的人、全職的學生、因爲身心障礙而無法工作的人、以及已經退休的人。爲了避免這樣的問題，首先必須仔細地思考你真正想要測量的事物，並且考量所有可能的受訪者的情況。舉例而言，如果你想要詢問有關就業狀況的問題，那麼你想要得到的資訊究竟是關於主要的工作或是所有的工作？只有全職的工作、或是全職和兼職工作都要算在內？只有領薪水的工作、或者是無薪資的或自願性質的工作也都要包括在內呢？

你應該要使答案的類別維持*均衡*（*balanced*）的狀況。不均衡的答案類別會造成某種類型的引導式問題。例如，像是在「市長在工作方面的表現如何：傑出、優秀、非常好、或是令人滿意？」這個問題中的答案類別就是不均衡的。不均衡問題的另外一種類型則是遺漏了一些資訊，例如下列這個問題：「在五位市長候選人當中，你支持哪一位：歐斯維果（Eugene Oswego）、或者是其他四位候選人中的某一位？」

你可以藉由提供相反的答案來使答案的類別維持均衡。人們很容易就可以看出像是「誠實」以及「不誠實」這類的詞彙具有不同的意義以及含意。要求受訪者去回答「某位市長是非常誠實、還算誠實、或者不是非常誠實」，與要求受訪者去回答「某位市長是非常誠實、還算誠實、既非誠實也非不誠實、有點不誠實、或是非常不誠實」是不一樣的。你提出問題的方式，會大幅度地影響到你所得到關於人們想法的資訊。除非你有明確的理由必須採取其他的作法，否則的話，你應該提供受訪

者在一條連續線上的兩個相反端點的答案選項[18]（請參閱表 10.1）。

受訪者回憶

我們經常會詢問受訪者關於過去的行爲或是事件。在回答調查問題的時候，每個受訪者在正確地回憶過去行爲以及事件的能力方面會有所差異。[19] 回憶過去的事件所需要花費的時間和努力，通常會超過我們給予受訪者回答一個調查問題的短暫時間。除此之外，人們正確地回憶過去事件的能力，也會隨著時間經過而快速地降低。受訪者或許能夠正確地回憶發生在兩個星期前的一起重大事件，但是幾乎沒有任何受訪者能夠正確地回憶發生在兩年前不具意義的微小事件。

調查研究者承認，人們的記憶比過去他們所以爲的還要更加不可靠。有許多因素會影響到人們的回憶能力：主題（具有威脅性的、或是符合社會期望的）、同時以及之後發生的事件、事件對於某個人的重要性、情境條件（問題的措辭方式以及訪談的形式）、以及受訪者內在一致性的需求。除此之外，回憶（例如，你居住的城市的市長名字是什麼？）也比辨認（例如，請看一下這張姓名的清單，並請指出那一個是你居住的城市的市長名字）還要更加困難。

受訪者回憶所可能造成的問題，並非意味著我們不能去詢問過去的事件，相反地，我們必須明確地以這個目的來撰寫調查問題，並且用小心謹慎的態度來詮釋結果。爲了要改善受訪者的回憶能力，我們可以提供特殊的指示以及額外的思考時間。我們也可以提供一些協助受訪者回憶的方式，像是一段固定的時間範圍或者是地點的參考訊息。與其問「去年冬天你多常參加體育活動？」，你應該說「我想要知道你去年冬天參加了多少次體育活動。讓我們逐月來看看。先回想去年十二月。你在十二月當中是否參加過任何需要付費購買門票的體育活動？現在，再想想一月。你在一月當中是否參加過任何需要收取門票費用的體育活動？」（請參閱範例 10.1 如何在調查中測量看電視的行爲）

當被問到過去的事件時，受訪者通常會**壓縮回答**（telescope），也就是一種壓縮時間的現象。他們能夠回想起一個事件，但是他們所認定的該事件發生的時間，卻可能比實際發生的時間更早（向後壓縮）、或者是更晚（向前壓縮）。有數種技術可以用來減少壓縮回答的問題（請參閱資訊箱 10.4 四種減少壓縮回答的技術）。

壓縮回答 在回答關於過去的事件時，調查研究的受訪者所出現的壓縮時間的狀況，像是把最近發生的事件誤以為是更早之前發生的、以及把很久之前發生的事件誤以為是最近才剛發生的等等。

誠實的答案

關於敏感主題的問題 我們有時候會詢問一些敏感的議題、或者是一些人們認爲會威脅到他們外在形象的問題，此種問題包括了關於性行爲、使用藥物或是酒精、心理健康狀況、違法行爲、或者是不受社會歡迎的行爲的各種問題。在遇到這種問題時，受訪者可能不願意提供回整以及真實的答案。在提出像是這類的問題時，我們必須調整詢問的方

表 10.1　撰寫調查問題時易犯的錯誤摘要

應該避免的錯誤	不佳的調查問題	改善後的調查問題
術語、俚語、縮寫	你昨晚是否一直灌啤酒灌到掛掉？	昨天晚上你大約喝了多少啤酒？
模糊不清	你常常外食嗎？	在典型的一個禮拜之間，你有多少餐是在外面用餐的，也就是在餐廳、自助餐、或者是其他用餐地點用餐的？
情緒性的語言以及聲望的偏誤	「備受尊敬的顧雷斯委員會以文件證明了，在我們所繳納的稅金當中，有數量驚人的三千五百億被完全地浪費在低劣的採購過程、糟糕的管理制度、草率的記帳、『有缺陷的』合約管理作法、浮濫的人事制度、以及其他揮霍的作法上面。對你而言，削減政治分肥的支出以及消除政府的浪費，是不是應該優先處理的事情？」*	對你而言，國會採取一些減少政府浪費的措施這件事有多重要？ 非常重要 有點重要 既非重要也非不重要 有些不重要 完全不重要
模稜兩可的問題	你支持或是反對提高社會安全福利以及增加軍隊支出？	你支持或是反對提高社會安全福利？你支持或是反對增加軍隊支出？
引導式的問題	你是否盡到愛國的責任並且參與了上一次市長選舉的投票？	在上個月舉行的市長選舉中，你是否有去投票？
超出受訪者回答能力的議題	兩年前，你每個月總共看電視多少小時？	在過去兩個星期裡，你覺得你一天大概看多少小時的電視？
錯誤的前提	你是在什麼時候停止毆打你的女朋友或是男朋友的？	你是否曾經對你的女朋友或是男朋友甩耳光、用力猛擊、或是打他（她）？
久遠的未來意圖	在你從大學畢業、找到工作、並且穩定下來之後，你會不會投資一大筆錢在股市裡？	在未來的兩個月之中，你是否有明確的計畫要在股市裡投資一筆錢？
雙重否定	你是否不認同那些不想要建造一座新的市立游泳池的人？	有人提案建立一座新的市立游泳池。你同意或是不同意這個提案呢？
不均衡的答案	你認為這家旅館的服務品質是傑出、極佳、優秀、或是良好？	你為這家旅館的服務品質打分數：傑出、非常好、差強人意、或是差勁。

*這是一個取自於一份郵寄問卷中的真實調查問題，該份問卷是由美國共和黨國會委員會於 1998 年 5 月寄給作者的，而這個問題同時也是一個模稜兩可的問題。

範例　10.1

如何在調查中測量看電視的行為

由普萊爾（Prior, 2009a, 2009b）所進行的兩項研究，可以當作一個例子，用來說明使用回憶式的調查問題來測量看電視的行為時所會遇上的困難之處。我們測量媒體使用行為的主要方式，就是在調查中請受訪者進行自我回報。在過去的十年裡，有將近五十項出現在主要學術期刊上的研究，都是以自我回報的媒體使用方式的調查結果來當作資料。不幸的是，人們的回憶不一定是正確的，而且在調查中可能會過度地誇大媒體使用行為。人們自我回報在過去一週裡收看電視新聞的時間，比媒體公司尼爾森根據媒體使用行為監測技術所發現的時間高了三倍之多！雖然大多數年齡層的群體都有過度回報的情況，普萊爾卻發現到，在 18 歲到 34 歲這個年齡層的群體中，過度回報的比例是最高的。在這個年齡層的群體中，有大約 35% 的人說他們每天都有收看電視新聞，但是尼爾森公司監測技術的結果卻顯示，只有 5%的人真的每天都有收看電視新聞。即使是年齡較大、回憶也比較正確的人們，其過度回報的比例也有兩倍之多。從受訪者在調查中如何進行回答的相關文獻當中，普萊爾檢閱了關於在調查中無法正確地回憶過去行為的三個解釋：滿意、錯誤的估計、以及社會期望。滿意（satisficing）一詞所指的是，人們之所以會有不正確的回憶，是因為他們缺乏動機、或者是沒有很努力地去搜尋自己的記憶。當人們沒有使用良好的記憶搜尋策略來回想過去的事件時，就會產生錯誤的估計（flawed estimates）。至於社會期望（social desirability）所指的，則是人們會去回報他們認為是適當的、或者是標準的答案。在以調查問題的格式進行了一系列的實驗之後，普萊爾發現幾乎沒有任何證據支持滿意以及社會期望這兩種解釋，至少在回憶收看電視新聞行為時是如此。即使給予受訪者額外的思考時間、告訴受訪者他們的答案很重要、並且之後又重覆問了第二次，人們仍然有過度誇大的情況。而當受訪者被告知其他人花了多少時間去收看電視新聞時，他們則是會戲劇性地改變答案，以使自己的答案與其他人相符合。然而，當受訪者得到一些協助他們回憶的方法時，極度誇張的情況就跟著減少了。當受訪者獲得一個「錨點」或「定位點」（“anchor”）、或者是一些額外的事實資訊以幫助他們回憶時，他們所作出的估計值就會比較正確一些。受訪者被問到下列這個問題：「接下來的問題，是關於在 CBS、ABC、以及 NBC 這三家電視台上所播放的全國晚間新聞。這些新聞和你居住地區的地方新聞、以及像是 CNN 或是福斯新聞頻道（Fox News channel）這類有線電視新聞網所播放的新聞不同。那麼，在過去一週裡，你有多少天收看了這些全國性的晚間新聞？」其中一組受訪者則聽到以下的介紹式敘述：「電視新聞的觀眾人數最近減少了很多。在平常的週一到週五的任何一天的晚上，只有不到十分之一的美國民眾會收看全國性的晚間新聞。」聽到這段話的受訪者，在回答問題時花了比較長的時間，並且所回報的看電視新聞時數也比較低。普萊爾的研究顯示出，如果受訪者在回答問題時有比較充裕的時間、而且在回憶的過程中有得到協助的話，他們在回答與回憶過去行為或事件有關的調查問題時，他們的答案就會比較正確一些。

式，而且在處理結果時也要特別地小心謹慎[20]（請參閱表 10.2）。

詢問到敏感議題的那些問題，其實是自我保護這個更大議題的其中一部分。我們大多數人都會努力試著呈現我們自己的正面形象給別人看到。我們可能會覺得羞恥、難爲情、或者是害怕回答真實的答案，或者，我們發現到，光是要誠實地面對我們自己的行爲，就足以讓我們覺得痛苦不已，更別提要去向別人承認這些行爲了。因此，當發生此種情況時，我們多多少少會刻意隱瞞或者是漏報那些我們想要隱藏、或者是我們認爲會違反社會規範的行爲或是態度。人們通常會傾向於隱瞞的事物，包括了疾病或是身心障礙（例如，癌症、精神疾病、性病等等）、從事違法或是偏差

資訊箱 10.4　四種減少壓縮回答的技術

1. *情境架構（situational framing）*　要求受訪者回憶一個特定的情境，並且詢問關於該情境的細節（「請告訴我在你結婚那天發生什麼事，並請從當天早晨開始敘述」）。
2. *分解（decomposition）*　詢問受訪者數個特定的事件，然後把這些事件加總起來（「上個星期你有沒有在販賣機買過任何東西？好，那麼上上個星期呢？你有沒有在販賣機買過任何東西？」）。
3. *重要事件定位（landmark anchoring）*　詢問受訪者在某個重大事件之前或是之後是否發生了什麼事（「在 2010 年的 6 月時，這兒曾經發生過一場大地震。你剛剛說的那件事是發生在這場地震之前還是之後呢？」）。
4. *有界限的回憶（bounded recall）*（用於追蹤調查）　詢問受訪者關於上次訪談之後所發生的事件（「我們上一次談話是在兩年前，而從那時到現在，你換過哪些工作？」）。

表 10.2　具威脅性的問題以及敏感的議題

主題	非常不舒服的百分比
手淫	56
性交	42
吸食大麻	42
吸食興奮劑以及鎮靜劑	31
喝醉	29
愛撫和親吻	20
收入	12
與朋友賭博	10
喝啤酒、葡萄酒、或是烈酒	10
快樂以及福祉	4
教育	3
職業	3
社交活動	2
一般休閒	2
體育活動	1

資料來源：改編自 Improving Interview Method and Questionnaire Design. Bradburn and Sudman. 1980. JosseyBass. ISBN 10: 087589402X

的行爲（例如，逃漏稅、吸毒、酗酒、從事不正當的性行爲等）、或者是他們的財務狀況（例如，收入、存款、以及債務）等等。

我們可以用四種方法來使人們在回答敏感主題時，比較願意提供誠實的答案：創造自在感以及信任感、使用強化過的措辭方式、建立一個減低敏感性的脈絡、以

及使用匿名的提問方式。接下來，讓我們來逐一討論這些方法。

1. *創造自在感以及信任感* 在提出問題之前，先創造出信任感以及一個舒適自在的情境。在開始進行訪談之前，我們可以再次明確地向受訪者保證調查的匿名性以及機密性，並且強調我們必須得到他們誠實的答案。我們也可以先詢問一些不具威脅性的問題、並且創造出信任感或是自在感，然後在這個「暖身階段」過後，再開始詢問敏感的問題。

2. *使用強化過的措辭方式* 你可以修改問題的措辭方式以減低威脅性。例如，如果你直接問：「你是否曾經做過順手牽羊的事？」，這個問題就帶有指控的意味，而且也使用了「順手牽羊」這個牽涉到某種非法行爲的情緒性字眼。若是想要問到與上述問題中同樣的行爲，你可以這麼問：「你是否曾經在未付錢的情況下，從商店中拿走任何東西？」這種問法只有描述了行爲的過程，沒有使用情緒性的字眼，而且還留下了一種可能性，也就是此種行爲可能是在可以接受的情況下發生的（例如，因爲一時不小心而忘記付錢）。

3. *建立一個減低敏感性的脈絡（desensitizing context）* 我們也可以藉由提供已經減低敏感性的脈絡資訊的方式，來減少問題的威脅性、並且讓受訪者比較容易針對敏感主題說出誠實的答案。方法之一就是先詢問一些比我們感興趣的行爲更加嚴重的其他行爲。例如，在回答關於順手牽羊的問題時，受訪者可能會有所遲疑，但是如果我們先提出一連串關於比較嚴重的罪行的問題（例如：持槍搶劫、闖空門）、然後再提出關於順手牽羊的問題時，這個問題感覺起來就比較沒那麼嚴重，而且受訪者也比較可能願意誠實地回答。

4. *使用匿名的提問方式* 提問的格式也會明顯地影響受訪者回答敏感問題的方式。允許較多匿名性的形式，像是自填式問卷或者是網路調查等等，會比那些像是面對面的訪談之類需要與另外一人進行互動的形式，更有可能讓受訪者願意誠實地回答敏感的問題。[21]

科技方面的創新發明，像是**電腦輔助自填問卷訪談**（computer-assisted self-administered interviewing, CASAI）以及**電腦輔助面訪**（computer-assisted personal interviewing, CAPI）等等，也都能夠讓受訪者擁有某種程度的匿名性。在以 CASAI 這項技術「訪談」受訪者的時候，受訪者會被要求去閱讀出現在電腦螢幕上的問題、或者是戴上耳機來聆聽這些問題。而受訪者回答問題的方式，則是去移動滑鼠點選答案、或是用鍵盤來輸入答案。即使當訪談者或者是其他人也都在同一個房間裡時，受訪者仍然與他人處於半隔離的狀態，並且只有和一個自動化的系統進行互動而已。而在使用 CAPI 這項技

電腦輔助自填問卷訪談（CASAI） 一種電腦輔助的技術，在使用此種技術時，受訪者在電腦螢幕上閱讀問題、或者利用耳機來聽問題，然後再移動電腦的滑鼠來選擇答案、或者是在鍵盤上輸入答案。

電腦輔助面訪系統（CAPI） 一種電腦輔助的技術，在使用此種技術時，訪談者先幫受訪者架設好一部筆記型電腦，並且隨時準備為受訪者提供必要的協助，而受訪者則是經由耳機來聽取問題以及／或者在螢幕上閱讀問題，然後輸入答案。

術時，受訪者則是使用一部筆記型電腦，而訪談者則隨時可以提供協助或是回答相關的問題。受訪者從耳機中聽取問題以及／或是在螢幕上閱讀問題，然後在訪談者沒有直接觀察的情況下輸入答案。在完成此種電腦輔助的訪談時，即使其他人也在場，受訪者們仍然會覺得他們保有了自己的隱私。[22]

而在面對面的訪談情境中，有一種詢問敏感問題的複雜方法，也就是**隨機回應技術**（randomized response technique, RRT）。此種技術所使用的統計觀念超過本書的層次，但是類似於在本章一開頭所描述的關於女性總統候選人的研究中所使用的方法。此種技術的基本觀念是使用已知的機率來推估未知的比例。以下是隨機回應技術的運作方式。一名訪談者向受訪者提出二個問題：其中一個具有威脅性（例如，「你施打過海洛因嗎？」），而另一個問題則不具有威脅性（例如，「你是在九月出生的嗎？」）。接下來則利用隨機的方法來選擇要回答的問題（例如，拋擲一枚硬幣，出現正面則代表關於海洛因的問題，而出現反面則代表關於出生日期的問題）。訪談者不知道被選中的問題是哪一個，而只是記錄下受訪者的答案（是或否）。利用隨機結果的機率（例如，在九月出生的人們的比例），我們就可以去推估敏感行爲的出現頻率有多高。

我們希望在敏感主題方面得到誠實的答案，並且也想盡量避免出現像在本章一開始所提及的研究中所描述的情況，也就是受訪者所給予的是不誠實但卻符合社會期望的答案。然而，**社會期望偏誤**（social desirability bias）是相當普遍的狀況。當受訪者刻意扭曲答案以符合一般社會規範時，就會發生社會期望偏誤的問題。人們通常傾向於過度誇大自己具有高度的文化水準（例如，經常閱讀、參加文化活動）、捐款給慈善機構、擁有美好婚姻關係、深愛自己的子女之類的事情。有一項研究發現到，在一項調查當中，那些回報說自己曾經捐款給慈善機構的受訪者，有高達 34% 的人其實根本沒做過這件事。[23] 由於一般社會規範告訴人們應該要在選舉中去投票，因此，很多人雖然沒有去投票，但是也會回報說他們有去投票。在美國，那些承受最大的投票壓力的人（例如，教育程度較高的、政黨傾向明顯的、以及極度虔誠並且被某個組織強烈要求去投票的人），往往是最可能誇大自己投票行爲的人。此種關於投票行爲的不實陳述，會「大幅地扭曲」了那些依賴自我回報的調查資料所進行的投票研究（Bernstein et al., 2001: 41）。

減少社會期望偏誤的方法之一，就是改變問題的措辭方式，讓受訪者覺得違反社會規範似乎不是那麼令人討厭的事、或者是給予受訪者一些「保全顏面」的替代選項。例如，貝里等人（Belli et al., 1999）在他們關於投票的問題中，加入了像是「許多人因爲沒有登記、生病、或者單純因爲沒有時間的緣故而沒有去投票」之類的敘述，因而減少了過度誇大投票行爲的狀況、並且讓受訪者得以「保全

隨機回應技術（RRT）　一種在調查研究中用來處理非常敏感主題的特殊技術，在使用此種技術時，受訪者所回答的問題是隨機挑選出來的，而且訪談者並不知道受訪者正在回答的是哪一個問題。

社會期望偏誤　在調查研究中所可能產生的一種問題，也就是受訪者不願意回答誠實的答案，而是回答「符合規範」的答案、或者是能被社會大眾接受的答案。

顏面」。他們提供了四個答案選項：「我在 11 月 5 日那天沒有去投票；我有想過要去投票，但是後來沒有去投票；我平常都會去投票，但是這次沒去投票；我很確定我在 11 月 5 日那天有去投票。」只有最後一個選項才是清楚、明確地代表受訪者有去投票。然而，由於以此種方式來陳述的關係，所以有比較多人願意承認自己並沒有去投票。

知識性問題　許多研究指出，大多數人無法正確地回答基本的地理問題、說出他們選出的領導人的名字、或者是辨認出一些重要的文件（例如：獨立宣言）。如果我們使用知識性問題來瞭解受訪者所具備的知識，我們必須要小心謹慎，因爲受訪者可能爲了不要讓自己顯得很無知而刻意說謊。[24] 知識性問題具有相當的重要性，因爲此種問題所處理的是人們作出判斷以及形成意見時所依賴的基礎，換句話說，此種問題可以讓我們知道，人們是不是依據不正確的資訊來形成他們的意見。

納多以及同事（Nadeau and colleagues, 1993）發現，大多數的美國人都嚴重地高估了少數民族在整體人口中所佔的百分比。只有 15%（正負 6%）的美國成人能夠正確地回答出在美國整體人口中有 12.1%是黑人。有超過一半以上的成人認爲黑人所佔的比例高於 30%。相同地，猶太人在美國整體人口中只佔了大約 3%，但是有超過半數（大約 60%）的美國民眾相信猶太人所佔的比例大約是 10%左右。由西格爾曼以及尼艾密（Sigelman and Niemi, 2001: 93）所進行的後續研究則發現，「黑人自己高估黑人人口比例的程度，至少和其他的受訪者一樣高」。有大約 15%的白人認爲黑人在美國人口中佔了半數之多，而有將近兩倍的黑人（大約 30%）也抱持此種看法。顯而易見的，對於美國人口的種族組成比例問題，許多美國人的看法都是錯誤的。

種族並不是唯一一個一般大眾抱持著錯誤看法的議題。例如，當我們詢問美國民眾關於政府在國外援助支出方面的問題時，有很大比例的民眾認爲政府支出的金額過於高昂。然而，如果我們改問這些受訪者他們認爲政府應該花費多少金額在國外援助上面時，這些受訪者所回報的金額數量，事實上卻高於政府目前所支出的相關金額。此種情況造成了一種兩難的困境。如果我們用某種方式來問這個議題，我們會發現民眾認爲支出過高，但是如果我們換成另一種方式來問的話，我們卻會發現民眾（間接地）認爲目前的金額數量少於應該支出的金額數量。此種兩難的困境其實並非只是國外援助這個議題所獨有的現象。在許多議題領域裡，像是大學支出、健康照護方案、對於窮人的協助等等，受訪者都會提出支持或是反對某個議題或是政策主張的意見，但是如果我們用另一種不同的方式來詢問相同的議題時，他們的立場卻會出現徹底翻轉的狀況。

此種兩難困境並非意味著我們無法以調查來取得有效的輿論測量結果，相反地，此種困境提醒我們，社會生活是相當複雜的，而且我們必須要付出努力以及勤奮工作，才能寫出良好的調查問題來瞭解人們的想法。如果我們漫不經心地去詢問人們的意見，我們可能只會得到膚淺的答案，因爲人們在回答時根本沒有經過仔細的思考、或者只是根據不正確的資訊來作出回答。或者，我們所得到的意見也可能

只是受訪者道聽塗說的看法、或者是他們在電視上所聽到的一些政治人物的「名言佳句」而已。

你或許會認爲，人們之所以會對美國的種族組成方式或是國外援助支出抱持著不正確的看法，是因爲這些資訊都超出了人們的日常經驗。然而，就算是在回答關於住在家庭裡的人口數目時，人們所給的答案也不一定是正確的。此種情況並不是無知所造成的，而且來自於人們日常生活的複雜性。有些人不會把一些邊緣性的人物（例如，一個因爲有事而暫時離開一星期的男朋友、一個在懷孕之後和家人大吵一架並且離家出走的成年女兒、或者是一個爲了金錢問題而和家人發生爭執並且憤而離開的叔叔）當作是家中的一份子，然而，這些邊緣性的人們或許並沒有另外一個永久的居住地點。如果我們去問這些人住在哪裡，他們或許會說他們仍然住在那些把他們排除在外的家戶裡，而且他們也打算要找時間搬回去繼續和其他人一起生活。[25]

在調查研究中，我們的目標是要取得正確的資訊（也就是說，關於一個人真正的想法、行爲、以及感覺的有效而且可靠的量數）。前導測試中的問題（將在本章稍後部分進行討論）可以幫助我們達成這個目標。前導測試可以讓我們得知問題的難易度是否適中。如果有 99%的受訪者都無法正確地回答問題時，我們就等同於一無所獲。我們在問句的用字遣詞方面也必須讓受訪者可以自在地承認他們不知道問題的答案，例如：「如果有的話，你聽說過多少有關……的事？」

我們可以利用受訪者完全不可能知道答案的**測謊題**（sleeper question），來確認受訪者是否誇大了他們的知識。例如，在一項關於受訪者能夠辨認出哪些民權運動領袖的研究中，研究者們加入了一個虛構人物的名字。結果，有 15%的受訪者「認識」這名虛構的人物，而這暗示著，在受訪者回報他們「認識」的實際領袖當中，有大約 15%其實是不認識的。另外一個方法則是在受訪者認出一個名字之後，向他們提出一個開放式的問題，像是「能不能請你說明一下關於這個人的一些事蹟呢？」（請參閱下一個部分：開放式以及封閉式的問題）。

條件式問題　有些問題只能適用於特定的受訪者，而且研究者應該避免對受訪者提出不相關的問題。所謂的**條件式問題**（contingency question）〔有時候也稱作是*篩選*（*screen*）或是*跳躍式*（*skip*）問題〕是指那些有兩個部分（或是更多部分）的問題。[26] 受訪者對於問題的第一部分的答案，決定了接下來受訪者將要回答兩個問題當中的哪一個問題。條件式問題可以辨識出與第二個問題有關的受訪者。根據對於第一個問題的答案，研究者可以指示受訪者或是訪談者前進到另一個問題、或者是直接跳過某些問題（請參閱資訊箱 10.5 條件式問題的範例）。

測謊題　在調查研究中一種詢問不存在的人物或事件的問題，目的在於確認受訪者是否誠實。

條件式問題　一種包括了兩個部分的調查問題，在此種問題中，受訪者對於第一個問題的答案，會將受訪者引導到問卷上的另外一個問題、或者是引導到更為具體而且相關的第二個問題。

資訊箱 10.5　條件式問題的範例

問題版本 1 （非條件式問題）

在過去一年當中，當你坐在汽車的後座時，你繫上安全帶的頻率有多高？

問題版本 2 （條件式問題）

你是否曾經有坐在汽車後座的經驗？

否〔請直接跳到下一個問題〕

是 → 當你坐在汽車後座時，你繫上安全帶的頻率有多高？

結果	總是繫上	從未繫上
版本 1	30%	24%
版本 2	42%	4%

在前導測試當中，研究者得知許多對問題版本 1 回答「從未繫上」的受訪者，其實從未有坐在汽車後座的經驗。問題版本 1 造成了一種模稜兩可的狀況，因為不論是從未坐過汽車後座的受訪者、或者是那些坐在後座但是從不繫安全帶的受訪者，都會回答「從未繫上」。而問題版本 2 由於使用了條件式的問題，因此可以讓這個問題更加清楚明確。

*資料來源：*改編自 Presser, *Evaluating Survey* Question-*naires*, Hoboken, NJ: Wiley. (2004). Reprinted by permission of John Wiley & Sons, Inc.

開放式以及封閉式問題

研究者一直在熱烈地爭論開放式以及封閉式問題的優缺點。[27] **開放式問題**（open-ended question）是一種允許受訪者可以提出任何答案的問題（例如：「你最喜愛的電視節目是什麼？」）。而**封閉式問題**（closed-ended question）則是在提出問題之後，只提供一組固定的答案選項讓受訪者去從中挑選（例如：「依你的看法，目前總統的表現是非常好、好、普通、或是不佳？」）。這兩種問題形式中的每一種都具有其優點以及缺點（請參閱表 10.3）。真正關鍵的議題並不是兩者之間孰優孰劣，而是對於特定的情境而言，哪一種問題形式才是最爲適當的。你應該要選擇開放式或者是封閉式問題，必須取決於研究的目的以及實際的限制。由於在使用開放式問題時，訪談者必須逐字寫下受訪者的回答、並且在之後必須耗費大量的時間來進行編碼的工作，因此，在許多的研究中，使用開放式問題就會比較不切實際。

開放式問題　調查研究中的一種問題類型，受訪者在回答此種問題時，可以自由地回答他們針對這個問題所想要提出的答案。

封閉式問題　調查研究中的一種問題類型，在回答此種問題時，受訪者只能從一組固定的答案中去作出選擇。

在大規模的調查當中，我們通常會使用封閉式問題，因爲此種問題對於受訪者以及研究者而言都是比較快速而且容易的，不過，每當我們強迫一個人在一組固定的、預先決定好的類別中去選出自己的信念以及感覺時，我們就有可能會遺失某些重要的資訊。若是想要得知一名受訪者的想法並且發現什麼才是他認爲重要的事物、或者是對於有許多答案類別的問題而言（例如：年齡），那麼，使用開放

表 10.3　封閉式與開放式問題

封閉式問題的優點

- 對於受訪者而言，此種問題比較容易也比較能夠迅速作答。
- 比較容易對不同受訪者的答案進行比較。
- 答案比較容易進行編碼以及進行統計分析。
- 答案的選項可以幫助受訪者澄清問題的意義。
- 受訪者比較有可能願意回答敏感的主題。
- 就問題本身而言，不相關或是造成混淆的答案會比較少。
- 對於比較不擅言語或是文字能力較差的受訪者，不會造成不利的情形。
- 比較容易複製。

封閉式問題的缺點

- 可能會暗示一些受訪者原本不會想到的意見。
- 就算是沒有意見或是相關知識的受訪者也能作答。
- 如果受訪者想要的答案不在選項之內，受訪者可能會覺得受到挫折。
- 如果有太多答案選項的話（例如，20 個選項），則有可能會造成混淆。
- 對於問題的錯誤詮釋可能不會被注意到。
- 受訪者的答案之間的區別可能不明顯。
- 有可能發生抄寫錯誤或是選錯答案的情況。
- 強迫受訪者針對複雜的議題選出簡化的答案。
- 強迫受訪者作出他們在真實世界中可能不會作的選擇。

開放式問題的優點

- 可能的答案數目是無限的。
- 受訪者可以詳細地回答，並且能夠確認以及澄清他們的答案。
- 能夠幫助我們找到預期之外的發現。
- 允許針對複雜議題提出適當的回答。
- 允許創意、自我表達、以及細節的豐富性。
- 可以揭露受訪者的邏輯、思考過程、以及參考架構。

開放式問題的缺點

- 不同的受訪者所提出的答案詳細程度不同。
- 受訪者的回答可能沒有關聯性、或者是被淹沒在大量毫無用處的細節當中。
- 進行比較以及統計分析會變得非常困難。
- 對回答進行編碼會有困難。
- 能言善道以及文字能力較好的受訪者具有優勢。
- 對於沒有方向的受訪者而言，問題可能會過於籠統空泛。
- 訪談者必須逐字記錄回答，這對訪談者是一件困難的事。
- 受訪者必須花費較多的時間以及精力去進行思考。
- 受訪者可能會對問題覺得恐懼。
- 受訪者的回答可能會在問卷中佔據大量的篇幅。

式問題就會是最適合的方式。

你也可以在一份問卷中同時結合開放式以及封閉式的問題，以減少僅使用某種問題格式所可能造成的缺點。此外，混合兩種問題格式的作法也可以提供改變步調的機會，並且有助於訪談者和受訪者之間建立和諧的關係。間歇性地以封閉式問題

部分開放式問題　一種調查研究的問題類別，也就是受訪者雖然也會有一組固定的答案選項可以從中挑選，但是在選項中還會加上一個「其他」的類別，好讓受訪者可以提出一個不同的答案。

來進行探問（probe）（也就是由訪談者提出後續追蹤的問題，後面的部分將會再進行討論），能夠揭露出一名受訪者的推理思考方式。要求訪談者經常地使用探問的技巧來詢問受訪者的想法，可以確認受訪者是否真的如你預期一般地理解調查問題的意義。然而，探問並不能被用來取代清楚撰寫的問題、或者是爲受訪者創造出一個良好的理解架構。除非是以仔細小心的方式來加以說明，否則探問或許會影響到受訪者的答案、或者是迫使沒有意見的受訪者去提出答案，然而，在彈性的或者是對話式訪談（conversational interviewing）（將在本章稍後部分進行討論）當中，則會鼓勵多使用探問的技巧。例如，對於「你上個星期是否爲了賺錢而做了任何工作？」這個問題而言，受訪者可能會有所遲疑，然後說「有」。訪談者接著探問：「你可以明確地告訴我你做的是什麼工作嗎？」受訪者可能會回答：「在星期二以及星期三，我花了幾個小時幫我的好友約翰搬家到他的新公寓裡去，而且他給了我 40 美金當作酬勞，但是除了這件事以外，我就沒做過其他有賺錢的工作了。」如果你的意圖是只要獲得一般性雇用工作的資訊，在探問之後就能夠發現到受訪者誤解了你原本的問題。我們也會使用**部分開放式問題**（partially open questions）（也就是一組固定的選項，其中最後再加上「其他」這個開放的選項），而此種問題允許受訪者可以提出除了固定選項之外的其他答案。

完全依賴封閉式問題可能會扭曲結果。例如，有一項研究針對「我們國家目前所面臨的重大問題是什麼？」這個問題比較了開放式和封閉式的版本，研究中的受訪者則依據問題的形式而對不同問題的重要性加以排序。如同舒曼和普雷瑟（Schuman and Presser, 1979: 86）所報告的：「幾乎所有的受訪者都是在由調查者所提供的優先順序框架中去進行排序，*不論此種順序是否符合他們本身的優先順序*」。而在另一份研究中，研究者分別以開放式問題以及封閉式問題來詢問受訪者認爲工作中的重要事項，在回答開放式問題的受訪者當中，有一半的人所回答的答案是不包括在封閉式問題的答案選項之中的。

在研究的早期或是探索階段中，開放式問題是非常具有價值的。而對於大規模的研究而言，我們可以在前導測試中使用開放式問題，並且依據開放式問題的答案來發展出封閉式的問題。

封閉式問題會迫使我們去作出許多決定。像是「我們應該提供多少個選項？」、「我們是否應該提供一個中間或者是中立的選項？」、「答案應該如何排列順序？」、「應該要包括什麼類型的答案選項？」等等。要回答上述這些問題並不是一件容易的事。例如，只有兩個答案選項的話顯然太少，但是超過七個選項的話又會太多。我們想要去測量有意義的區別，而不是無止盡地去仔細畫分這些區別。較爲明確的答案選項可以產生較多的資訊，但是太多細節又會很容易造成受訪者的混淆。例如，把「你是否滿意你自己的牙醫？」（問題的答案只有是或否）這個問題重新敘述爲「你對自己的牙醫的滿意程度有多高：非常滿意、有點滿意、有些不滿意、完全不滿意？」之後，我們就可以得到比較多的資訊，而且也給受訪者比較多的選擇。

中立的立場、猶疑者、以及選擇性拒絕

當研究者無法從每一位受試者那兒得到有效的回答時，就會使得一項調查的效力變弱。受訪者在作答時，可能會有三種產生無效回答的情況。

1. *搖擺的意見*　此種意見牽涉到錯誤地誇大某種立場，就像是在社會期望偏誤中的情形一樣；或者是錯誤地隱瞞或是保留某種立場，就像是在回答敏感性的主題時的情形一樣。
2. *錯誤的正向答案*　此種狀況來自於受訪者選擇了支持某種態度的立場，但是卻根本完全不瞭解該項議題、而且對於該議題其實根本沒有任何真正的意見或是看法。
3. *錯誤的負向答案*　當受訪者拒絕回答某些問題，或者是雖然該名受訪者實際上擁有相關資訊或是有特定意見、但是卻不願說出來時，就會發生此種情況。

上述這三種類型的回答彼此之間有重疊的部分。第一種類型牽涉到某種傾向於一種規範性立場的不實回答，第二種則是以胡亂猜測來取代嚴肅的回答，而最後一種類型則是對調查採取一種部分的以及選擇性的不回應態度。[28]

中立的立場　調查研究者之間經常會產生爭議的焦點之一，就是對於缺乏相關知識、或者是沒有特定立場的受訪者，研究者是否應該提供一種中立的立場以及一個「沒有意見」的選項。[29]

有些人不贊成提供一個中立的或是中間的立場，以及「沒有意見」的選項，他們比較支持去強迫受訪者選出一個答案。[30] 抱持此種觀點的研究者認爲，受訪者會出現所謂的**滿意**（satisficing）現象，換句話說，他們會選擇「沒有意見」或者是一個中立的答案，以避免在作答時必須花費一些認知方面的努力。那些支持此種立場的人認爲，教育程度較低的受訪者雖然實際上有特定意見，但是卻可能會選擇「沒有意見」的選項。因此，他們認爲就算強迫這樣的受訪者選出答案，也不會降低資料的品質。

其他持相反意見的人則認爲，研究者最好還是提供一個中立的選項（也就是「沒有意見」），因爲人們通常在回答問題時是爲了要讓他人高興、或者是不想讓自己看起來顯得很無知。因此，若是強迫受訪者選出答案，他們可能會對一些虛構的議題、物體、以及事件提出錯誤的意見。藉由提供一個沒有特定態度的選項（中間或是沒有意見），我們就可以看出哪些受訪者是沒有特定意見的，並且把這些受訪者和那些確實有特定意見的受訪者區隔開來。

游移者　調查問題可以利用三種類型的態度問題，處理無意見這個議題：標準形式、半過濾式、以及全過濾式問題（請參閱資訊箱 10.6 標準形式、半過濾式、以及全過濾式問題）。

滿意　在回答調查問題時，避免花費認知方面的努力，並且只提出最不必花費力氣的回答，以滿足調查問題或者訪談情境的最低限度要求。

資訊箱 10.6 標準形式、半過濾式、以及全過濾式問題

標準形式問題

以下是一個關於其他國家的問題。你同意或是不同意以下的敘述？「俄羅斯的領導人基本上希望與美國和平共處。」

半過濾式問題

以下是一個關於其他國家的敘述：「俄羅斯的領導人基本上努力地試圖與美國和平共處。」關於這句敘述，請問你同意、不同意、或者是沒有意見？

全過濾式問題

以下是一個關於其他國家的敘述，但並非每個人都會對這句敘述有意見。如果你沒有任何意見的話，請如實說出來。敘述如下：「俄羅斯的領導人基本上努力地試圖與美國和平共處。」請問你對這句敘述是否有任何意見？沒有（請跳到下一題），有（請繼續作答）。請問你同意或是不同意？

不同問題形式所得到的答案範例

	標準形式(%)	*半過濾式(%)*	*全過濾式(%)*
同意	48.2	27.7	22.9
不同意	38.2	29.5	20.9
無意見	13.6*	42.8	56.3

*表示主動提出

資料來源：改編自 Schuman and Presser (1981). *Questions and Answers in Attitude Surveys: Experiments in Question From, Wording, and Context* (116-125). Academic Press. With permission from Elsevier. 標準形式是取自於 1978 年秋季，而半過濾以及全過濾式則是取自於 1977 年 2 月。

標準形式問題 一種調查研究中的問題形式，在此種形式中，答案的類別並不包括「沒有意見」或者是「不知道」之類的選項。

半過濾式問題 一種調查研究的問題形式，其中的答案選項包括了「沒有意見」、「不確定」、或者是「不知道」。

全過濾式問題 一種調查研究的問題形式，此種問題形式首先詢問受訪者對於某個主題是否有意見或是相關瞭解，然後再針對那些有特定意見或是瞭解的人更進一步地詢問關於該主題的其他問題。

標準形式問題（standard-format question）不提供「不知道」的選項，受訪者必須主動提出此種回答。**半過濾式問題**（quasi-filter question）則是提供了「不知道」的選項。至於**全過濾式問題**（full-filter question）則是一種特殊類別的條件式問題。此種問題首先會詢問受訪者是否有任何意見，然後再深入地詢問那些回答說有意見或看法的人。

在這三種問題形式背後的邏輯是，許多受訪者在沒有「無意見」這個選項的情況下，就會選出一個答案。但是當我們提供「無意見」這個選項時，他們就會選擇這個答案，或者，當我們直接地詢問他們時，他們也會說他們沒有任何意見。此種受訪

者被稱作是**游移者**（floaters），因爲他們總是在回答一些他們瞭解而且有相關知識的問題、以及回答他們不瞭解也沒有相關知識的問題之間「游移不定」。只要對問題的措辭方式作出一點小小的改變，就可能會使他們改變答案。半過濾式與全過濾式問題可以幫助我們篩選這些游移者。對於那些連不存在的議題都給予回答的受訪者，過濾式問題或許無法把這些受訪者全部排除掉，但是卻可以減少此種問題。

當我們提供中間立場的選項時，中間選項的游移者將會選擇這個答案，但是如果我們沒有提供此種選項時，他們就會選擇其他答案。此種受訪者對於某個議題可能有矛盾的感受、或者是沒有強烈的感覺。另外，也可能會有輕微的**時近效應**（recency effect），也就是說，受訪者傾向於選擇最後一個答案。時近效應所顯示的是，我們應該把答案選項在一條連續線上依序排列呈現，並且把中立的立場放在中間的位置上。

態度有兩個面向：方向性（支持或是反對）以及強度（態度強烈或是微弱）。例如，有兩位受訪者都反對墮胎。其中一位受訪者非常堅持自己的立場而且強烈反對墮胎，而另外一位的意見就比較微弱，而且有點搖擺不定。如果我們只有提出「同意／不同意」這兩個答案選項時，這兩位受訪者的答案可能會完全一樣，然而，我們可以利用提供更多選項的方式（強烈同意、同意）、或者是提出條件式問題（先詢問受訪者同意或是不同意，然後再詢問受訪者對於自己答案的支持程度有多強烈）的方式，去瞭解受訪者意見的兩個面向。

選擇性拒絕　前面已經提過了關於滿意的議題，也就是受訪者會因爲想要偷懶的原因而選擇沒有意見或者是一個中立的答案，而除了這個議題之外，有些受訪者也可能會拒絕回答某些問題。在牽涉到某種敏感的議題時，經常會發生此種狀況，也就是受訪者寧可拒絕回答問題，也不願選擇一個在社會上可能被認爲是不適當的答案。

例如，在 1962 年時，有超過三分之一的美國人，拒絕回答一個關於種族融合的敏感性問題。當許多的受訪者沒有回答問題時，如果這些不願作答的受訪者實際上抱持著特定意見的話，那麼研究發現就可能具有誤導性。例如，如果那些反對種族融合的受訪者都回答「不知道」的話，那麼研究結果就會比所有的受訪者都誠實作答時的情況更加偏向於支持種族融合。在對這些未作答的情況進行調整之後，貝林斯基（Berinsky, 1999）發現到，支持種族融合的美國民眾比例，從 49.4%下滑到 34.9%。他警告說（p. 1225）：「受訪者在這個調查訪談中所表達的意見，並不必然等同於當他們瞭解了這個調查問題時所真正抱持的意見。」

同意／不同意，排序或評分？　測量價值和態度的調查研究者，一直在爭論兩個關於提供給受訪者的答案的議題。[31] 問卷裡的題目是否應該先提出一句敘述、並且詢問受訪者是否同意或不同意這句敘述，或者是應該要提供受訪者其他特定的選項？問

游移者　一種在調查研究中的特殊受訪者，此種受訪者雖然對於回答某個調查問題沒有任何知識或是意見，但是卻仍然去選出答案，而且他們所給的答案通常是不一致的。

時近效應　一種發生在調查研究中的結果，也就是受訪者會選擇最後一個出現的答案，而不是去認真地思考所有的答案選項。

卷是否應該包含一組項目、並且要求受訪者去加以評分（例如：贊成、不贊成），或者，問卷是否應該給受訪者一張項目的清單、並且強迫受訪者把這些項目加以排序（例如：從最喜歡到最不喜歡）？

提供受訪者明確的選項是最好的作法。例如，與其詢問受訪者「你同意或不同意下列這句話：『男性比較適合治理國家』？」，我們應該改問「你認爲男性比較適合治理國家、女性比較適合治理國家、或者是兩性都同樣適合治理國家？」教育程度較低的受訪者傾向於同意一項敘述。明確地強迫受訪者去選出答案，可以鼓勵他們進行思考，並且可以避免*反應心向*（*response set*）的偏誤—也就是有些受訪者會有一種同意所有敘述的傾向（你已經在第七章中讀過關於反應心向的問題了）。

被詢問到關於價值問題的調查受訪者，通常幾乎不會顯示出任何的變異性，而且他們的回答也會集中在兩個極端。解決此種問題的方法之一，就是使用一種「先排序然後再評分」（"rank-then-rate"）的程序。我們首先要求受訪者把價值排序，順序是從最重要的到最不重要的。接著，我們要求他們對每一個價值給予評分。例如，受訪者先以重要性對各種不同的價值（例如：世界和平、個人財富、家庭安全等等）進行排序。接著，請受訪者依照完全不重要到極度重要的次序，爲每個價值指定一個從 1 到 10 的數值，數值越高代表重要性越高。某位受訪者可能把世界和平的次序排在個人財富的前面，但是當被要求去爲世界和平或是世界和平對個人的重要性加以評分時，該位受訪者可能會給世界和平 4 分，但是卻給個人財富 8 分。[32]

記住，我們必須以公平的方式來提出所有的選項，並且不能去引導受訪者選擇某個特定的選項或是答案。例如，如果你提出的問題是「你支持節約能源的法案嗎？或者是你會因爲此種法案難以執行而反對此種法案呢？」，而不是直接問「你支持或是反對節約能源的法案？」的話，你就是等同於提出了一個反對節約能源法案的引導式問題。這就是爲什麼我們必須去要求受訪者以排序的方式去在選項之間作出選擇（例如，請告訴我你的第一選擇、第二選擇、第三選擇）、而不是要求他們在一條想像的連續線上對每個項目進行評分（例如：這些選項中的哪一個是最好的？）的原因。受訪者可能會把數個項目都評爲同樣的高分，但是如果我們要求他們把所有的項目彼此互相比較並且進行排序的話，他們就必須把這些項目依照一個等級次序而加以排列出來。[33]

把一個回答量表加上數字，能夠對受訪者有所幫助，並且給他們一些瞭解的線索。當我們把變項概念化爲相對的選項時，分別以正數和負數來代表兩個極端（例如：+5 以及−5）是最好的作法。如果我們將變項概念化爲一個單一的連續概念的話，使用一連串的正數（例如：從 0 到 10）則會是最好的作法。要再強調一次的是，我們所採取的作法，會引導我們應該如何去安排問題以及各個問題的答案選項。

視覺的呈現方式，包括使用顏色、符號、以及圖片等等，會影響到受訪者對於問卷的反應。有時候，視覺的呈現方式會比問題措辭方式的變化造成更大的影響。受訪者傾向於把一組回答的中間部分詮釋爲一個典型的或是中間的選項、把在問卷上的空間接近性當作是在暗示意義的相似性、把在一份直立式清單最上面的幾個項目認爲是最符合研究者期望的選項、並且把答案之間的空間差異或者是不同的顏色

當作是在暗示意義方面一些更爲重要的差異。除此之外，受訪者通常會覺得把答案類別以直立方式排列，會比以水平方式排列時更爲清晰易懂、更不易造成混淆。[34]

問題的形式以及問題的設計方式可能會影響到我們所得到的結果。洛克伍德、桑斯特、以及狄爾曼（Rockwood, Sangster, and Dillman, 1997）曾經詢問過大學生每天念書多少個小時。其中有些學生所得到的是一組「數字較低」的五個答案選項，範圍從一天 0.5 個小時到超過 2.5 個小時。而其他學生所得到的是一組「數字較高」的五個答案選項，範圍從每天低於 2.0 個小時到超過 4.5 個小時。在得到「數字較低」答案選項的學生當中，有約 28%的人說他們一天念書超過 2.5 個小時。但是在得到「數字較高」答案選項的學生當中，有 69%的人說他們一天念書超過 2.5 個小時。顯而易見地，答案選項的安排方式，會影響到受訪者所選擇的答案。研究者們也比較了相同問題以及答案選項的不同調查形式。他們寄了郵寄問卷給某些學生，並且也用電話訪談了其他的學生。他們所得到的答案也會隨著調查形式而有所變化。在接到郵寄問卷、並且其答案選項是「數字較低」的五個答案那一組的學生當中，有 23%的學生說他們每天念書超過 2.5 個小時。而在接受電話訪談、而且答案選項爲「數字較低」的那一組學生當中，則有 42%的人說他們每天念書超過 2.5 個小時。而當答案選項爲「數字較高」的五個選項時，不論是郵寄問卷或是電話訪談，其結果則大約相同。在同樣的一項研究中，研究者另外使用了兩組類似的「數字較高」以及「數字較低」的答案選項來詢問學生他們每天看電視的小時數，並且也比較了郵寄問卷以及電話訪談的結果。而對於看電視這個主題而言，答案類別的範圍高低、或者是調查的形式，對答案都沒有太大的影響。

這項研究告訴我們三件事情：受訪者會依賴問題中的答案類別的範圍來當作指引；相較於像是訪談這種匿名性較低的調查形式，在像是郵寄問卷這種匿名性比較高的調查形式中，受訪者所給的答案會比較誠實一些；而且對於某些主題而言，答案類別以及調查的形式對於受訪者的回答會造成比較大的影響。[35]

措辭的議題

在製作問卷的時候，我們會遭遇到兩個與措辭有關的議題。第一個議題是我們在之前已經討論過的，也就是要使用簡單的詞彙以及文法，以使混淆的情況減至最少。第二個議題則牽涉到特定詞彙所產生的效果，而這個議題比較微妙、難以捉摸，因爲我們無法事先得知某個詞彙是否會影響到受訪者作答的方式。[36]

目前已經有大量的研究文獻記錄了在「禁止」（forbid）以及「不允許」（not allow）之間的差異，而此種差異正好可以用來說明這個問題。這兩個詞彙的意義是完全相同的，可是許多人比較願意「不允許」某件事、而不是「禁止」某件事。一般而言，教育程度較低的受訪者會比教育程度較高的受訪者更容易受到微小的措辭差異所影響。

除此之外，某些字眼會引起情緒反應，而且我們才剛剛開始瞭解這個問題。例如，史密斯（Smith, 1987）發現到，根據一個問題所詢問的內容是關於支出經費「去幫助貧窮的人們」或是「提供福利」，在美國的調查結果當中就會造成極大的差異（例

如，支持的人數可能是兩倍之多）。他認爲，對於美國人而言，「福利」這個詞彙具有強烈的負面意涵（會讓人們聯想到懶惰的人、浪費以及昂貴的計畫等等），因此研究者最好避免使用這個詞彙。

我們也可以利用從一個表面上看似毫無爭議的問題當作例子，來說明可能的**措辭效應**（wording effects）。彼得森（Peterson, 1984）檢視了四種詢問年齡的方式：「你幾歲？」「你的年齡多少？」「你出生於哪一年？」以及「你的年齡層是介於 18 到 24 歲之間、25 到 34 歲之間、或是……？」他利用出生證明的記錄來檢查確認回答的正確性，並且發現到，根據所使用的問題形式的不同，有 98.7%到 95.1%的受訪者會給予正確的答案。他同時也發現到，對於錯誤最少的問題形式而言，受訪者拒絕回答的比例是最高的；而對於錯誤最多的問題形式而言，受訪者拒絕回答的比例則是最低的。這個例子顯示出，在一個不具爭議性的事實性問題中的錯誤，可能會隨著些微不同的措辭變化而有所改變。

問卷設計的議題

調查或是問卷的長度 一份問卷的篇幅應該要多長、或是一次訪談的時間應該要多久才算適合呢？[37] 我們比較偏好篇幅較長的問卷或是時間較久的訪談，因爲這樣的成本效益會比較高。在一位受訪者被抽樣出來、被聯絡上、並且已經完成其他的問題時，額外增加幾個問題的成本是相當低的。問卷和訪談沒有所謂絕對適合的篇幅或是時間長度，適當的篇幅以及時間長度完全要取決於調查的形式（稍後將進行討論）以及受訪者的特質。5 分鐘的電話訪談幾乎不會造成任何問題。郵寄問卷的形式則可能會有比較多變數。對於一般大眾而言，一份簡短的問卷（大約 3 頁左右）是比較適當的。雖然有些研究者曾經成功地使用過長達 10 頁的問卷（裡面大約有 100 個問題），但是，隨著問卷的篇幅越長，受訪者回應的意願也會跟著明顯地降低。對於受過高度教育的受訪者以及一個非常顯著的主題而言，或許有可能採用長達 15 頁左右的問卷。面對面的訪談可能會持續一段很長的時間，即使是長達一個小時也是很常見的狀況。在特殊情境中，研究者甚至有可能進行長達 3 到 5 個小時的面對面訪談。

問題順序或是次序 我們會面臨三個與問題順序有關的議題：整份問卷的組織方式、問題的順序效應、以及脈絡效應。

1. *問卷的組織方式* 一般而言，你在爲問題排出順序的時候，應該要盡量將受訪者的不舒服感以及混淆的情況減至最低。一份問卷包括了開頭部分、中間部分、以及結尾部分的問題。在解釋該項研究的簡短介紹之後，最好的作法是讓受訪者覺得開頭部分的問題是令人愉快的、有趣的、而且容易作答的。此種情況有助於使受訪者對整份問卷感到自在舒服，因此，你應該盡量避免在一開始的部分就去詢問一些沉悶的背景問題、或者是敏感的問題。至於中間部分的問題，則應該依照共同的主題

措辭效應 在調查研究中，當使用某個特定的詞彙或是字眼，會強烈影響到某些受訪者回答問題的方式時，就是出現了此種效應。

來組織安排，若是把不同主題的問題混合在一起，只會徒然造成混淆而已。因此，你應該把相同主題的問題放在一起，並且用一段簡短的敘述（例如：「現在我想要詢問你關於居住方面的問題」）來介紹這個部分，以引導受訪者回答這部分的問題。盡量使問題的主題能夠順暢地、合乎邏輯地銜接在一起，並且以能夠幫助受訪者回憶或是感到舒服的方式來組織安排問題。不要以關於敏感議題的問題來作爲結尾，而且在最後的部分一定要說「謝謝您」。

2. *順序效應*　在問卷裡面的問題的出現順序，也可能會影響到受訪者的答案。[38] 對於缺乏強烈看法的受訪者、教育程度較低的受訪者、以及年紀較大或是記憶力較差的受訪者而言，此種**順序效應**（order effects）所產生的影響似乎是最爲強烈的。[39] 例如，以對於單身女性墮胎問題的支持程度而言，如果這個問題是出現在「當胚胎有嚴重缺陷時，墮胎是否可以接受」這個問題的後面、而不是單獨出現或是出現在關於胚胎缺陷問題的前面時，對於單身女性墮胎的支持度通常就會上升。關於順序效應的經典案例，請參閱「資訊箱 10.7 問題順序效應」。

對於先前問題的答案，也有可能會透過下列兩種方式影響到對於後來問題的答案：透過問題的內容（也就是議題）、以及透過受訪者的回答。例如，你詢問了一名學生：「你是否支持或是贊成學生的教育捐獻（educational contribution）？」。這個問題的答案會因爲之前的問題而有所變化。如果前一個問題是「美國的學生平均繳多少學費？」，受訪者就會把「捐獻」詮釋爲學生必須支付的金錢。如果前一個問題是「瑞典政府在學生身上花了多少錢？」，那麼受訪者就會把「捐獻」詮釋爲政府所支出的經費。對於先前問題的答案也可能會影響到後來的答案，因爲受訪者在已經回答一部分的問題之後，就會假定之後的問題應該沒有重疊的部分。例如，你先問了受訪者一個問題：「你的太太近況如何？」，而下一個問題是：「你的家人近況如何？」如此一來，大多數的受訪者會假定第二個問題所指的是除了太太以外的其他家人，因爲他們在前面已經回答過關於太太的問題了。[40]

3. *脈絡效應*　調查研究者已經在調查中觀察到強大的**脈絡效應**（context effects）。[41]「脈絡所包含的不只是某個問題對另一個問題所造成的影響而已，而是還包括了訪談者、訪談的情境、以及歷史情境所造成的效果…目前，我們尚無法清楚地掌握問卷的脈絡效應以及對於調查的回答效應之間有什麼樣的關聯性」（Schuman, 1992:18）。相對於電話調查而言，脈絡效應在郵寄問卷調查中所產生的影響是比較明顯的，因爲在郵寄問卷調查中，受訪者可以看到所有的問題。[42]

你可以用兩種方法來處理關於脈絡效應的問題。你可以使用**漏斗順序**（funnel sequence）來排列問題，換句話說，你可以先詢問一般性的問題，然後再詢問特定具體的問題（例如，先詢問關於健

順序效應　調查研究中的一種效應，意思是指某些已經先問過的主題或是問題，影響到受訪者對於後續問題的答案。

脈絡效應　在調查研究中的一種效應，當調查的整體氣氛、背景、或者是受訪者所聽到的主題，影響了他們對於後續問題意義的詮釋方式時，就是發生了脈絡效應。

漏斗順序　把一份問卷上的調查研究問題，從一般性問題到特定具體問題的方式來加以組織。

實例箱 10.7　問題順序效應

問題一

「你認為美國是否應該讓來自於其他國家的共產主義報社記者來此地採訪，並且把他們所採訪的新聞寄回報社去刊登出來？」

問題二

「你認為像是俄羅斯之類的共產主義國家是否應該讓美國的新聞記者去入境採訪，並且把他們所採訪的新聞寄回美國去刊登出來？」

回答「是」的百分比

先聽到的問題	*問題一回答「是」（共產主義的記者）*	*問題二回答「是」（美國的記者）*
問題一	54%	75%
問題二	64%	82%

先回答第一個問題的脈絡，影響了對第二個問題的答案。

*資料來源：*改寫自 Schumna and Presser (1981). *Questions and Answers in Attitude Surveys: Experiments in Question From, Wording, and Context*, p. 29. New York: Academic Press. With permission from Elsevier.

康的一般性問題，然後再詢問關於特定疾病的問題）。或者，你也可以把受訪者分成兩組，然後給其中一組的問題是以某種方式排列，而給另外一組的問題則是以另外一種方式來排列，之後再檢視比較兩組的結果，看看問題的順序是否會造成任何差異。如果你確實發現了問題順序效應的話，那麼，哪一種問題順序才能讓你得知受訪者真正的想法呢？可惜的是，你可能永遠也無法確定這個問題答案到底是什麼。

例如，在幾年之前，我的學生曾經針對兩個主題進行過一項電話調查：對於犯罪的關心程度、以及對於一條反酒醉駕車的新法令的態度。其中有一半由隨機方式選出的受訪者先聽到關於酒醉駕車的新法令的問題，而另外一半的受訪者則是先聽到關於犯罪的關心程度的問題。我檢視了調查結果，以確定是否有任何的脈絡效應－也就是因爲主題順序不同而產生的差異。我發現到，先被問到反酒醉駕車法令問題的受訪者，對於犯罪的關心程度會低於另外一組先被問到犯罪問題的受訪者。相同地，先被問到反酒醉駕車法令問題的受訪者，也比先被問到犯罪問題的受訪者更加支持反酒醉駕車的法令。第一個主題建立了受訪者在回答第二個主題時所依據的脈絡。在我們向受訪者詢問了關於一般犯罪的問題、而且他們也思考過了暴力犯罪事件之後，酒醉駕車對他們而言就可能會變成一個似乎比較不是那麼重要的議題。對比之下，在我們詢問過關於把酒醉駕車當作是犯罪行爲的問題之後，受訪者則可能會對一般的犯罪行爲覺得比較不是那麼關心。

我們必須記住，當問題的意義越模稜兩可的時候，脈絡效應的影響就會越大，這是因爲受訪者會依賴脈絡來詮釋以及理解問題。之前針對相同主題所提出的問題、以及在某個問題之前剛剛出現的問題，都可能具有強大的脈絡效應。例如，薩德曼等人（Sudman et al., 1996: 90-91）曾經使用三種方式去詢問受訪者對於政治的關心程度，並且對這三種方式所得到的結果進行比較。當研究者單獨提出這個問題

時，有大約 21%的受訪者說他們「偶爾」或是「幾乎不」關心政治。當研究者先詢問受訪者關於他們選出的民意代表最近所做的事情、然後再提出對於政治關心程度的問題時，回答說幾乎不關心政治的受訪者比例幾乎變成兩倍之多（大約 39%）。這可能是因爲先提出的知識性問題使得受訪者覺得他們對於政治方面的議題所知不多。當研究者在上述兩個問題之間，加入了關於選出的民意代表對於該地區所提供的「公共關係工作」數量的問題時，有 29%的受訪者回答說他們並不關心政治議題。這個新加入的問題給了受訪者無法回答第一個問題的藉口，也就是他們可以把他們對於政治議題的無知歸咎於民意代表身上。由於問題的脈絡可能會造成影響，因此研究者必須時時刻刻注意到脈絡的問題：「提升對於問題的瞭解，並不是問題措辭方式的唯一功能而已。受訪者會利用問題的脈絡所提供的資訊，來決定問題所打算要傳達的意義是什麼」（Sudman et al., 1996: 69）。

排版以及格式　有兩個關於排版以及格式的議題：問卷整體的實體排版方式，以及問題和作答的格式。

問卷排版　對於訪談者以及受訪者而言，問卷的排版方式都是非常重要的。[43] 問卷的版面應該要清楚、整齊、而且容易閱讀。你應該在問卷裡加上辨識的資訊（例如：組織的名稱），並且爲每個問題都加上編號。絕對不要把所有的問題都擠在一起、或者是採用容易令人產生混淆的組織方式。如果你爲了在郵資或是印刷上面省下一些錢，到最後可能會因爲回收率過低而造成效度變差、或者是造成訪談者以及受訪者的混淆，結果使得整體成本反得變得更高。使用高品質的圖表、在問題之間留下適當的間距、以及使用良好的排版方式，可以讓問卷呈現出專業的感覺，如此一來，將會鼓勵受訪者願意正確地、完整地完成整份問卷，並且有助於受訪者順暢地完成問卷。如果你所使用的是訪談的形式，那麼你應該要製作一個*封面*（*face sheet*）（我們將在第十三章中討論關於封面的問題）來當作問卷的一部分，以作爲後續的行政作業之用。封面應該要包含訪談的時間以及日期、訪談者的姓名、受訪者的識別號碼、以及訪談者的評論與對該次訪談的觀察等資訊。

你也應該要給予訪談者和受訪者明確的指示，而在給予指示時，最好的作法是以和問題不同的格式（例如，以不同的顏色或是字體）將這些指示印出來，以和問題之間有所區隔。這樣子可以幫助訪談者辨別哪些部分是要去詢問受訪者的問題，而哪些部分又是單獨針對訪談者所提供的指示。

對於郵寄以及網路問卷而言，排版方式是極爲重要的，因爲在此種調查形式中，並沒有一位友善的訪談者與受訪者進行互動。相反地，真正能夠說服受訪者完成問卷的，是問卷的整體外觀以及版面。

在使用郵寄問卷來進行調查時，你應該在最前面的部分附上一封禮貌的、專業的說明信函，並且標示研究者的身分以及有問題時可以諮詢的電話號碼。各種微小的細節也是非常重要的。如果受訪者收到的是上面只有標明收件人的大宗郵寄的牛皮紙信封、或者是問卷與回郵信封的尺寸完全不合時，他們就會對這項調查覺得反感。

網路調查目前仍然是比較新穎的作法，而且研究者們也還在學習什麼樣的設計

特色是最具有效果的，但是視覺設計的細節仍然是相當重要的（請參閱本章稍後部分針對網路調查所進行的討論內容）。

問題格式 你必須決定問題以及作答的格式。受訪者應該是要圈選他們的答案、在方塊裡打勾、勾選小點、或者是在空格中寫上答案？基本原則是要讓作答方式既清楚而且不會造成混淆。在方塊或是括弧中打勾、以及圈選數字的作答方式，通常是最爲清楚的方式。除此之外，以垂直的方式、而不是水平的方式列出答案，會使受訪者更加容易看到所有的答案（請參閱資訊箱 10.8 問題格式的範例）。針對條件式問題，你應該要加上箭頭以及相關的說明。視覺上的輔助也是相當有幫助的。例如，當詢問到受訪者對於某人的感覺是溫暖或是冷漠時，提供一個像是溫度計一樣的圖形，將會有助於他們回答問題。在利用相同的答案類別來呈現一系列的問題時，**矩陣式問題**（matrix question）〔又稱爲*格狀問題*（*grid question*）〕則是一種簡潔的方式，此種方式不但可以節省空間，而且也讓受訪者以及訪談者更加容易在相同的答案類別中去記下答案。

未回應 若是研究者無法從每一個被抽樣的受訪者那兒得到有效的回答，就會削弱一項調查的效力。除了調查研究之外，人們也會被要求去回應許多由慈善機構、市場行銷公司、以及候選人後援會之類的機構所進行的調查。一般而言，慈善機構以及市場行銷公司的回應率都會比較低，而政府機構的回應率則比較高。未回應的狀況可能會是一個重大的問題，因爲如果在被抽樣的受訪者中有極高比例的人不願意回應的話，調查結果可能就無法進行類推，尤其是當那些不回應的人與那些回應的人之間存在著明顯差異的時候。

幾乎所有的人都會在某個時刻完成調查，而且在 1960 年代之後，大型報紙上關於調查或是民意調查結果的報導也迅速地成長。到了 1970 年代時，似乎每一天都會有一則新聞報導引用了某項調查或是民意調查的結果。隨著調查越來越常被使用，回應率也跟著開始逐漸下滑。在調查當中未回應的比例差異很大，而對於學術組織而言，未回應的比例範圍大約是從 25%到 33%之間。而在美國，大型學術調查的未回應比例，從在 1950 年代時少於 10%，到 1980 年代時已經逐漸上升到大約 25%。在大多數的國家裡，社會大眾對於調查研究的參與程度已經逐年下降，其中荷蘭民眾拒絕作答的比例是最高的，而在美國最高則可能達到 30%左右。[44] 然而，在商業的民意調查（洛普、蓋洛普、以及 CBS 等等）、以及選舉的民意調查中的未回應比例，則可能會更高，有時候甚至可能高達 50%左右。

研究者已經發現到，有一群人數逐漸增加的「強硬派」（"hard core"），這群人會拒絕所有的調查。除此之外，由於人們普遍認爲調查的數量實在太多了，因此一般的調查參與率也已經降低了。造成調查參與率下降的其他原因，則包括了對於陌生人的恐懼、越來越忙亂的生活型態、個人隱私受到侵犯、以及對於權威越來越不信任等等。除此之外，由於有一些公司濫用調查來推銷產品或是說服民眾、有一些問卷

矩陣式問題 一種調查研究的問題格式，在此種格式中，會以一種簡潔的方式把共用相同答案類別的一組問題歸類在一起。

資訊箱 10.8　問題格式的範例

水平以及垂直答案選項的範例

你認為要離婚很容易、很困難、或者是還好？

○很容易　○很困難　○還好

你認為要離婚很容易、很困難、或者是還好？

○很容易
○很困難
○還好

矩陣式問題格式的範例

	非常同意	*同意*	*不同意*	*非常不同意*	*不知道*
老師講課速度太快	○	○	○	○	○
我在這門課學到很多	○	○	○	○	○
測驗非常簡單	○	○	○	○	○
老師講了很多笑話	○	○	○	○	○
老師上課很有條理	○	○	○	○	○

一些答案類別選項的範例

極佳、良好、還可以、差勁
同意／不同意
支持／反對
非常同意、同意、有點同意、有點不同意、不同意、非常不同意
太多、太少、剛剛好
更好、更差、差不多
頻繁、經常、很少、從來沒有
總是、經常、有時候、幾乎沒有、從來沒有
有可能、不太可能、沒有差別
非常有興趣、有興趣、沒興趣

的設計品質過於低劣、以及對於調查的解釋不足或是不恰當等等原因，也使得民眾拒絕參與合法的、嚴肅的調查的情況越來越普遍。

在社會中最有興趣的、最能夠取得資訊的、以及最爲活躍的成員，通常比較願意參與調查，而這也意味著未回應的狀況可能會傷害到調查的效度，並且將整體人口中某個特定部分的人們排除在調查結果之外。在美國，不願回應的受訪者通常是年輕的非白人男性，以及教育程度較低的人。

未回應的比率具有五個成分（請參閱資訊箱 10.9 關於回應率的混淆狀況）：[45]

1. *尋獲*（*location*）研究者是否能夠尋獲一名被抽樣出來的受訪者？
2. *接觸*（*contact*）一名被尋獲的受訪者是否在家？或者，在研究者多次嘗試之後是否能夠聯絡上該名受訪者？

資訊框 10.9　關於回應率的混淆狀況

人們對於回應率通常會有一些感到困惑或是混淆的狀況，這是因為整體回應率是由五個成分中的回應成功率來決定的，而且其中每一個成分都各自有自己的比率：

尋獲率：在抽樣架構中的受訪者被成功尋獲的百分比率。

接觸率：在被尋獲的受訪者當中，研究者能夠成功接觸的百分比率。

符合資格率：在接觸到的受訪者當中，符合資格的百分比率。

合作率：在接觸到而且符合資格的受訪者當中，同意參與調查的百分比率。

完成率：在願意合作的受訪者當中，完成整項調查的百分比率。

整體回應率：在一開始的抽樣架構的所有受訪者當中，被成功地尋獲、進行接觸、符合資格、同意參與調查、並且完成整份問卷的受訪者的比率。

例如，研究者從一個抽樣架構中的 1,000 名受訪者開始，以電話或是地址尋獲了其中的 950 名受訪者，然後成功地接觸了其中的 800 人（藉由訪談或是成功地郵寄問卷的方式），並且確認其中有 780 人符合資格（也就是說，符合基本的標準、會說某一種語言、而且具備正常的心智能力等等）。接著，研究者發現其中有 700 人願意合作填寫問卷或是進行訪談，而且最後有 690 人完成了整份問卷或是訪談。此種情況產生了下列的比率：尋獲率：95%；接觸率：84.2%；符合資格率：97.5%；合作率：89.8%；完成率：98.6%：整體回應率：69%。整體回應率的計算方式是所有個別比率的乘積，也就是. 95 × .842 × .975 × .898 × .986 = .690

3. *符合資格*（*eligibility*）接觸到的受訪者在年齡、種族、性別、以及國藉等等方面是否符合調查的目的？
4. *合作*（*cooperation*）符合資格的受訪者是否願意接受訪談、或者是填寫問卷？
5. *完成*（*completion*）一名願意合作的受訪者是否在結束前就停止作答、或者是開始對大多數的問題都回答「不知道」或是「無意見」？

若是想要提升整體的回應率，我們就必須減少每一種類型的未回應狀況。

尋獲率　要提升尋獲率，意味著要使用更好的抽樣架構以及地圖或者是電話號碼簿。若是想要提升接觸受訪者的機會，我們就必須重覆打許多通電話、在一天之間的不同時間打電話、並且拉長進行接觸的時間。在美國，無法成功接觸受訪者的情況可能和以下幾種因素有所關聯：高人口密度、城市中的貧民區、非屋主所居住的住宅（也就是出租）、高犯罪率、高比例的少數民族人口、存在著一些實體的障礙物（也就是像籬笆、鐵窗、一些寫著內有惡犬或是禁止進入的告示牌等等）、以及獨居的成年人或者是沒有任何年幼兒童的家戶等等。雖然擁有較高收入以及教育程度較高的民眾可能會比較容易尋獲並且進行接觸，但是此種受訪者就算被尋獲了，也比較不願意合作。正如同葛洛夫斯以及庫伯（Groves and Couper, 1998: 130）所觀察到的，「我們在我們的資料中找到了對下列這個想法的支持性證據：高社經地位（socio-economic status, SES）家庭裡的人們，比在低社經地位團體裡的人們更不願意合作參與調查。」

雖然目前顯示來電號碼（caller ID）的情況已經越來越普遍，但是仍然只有少數的民眾會利用來電號碼以及電話過濾技術來明顯地阻擋調查研究的進行。[46]

接觸率　在未回應或是拒絕參與的情況中，有一個相當關鍵的時間點會出現在訪談者與受訪者剛剛進行接觸的時候。當受訪者相信調查的主題或是結果對他或她是有利的（也就是受訪者覺得很感興趣、或者是會爲他們帶來直接的益處）、或者是當訪談者在對受訪者所進行的介紹當中使用了「量身打造」（"tailoring"）的技術（稍後將進行討論）或是指供一個小型的誘因（例如：一些金錢方面的獎勵）時，合作率就會因此而增加。

符合資格率　我們可以利用一些方法來增加符合資格的比率，像是建立嚴謹的受訪者篩選方式、使用較好的抽樣架構定義、並且雇用能說多種語言的訪談者等等。在減低受訪者拒絕參與調查的狀況方面，我們可以在訪談之前先寄送通知信件、使用一些誘因或是獎勵（也就是像前面剛剛提過的，一份小禮物或是一些金錢的獎勵）、調整訪談者的行爲以及敘述方式（也就是與受訪者進行眼神的接觸、表現出真誠的態度、解釋抽樣方式或是調查的目的、強調訪談的重要性、清楚說明對於機密性的承諾與保證等等）。我們也可以使用替代的訪談者（也就是具有不同人口統計學特質的訪談者，像是不同的年齡、種族、性別、國藉等等）、使用替代的訪談方法（像是電話訪談以及面對面訪談等等）、或者是接受在同一戶家庭裡的其他成員來當作受訪者等等。

合作率　由於研究者逐漸使用新聞型式的信件以及親自打電話的方式、而不是使用標準學術信件，因此在都市內貧民區的居民、低收入的人們、以及少數種族的成員的合作率也逐漸地增加了。對於政府以及社會服務機構感到悲觀的人、以及覺得自己遭到誤解的人，在訪談者以他們能夠感同身受的方式解釋調查的本質之後，通常都會願意參與調查。[47]

如同之前已經提過的，在所有形式的調查中，一些事先給予的誘因都能增加受訪者的合作率，而且對於調查結果或是未來的參與率似乎都不會造成任何的負面影響。例如，布瑞姆（Brehm, 1994）就發現到，在沒有事先進行接觸的情況下，有 71%的受訪者願意合作，但是在有事先進行接觸（以信件告知）以及提供誘因（提供 1 美元的獎勵）的情況下，受訪者的合作率就會提升到 78%，而且受訪者也會比較願意侃侃而談。除此之外，受訪者們也不會覺得研究者提供不同的獎勵金額是不公平的作法。[48]

研究者過去一向把受訪者當作是已經具有完整的態度、信念、以及意見的成熟個體，而且當受訪者在調查中接受詢問時，可以隨時利用這些態度、信念、以及意見來回答問題。不過，現在我們已經不再抱持著這樣的看法，而是把調查視爲牽涉到數個不同的過程。第一個過程就是要引起受訪者願意合作的動機，如此一來，受訪者才會願意完全地參與整個調查過程。第二個過程是要去協助受訪者正確地詮釋調查問題，以及幫助受訪者從記憶或是過往的經驗中去找出適合的以及正確的答案。第三個過程則是幫助受訪者恰當地回答、或者是提出適合的答案（請同時參閱本章前面部分所提過的範例 10.1）。

有兩個相關的理論能夠幫助我們解釋此種合作動機的過程。社會交換理論（social exchange theory）、或者是整體設計方法（total design method）（請參閱Dillman, 1978, 2000）把正式的調查視爲一種社會互動的特殊類型。換句話說，受訪者所表現出的行爲，是根據他或她預期在合作之後所應該得到的事物來決定的。爲了增加回應率以及正確性，我們必須將合作所造成的負擔減到最少的程度，而方法是讓受訪者很輕鬆地就可以參與調查，並且利用提供好處的方式（也就是受到尊重的感覺、物質的誘因、以及情緒的獎勵等等）來使合作所得到的獎勵最大化。

槓桿顯著性理論（leverage saliency theory）認爲，顯著性以及興趣／動機會因受訪者而異。對於調查過程中的特定面向（例如，時間長度、調查的主題、贊助者等等），不同的人所評估的價值（不論是正面的或是負面的）也會有所不同。爲了要使調查的合作率達到最高的程度，我們必須要在調查過程的早期階段就確認並且呈現出得到正面評價的面向。而這種情況具有二個實際層面的意涵：贊助者以及量身打造。*贊助者*（*sponsorship*）所指的是進行調查或是提供調查經費的組織。而**量身打造**（tailoring）則是指訪談者針對特定的受訪者而去調整在介紹中所說的內容，以突顯出訪談者認爲可以鼓勵受訪者合作的部分。要達到量身打造的目的，訪談者必須先接受訓練，以對各種不同的家戶類型以及關心的事項保持敏感度，如此一來，他們才能夠「讀出」整個背景環境以及各種不同的言語與非言語的線索。訪談者應該要能夠快速地轉換到替代的腳本，以說服受訪者參與調查，並且爲特定的受訪者量身打造特定的說服方式。[49]

完成率 狄爾曼（Dillman, 2000: 252）觀察到，如果有人親自把問卷交給受訪者，而不是在門口撿到問卷或者是以郵寄的方式收到問卷的話，受訪者自填式問卷的完成率就會提高。狄爾曼利用派人親自遞送問卷給受訪者、寄送兩次後續的提醒信件、以及加上完成問卷的金錢誘因等方式，使得回應率提高到77%（當其中有一項或多項技術沒有被使用時，回應率則是在53%到71%之間）。

整體回應率 已經有大量的文獻資料，檢視了如何增加郵寄問卷的回收率的問題（請參閱資訊箱 10.10 提高郵寄問卷回收率的十種方法）。[50]

曾有研究者從在1940年到1988年間出版的25份期刊中，找出 115 篇主題爲郵寄調查回收的文章，並且針對這些文章進行了後設分析（meta-analysis）（我們已經在第五章中解釋過後設分析的概念了），分析的結果顯示，說明信函、問卷頁數少於四頁、附上郵資的回郵信封、以及附上少量的金錢報酬等等，都有助於增加問卷的回收率（Yammarino et al., 1991）。另外一個對郵寄調查以及網路調查進行比較的後設分析則發現，郵寄調查的回收率比較高。大學生受訪者比較願意參與網路

槓桿顯著性理論 一個關於調查研究合作率的假設，該假設認為，對於不同的受訪者而言，調查訪談中的顯著面向也會有所不同，而且受訪者會依據訪談中的不同特定面向來決定是否要與研究者合作。

量身打造 要求訪談者找到在訪談中讓受訪者覺得顯著以及有正面評價的面向，並且強調這些面向，以增加受訪者在調查研究訪談中的合作意願。

資訊箱 10.10　提高郵寄問卷回收率的十種方法

1. 在問卷上標明為寄給某個特定的人士，而不是只有標明「收件人」而已，並且以限時信件的方式寄出。
2. 在問卷的最前面附上一封仔細撰寫、標明日期的說明信函。在說明信函中禮貌地請求受訪者的合作、保證機密性、解釋調查的目的、並且提供研究者的姓名以及電話號碼。
3. 一定要附上貼有郵資的、寫好地址的回郵信封。
4. 問卷應該要有簡潔整齊的、吸引人的排版方式，而且頁數應該合理。
5. 問卷應該以專業方式印刷、容易閱讀、而且有明確的指示或說明。
6. 對於沒有回覆的受訪者，要寄發兩次提醒信函。第一封提醒信函的寄送時間是在寄出問卷大約一個星期之後，第二封則在寄出第一封的一個星期之後。禮貌性地再次請求受訪者的合作，並且詢問是否需要再寄一次問卷。
7. 不要在重要假期的期間寄送問卷。
8. 不要把問題印在紙張的背面，而應該在背面留下空間，並且請受訪者寫下一般性的評論或是意見。
9. 如果贊助者位於當地地區、而且被認為是具有合法地位的機構（例如：政府機關、大學、大型的企業等等），則有助於提高回收率。
10. 如果可能的話，在信中附上小額的金錢誘因（美金 1 元）。

調查，但是其他的受訪者（例如，醫生、教師、消費者等等）則比較偏好郵寄調查。對於網路調查而言，後續提醒的效果似乎略遜於郵寄調查（Shih and Fan, 2008）。而許多研究者所建議的技術，則是依循著整體設計方法，並且有助於使受訪者能輕鬆地完成調查、並且也覺得相當有趣。

調查的類型：優點與缺點

郵寄以及自填式問卷

優點　我們可以把問卷直接拿給受訪者或者是郵寄給他們，然後由受訪者自行閱讀說明以及問題。並且記下他們自己的答案。就算是單獨一位研究者，也可以用極低的成本獨立進行此種類型的調查，而且還能涵蓋廣泛的地理區域。受訪者可以在方便的時候完成問卷，而且如果有必要的話，也能去確認自己所寫的答案。郵寄問卷具有匿名性，而且也能避免訪談者的偏誤。郵寄問卷非常有效，而且如果受訪對象具有較高的教育程度、對於調查主題或者是調查組織有強烈的興趣，那麼也會有不錯的回收率。

缺點　由於許多人不一定會完成問卷並且把郵寄問卷寄回去，因此郵寄問卷的最大問題就是回收率過低。大多數的問卷會在兩週之內回收，但是剩下的問卷則可能在

長達兩個月的時間之內陸陸續續地寄回來。我們可以針對未回應的受訪者寄發提醒信函，以提高問卷的回收率，但是此種作法也會增加資料蒐集的時間以及成本。

除此之外，我們也無法去掌控受訪者完成問卷時的條件或是狀況。例如，在同時寄回來的兩份問卷中，其中一份問卷可能是由一群在狂歡派對中喧鬧的人們所完成的，而另外一份則可能是由一位熱心積極的受訪者所填寫的。而且，當受訪者在填寫問卷時，沒有任何人在場去澄清問卷上的問題、或者是在受訪者作答不完全時去探問更多的資訊。也可能會有除了被抽樣的受訪者之外的人（例如：配偶、新的成員），在研究者不知情的情況下打開信封並且完成了問卷。我們也無法親自觀察受訪者對於問題的反應、受訪者的生理特質、或者是填答時的整體情境。例如，一位 70 歲的、單獨住在農場裡的貧困老婦人，可能會謊稱說她是一名成功的 40 歲亞裔醫生，目前住在城裡，而且育有三名子女。雖然此種極端的謊言是相當罕見的，但是研究者仍然有可能會對某些嚴重的錯誤渾然不覺。此外，不同的受訪者完成問卷的時間可能會間隔數個星期之久、或者是不按照問題的順序作答。作答不完整的問卷也可能會是一個嚴重的問題。

郵寄問卷的形式也限制了我們所能使用的問題。那些需要有視覺輔助的問題（例如：看著這張圖片，並且告訴我你看到了些什麼）、開放式的問題、許多條件式的問題、以及複雜的問題，在大多數的郵寄問卷裡都是不能使用的。同樣的，郵寄問卷也不適合那些不識字或者是識字不多的人（請參閱表 10.4）。

電話訪談

優點 電話訪談是一種很受歡迎的調查方法，因爲有大約 95%的人口都可以透過電話來進行接觸。在進行電話訪談時，會由一位訪談者打電話給一名受訪者（通常是在家裡），詢問一些問題，並且將答案記錄下來。研究者可以從名單上或者是電話號碼簿中、或者是利用隨機數碼撥號（RDD）的方式去抽樣出要進行訪談的受訪者，並且能夠很快地聯絡到在地理上距離很遙遠的受訪者。一群訪談者可以在數天之內就訪談到在一個國家中的 1,500 名受訪者，並且只要回撥數次電話，就能達到高達 80%的回應率。由於訪談者必須花時間去進行訪談，因此電話調查的花費會比郵寄問卷高昂。一般而言，電話訪談是一個具有彈性的方法，不但具有面對面訪談的大部分優點，而且成本也遠比面對面訪談低很多。訪談者可以掌控問題的順序，並且也能使用一些探問的技巧。某個特定的受訪者會被挑選出來，而且也有可能單獨回答所有的問題。我們可以得知受訪者回答問題的時間，而且也能有效地運用條件式問題。

電腦輔助電話訪談（CATI）
一種在進行電話訪談時所使用的電腦技術，在使用此種技術時，訪談者坐在電腦螢幕以及鍵盤前面，念出顯示在螢幕上的問題，並且直接將受訪者所回答的答案輸入電腦。

大多數的研究者會在電話訪談中使用電腦輔助的科技，而我們在此要討論其中的兩種技術。**電腦輔助電話訪談**（computer-assisted telephone interviewing, CATI）是一種被廣泛使用的技術，[51] 在使用此種技術時，訪談者會坐在電腦前面，而且由

表 10.4　調查的類型以及其特性

特性	調查的類型			
	郵寄問卷	*電話訪談*	*面對面訪談*	*網路調查*
執行方面的議題				
成本	便宜	中等	昂貴	最便宜
速度	最慢	快速	緩慢到中等	最快速
長度（問題數）	中等	簡短	最長	中等
回應率	最低	中等	最高	中等
研究控制				
探問的可能性	否	是	是	否
特定的受訪者	否	是	是	否
問題順序	否	是	是	是
只有一名受訪者	否	是	是	否
視覺觀察	否	否	是	是
不同問題的成功與否				
視覺輔助	有限制	無	是	是
問放式問題	有限制	有限制	是	是
條件式問題	有限制	是	是	是
複雜的問題	有限制	有限制	是	是
敏感的問題	部分	有限制	有限制	是
偏誤的來源				
社會期望	無	部分	較嚴重	無
訪談者偏誤	無	部分	較嚴重	無
受訪者的閱讀能力	有	無	無	部分

電腦來進行撥號。訪談者會戴著耳機以及麥克風，並且念出在電腦螢幕上所顯示的問題給特定的受訪者聽，然後透過電腦鍵盤來輸入受訪者的答案。電腦程式會控制接下來出現的問題，並且允許使用複雜的條件式問題。CATI 可以加快整個訪談過程，並且減少訪談者可能犯下的錯誤。原本在一般訪談中有兩個分開的步驟，也就是由訪談者在紙上寫下答案、然後再由其他人把這些資訊輸入電腦裡，而在使用 CATI 時則可以省略這些步驟，並且加快資料蒐集的速度。訪談者會戴著耳機及麥克風，並念出電腦螢幕上所顯示的問題給特定的受訪者聽，然後透過電腦鍵盤輸入受訪者的答案。電腦程式會控制接下來出現的問題，且允許使用複雜的條件式問題。CATI 可以加快整個訪談過程，減少訪談者可能犯下的錯誤。原本在一般訪談中有兩個分開的步驟，也就是由訪談者在紙上寫下答案、然後再由其他人把這些資訊輸入電腦裡，

互動語音回覆（IVR） 一種在電話訪談中所使用的電腦技術，在訪談時，受訪者會聽到由電腦所自動讀出的問題，並且使用按鍵式電話或者是聲控軟體來輸入自己的答案。

而在使用 CATI 時則可以省略這些步驟，更加快了資料蒐集的速度。

互動語音回覆（interactive voice response, IVR）包括了數個透過電話技術來操控的電腦自動系統，而且也被廣泛地使用於市場行銷當中。在使用 IVR 時，受訪者透過電話傾聽問題以及答案選項，並且利用按鍵式電話或者是藉由聲控的方式（電腦會使用語音辨識軟體以辨識受訪者所說出的答案）來輸入答案。相較於使用真正的訪談者，IVR 可能會具有一些優勢，像是資料蒐集的過程迅速而且自動化、不會有訪談者在念問題或是記錄答案時所可能發生的錯誤、而且具有高度的匿名性。有些使用 IVR 的訪談者，會先利用一位真人訪談者來充當受訪者並且把器材先架設好，然後再把問題念出來並且錄音下來。對於非常簡短而且非常簡單的調查而言，使用 IVR 技術可能會非常成功。

缺點 IVR 的回應率常出現急速下降的現象（可能會有高達 40%的受訪者沒有完成整份問卷）。[52] 稍微較高的成本、以及有限的訪談時間長度，也都是 CATI 和傳統電話訪談所共有的缺點。除此之外，受訪者有時候可能不方便接電話，而且也無法去接觸到那些根本沒有電話的受訪者。利用訪談者來進行訪談的作法會減低匿名性，但是也會引進潛在的訪談者偏誤。在電話訪談中很難使用開放式的問題，而且也不可能提出一些需要有視覺輔助的問題。訪談者只能注意到嚴重的干擾（例如，背景的噪音），以及受訪者的語調（例如，憤怒或是輕率）或是猶豫不決的狀況。

調查研究者是在人們只有傳統有線電話時發展出電話訪談技術的，因此，自從 2000 年以後，行動電話使用者日漸增加的情況，已經變成一項不得不面對的議題。截至 2006 年爲止，在美國的 18 歲到 24 歲的成人之中，大約有四分之一的人家中只有行動電話而沒有傳統的有線電話，因此無法被涵蓋在目前的 RDD 有線電話抽樣程序之中。僅有使用行動電話的人口數有可能會不斷增加，而這也暗示了研究者在抽取樣本時，必須把那些使用行動電話以及傳統有線電話的受訪者同時結合在一起。相較於傳統有線電話的調查，行動電話調查的回應率通常比較低、拒絕率也比較高、而且之後被說服去參與調查的比率也比較低。早期的研究提供了一些關於行動電話調查的建議，像是在週一到週五的晚上打電話、允許行動電話比傳統有線電話響得更久、特別注意到任何暗示不方便接受訪談的線索（例如，受訪者正在開車）、必須安排回撥電話的時間表、以及決定在重新接觸某個行動電話號碼之前要等待多久的時間等等。[53]

面對面訪談

優點 面對面訪談擁有最高的回應率，而且也允許使用最長而且最複雜的問卷。面對面訪談擁有電話訪談的所有優點，並且也允許訪談者去觀察環境，以及使用非言語的溝通與視覺輔助。受過良好訓練的訪談者能夠詢問任何類型的問題，並且運用

廣泛的探問技巧。

缺點　成本過高是面對面訪談最大的缺點。爲了訪談所必須進行的訓練、旅途奔波、監督指導、以及人事等等的成本可能會非常高昂。而在面對面訪談當中，發生訪談者偏誤的可能性也是最大的。訪談者的外表、說話音調、以及問題的措辭等等，都有可能會對受訪者造成影響。除此之外，在電話訪談中一般都會有一名督導者在一旁傾聽訪談的內容，而在面對面訪談中則沒有此種措施，因此對於訪談者的督導程度也會比較低一些。[54]

電腦輔助面訪（CAPI）（已經在本章之前的內容中描述過了）是一種在詢問敏感議題時所使用的面對面調查的變化形式。在進行 CAPI 時，訪談者會帶著一部筆記型電腦，而受訪者則是在筆記型電腦上完成問卷。訪談者的功能在於鼓勵受訪者完成整份問卷，並且適時地針對一些受訪者不瞭解的問題進行澄清。

網路調查

一直到 1990 年代晚期，一般大眾才能夠廣泛地使用網際網路以及電子郵件。例如，在 1994 年時，只有 3%的美國民眾在家裡或是工作場所裡有電子郵件信箱。而到了 2007 年時，有 62%的家戶同時擁有電子郵件信箱以及網路連線。而到了 2012 年時，有些機構推估美國將有 77%的家戶擁有網路連線。而在其他國家裡，擁有網路連線的家戶比例則更高，例如，在南韓是 97%、荷蘭是 82%、香港是 81%、加拿大是 79%、而日本則是 77%。[55]

優點　網路調查或是電子郵件調查都非常快速、而且也非常便宜。此種調查也允許有彈性的設計，並且也能使用視覺影像、甚至是聲音或是影片。網路調查的兩種方式分別是靜態的以及互動式的。靜態的網路調查或是電子郵件調查，就像是呈現一張紙給受訪者一樣，只不過是在電腦螢幕上呈現的。而互動式的網路或是電子郵件調查，則有一些條件式問題，而且可能會依據之前的答案而呈現不同的問題給不同的受訪者來作答。

缺點　網路調查有一個比較特殊的缺點，就是此種調查既便宜又容易進行。正如同魏斯柏格（Weissberg, 2005: 38）所說的：「由於在網路上舉行一項民意調查根本不必花到多少錢，因此有許多團體在網路上進行民意調查時，幾乎都不會注意到品質的問題。」網路調查有三個缺點或是值得關切的問題：涵蓋範圍、隱私與驗證、以及設計的議題。第一個問題涉及到抽樣方式，以及在接觸與使用網路方面的不平等現象。年紀較大的、教育程度較低的、低收入的、以及住在鄉間地區的民眾通常比較沒有機會接觸到網路，而且大多數目前沒有接觸網路的人也說他們在未來沒有去架設網路連線的打算。除此之外，也有許多人同時擁有多個電子郵件信箱。

第二個問題則牽涉到保護受訪者的隱私。必須輸入密碼或是個人識別號碼的安全網站、以及高度的保密措施，都有助於保護受訪者的隱私。受訪者的驗證措施也是必須的，因爲這樣才能確保只有被抽樣的受訪者才能參與調查，而且每個受訪者

只參加了一次。

第三個問題涉及到設計的複雜度以及彈性。各種不同的網路軟體與硬體組合之間的相容性，必須事先加以確認。我們現在才剛剛開始瞭解在設計網路調查時最爲有效的方法。例如，最好的作法就是在螢幕上一次只顯示一個或是少數個問題，而且整個問題應該要以一致的格式一次就在螢幕上完整顯示出來，並且以下拉式的選單來提供答案選項。此外，最好能夠顯示出一個進度表（當作鼓勵的動機），像是一個時鐘、或是指出目前進度的符號（也就是目前受訪者已經完成多少部分，而且還有多少部分的問卷還沒完成）。另外一個建議的作法是保持整體頁面看起來簡單清爽（不要使用太多不同的顏色以及字體），並且盡量維持一致性。任何關於電腦的動作（例如，使用下拉式選單）都必須要有非常清楚的說明，而且也都應該包含「點選此處」的指示。除此之外，最好也能夠讓受訪者可以輕易地在不同問題之間來回移動。在網路調查中，如果能夠提供詳細的問題、而且在開放式問題中提供較大的空間讓受訪者來填寫答案，都有助於鼓勵受訪者寫下較長而且較爲完整的答案。另外也相當重要的是，研究者應該使用專屬的伺服器，以避免技術方面的小毛病或是「故障」情況，並且要確保有足夠的頻寬以處理所有對於伺服器的連線請求。[56]

特殊情況

特殊調查的種類有相當多種，其中之一是對於組織（例如：企業、學校等等）的調查。我們可能會去撰寫問題以詢問關於某個組織的資訊，但是同時也是爲了得知組織中的哪一個人擁有我們所需要的資訊。清楚說明調查的意義與重要性，也是不可或缺的，因爲在組織裡的官員通常會接到許多獲取資訊的請求，但是他們不見得會全部都照單全收。

在調查白領菁英時，必須要有特殊的技術才能成功。[57] 在企業裡握有大權的領導人、媒體、以及政府等等，都是難以進行接觸的對象。助理經常會攔截郵寄問卷、或者是限制進行面對面或是電話訪談的機會。取得接觸機會的方法之一，就是請一個備受敬重的人物或是組織打一通電話、或者是寄一封介紹信。在約好碰面時間之後，必須由研究者本人親自進行訪談，而不是由其他的訪談者來進行訪談。一般而言，在進行面對面訪談時，較高比例的開放式問題，會比全部都是封閉式問題時更爲成功。在進行此種調查時，機密性是極爲重要的問題，因爲菁英們通常會擁有一些其他人所不知道的資訊，而且他們對於被指認出是提供特殊資訊的來源這件事是非常敏感的。

時間預算調查 一種特殊類型的調查，在進行此種調查時，受訪者會在一段時間之內去詳細地記錄他們所從事的活動以及持續的時間。

時間預算調查（time budget survey）是一種特殊的類型的調查，主要是用來研究人們如何分配他們的時間。關於都市計畫、勞力性別分工、生活品質、大眾媒體的使用方式、以及休閒時間的使用方式等等研究，都會使用到時間預算調查。時間預算

調查的受訪者同意在數天的時間之內去詳細地記錄自己的活動，而且通常是把這些活動記在日記裡，並且每隔 10 到 15 分鐘的時間間隔就記下自己所從事的活動。例如，在大約十年之前，有數個任教於我的大學裡的教授，被要求去參與一項時間預算調查。這項調查是由一些政府官員所發起的，因爲他們想要得知教授們投入多少時間在學術工作的活動上面。這些參與調查的教授詳細地填寫了日記，並且在兩個星期的時間之內，記錄了自己每隔 15 分鐘在家裡或是在大學裡所從事的活動。結果，這些政府官員們認爲教授們的工作量太少。根據其他數十個此種時間預算調查的結果，如果把所有的會議、社區服務活動、研究工作、準備以及規畫課程、編寫以及批改考卷、評估報告、爲學生提供建議、以及直接教學的時間全部加總起來的話，大多數的教授每週的工作時數大約是 55 到 60 個小時。順便一提的是，大學部的學生傾向於認爲教授們每週只有花費大約 40 個小時在學術工作上面而已。[58]

成本

如果我們把所有的成本（根據所使用的調查類型，相關的成本會有所不同）都考量在內的話，專業品質的調查研究將會是非常昂貴的。有一個簡單的公式是，在郵寄問卷中每花費 1 美元的成本，則在電話訪談調查中就必須花費 5 美元、而在面對面訪談中則必須花費大約 25 美元。但是在網路調查中，若是把前置作業的準備時間排除在外的話，則是幾乎不必花費任何的成本。

成本的變動差異性非常大。[59]除了一般的支出之外，花費最高的就是人力成本，其中包括雇用專業的工作人員（負責製作問卷並且進行前導測試）、雇用文書人員以及訪談者、以及訓練訪談者的費用。剛接觸到調查研究的研究者通常會低估了所需要花費的所有成本以及時間。在 2008 年時，把一份兩頁的郵寄問卷寄給 200 名受訪者需要花費我大約 2,500 美元，也就是平均一個受訪者要花費大約 8.30 美元。這還不包括撰寫以及檢查問卷的支出、或者是以統計方式分析資料的費用。而在回收率爲 60%的情況下（只有回收 180 份問卷），每一份填寫完畢的問卷的真正成本則是接近於 13.90 美元。

專業調查機構完成一次 15 分鐘的電話訪談，通常要價大約 75 美元或是更多。進行一項面對面訪談的研究所必須花費的成本則是更加高昂。視訪談時間長短以及旅行的支出而定，由專業人士來完成一次面對面訪談可能必須花費超過 200 美元以上。以最極端的狀況而言，針對 1,000 名散居各地的一般大眾受訪者所進行的一項面對面調查，其成本可能會高達 30 萬美元以上，並且必須持續長達一年的時間才能完成。而在另一種極端的狀況中，由教師自行影印一份只有一頁的自填問卷、並且發給在一所學校裡的 100 學生來作答，若是不計入教師的時間以及勞力的話，其成本可能會極爲低廉。而且這名教師或許能夠在短短一週的時間之內自行準備並且發放問卷、蒐集答案、並且把結果製作成簡單的分析表格。

調查訪談

在過去數十年當中，我們瞭解訪談錯誤的過程，可以分爲三個演變的階段。在 1960 年代以及 1970 年代時，我們的焦點在於如何防止錯誤，這是因爲受訪者並未完全認真地投入於調查訪談的情境。爲了提升調查訪談的品質，我們會告訴訪談者要去對受訪者強調完整以及正確的答案的重要性、或者是去塑造適當的受訪者行爲。到了 1980 年代至 1990 年代時，提升訪談品質的方式，轉移到了將訪談者的行爲進行標準化。我們會仔細地訓練受訪者，要求他們按照書面的資料正確無誤地念出每一個調查問題、使用中立的探問技巧、逐字記錄受訪者的答案、並且要保持非常客觀的立場與態度。我們所強調的，是要把每一次的訪談情境都變成完全相同的經驗。

標準的訪談是以**天真的假定模型**（naïve assumption model）爲基礎的（請參閱 Foddy, 1993: 13）。我們所努力去追求的，是要去減少存在於進行調查的實際經驗、以及該模型所表達的理想調查情境之間的任何差距（請參閱資訊箱 10.11 調查訪談的天真的假定模型）。

到了 2000 年時，有些研究者主張要揚棄標準化的方式，並且使用一種替代的訪談形式，也就是一種有彈性的、或者是**對話式訪談**（conversational interview），而此種訪談方式則是以合作的邂逅模型爲基礎的（在本章的稍後部分將會進行討論）。訪談被視爲是一種社會情境，在此種情境中，受訪者必須去詮釋調查問題的意義。訪談者必須與受訪者合作、或者是去協助受訪者，如此一來，受訪者才能正確地掌握研究者在某個問題中的意圖或是目的。對於一些關於複雜議題的問題、或者是受訪者比較難以表達其想法的問題，訪談者會主動積極地去改善這些問題的精確性。

大多數專業的調查研究者仍然仰賴標準化的訪談，並且質疑對話式訪談的有效性。他們認爲訪談者效應將會扭曲受訪者的答案、或者是使受訪者的答案帶有偏見。然而，這兩種訪談的方法都有其支持者。標準化訪談方法的支持者們相信，更爲精緻的調查問題措辭方式，能夠解決任何受訪者所可能產生的錯誤詮釋。而對話式訪談的支持者們，則是強調社會互動的流動性本質、以及不同的受訪者所抱持的不同社會現實以及理解。這些支持者認爲，訪談的真正目的在於對問題背後的調查研究者的意圖創造出共同的詮釋，而不是單純地去重覆問題中完全相同的字詞。爲了要在不同的受訪者之間達到共同的詮釋，訪談者或許必須以不同的方式來詢問某些受訪者。因此，或許只有受過高度訓練的、嫻熟社交技巧的、而且深深瞭解研究者對於每個問題的意圖的訪談者，才能在許多不同的受訪者之間，針對研究者的意圖達成共同的理解。對於標準化訪談以及對話式訪談之間的分歧情況，我們可以追溯回到社會科學中的實證主義取向以及詮釋取向之間的不同假

天真的假定模型　一種特定的標準化調查研究類型，在此種模型中沒有任何溝通方面的問題，而且受訪者的答案也能夠完全地符合他們的想法。

對話式訪談　一種以合作的邂逅模型為基礎的、具有彈性的訪談技術，在使用此種技術時，訪談者會依據特定的受訪者的理解情形來調整調查問題，但是仍然會維持研究者在設計每個問題時的意圖或是目的。

資訊箱 10.11　調查訪談的天真的假定模型

1. 研究者已經把所有要被測量的變項都加以清楚地概念化了。
2. 問卷沒有任何措辭效應、問題順序效應、或者是其他相關方面的效應。
3. 受訪者得到鼓勵，並且願意回答所有的問題。
4. 受訪者擁有完整的資訊，並且能夠正確地回憶所有的事件。
5. 受訪者能夠以研究者預期的方式正確無誤地理解問題。
6. 如果受訪者不知道研究假設的話，他們的答案將會更為真實可靠。
7. 如果受訪者沒有得到任何線索或是暗示的話，他們的答案將會更為真實可靠。
8. 訪談的情境以及特定的訪談者對於受訪者的答案不會造成任何影響。
9. 訪談的過程對於受訪者的信念或是態度不會造成任何影響。
10. 在訪談當中，受訪者的行為會完美地與他們的言語回應相契合。

定（你已經在第四章中讀過相關的討論了）。[60]

訪談者的角色

在許多環境以及情境當中，人們都可能會利用訪談來蒐集資訊。例如，雇主會去面試或訪談前來應徵的求職者、醫護人員會訪談病患、心理健康專業人士會訪談當事人、社會服務工作者會去訪談有需要的人們、記者會去訪談政客以及其他人、警察會去訪談目擊者與犯罪的受害者、以及脫口秀主持人會去訪談名人等等（請參閱資訊箱 10.12 非研究訪談的類型）。調查研究的訪談是一種特殊類型的訪談，而且就如同大多數的訪談一樣，此種訪談的目標是要從其他人那兒獲取正確的資訊。[61]

訪談是在兩個陌生人之間的一種短期的、次級的社會互動，而且其明確的目標就是其中一人要從另外一人那兒獲取特定的資訊。在此種互動中的社會角色就是訪談者與受訪者。互動的形式是結構式的對話，而在此種對話中，訪談者會提出一些事先安排好的問題、受訪者針對問題提供自己的答案、然後再由訪談者記錄下這些答案。此種情況與一般的對話有數種不同之處（請參閱表 10.5）。

訪談者經常發現到，受訪者對於調查受訪者的角色感到陌生，而且「受訪者對於他們所被要求的事物通常沒有清楚的概念」（Turner and Martin, 1984: 282）。結果，受訪者可能會利用一個他們所熟悉的角色來取代受訪者的角色（例如，一次親密的對話或者治療的療程、一種完成官方表格的練習、對於政策選擇的公投、一種測驗的情境、或者是一種訪談者試圖誘使受訪者上當的欺騙形式等等）。後續的研究發現到，即使是對於一項設計完善而且專業的調查而言，通常也只有大約一半的受訪者能夠如研究者所預期般地去正確地理解問題。受訪者通常會將問題重新詮釋，以使問題能夠適用於他們自己本身的個人情境、或者是使這些問題變得較爲容易回答。[62]

訪談者的角色很難拿捏。他們侵佔了受訪者的時間以及隱私，尋求受訪者的合作以及建立和諧的關係以獲取一些資訊，但是對受訪者而言，提供這些資訊可能並

資訊箱 10.12 非研究訪談的類型

1. *工作面試* 雇主詢問一些開放式的問題，以蒐集關於某個工作應徵者的資訊，並且觀察該名應徵者如何表現自己。應徵者（受訪者）主動進行接觸，並且試圖呈現出正面的自我形象。雇主（訪談者）則試圖去發現應徵者真正的才能以及缺點。面試的氣氛可能是嚴肅而且帶有判斷意味的，因為雇主握有接受或是拒絕應徵者的權力。此種情況通常會造成緊張的關係以及有限的信任感。雙方的目標可能彼此衝突，而且任何一方都有可能會使用一些欺騙的手段。除此之外，面試的結果是可能會被公開的。
2. *協助式的訪談* 一位輔助的專業人員（諮商者、律師、社工人員、醫師等等）蒐集關於某位當事人所遇到的問題的資訊，包括背景以及現況等等。輔助的專業人員（訪談者）利用這些資訊來瞭解當事人（受訪者）的問題，並將問題轉化為解決問題的專業術語。訪談的氣氛是認真而且關心的。雙方之間通常沒有太多緊張的感受，而且具有高度的互信。雙方的共同目標是要解決當事人的問題，而且幾乎不會用到任何欺騙的手段。訪談的結果通常是保密的。
3. *新聞訪問* 新聞記者會從名人、新聞人物、目擊者、或者是背景人士那兒蒐集資訊，以便於之後用來建構一個有新聞價值的故事。新聞記者（訪談者）會使用各種不同的技巧以試圖從新聞來源（受訪者）那兒獲得新奇的資訊（其中有些資訊可能並不是那麼容易取得）、以及「值得引述的引言」。新聞記者會以選擇性的方式將訪談所得到的資訊與其他的資訊結合在一起，而且這些作法通常都是受訪者無法去控制的。訪談的氣氛、信任的程度、以及彼此之間的緊張關係可能會有非常大的差異。雙方的目標可能相當分歧，而且彼此都有可能會使用一些欺騙的手段。訪談的結果並不是保密的，而且還可能會廣為流傳。
4. *審問或是調查式的訪談* 犯罪司法官員、稽核人員、或是其他官方人員嚴肅地詢問一些問題，以從一名被控告的人或者是其他擁有關於犯罪行為資訊的人們那兒獲得資訊。官員（訪談者）將會利用這些資訊當作證據，以提出對於某人（有可能正是受訪者）的訴訟。由於彼此之間的不信任，因此通常會有高度緊張的關係。雙方的目標可以說是南轅北轍，而且彼此通常都會運用欺騙的手段。訪談結果幾乎都會被公開，而且可能會成為正式公開記錄的一部分。
5. *娛樂式的訪問* 一名節目主持人提出一些評論，並且對名人或其他人詢問了一些開放式的問題，這些受訪者可能會偏離主題或者是開始一段獨白。主要的目標是要引起觀眾或是聽眾的興趣、造成娛樂的效果、或者是讓他們開心。一般而言，在此種訪談中真正重要的通常是雙方所展現出來的風格，而不是任何被揭露出來的資訊。主持人（訪談者）所尋求的是觀眾或聽眾的立即回饋或是反應，而名人們（受訪者）則是試著要增加自己的名氣或是個人聲望。整體氣氛是輕鬆愉快的、沒有太多緊張的感受、而且信任感也相當高。彼此之間的某些目標是一致的。雙方或許會欺騙彼此、或者是一起合作去欺騙觀眾或聽眾。在此種情境中，訪談的結果幾乎是完全公開的。

人們可以混合不同類型的訪談方式，而且人們通常會交錯使用各種不同的類型。例如，一個採取社會控制角色、而非輔助角色的社工人員，就可能會進行調查式的訪談。或者，一名想要幫助犯罪受害者的警察也有可能使用協助式的訪談，而不是使用審問的訪談方式。

表 10.5　一般對話以及結構式調查訪談之間的差異

一般對話	調查訪談
1. 相對而言，來自於每個參與者的問題以及答案是相當平衡的。	1. 在大部分的時間裡，都是由訪談者提出問題，而受訪者則是回答問題。
2. 可以公開地交流感受及意見。	2. 只有受訪者揭露個人的感受以及意見。
3. 可以陳述自己的判斷，並且可以嘗試去說服別人接受某個特定的觀點。	3. 訪談者不作任何判斷，而且也不會試圖去改變受訪者的意見或是信念。
4. 個人可以揭露深層的內心感受以博取他人的同情，或者是當作一種具有治療功能的情緒釋放。	4. 訪談者試圖得到對於特定問題的直接答案。
5. 例行性的回應是相當普遍的（像是「嗯…」、搖頭、「你還好嗎？」「我很好」等等）	5. 訪談者避免作出一些會影響受訪者的例行性回應，並且尋求得到真誠的答案，而不是例行性的回答。
6. 參與者彼此交換資訊，並且更正一些他們發現的事實方面的錯誤。	6. 幾乎所有的資訊都是由受訪者所提供的。訪談者不會去更正受訪者在事實方面的錯誤。
7. 不同的話題會來來去去，而且每一個人都可以帶進新的話題。對話的焦點可以轉移方向、或者是岔開到比較不相關的議題上面。	7. 訪談者掌控了主題、方向、以及步調。訪談者會使受訪者專心於問題上面，而且會排除不相關的其他議題。
8. 情緒的調性可能會從幽默轉為歡樂、感動、哀傷、或者是憤怒等等。	8. 在整個訪談過程中，訪談者試圖去維持一種一致的溫暖調性，但是仍然會維持認真而且客觀的調性。
9. 人們可以迴避或者是忽略某些問題，或者是說出輕率的、不確定的答案。	9. 受訪者不應該迴避問題，而且應該提供真實的、深思熟慮過的答案。

資料來源：改編自 Gorden (1980: 19-25)以及 Sudman and Bradburn (1983: 5-10)

沒有什麼直接的利益。訪談者可能必須解釋調查研究的本質，或者是在訪談當中提示一些關於社會角色的線索。與此同時，訪談者必須維持中立以及客觀。他們努力減少尷尬、恐懼、以及懷疑，如此一來，受訪者才能自在地揭露資訊。好的訪談者會去密切注意社會訪談的步調以及方向，並且會留心答案的內容以及受訪者的行爲。

調查訪談者不作任何判斷，而且不論是以言語或是非言語的方式，都不會揭露出他們的意見。例如，如果受訪者給了一個令人震驚的答案（例如：「我因爲毆打我那個才剛出生不久的小女兒、而且還用香煙去燙她，所以曾經被逮捕三次」），訪談者不能表現出震驚、驚訝、或者是鄙視的表情，而是要以一種不帶任何情感的方式來對待這個答案。訪談者要盡可能地去幫助受訪者，讓他們覺得自己可以放心地提供任何真實的答案。如果受訪者詢問訪談者的意見，訪談者應該要禮貌地引導受訪者，並且告知這樣的問題是不恰當的。例如，如果受訪者問說：「你認爲呢？」，訪談者或許可以這麼回答：「在此我們真正感興趣的是*你*的想法，我的想法如何其實並不重要。」

訪談者協助定義情境，並且確保受訪者擁有必要的資訊、瞭解研究者對於他們的期望、提供相關而且認真的答案、並且具有合作的動機。訪談者必須做的工作比受訪者還要多。面對面的訪談者只有花費大約 35%的時間在進行訪談，有大約 40%的時間是花費於找到正確的受訪者、15%的時間用於旅途上、以及 10%的時間用於研究調查資料以及處理行政工作與記錄細節等。[63]

訪談的階段

從進行簡介開始，訪談的過程必須經過數個階段。在面對面的訪談時，訪談者先進入受訪者的家門、出示授權的證件、消除受訪者的疑慮、並且獲得受訪者的合作。訪談者也會作好準備，以面對像是「你如何挑到我的？」、「這對我有什麼好處？」、「我不瞭解這些」、「這到底是什麼？」之類的反應。訪談者必須解釋爲何一定要訪談某個特定的受訪者，而不是任何一個替代的人。

訪談的主要部分包括了提出問題以及記錄答案。在標準的訪談（非對話式訪談）中，訪談者會一字不差地念出問卷上的內容，不會增加或是刪減任何字詞，不會重新說明或覆述，並且依照順序提出問題，不會回到之前的問題或是跳過某些問題。訪談者會以舒適自在的步調來進行訪談，並且會給予一些不具指示性的回饋來維持受訪者的興趣。

除了提出問題之外，訪談者還必須正確地記錄答案。對於封閉式的問題而言，這是一件相當簡單的工作，因爲訪談者只需要勾選正確的空格即可。但是對於開放式的問題而言，訪談者的工作就會變得比較困難了。訪談者必須仔細傾聽、清楚地寫下答案、並且在不修正文法或是俚語的情況下逐字逐句地記下受訪者所說的話。更加重要的是，訪談者絕對不能只有將受訪者的答案摘要記錄下來、或者是重新釋義，因爲此種作法會造成資訊的流失或者是扭曲受訪者的答案。例如，假設有一位受訪者說：「我真的很擔心我女兒心臟的問題。她也不過才十歲而已，但是上樓梯就已經有困難了。我不知道當她長大之後能夠做什麼。心臟手術對她而言風險太高，而且又太過昂貴。她必須學著和這個問題和平共處下去。」如果訪談者只寫下「擔心女兒的健康情況」的話，就會喪失了許多資訊。

訪談者知道如何以及何時使用**探問**（probe），這是一種中性的請求，目的在於請受訪者澄清一個模糊的答案、完成一個未完成的答案、或者是獲得相關的回答。訪談者辨識出不相關或是不正確的答案，並且在必要的時候使用探問的技巧。[64] 探問有許多種不同的類型。例如，暫停 3 到 5 秒的時間通常會很有效，而非言語的溝通（例如，把頭傾向一邊、挑高眉毛、或者是眼神接觸等等）也很有用。除此之外，訪談者可以重述問題或是答案，然後暫停。訪談者也可以提出一個中性的問題，像是「還有其他的原因嗎？」、「可以再多說一些嗎？」、「你的意思是指什麼呢？」、「你能再多作一些解釋嗎？」之類的問題（請參閱資訊箱 10.13 探問以及記錄封閉式問題完整答案的範例）。

受訪者詮釋一些簡單問題的方式，通常會和調查設計者原先所想的不一樣。例如，「不正確的回報並不是一種回應傾向、或者是一種不誠實的傾向。在某個場合中、或者是對特定問題坦誠以對的人，在其他時候或是面對其他問題時，就不見得會同樣地坦率」（Wentworth, 1993: 130）。

探問　由訪談者所提出的後續問題，目的在於當受訪者的答案不清楚或是不完整時，去引導出一個適當的答案。

像是對話式訪談這種減少誤解的技術，與標準

資訊箱 10.13　探問以及記錄封閉式問題完整答案的範例

訪談者提出的問題：你所從事的工作是什麼？

受訪者的回答：我在通用汽車工作。

探問：你在通用汽車的工作是什麼？你所負責的是什麼類型的工作？

訪談者提出的問題：你失業多久了？

受訪者的回答：已經很久了。

探問：你能不能更加精確地告訴我，你目前的失業狀況是從什麼時候開始的？

訪談者提出的問題：以國家整體來考量的話，你認為明年的景氣會變好、變差、或者是有其他的狀況？

受訪者的回答：可能變好、也可能變差，這要看情況，誰曉得呢？

探問：你覺得有可能發生什麼樣的狀況呢？

記錄封閉式問題的答案

訪談者提出的問題：在一份由 1 分到 7 分的量表上，1 分表示非常贊成死刑，而 7 分表示非常反對死刑，那麼，你對於死刑的看法與評分為何？（贊成）1__2__3__4__5__6__7__（反對）

受訪者的回答：大約在 4 分左右。我認為所有的殺人兇手、強暴犯、以及暴力罪犯都應該被判死刑，但是我不贊成像是偷車之類的較輕罪行也要判死刑。

化的訪談模型是不同的。除了可能會引進偏誤之外，對話式訪談需要花費更多時間，而且對於訪談者的訓練也要更加嚴格。然而，如同康拉德以及修伯爾（Conrad and Schober, 2000: 20）已經觀察到的，「當某些訪談者的行爲（關於問題意義的討論）*比較不*一致時，受訪者的理解可能反而會變得更爲一致——而且他們的答案也比較能夠進行比較。」吊詭的是，只要能夠改善受訪者在詮釋調查問題以及回應時的一致性，非標準化的訪談就能提高其信度。

面對著此種複雜性以及可能的扭曲情況，認真的調查研究者應該要怎麼做才對？至少，我們應該在封閉式問題之外，另外補充開放式問題以及探問。開放式問題需要花費比較多的時間、必須由受過良好訓練的訪談者來進行、而且所產生的答案可能比較不一致並且更加難以量化。以天真的假定模型爲基礎的固定式答案問卷，所意味的是一種較爲簡單的、機械式的作答方式，但是此種方式卻與在許多情境中所發生的狀況不同。當我們深入地去探究訪談者偏誤、文化意義、以及訪談作爲一種社會互動等等問題時，我們就能夠更加瞭解社會研究的質化與量化取向之間如何彼此互補。在所有的研究中，我們都會盡力去排除那些會造成訪談者偏誤以及受訪者混淆的來源。在過去的十年當中，量化調查的研究者發現到，在瞭解人們如何在不同的社會環境中去建構意義時，質化研究者提供了相當寶貴的洞見。

訪談的最後階段是結束階段，也就是訪談者感謝受訪者的合作，然後離開。訪談者通常會找一個安靜的、私密的場所去整理、編輯問卷，並且記錄其他的細節，像是訪談的日期、時間、以及地點等。此外，訪談者通常也會概略描繪受訪者的外

觀以及訪談的情境，包括受訪者的態度（例如：嚴肅認真的、憤怒的、或者是笑容滿面的）、以及任何不尋常的狀況（例如：「在進行到第 27 題問題時，電話響了起來，而受訪者講了大約 4 分鐘的電話，然後才繼續進行訪談」）。訪談者會記下在訪談過程中發生的任何打斷訪談的事件（例如：「受訪者十幾歲的兒子進到房間裡，坐在房間的另一頭，打開電視收看棒球比賽，而且電視的音量開得很大」）。訪談者也會記錄自己的個人感受以及任何感到可疑的事情（例如：「當被問到自己的婚姻狀況時，受訪者變得有點緊張而且坐立不安」）。

訪談者的訓練

在進行大規模的調查時，必須雇用多個訪談者。[65] 若是想要進行具有專業品質的訪談，就必須仔細地挑選訪談者，並且爲他們提供嚴謹的訓練。正如同任何雇用員工的情境一樣，若要維持一致的高品質表現，適當的薪資以及良好的管理都是相當重要的。不幸的是，專業的訪談未必都能提供適當的收入以及常態性的工作。在過去，大多數的訪談者都是願意接受不規律兼職工作的中年女性。良好的訪談者是親切的、誠實的、精確的、成熟的、負責任的、相當機靈的、穩定的、而且積極的。他們擁有不具威脅性的外表、擁有和不同類型的人們打交道的經驗、並且既冷靜又善與人相處。如果調查中牽涉到在高犯罪率地區進行訪談的話，訪談者就需要有額外的保護。

我們可能必須考量到訪談者的身體外觀、年齡、種族、性別、會說的語言、而且甚至是聲音（請參閱本章稍後部分對於訪談者偏誤所進行的討論）。例如，歐克森柏格以及同事（Oksenberg and colleagues, 1986）曾經進行過一項研究，在研究中雇用了來自於同質社會背景的、受過訓練的女性電話訪談者，而研究者們發現，那些音調較高、音調變化較多的訪談者，以及那些聲音較大、說話速度較快、發音清楚、而且聲音聽起來比較愉悅的訪談者，被拒絕的次數明顯比較低。大多數對專業訪談者所進行的訓練課程會持續兩週左右，而這些課程通常包括上課以及閱讀、觀摩專業的訪談者、在辦公室以及田野中進行模擬訪談並加以記錄與接受評判、許多的訪談練習、以及角色扮演。訪談者在這些課程當中學習關於調查研究的內容，以及訪談者所扮演的角色。除此之外，他們也會去熟悉問卷以及每個問題的目的，但是他們並不知道預期的答案爲何。

雖然訪談者主要是獨立進行工作，但是在擁有多位訪談者的大規模調查當中，研究者通常會再雇用一名管理訪談者的督導員。督導員熟悉訪談的地理位置、協助處理問題、監督訪談者、並且確保訪談者能夠準時完成工作。而在電話訪談中，督導員則是協助打電話、確認訪談者抵達以及離開的時間、並且監督訪談的電話內容。在面對面的訪談中，督導員會去確認訪談者是不是真的去進行了訪談，而這意味著督導員可能會打電話給受訪者，或者是寄一封確認的明信片給被抽樣的受訪者。督導員也可以確認回收率以及未完成的問卷，以瞭解訪談者是否取得受訪者的合作，此外，督導員或許會重新訪談一個小型的次樣本並且分析答案、或者是觀察訪談過程，以瞭解訪談者是否正確地提出問題並且記錄答案。

訪談者偏誤

調查研究者會規定訪談者的行爲以減少偏誤。理想上，特定訪談者的行動應該不會影響到受訪者作答的方式，而且受訪者的回答也應該不會因爲不同的訪談者而有所改變。對於訪談者所規定的行爲，不僅僅只有給予指示並且要求訪談者正確無誤地讀出每個問題而已，而且訪談偏誤也可能有多種不同的形式（請參閱資訊箱 10.14 訪談偏誤的六種類別）。

我們仍然還在學習那些會影響調查訪談的因素。我們知道訪談者期望可能會造成明顯的偏誤。那些預期訪談會難以進行的訪談者可能會造成此種偏誤，而那些預期得到某些特定答案的訪談者，則更有可能會造成此種偏誤（請參閱圖表 10.1）。要達到適當的訪談者行爲以及正確無誤地念出問題，或許會相當困難，但是仍然可能還有其他形式的訪談偏誤。

訪談者偏誤也有可能是來自於對受訪者的年齡以及種族所產生的期望。在一項大型的美國國家調查中，研究者得知，訪談者經常認爲黑人受訪者是比較不聰明的，並且認爲年輕的受訪者是比較不聰明的，而且也比較缺乏相關的資訊。因此，爲了要減少此種在調查結果中的偏誤，就必須要對訪談者進行更爲嚴謹、而且更好的訓練。[66]

訪談的情境也可能會影響答案。例如，依據是在家裡或是在學校進行訪談的不同，高中學生的答案可能就會有所差異。有其他人在場通常也會影響答案，所以我們一般不希望有其他人出現在現場。[67] 舉例而言，齊普以及托斯（Zipp and Toth, 2002）發現，當配偶也在訪談現場中的時候，夫妻雙方在許多關於態度的問題上面會有比較大的一致性，也就是說，太太會修正她們的答案以配合先生的答案，而先生的答案則幾乎沒有太大的改變。

訪談者外顯的身體特質，包括種族以及性別等等，也都可能會影響受訪者的答案，尤其是在關於種族或是性別的問題方面。例如，非裔美國人以及西班牙裔的美

資訊箱 10.14　訪談偏誤的六種類別

1. *受訪者的錯誤*　受訪者忘記某些資訊、感到難為情、誤解問題、或是因為有其他人在場而說謊。
2. *不經意的錯誤或是訪談者過於草率*　訪問了錯誤的受訪者、念錯問題、略過問題、以錯誤的順序念出問題、把問題的答案記錄錯誤、或是誤解了受訪者的意思。
3. *訪談者蓄意破壞*　訪談者故意更改答案、略過或改寫問題、或是選擇了替代的受訪者。
4. *由於訪談者的期望而產生的影響*　因為受訪者的外表、生活情況、或是其他的答案，而使得訪談者預期受訪者會給予某些特定的答案。
5. *訪談者未能進一步探問*　訪談者沒有進一步探問，或以適當的方式進行探問。
6. *訪談者對於受訪者所造成的影響*　受訪者因為訪談者的外表、音調、對於答案的反應、或是作出在訪談程序之外的評論等等因素而受到影響。

圖表 10.1 訪談者的特質可能會影響答案

訪談者期望效應的範例

女訪談員本身的狀況	*回答由先生買大部分傢俱的女性受訪者的比例*
先生買大部分的傢俱	89%
先生沒有買大部分的傢俱	15%

種族外表效應的範例

	回答「是」的比例	
訪談者	*你認為政府機構的工作中有太多猶太人嗎？*	*你認為猶太人的勢力太過龐大嗎？*
外表像猶太人而且具有像猶太人的姓名	11.7%	5.8%
只有外表像猶太人	15.4%	15.6%
外表看起來不像猶太人	21.2%	24.3%
外表不像猶太人而且姓名也不像猶太人	19.5%	21.4%

*注意：*受訪者所具有的種族刻板印象可能會影響他們在訪談中作答的方式。

*資料來源：*改編自 *Interviewing in social research* by Herbert H. Hyman with William J. Cobb et al.; foreword by Samuel A. Stouffer. ©1954, 1975 University of Chicago Press, p. 153.

國人，會因爲訪談者在外表上明顯的種族差異，而對那些與種族相關的議題表達出不同的政策立場。甚至就連在電話訪談當中，如果受訪者察覺到關於訪談者種族身分的一些線索時，也會發生此種狀況。一般而言，具有和受訪者相同種族背景的訪談者，會比來自於不同背景的訪談者更有可能得到較爲正確的答案。在一些像是性行爲之類的敏感議題、以及對於與性別相關的集體行動或是性別平等的議題方面，性別也會影響訪談的結果。然而，如同魏斯柏格（Weisberg, 2006: 61）所提到的：「訪談者配對在美國是很罕見的作法，除非是當研究者確實有必要去雇用一個能夠說另一種語言的訪談者的時候……在某些其他的國家中，像是阿拉伯國家，訪談配對是比較必要的作法，因爲在這些國家中，以某個性別的訪談者去和另外一個性別的受訪者談話，會被認爲是不恰當的。」[68]

訪談者的特質可能會以許多方式來影響答案。例如，當訪談者是一位身心障礙者時，受訪者在自我回報的「快樂」程度方面，會比他們在單獨完成一份自填問卷時還要更低。顯而易見的，這是因爲受訪者不想要讓人覺得和訪談者相較之下，他們聽起來太過於幸福。然而，如果有一名身心障礙者在同一個房間裡，當受訪者在完成一份自填問卷時，他們所回報的快樂程度則會比較高。很顯然的，相較於一個沒有任何直接提醒受訪者其他人的生活狀況的情境，當一名身心障礙者在場時，受訪者會覺得相較之下自己的處境好多了。[69] 若是受訪者會因爲訪談者的特質而對同

樣的問題提出不同的答案時，此種情況就會威脅到研究的代表性信度（representative reliability）（你已經在第七章中讀過關於代表性信度的內容了）。

合作的邂逅模型　一種特殊類型的調查訪談，在此種訪談中，受訪者以及訪談者一同合作，以達到研究者所預期的調查問題的意義，並且產生對於問題的正確回答。

文化意義以及調查訪談

對於調查誤差以及訪談偏誤所進行的研究，已經讓我們能夠更加瞭解人們如何創造意義以及達到文化理解的方式。[70] 當同樣一個詞彙的意義以及暗示，會因爲社會情境、說話的人、說話的方式、以及說話者以及傾聽者之間的社會距離而產生差異時，我們就會陷入困境之中。當受訪者以錯誤的方式去詮釋調查研究的本質，並且在問題的措辭或者是訪談者的微妙動作當中去尋找如何回答的線索時，調查研究就會變得很複雜。除此之外，「重要的是不要忽略了這個事實：訪談的情境本身、與其他可以表達出態度的情境是完全不同的。因此，我們不應該預期受訪者在訪談中所表達出來的態度，會與他們在其他社會脈絡中所表達出來的態度完全一致。」[71]

我們面臨了一個進退兩難的處境：一位努力以中立以及一致的方式來進行訪談的訪談者，將能夠減少那些由於個別訪談者行爲所造成的、會威脅到信度的偏誤類型。然而，根據詮釋的或者是批判的社會科學研究者（其中也包括了女性主義研究者）的看法，此種作法卻可能反而會造成其他的問題（請參閱資訊箱 10.15 訪談：實證主義以及女性主義取向）。[72]

非實證主義的研究者主張，意義是在社會脈絡中被創造出來的，因此，標準的調查問題措辭方式，將無法對所有受訪者都產生完全相同的意義。例如，有些受訪者在表達感受的時候，是利用說故事的方式，而不是去回答一些有固定答案的簡單問題。非實證主義的研究者提倡使用調查情境的**合作的邂逅模型**（collaborative encounter model）。此種模型把所有人與人之間的邂逅都視爲是高度動態的、複雜的共同互動，在此種互動中，即使是微小的、不經意的回饋（例如，說「嗯……」、大笑、微笑、點頭等等）也都會造成影響，並且因此建議使用對話式訪談。當受訪者對於那些固定選項的問題提出一些其他的資訊時，合作的邂逅模型也允許訪談者去結合此種資訊。然而，標準化的訪談是禁止此種資訊的，並且還會將此種資訊視爲是錯誤，因爲此種資訊無法符合於一個預先設定的、標準化的格式。

根據合作的邂逅模型，在複雜的人類互動中，人們會對簡單的問題賦予詮釋性的意義。例如，我的鄰居問我一個簡單的問題：「你多久整理一次草坪？」我可能會以下列數種方式來詮釋他這個問題：

> 我自己多久整理一次草坪（相對於請其他人幫我整理草坪）？
> 我多久割草一次（相對於只是把割草機推過草坪上去切碎落葉而已）？
> 我多久把整片草坪通通整理一次（相對於只有去整理那些草長得比較快的區域）？
> 我在一整個季節裡、一個月裡、或是一個星期裡整理草坪幾次？
> 我在大部分的季節裡多久整理草坪一次（相對於去年當我的割草機壞了好幾次，而且氣候非常乾燥、草也長得比較少，因此我比較不常整理草坪）？

資訊箱 10.15　訪談：實證主義以及女性主義取向

在本章中，我們大部分所討論的都是調查研究訪談的實證主義取向。在理想的調查訪談當中，訪談者會隱瞞自己本身的感受以及信念。訪談者應該要力求客觀以及中立，如此一來，就算是以另外一名訪談者來代替，也有可能會得到完全相同的答案。

女性主義研究者對於訪談的看法則非常不同。女性主義訪談類似於質化訪談（我們將在第十四章中進行討論）。歐克蕾（Oakley, 1981）批評實證主義調查訪談是男性典範的一部分，因為此種訪談是一種社會情境，在此種情境中，訪談者一方面要壓抑自己、不能表達出個人的感受，另一方面則是要去控制以及支配受訪者。因此，訪談是操縱的以及工具性的。訪談者以及受訪者變成純粹只是獲得客觀資料的工具而已。

女性主義研究的目標雖然各自不同，但是仍有兩個共同的目標，也就是讓女性的主觀經驗更加為人所知、以及增加受訪者對於研究過程的參與程度。女性主義訪談的特徵包括下列幾項：

- 偏好無結構的以及開放的形式。
- 偏好對同一位受訪者進行多次訪談。
- 創造社會連結以及建立信任的社會關係。
- 訪談者本身也可以揭露個人的經驗。
- 藉由低調處理專業地位的方式，避免去控制受訪者，並且鼓勵彼此之間的平等性。
- 仔細傾聽，而且訪談者在情緒上能夠與受訪者感同身受。
- 受訪者導向，而不是研究者導向或是問卷導向。
- 鼓勵受訪者以他們覺得最自在的方式來表達自己——例如以說故事或是談論一些題外話的方式。
- 創造一種彰權賦能的感受（sense of empowerment），以及女性之間的團體意識。

在幾秒鐘之內，我作出了詮釋並且說出答案，但是由於在我自己和鄰居之間所進行的是開放式的、持續進行的互動，因此我可以去要求他澄清問題，並且提出幾個後續的問題，以幫助我們兩人達到相互的理解。

調查訪談的互動與一般的對話是不同的。標準的調查研究訪談是一種人爲的互動，此種互動以相同的方式來對待各種不同的受訪者，以控制溝通的情境並且產生一個一致的量數。一般的互動則是包含了偵測以及修正錯誤詮釋的內在特徵，換句話說，一般的互動所依賴的是一些細微的差異以及彼此之間的互相遷就。在一般的對話當中，人們達到社會意義的方式，是利用脈絡中的線索、調整對於特定人們的互動方式、以及依賴一個文化架構（通常是以種族、階級、性別、地區、或者是宗教爲基礎）。一般對話的流動性互動是可以進行自我調整的，因爲不同的人們不必然會對相同的詞彙、俚語、以及問題賦予相同的意義。例如，男性以及女性會以不同的方式來回報自己的健康狀況。當一個男性說他的健康狀況非常好的時候，他所指的意義可能與對同一個問題提出相同答案的女性完全不同。由於把人類的互動加以

標準化了，因此標準的調查訪談就排除了一般對話中的一些特徵，而這些特徵可以提供自我修正、在不同的人們之間促進建構一個共享的意義、並且增加人們之間的相互理解。[73]

前導測試以及認知訪談

在正式實施之前，先對調查訪談以及問卷進行前導測試是相當重要的。在調查過程當中，對前導測試以及認知處理模型所進行的系統性研究，已經幫助我們能夠更加清楚地瞭解調查過程。我們已經瞭解到，回答調查問題的過程可以分爲下列數個步驟：詮釋以及理解問題、回憶起相關的資訊、整合以及評估資訊、以及選擇一個回應類別。有一個比較新的研究領域，該領域被稱作認知測驗或是**認知訪談**（cognitive interviewing），而在這個新的領域中，我們所研究的主題，正是受訪者在前導測試的情境中去回答問題的方式。[74]

在進行認知訪談時，我們會要求一小群的前測參與者以口頭的方式回報他們在回答問題時的想法，以幫助我們發現一份尚在發展中的問卷裡的潛在問題。認知訪談提供了一扇窗，讓我們得以一窺受訪者在回答調查問題時的想法以及所面臨的問題。認知訪談者會進一步探問關於作答過程的額外資訊，而我們則能夠使用此種資訊來改善問卷或者是訪談的過程（請參閱資訊箱 10.16 以前導測試來改善問卷的方法）。

另外一個相關的發展趨勢，則是利用俗民方法論（請參閱第十三章的相關討論內容）以及對話分析，將訪談過程當作是一種特殊類別的社會互動以及言語事件來加以研究。這些方法支持合作的邂逅模型，並且也建議把非標準化的訪談行爲，像

認知訪談　一種運用於前導測試中的技術，在使用此種技術時，研究者會先訪談受訪者並且詢問關於他們的思考過程的問題、或者是要求受訪者在回答調查問題時「把心中的想法說出來」（“think out loud”），以此種方式來得知關於一份問卷的潛在問題並且加以改善。

資訊箱 10.16　以前導測試來改善問卷的方法

1. *說出對於訪談的想法*　在回答每道問題的過程中，要求受訪者解釋自己的想法。
2. *回顧式訪談以及目標式探問*　在完成一份問卷之後，受訪者向研究者解釋自己在選擇每個答案時所使用的過程或方法。
3. *專家評估*　由一群經驗豐富的調查研究者組成一個獨立的專門小組來檢閱以及批評問卷。
4. *行為編碼*　研究者密切地監督訪談過程，通常是利用聽錄音帶或者是觀看錄影帶的方式，以找出念錯題目、猶豫、被漏掉的指示、無回應、拒絕、困惑的表情、以及不符合任何作答選項的答案等等情形。
5. *田野實驗*　研究者在田野環境中進行問卷項目的替代形式，並且比較其結果。
6. *短片以及簡報*　研究者先讓訪談者以及受訪者觀看簡短的、人為的「與真實狀況相仿的」情境，並且詢問他們將會在問卷上使用何種答案的類型。

*資料來源：*Dillman and Redine (2004), Fowler (2004), Martin (2004), Tourangeau (2004a, 2004b), van der Zouwen and Smit (2004), and Willis (2004).

是受訪者的疑問、以及比較輕微的訪談者回饋（說「嗯……」、大笑、微笑）等等，當作是更進一步瞭解訪談過程的機會。[75]

合乎倫理的調查

就如同所有的社會研究一般，我們可以採取合乎倫理或是不合乎倫理的方式來進行調查。在調查研究中的一個重大的倫理議題，就是對於隱私的侵犯。[76] 所有的人都擁有隱私權，而且受訪者有權去決定何時以及對誰揭露自己的個人資訊。當我們詢問了關於親密行爲以及個人信念的問題時，就可能會侵犯了受訪者的隱私。在受訪者被問到此類的問題時，若是處於一個互相尊敬以及信任的自在脈絡當中，就有可能會正確而且誠實地提供此種私人的資訊。當受訪者相信我們是爲了正當的研究目的而想要得到認真的答案時、以及當受訪者相信他們的回答會維持在保密狀態時，就最有可能會回答所有的問題。我們必須以尊重的態度來對待所有的受訪者、減低不舒服的感受、並且保護調查資料的機密性。

第二個與倫理有關的議題，則是牽涉到受訪者的自願參與。不管在任何時間，受訪者都可以同意回答問題或者是拒絕參與調查。在參與一項研究時，受訪者必須給予「知情後的同意」。我們依靠受訪者的自願合作，而且也必須以一種謹慎的方式來提出設計完善的問題、以尊重的方式來對待受訪者、並且謹守保密的原則。

第三個與倫理有關的議題，則是對於調查的濫用以及僞調查。由於調查相當普遍而且受到歡迎，有一些組織以及人們因而利用調查來誤導他人。所謂的**僞調查**（pseudosurvey），指的是一種特定的調查形式，其真正的目的在於說服某人去做某些事情，而不是要去得知關於受訪者的任何資訊。一些招搖撞騙的人利用進行調查的名義進行僞裝，以侵犯受訪者的隱私、進入他人的家中、或者是利用調查的名義來推銷東西。僞調查的例子之一，是發生在 1994 年的美國選舉活動中的一種「抑制性民調」（“suppression polls”），在此種民調中，一個不知名的調查組織打電話給潛在的選民，並且詢問該位選民是否支持某個特定的候選人。如果這位選民支持該位候選人的話，訪談者會接著問，如果他或她知道這名候選人曾經有過一些不良紀錄（例如，曾經因爲酒醉駕車而被逮捕、施打過非法藥物、提高了在監獄中被判刑的罪犯的工資等等）的話，是否還會支持這名候選人。這個訪談的目的不在於調查候選人的支持度，相反地，其目的在於找出某位候選人的支持者，然後試圖去抑制其投票行爲。我曾經接過此種電話，另外有一名競選州長失利、而且是抑制性民調攻擊對象的候選人也曾經接過此種電話。然而，到目前爲止並未有任何人因爲運用了此種選戰手法而遭到起訴。

僞調查　一種僞造的、欺騙的、像是調查的工具，雖然此種工具運用了調查訪談的形式，但是其真正的目的只不過是在於說服受訪者而已。

另外一個與倫理有關的議題，則是不當使用調查結果，或者是使用了設計不良的、或者是被蓄意操縱的調查。人們或許會要求得到一些調查結果無

資訊箱 10.17　報導調查研究時應該要納入的十個項目

1. 所使用的抽樣架構（例如：電話簿）。
2. 進行調查的日期。
3. 樣本所代表的母體（例如：美國成人、澳洲的大學生等）。
4. 蒐集資訊的樣本大小。
5. 抽樣方法（例如：隨機抽樣）。
6. 所詢問的問題的確切措辭。
7. 進行調查的方法（例如：面對面訪談、電話訪談等）。
8. 贊助以及主持該項調查的組織（也就是提供經費以及進行調查的組織）。
9. 回應率以及所接觸的受訪者中真正完成整份問卷的百分比。
10. 當特定問題的結果被報導出來時，列出任何遺漏的資訊或者是「不知道」的答案。

法提供的答案，或者，人們也可能無法理解調查資料的限制。除此之外，設計以及準備調查的人，也有可能缺乏關於進行合理調查的充分訓練。以草率的或是設計不良的調查爲基礎所作出的政策決定，可能會導致浪費以及造成人們的困境。由於有這些不當使用調查的狀況，因此，瞭解調查研究的複雜性、以及只進行在方法論方面設計完善的調查研究，就是很重要的事。

另外一個議題，則是大眾媒體對於調查結果的報導，也可能造成濫用的情況。[77] 即使人們在閱讀調查結果時可能幾乎不會去注意到調查的細節，我們仍然應該要總是都把關於調查的細節包含在調查結果之內（請參閱資訊箱 10.17 報導調查研究時應該要納入的十個項目），以減少調查研究被誤用的機會，並且提高人們的意識，以對缺乏此類資訊的的調查提出質疑。在大眾媒體關於調查的報導中，有高達 88% 的報導都沒有揭露進行調查的研究者的身分，而且只有 18%的報導提供了關於調查進行方式的細節。[78] 我們強烈要求媒體應該要納入此種資訊，尤其是因爲媒體報導調查研究的次數，遠高於報導其他類型社會研究的次數。

目前仍然沒有任何品質控制的標準，可以用來規範在美國媒體上關於民意請查或者是其他調查的報導。有將近五十年之久的時間，專業的調查研究社群一直在努力試圖要求媒體只報導那些具備適當的科學樣本、嚴謹的訪談者訓練以及監督、令人滿意的問卷設計、資料向大眾公開、以及調查機構具有公正性的調查研究結果，但是卻徒勞無功。[79] 不幸的是，媒體在進行報導時，似乎總是不分青紅皂白地把有偏差的、誤導的調查結果，以及來自於嚴謹的、專業的調查結果通通混爲一談而加以報導。因此，也就難怪大眾會產生混淆，而且甚至對所有的調查都產生不信任的感覺了。

結論

在本章中，你學到了關於調查研究的知識。調查是最被廣泛使用的社會研究技術。你也學到了關於撰寫良好調查問題的一些原則，當你在撰寫調查問題時，有許

多要避免以及要包含在內的事物。本章也提出了各種不同類型的調查研究的優點以及缺點，而且也提到了進行訪談，尤其是面對面的訪談，可能會是相當困難的事。

雖然本章的焦點在於調查研究，但是我們在其他類型的量化研究中（例如：實驗研究）也會使用問卷來測量變項。調查是一種獨特的技術，而且因爲在調查中通常會使用隨機抽樣方法，因此也經常被稱爲*樣本調查*（*sample survey*）。調查是一種過程，也就是去詢問許多人相同的問題，然後去檢視他們所回答的答案。

雖然調查研究者試圖將錯誤減到最低的程度，但是在調查資料中仍然還是可能會有一些錯誤。在調查中的不同錯誤可能會使得彼此所造成的影響更爲嚴重。例如，錯誤可能來自於抽樣架構、受訪者無回應的情況、問題的措辭或是順序、以及訪談者偏誤等等。不過，不要因爲這些可能發生的錯誤而使你對於運用調查感到卻步。相反地，你應該學習以非常謹慎小心的態度來設計調查研究，而且在把調查結果加以類推時，也應該要抱持著慎重的態度。

問題回顧

1. 調查中經常問到的六種類型的資訊是什麼？請針對每一種類型舉出一個和書中範例不同的例子。
2. 爲什麼調查會被稱爲是*關聯性*的研究？調查與實驗之間有什麼差異之處？
3. 在撰寫問題時有十個應該要避免的事項，請列出其中五項。
4. 什麼樣的主題通常會讓受訪者覺得具有威脅性？研究者應該如何詢問這些問題？
5. 開放式以及封閉式問題的優點與缺點爲何？
6. 什麼是全過濾式、半過濾式、以及標準形式問題？這些問題與游移的受訪者之間有何關聯？
7. 標準訪談以及對話式訪談之間有何差異？彼此的相對優點與價值爲何？
8. 什麼是認知訪談？認知訪談如何改善調查研究的品質？
9. 在何種情況之下，使用郵寄問卷、電話訪談、網路調查、以及面對面訪談會是最佳的作法？
10. 什麼是 CATI 和 IVR？這兩種技術何時會有幫助？這兩種技術和 CASAI 或者是 CAPI 之間有何差異？

注解

1. See Carr-Hill (1984b), Denzin (1989), Mishler (1986), and Phillips (1971) for criticisms of a strict positivist approach in surveys.
2. "Why" questions require special techniques. See Barton (1995) and Wilson and colleagues (1996).
3. The history of survey research is discussed in Converse (1987), Hyman (1991), Marsh (1982:9–47), Miller (1983:19–125), Moser and Kalton (1972:6–15), Rossi and colleagues (1983), Sudman (1976b), and Sudman and Bradburn (1987).

4. See Bannister (1987), Blumer (1991a, 1991b), Blumer et al. (1991), Camic and Xie (1994), Cohen (1991), Deegan (1988), Ross (1991), Sklar (1991), Turner (1991), and Yeo (1991). Also see R. Smith (1996) on how political ideological conflicts and private foundations affected the development of survey research.

5. See Scheuch (1990) on national surveys conducted in various countries.

6. See Converse (1987:383–385), Statistical Abstract of the United States, and Rossi et al. (1983:8).

7. See Rossi et al. (1983:10).

8. See Bayless (1981) on the Research Triangle Institute.

9. Some organizations include the American Association for Public Opinion Research, founded in 1947. The Council of American Survey Research Organization is an organization for U.S. commercial polling firms and the World Association of Public Opinion Research is an international organization for commercial polling. See Bradburn and Sudman (1988).

10. Bishop et al. (1983, 1984, 1985), Bradburn (1983), Bradburn and Sudman (1980), Cannell et al. (1981), Converse and Presser (1986), Groves and Kahn (1979), Groves et al. (2000), Groves and Couper (1998), Hyman (1991), Lacy (2001), Lyberg et al. (1997), Schacter (2001), Schuman and Presser (1981), Schwarz and Sudman (1992, 1994), Sniderman and Grob (1996), Sudman and Bradburn (1983), Sudman et al. (1996), and Tanur (1992).

11. For a discussion of pilot testing techniques, see Bishop (1992), Bolton and Bronkhorst (1996), Fowler and Cannell (1996), and Sudman et al. (1996).

12. On the administration of survey research, see Backstrom and Hursh-Cesar (1981:38–45), Dillman (1978:200–281; 1983), Frey (1983:129–169), Groves and Kahn (1979:40–78, 186–212), Prewitt (1983), Tanur (1983), and Warwick and Lininger (1975:20–45, 220–264).

13. Similar lists of prohibitions can be found in Babbie (1990:127–132), Backstrom and Hursh-Cesar (1981:140–153), Bailey (1987:110–115), Bradburn and Sudman (1988:145–153), Converse and Presser (1986:13–31), deVaus (1986:71–74), Dillman (1978:95–117), Fowler (1984:75–86), Frey (1983:116–127), Moser and Kalton (1972:318–341), Sheatsley (1983:216–217), Sudman and Bradburn (1983:132–136), and Warwick and Lininger (1975:140–148).

14. Binson and Catania (1998), Foddy (1993), and Presser (1990).

15. Sudman and Bradburn (1983:39) suggest that even simple questions (e.g., "What brand of soft drink do you usually buy?") can cause problems. Respondents who are highly loyal to one brand answer the question easily.

16. See Schaeffer (2000) and Sudman et al. (1996:197–226).

17. See Dykema and Schaeffer (2000).

18. On using a continuum, see Ostrom and Gannon (1996).

19. See Abelson and associates (1992), Auriat (1993), Bernard et al. (1984), Croyle and Loftus (1992), Gaskell et al. (2000), Krosnick and Abelson (1992), Loftus et al. (1990), Loftus et al. (1992), Pearson and Dawes (1992), Sudman et al. (1996), and Weisberg (2005:76–81, 127).

20. See Bradburn (1983), Bradburn and Sudman (1980), and Sudman and Bradburn (1983) on threatening or sensitive questions. Backstrom and Hursh-Cesar (1981:219) and Warwick and Lininger (1975:150–151) provide useful suggestions as well. Fox and Tracy (1986) discuss the randomized response technique. Also see DeLamater and MacCorquodale (1975) on measuring sexual behavior and Herzberger (1993) on sensitive topics.

21. For studies on survey format and answer honesty, see Holbrook et al. (2004), Johnson et al. (1989), Schaeffer and Presser (2003:75), and Tourangeau et al. (2002).

22. See Couper et al. (2003), DeMaio (1984), and Sudman and Bradburn (1983:59).

23. For more on surveys with threatening or sensitive topics and computer-assisted techniques, see Aquilino and Losciuto (1990), Couper and Rowe (1996), Johnson et al. (1989), Tourangeau and Smith (1996), and Wright et al. (1998).

24. For a discussion of knowledge questions, see Backstrom and Hursh-Cesar (1981:124–126), Converse and Presser (1986:24–31), Sudman and Bradburn (1983:88–118), and Warwick and Lininger (1975:158–160).

25. On how "Who knows who lives here?" can be complicated, see Martin (1999) and Tourangeau et al. (1997).

26. Contingency questions are discussed in Babbie (1990:136–138), Bailey (1987:135–137), deVaus

(1986:78–80), Dillman (1978:144–146), and Sudman and Bradburn (1983:250–251).

27.For further discussion of open and closed questions, see Bailey (1987:117–122), Converse (1984), Converse and Presser (1986:33–34), deVaus (1986:74–75), Geer (1988), Moser and Kalton (1972:341–345), Schuman and Presser (1979; 1981:79–111), Sudman and Bradburn (1983:149–155), and Warwick and Lininger (1975:132–140).

28.See Gilljam and Grandberg (1993). Moors (2008) notes that generally five versus six choices are equally effective in statistical tests but six is sometimes better, and the "optimal" solution depends on the content of the survey items.

29.For a discussion of the "don't know," "no opinion," and middle positions in response categories, see Backstrom and Hursh-Cesar (1981:148–149), Bishop (1987), Bradburn and Sudman (1988:154), Brody (1986), Converse and Presser (1986:35–37), Duncan and Stenbeck (1988), Poe et al. (1988), Sudman and Bradburn (1983:140–141), and Schuman and Presser (1981:113–178). For more on filtered questions, see Bishop et al. (1983, 1984), Bishop et al. (1986), and Weisberg (2005:134-136).

30.See Krosnick et al. (2002), Schaefer and Presser (2003:79–80), and Tourganeau (2004:786).

31.The disagree/agree versus specific alternatives debate is discussed in Bradburn and Sudman (1988:149–151), Converse and Presser (1986:38–39), Schuman and Presser (1981:179–223), and Sudman and Bradburn (1983: 119–140). Backstrom and Hursh-Cesar (1981:136–140) discuss asking Likert, agree/disagree questions.

32.See McCarty and Shrum (2000) and Narayan and Krosnick (1996).

33.The ranking versus ratings issue is discussed in Alwin and Krosnick (1985), Krosnick and Alwin (1988), and Presser (1984). Also see Backstrom and Hursh-Cesar (1981:132–134) and Sudman and Bradburn (1983:156–165) for formats of asking rating and ranking questions.

34.For more on specific design issues, see Christian and Dillman (2004), Dillman and Redline (2004), Kaplowitz et al. (2004), Ostrom and Gannon (1996), Schwarz et al. (1991), and Tourangeau et al. (2004).

35.See Dillman (2000:32–39) and Dillman and Christian (2005) for discussion.

36.For a discussion of wording effects in questionnaires, see Bradburn and Miles (1979), Peterson (1984), Schuman and Presser (1981:275–296), Sheatsley (1983), and Smith (1987). Hippler and Schwarz (1986) found the same difference between forbid and not allow in the Federal Republic of Germany.

37.The length of questionnaires is discussed in Dillman (1978:51–57; 1983), Frey (1983:48–49), Herzog and Bachman (1981), and Sudman and Bradburn (1983:226–227).

38.For a discussion of the sequence of questions or question order effects, see Backstrom and Hursh-Cesar (1981:154–176), Bishop et al. (1985), Bradburn (1983:302–304), Bradburn and Sudman (1988:153–154), Converse and Presser (1986:39–40), Dillman (1978:218–220), McFarland (1981), McKee and O'Brien (1988), Moser and Kalton (1972:346–347), Schuman and Ludwig (1983), Schuman and Presser (1981:23–74), Schwartz and Hippler (1995), and Sudman and Bradburn (1983:207–226). Also see Knäuper (1999), Krosnick (1992), Lacy (2001), and Smith (1992) on the issue of question-order effects.

39.A study by Krosnick (1992) and a meta-analysis by Narayan and Krosnick (1996) show that education reduces response-order (primacy or recency) effects, but Knäuper (1999) found that age is strongly associated with response-order effects.

40.This example comes from Strack (1992).

41.For additional discussion of context effects, see Schuman (1992), Smith (1992), Todorov (2000a, 2000b), and Tourangeau (1992).

42.Tarnai and Dillman (1992) discuss how the method of survey affects context effects.

43.For a discussion of format and layout, see Babbie (1990), Backstrom and Hursh-Cesar (1981:187–236), Dillman (1978, 1983), Mayer and Piper (1982), Sudman and Bradburn (1983:229–260), Survey Research Center (1976), and Warwick and Lininger (1975:151–157).

44.For a discussion, see Couper et al. (1998), de Heer (1999), Keeter et al. (2000), Sudman and Bradburn (1983:11), and "Surveys Proliferate, but Answers Dwindle," New York Times (October 5, 1990), p. 1. Smith (1995) and Sudman (1976b:114–116) also discuss refusal rates.

45.For additional discussion of nonresponse and refusal rates, see Backstrom and Hursh-Cesar (1981:140–141, 274–275), DeMaio (1980), Frey (1983:38–41), Groves and Couper (1998), Groves and Kahn (1979:218–223), Martin (1985:701–706), Nederhof (1986), Oksenberg et al. (1986), Schuman and Presser (1981:331–336), Sigelman (1982), Stech (1981), Sudman and Bradburn (1983), and Yu and Cooper (1983). For a discussion of

methods for calculating response rates, see Bailey (1987:169), Dillman (1978:49–51), Fowler (1984:46–52), and Frey (1983:38).

46.Link and Oldendick (1999) examined telephone screening.

47.See Pottick and Lerman (1991) for a discussion of the study.

48.Introductions and incentives are discussed in Brehm (1994), Couper (1997), De Leeuw et al. (2007), Goldstein and Jennings (2002), Singer (1999), Singer et al. (1998), Singer et al. (1999), Singer et al. (2000), and Trussell and Lavrakas (2004). Dillman et al. (1996) discuss mandatory appeals.

49.Tailoring is discussed in Brehm (1994), Groves and Couper (1996, 1998, 2004), and Groves, Presser, and Dipko (2004).

50.On increasing mail questionnaire return rates, see Bailey (1987:153–168), Church (1993), Dillman (1978, 1983), Fox et al. (1988), Goyder (1982), Heberlein and Baumgartner (1978, 1981), Hubbard and Little (1988), Jones (1979), and Willimack and colleagues (1995).

51.CATI is discussed in Bailey (1987:201–202), Bradburn and Sudman (1988:100–101), Freeman and Shanks (1983), Frey (1983:24–25, 143–149), Groves and Kahn (1979:226), Groves and Mathiowetz (1984), and Karweit and Meyers (1983).

52.See Tourangeau et al. (2002), Tourangeau (2004a:791–792), and Weisberg (2005:30–37).

53.On cell phone survey interviewing issues, see Brick et al. (2007), Lavrakas et al. (2007), and Link et al. (2007).

54.For comparison of surveys, see Backstrom and Hursh-Cesar (1981:16–23), Bradburn and Sudman (1988:94–110), Dillman (1978:39–78), Fowler (1984:61–73), and Frey (1983:27–55).

55.For discussions of Web and e-mail surveys, see Birnhaum (2004), Couper (2000), Couper (2008), Couper et al. (2001), Fox and associates (2003), Koch and Emrey (2001), and Tourangeau (2004a:792–794). On Internet usage see "Internet Use Triples in Decade, U.S. Census Bureau Reports," June 3, 2009 [http://www.census.gov/Press-Release/www/releases/archives/communication_industries/013849.html] and "Broadband Internet to Reach 77 Percent of Households by 2012," TMC net, July 29, 2008 [http://www.tmcnet.com/voip/ip-communications/articles/35393-gartner-broadband-internet-reach-77-percent-households-2012.htm].

56.See Couper, Conrad, and Tourangeau (2007); Couper (2008); Dillman (2000:376–400); and Smyth et al. (2009).

57.Elite interviewing is discussed in Dexter (1970). Also see Galaskiewicz (1987), Useem (1984), Verba and Orren (1985), and Zuckerman (1972).

58.On time budget surveys, see Andorka (1987), Bittman and Wajcman (2000), ERIC (1976), Hornsby-Smith (1974), Jordan and Layzell (1992), Mattingly and Bianchi (2003), Meyer (1998), Milem et al. (2000), and Wiedmer (1993) for faculty hours.

59.Dillman (1983) and Groves and Kahn (1979:188–212) discuss costs.

60.See Schwartz (1996), Maynard et al. (2002), and Weisberg (2005:72–91).

61.For more on interviewing, see Brenner et al. (1985), Cannell and Kahn (1968), Converse and Schuman (1974), Dijkstra and van der Zouwen (1982), Foddy (1993), Gorden (1980), Hyman (1975), Moser and Kalton (1972:270–302), and Survey Research Center (1976). For a discussion of telephone interviewing, see Frey (1983), Groves and Mathiowetz (1984), Jordan et al. (1980), and Tucker (1983).

62.See Turner and Martin (1984:262–269, 282).

63.From Moser and Kalton (1972:273).

64.The use of probes is discussed in Backstrom and Hursh-Cesar (1981:266–273), Foddy (1995), Gorden (1980:368–390), Hyman (1975:236–241), Schober and Conrad (1997), and Smith (1989).

65.On interviewer training, see Backstrom and Hursh-Cesar (1981:237–307), Billiet and Loosveldt (1988), Bradburn and Sudman (1980), Oksenberg et al. (1986), Singer and Kohnke-Aguirre (1979), and Tucker (1983). Olson and Peytchev (2007) found negative effects from more interviewer experience, suggesting interviewers become sloppy or rush as they gain more experience.

66.See Leal and Hess (1999).

67.See Bradburn and Sudman (1980), Pollner and Adams (1997), and Zane and Matsoukas (1979).

68.See Anderson et al. (1988), Bradburn (1983), Catania et al. (1996), Cotter et al (1982), Finkel et al. (1991), Gorden (1980:168–172), Kane and MacAulay (1993), Reese et al. (1986), Schaeffer (1980), Schuman and Converse (1971), and Weeks and Moore (1981). Davis (1997) found that when African Americans are interviewed by Whites, they put "self-imposed limits on free expression" and are less likely to say that Whites keep Blacks down or that Blacks do not have the power to affect change.

69.Sudman et al. (1996:74–76).

70.See Bateson (1984), Clark and Schober (1992), Foddy (1993), Lessler (1984), and Turner (1984).

71.From Turner and Martin (1984:276).

72.See Briggs (1986), Cicourel (1982), and Mishler (1986) for critiques of survey research interviewing.

73.For additional discussion of ordinary conversation and survey interviews, see Beatty (1995), Conrad and Schober (2000), Groves el al. (1992), Moore (2004), Schaeffer (2004), Schober and Conrad (2004), Smith (1984), and Suchman and Jordan (1992).

74.On cognitive interviews .see Conrad and Blair (2009), Willis, (2004, 2005), and van der Zouwen and Smit (2004).

75.See Maynard et al. (2002), Maynard and Schaeffer (2004), Moore (2004), Schaeffer (2004), Schober and Conrad (2004), and Willis (2005) on pilot testing methods such as the cognitive interview and related techniques.

76.For a discussion of ethical concerns specific to survey research, see Backstrom and Hursh-Cesar (1981:46–50), Fowler (1984:135–144), Frey (1983:177–185), Kelman (1982:79–81), Marsh (1982:125–146), Miller (1983:47–96), and Reynolds (1982:48–57). Weisberg (2005:311-324). The use of informed consent is discussed in Singer and Frankel (1982) and in Sobal (1984).

77.On reporting survey results in the media, see Channels (1993) and MacKeun (1984).

78.See Singer (1988).

79.From Turner and Martin (1984:62).

第 11 章

非反應式研究以及次級分析

非反應式研究
內容分析
現存的統計數據／文件
調查資料的次級分析
推論的議題以及理論考驗
倫理的考量
結論

在某些研究情況中，若是單獨使用訪談或是問卷的話，會留下一些無法回答的替代性解釋。在此處所強調的那些較不普遍的測量類型的目的，就是去支撐這些弱點，並且提供相關的情報來評估對於效度的威脅。使用這些測量方式的花費很高，但是此種取向對於調查者而言則是更加吃力的。

——Eugene Webb et al., *Nonreactive Measures in the Social Sciences*, pp. 315-316

貝姆—莫拉維茲以及馬斯特洛（Behm-Morawitz and Mastro, 2008）探討了媒體曝光對於年輕成人的影響，尤其是一種「辣妹」（“mean girls”）的風潮（譯註： mean girls 在英文中原意是指脾氣暴躁的、難以駕馭的女孩，而這個詞被當作電影片名，中名譯為《辣妹過招》，但在大陸所譯的片名則為《賤女孩》，為了符合電影片名，而且不過於鄙俗，故在此譯為「辣妹」，）。「辣妹」一詞來自於一部 2004 年上映的電影，所指的是那些藉由表現出社會侵略性而獲得獎勵以及感到樂趣的年輕女孩。作者們在網路上進行了搜尋，發現在 1995 年到 2005 年之間，總共上映了 90 部關於美國青少年的電影。所謂的青少年電影，是指裡面的角色都是以青少年為主，而且其行銷的目標觀眾也都都是青少年。從這 90 部青少年電影裡，他們挑選出其中票房收入最高的前 20 部電影來當作樣本。作者們利用樣本以外的其他青少年電影，來對三名編碼者進行了 48 個小時的訓練。編碼者學習如何去辨識主要角色以及次要角色、社會合作的行為（幫助朋友、解決糾紛）、社會侵略的行為（辱罵他人、排擠他人）、以及這些行為所產生的正面或是負面後果。在蒐集資料時，編碼者在被挑出來的 20 部電影中找出了 139 位主要以及次要角色，其中大部分（87%）的年齡都是介於 15 到 18 歲之間。其中有稍微超過一半（55%）的角色是女性。編碼者辨識出了 337 件與社會侵略行為有關的事件、以及 534 件與社會合作行為有關的事件。在對資料進行了統計分析之後，作者們在電影中發現到，「不論是男性或女性，在對那些明顯地比較容易得到獎勵的女性進行社會侵略行為之後，通常比較可能因此而得到獎勵、而不是懲罰」（p. 136）。作者們接著針對 136 名大學生（年齡大約是 19 到 20 歲）進行一項調查。他們發現，最常觀看青少年電影、以及對電影中的角色認同度最高的大學生，比較可能相信社會侵略行為會得到獎勵，也就是會使他們在同儕之間受歡迎的程度變高。

非反應式研究　一種社會研究的類型，在此種研究中，被研究的人們並不知道自己正在被研究的事實。

非侵擾性測量　非反應式測量的另外一種名稱，此名稱所強調的是，由於研究者所使用的測量方法並未侵優了被研究的人們，因此他們並不知道自己正在被研究者進行研究之中。

實驗以及調查研究都是反應式的（reactive），換句話說，我們所研究的人們都知道他們正在被研究的事實。而在本章裡，我們所探討的是**非反應式研究**（nonreactive research），也就是我們所研究的人們並不知道他們正在被研究。我們將會討論四種非反應式的技術，這些技術通常是依賴實證主義的原則，不過，詮釋取向以及批判取向的研究者也會使用這些技術。我們首先要探討的是一些有創意的非反應式測量，然後接著探討*內容分析*（*content analysis*）。而剩下的另外兩種技術，*現存的統計數據*（*existing statistics*）以及*次級分析*（*secondary analysis*），所指的則是由政府文獻或者是之前的調查中去蒐集資訊。雖然這些資料在第一次被蒐集時可能是反應式的，我們仍然可以在無需使用其反應式效果的情況下來處理新的問題。

非反應式研究

當我們注意到某件事暗示了一個我們感興趣的變項時，非反應式的研究就開始了。當我們採取非反應式或者是**非侵擾性測量**（unobtrusive measures）（也就是不會對研究對象造成侵擾或是干擾的測量方式）時，由於我們所研究的人們並不知道自己正在被研究，因此會以「自然的方式」留下他們的社會行爲或是行動的證據。如此一來，我們就可以在不干擾研究對象的情況下，以這些證據來推論出行爲或是態度。未被注意到的觀察（unnoticed observation）也是一種非反應式的測量，此種測量是以間接的方式來指明某個構念。

舉例而言，羅森布魯姆等人（Rosenbloom et al., 2009）以非侵擾的方式去觀察以及記錄了以色列境內的兩個城市（人口數爲 300,000）、兩個市鎮（人口數爲 3,000）、以及兩個小村莊（人口數爲 800）裡的 1,062 名駕駛人的資訊。他們發現了五種類型的交通違規狀況：（1）沒有繫上安全帶；（2）沒有使用兒童安全座椅；（3）在開車時使用行動電話；（4）沒有遵守「禮讓行人」等交通標誌；（5）在非汽車停等區域停車。根據在都市裡的匿名性，也就是人們通常彼此互不認識的狀況，他們提出一個假設：交通違規狀況比較可能出現在都市區裡。然而，他們卻發現到，在市鎮以及村莊裡的交通違規狀況比在都市中還多，而且男性違規的情況也比女性嚴重許多。由於研究者所觀察的駕駛人根本不知道自己是研究中的一部分，因此這項研究就屬於非反應式的研究。

不同類型的非反應式或是非侵擾性觀察

非反應式測量有許多種不同的類型，而且研究者也已經發明了許多有創意的方法來間接地測量社會行爲（請參閱範例 11.1 在墓碑上發現資料）。由於這些測量方

範例　11.1

在墓碑上發現資料

佛斯特以及同事（Foster and colleagues, 1998）在伊利諾州某個地區的十個墓園裡，檢視了從 1830 年到 1989 年之間的墓碑。在 2,028 個墓地裡，他們找到超過 2,000 位被埋葬者的出生日期、死亡日期、以及性別。研究者們得知，在這個地區的一些趨勢與全國性的趨勢有所不同。他們發現到，該地區的女性懷孕有兩個高峰期（春天以及冬天），10 歲到 64 歲的女性死亡率高於男性，以及年輕人通常在夏季末期過世、而老人則通常在冬季末期過世等等。

法除了都是非反應式的以外，幾乎沒有任何共通之處，因此，我們學習這些測量方法的最佳方式，就是直接去探討許多實際的範例。其中一種類型是**耗損測量**（erosion measure），此種方法所考量的是表面的磨損以及衰退。另外一種類型則是**附加測量**（accretion measure），而此種方法所研究的則是被遺留下來的事物。[1]

研究者已經檢視了在不同歷史時期中的家族畫像，以瞭解座位安排模式如何反映出在某個家族裡的性別關係。都市人類學家也會去檢視垃圾場的垃圾，並且從被丟棄的物品中（例如，酒瓶表示飲用酒類的頻繁程度）去得知人們的生活型態。根據從人們丟棄的垃圾中所得到的資訊，研究者發現人們在酒類的飲用量方面低報了大約 40%到 60%左右（Rathje and Murphy, 1992: 71）。研究者也研究了駕駛人的收聽習慣，方式是趁著車子被送去維修時，去確認收音機轉到哪一個頻道。研究者也會去注意在博物館不同區域裡的地板磁磚的磨損程度，以測量人們對於博物館中不同的展覽品的興趣多寡。研究者們也研究了高中男女廁所裡的塗鴉的差異性，以瞭解塗鴉主題在性別方面的差異。有一些研究者則是去檢視了高中的畢業紀念冊，以比較那些在後來罹患精神疾病以及沒有精神問題的人，在高中時期所從事的活動有無任何差異。研究者也注意到那些貼在汽車保險桿上的、支持不同政治候選人的小貼紙，以瞭解某位候選人的支持者是否比反對者更有可能遵守交通規則。研究者甚至透過觀察水壓變化的方式，來測量民眾的收視習慣，因爲當電視上正在播放廣告時，人們通常會趁這個機會去上廁所。[2]（請同時參閱資訊箱 11.1 非反應式測量的範例）

記錄以及文件

在進行非反應式的測量時，我們所依循的是量化測量的邏輯，不過質化研究者有時候也會使用非反應式的觀察。你首先必須把一個構念加以概念化。接著，你把這個構念和一個非反應式的量數連結在一起。變項的操作性定義，就是你如何以系統性的方式來記錄觀察到的資訊。

如同其他的研究一樣，除了你感興趣的構念之外，你必須把所有進行觀察的理由都排除掉。例如，

耗損測量　一種針對表面的磨損以及衰退痕跡所進行的非反應式測量，而此種表面的磨損是由於人們的活動所造成的。

附加測量　針對人們活動的遺跡或者是人們留下的事物所進行的非反應式測量。

資訊箱 11.1　非反應式測量的範例

物質的痕跡

*耗損：*磨損意味著經常使用。

*範例：*研究者檢查了在某家托兒所裡同時採購進來的一批玩具。那些磨損較為嚴重的玩具，就代表兒童對於這些玩具比較感興趣。

*附加：*物理證據的累積暗示了行為。

*範例：*研究者檢視了在男女宿舍的垃圾桶以及回收桶裡的鋁罐的飲料品牌，而此種資訊能夠指出不同性別所偏好的飲料品牌以及類型。

文獻資料

*定期更新的記錄：*定期出版的公共刊物或許可以揭露出許多資訊。

*範例：*研究者檢視了在婚姻記錄中的新娘與新郎的年齡。在不同地區之間的差異，顯示了在某些地區裡，男性比較偏好與比自己年輕的女性結婚。

*其他記錄：*不定期的、或者是私人的記錄也能夠揭露出許多資訊。

*範例：*研究者發現到，在某一所大學裡，雖然 10 年來學生的數目相當穩定，但是校長辦公室所採購的紙張數量卻逐年增加。此種用紙量明顯增加的情況，顯示了在行政方面的文書作業量正在逐漸增加當中。

觀察

*外表：*人們的外表可以顯現出一些社會因素。

*範例：*研究者觀察學生的穿著，以瞭解在校隊贏了比賽或是輸了比賽之後，他們是否會更加願意穿著校服。

*計算行為次數：*計算人們做某件事情的次數，也可能獲得一些資訊。

*範例：*研究者計算男性與女性在「停止」標誌前面完全停下車和慢慢滑行的人次。此種資訊可以顯示出男性與女性在駕駛行為方面的差異。

*時間長短：*人們花多長的時間去做某件事，可以指出其專注的程度。

*範例：*研究者去測量了男性與女性在一幅裸男的畫作以及另外一幅裸女的畫作前面停留的時間長度。時間長短可能顯示出兩性在看到同性或是異性裸體時所感受到的尷尬或是興趣的程度。

假設你的構念是顧客對於陳列在店面的十項產品感到興趣的程度。你的操作性定義是這十項產品中的每一項所吸引的顧客流量，而測量的方法則是利用一個隱藏的攝影機來測量顧客流量。你必須澄清顧客流量的定義是什麼（例如：因爲產品陳列的位置靠近出口，所以造成更多人潮經過；人們所看的是除了這些產品以外的東西；人們剛好在產品前面停下來穿上外套、而不是在看這些產品；陳列產品的樓層剛好是通往另外一家百貨公司的通道；或者流量只不過代表了一個適合擺放產品的好位置而已）。接著，你以系統性的方式考量了影片所錄到的內容：把影片和店裡其他位

置的影片進行比較、注意一下觀看陳列產品的顧客們、注意顧客走過產品時的速度或者是停下來的時間長短、以及計算有多少顧客轉頭看著這些產品等等。除此之外，你也應該以固定的時間長度（例如，每個小時、每天、每周等等）來記錄結果。

內容分析

內容分析的定義

在內容分析研究中，你將會去蒐集並且分析文本的內容，而這裡所謂的內容可以是文字、意義、圖片、符號、想法、主題、或者是任何溝通的訊息。**文本**（text）所指的是任何寫下來的、視覺上的、或者是語言上能夠當作進行溝通的媒介物，其中包括了書籍、新聞或是雜誌的文章、廣告、演講、官方文件、電影或是影片、音樂歌詞、照片、衣物、網站、或者是藝術作品。在第二章裡，你曾經讀過一項內容分析的研究，該項研究的主題是在 1999 年的科倫拜高中槍擊案之後的校園槍擊案件。而在本章一開頭部分關於「辣妹」的研究，也是一個內容分析研究的範例。（請同時參閱範例 11.2 抗老化產品網站所傳達的訊息是什麼？）

內容分析已經出現了大約一個世紀之久，而且也被運用於許多領域之中，其中包括了文學、歷史學、新聞學、政治學、教育、以及心理學等等。在 1910 年德國社會學學會（German Sociological Society）的首次會議當中，韋伯（Max Weber）就建議使用內容分析法來研究報紙。[3] 在量化內容分析中，你運用客觀的以及系統化的計數和記錄程序，以產生對於某個文本中內容的數值描述。除此之外，內容分析也有質化或者是詮釋取向的版本（請參閱資訊箱 11.2 質化研究者如何研究文件或是統計報告）。不過，本書的焦點仍然是以文本內容中的量化資料爲主。

內容分析之所以是非反應式的，是因爲研究者完全不知道在文本中置入文字、訊息、或是象徵符號以去和讀者或是接受者溝通的過程。舉例而言，身爲本書的作者，我寫下文字或是畫出圖表，以把研究方法的內容傳達給身爲學生的你。我在撰寫這本書的時候、以及你閱讀這本書的時候，都不會知道或者是意圖讓這本書被其他人拿來進行內容分析。

內容分析讓你看到並且揭露在一個溝通來源（也就是一本書、一篇文章、或是一部電影）中的內容（也就是訊息、意義、或者是象徵符號）。你深入探討以及發現內容的方式，與你平常閱讀一本書或者是觀賞一個電視節目的方式是不一樣的。內容分析能夠以客觀的、量化的方式來證明，根據非系統化觀察結果所得到的感覺是否是真實的。內容分析能夠產生關於文本的可重覆的、精確的結果。在你蒐集了資料之後，你接著必須以統計方式來分析這些資料，你所使用的方式就和一名實驗者或是調查研究者所使用的方式是完全相同的。

文本　一種對於溝通媒介的廣泛名稱，而內容分析則是去測量其中的象徵性意義。

範例 11.2

抗老化產品網站所傳達的訊息是什麼？

就像性別歧視以及種族歧視一樣，年齡歧視（ageism）也需要有一組正面的以及負面的刻板印象以及訊息，以強化權力關係、不平等狀況、以及社會特權等等。卡拉珊媞（Calasanti, 2007）檢視了一些網站，以瞭解這些網站的行銷手法所傳遞給年長消費者的訊息是什麼。她找到了96個抗老化的網站來當作樣本，並且把每一個網站的圖片以及文字加以編碼成為各種類別。編碼後的類別包括了老年／老化的問題；對付老化問題的解決之道；老年的性別面向；網站上顯示的老化身體的各種面向；以及關於階級、種族、以及性別的傾向。年齡歧視的一個關鍵訊息，就是如果你能夠修復你的身體以阻止老化，那麼你就應該這麼做，否則的話，你就會被人忽視、或者是成為一名失敗者。卡拉珊媞發現，抗老化的廣告不只提倡各種隱藏身體老化跡象的方法，而且一貫地把理想的個人形象顯示為一個特定的種族（白人）、階級（中產階級或是更高的階級）、以及性傾向（異性戀）。這些廣告中所展現的男性，在運動比賽或是工作中都是佔優勢的，而且在性方面是有吸引力的。而在廣告當中所顯示的老化女性的形象，則是正在試圖引誘性伴侶、想和年輕女性們競爭、而且在性方面只能順從於男性。這些廣告所暗示的意義是，人們無法同時既年老又擁有某個特定的性別，至少以白人、中產階級、異性戀的理想情況而言正是如此。這些廣告傳達給人們的訊息是，當人們看起來年紀較老時，就意味著失去性別以及性吸引力（也就是說，變成了一個中性的、沒有性別的人，既不是男性、也不是女性），而且只有藉由讓自己看起來更為年輕的方式，他們才能恢復這些受到社會所重視的有利條件。

適合進行內容分析的主題

內容分析被運用於許多不同的用途之上，像是研究流行歌曲中的主題以及在聖歌中的宗教符號、報紙報導的主題的趨勢以及報紙中社論的意識型態論調、在教科書或是電影中的性別角色刻板印象、不同種族的人們出現在電視廣告以及節目中的頻率、對於開放式調查問題的回答、戰爭時期的敵方宣傳刊物、流行雜誌的封面、自殺者遺言中透露的人格特質、廣告訊息中的主題、以及在對話當中的性別差異等等。

你以內容分析爲基礎所作出的概化推論，會受限於文化溝通本身。內容分析無法決定某個陳述的真實性、或者是去決定文學作品的美學特質。內容分析可以揭露出文本的內容，但是無法詮釋該內容的重要性。因此，你應該要直接去檢視文本。何斯提（Holsti, 1968a: 602）曾經警告過：「內容分析也許可以被視爲對於文件進行主觀檢視時的補充物、但並不是替代物。」

內容分析對於以下三種類型的研究問題是很有用的：涉及大量文本的問題、距離可能相當遙遠或是散佈各處的內容、以及難以利用一般觀察來瞭解或是記錄的內容。你可以利用抽樣以及多位編碼者的方式來測量大量的文本（例如：在20年之間的報紙文章）。你也可以研究「有相當距離」（"at a distance"）的主題（像是在一個敵對國家中的廣播內容）、或者是散佈各處的主題（像是由同一位導演在20年之間所拍攝的15部影片中的共同主題）。最重要的是，內容分析能夠揭露一些在文本中難以用一般觀察來看到的訊息，甚至是文本的創造者、或者是閱讀文本的人可能也

資訊箱 11.2　質化研究者如何研究文件或是統計報告

運用詮釋或批判取向的質化研究者，也會研究一些有統計資訊的文件以及報告，但他們在進行研究時，所使用的方法和實證主義者是不一樣的。他們認為文件及統計報告，都是傳遞社會意義的文物或是媒介。他們把文件視為屬於某個範圍的其他文物（例如，紀念碑、日記、音樂的樂譜、採購清單、影片、相片、繪畫、工程圖、網頁等等），同時這些文物承載著意義。例如，一份建築的地面規劃圖就是一份表達空間安排的文件，也傳達了社會意義。有一些辦公室就座落於一些令人羨慕的地點、並搭配大片的窗戶，而其目的就是設計給位於某些高階工作職位的人士所使用。

質化研究者並不會把一份文件或統計報告，當成是一個包容了內容的中性容器，而是會檢視其創造、散佈、以及接收等等過程的更大的脈絡。與建構主義者的觀點（我們已經在第四章中討論過了）一致的是，質化研究者所強調的是從一份文件的創造（包括其創造者的意圖）、到該份文件被各種不同的接收者／消費者接收的整體過程，然後把該份文件置入於一個社會脈絡之中。簡而言之，他們把文件或報告當作是一件本身就承載著社會意義的文物。因此，雖然質化研究者或許會檢視一份文件或報告的內容，但是他們卻不會把自己僅僅限制在內容上。

質化研究者強調，人們是在意義以及文字或數字的基礎上，進行思考以及互動的。例如，在某份文件中的內容，可能傳達了醫療訊息給健康照護的員工、給予某個人得到社會服務的許可、販售產品給消費者、提供官員們發生問題的地理區域的資訊、或允許／防止某個人進入國境。不同的人可能在不同的時間、以不同的方式，運用完全相同的一份文件，而「閱讀」或詮釋文件的過程，通常取決於訓練以及依循著一定的規則。例如，人們學會了在醫療報告、統計報告、或是護照當中應該尋找什麼樣的訊息。看著相同一份文件的人們，可能會看到不同的事物、依循不同的規則、或把該份文件運用於不同的目的上（例如，取得保險理賠或開立醫療處方、考驗一個假設或分配經費給一座新的公共建築物、允許某人進入一個國家、或兌現一張支票等等）。質化研究者會仔細考慮一份文件及其內容的多重面向。例如，一篇雜誌文章可能傳達了能夠取悅讀者的內容、可能是作者用來建立名望的工具、可能會引發公共爭議、或是一種提高雜誌銷售量的方式等等〔請參閱葛理斯渥德（Griswold, 1987, 1994）以及普萊爾（Prior, 2003）對於文物以及文件相關研究的評論〕。

都沒有察覺到文本內所具有的主題、偏誤、或者是特質。例如，學齡前兒童繪本的作者可能不是有意識地故意以傳統刻板印象的性別角色來描繪書中的兒童人物，但是透過內容分析的方式，研究者已經揭露了此種高度性別刻板印象的狀況。[4] 另外一個例子，則是在單一性別團體中的對話。雖然人們可能沒有察覺到，但是在相同性別團體中，女性比較常會談論人際關係以及社會關係的問題，而男性則比較常會談論成就以及具侵略性的主題。[5]

測量以及編碼

如同在大多數的量化研究中一樣，謹慎小心的測量在內容分析中也是非常重要的。你必須把冗長而且模糊的象徵性溝通內容，轉化成爲精確的、客觀的、量化的資料。爲了要達到這個目的，你必須非常謹慎小心地設計以及記錄編碼的程序，以使他人可以複製研究過程。例如，你想要瞭解在電視劇中以負面刻板印象的方式來描繪年長者角色的頻率有多高。你必須發展出「年長者負面刻板印象」這個構念的測量方式，此種概念化的結果可能會是一份清單，上面列出了一些關於年長者的刻板印象或者是負面的概括看法（例如，衰老、健忘、暴躁易怒、脆弱、重聽、動作遲緩、生病、住在安養之家裡、沒有活動力、保守等等），這些印象可能正確地反映出年長者的狀況，但也可能沒有。例如，如果只有 5%超過 65 歲以上的年長者住在安養之家裡，但在電視節目裡，超過 65 歲以上的年長者中卻有高達 50%的人被描繪爲住在安養之家裡，這就是支持負面刻板印象的證據。[6]

在內容分析研究中，你會以一個**編碼系統**（coding system）來將構念加以操作化。編碼系統是一組指示或是規則，描述了如何觀察以及記錄來自於文本的內容。你將這個系統加以量身打造，以適合於你正在研究的文本或是溝通媒介的類型（例如：電視劇、小說、雜誌廣告中的圖片等等）。除此之外，編碼系統也取決於你的分析單位。

在內容分析中所使用的分析單位可能會有非常大的差異。分析單位可能是一個字、一個片語、一個主題、一段情節、一篇報紙文章、或是一個角色等等。在本章一開頭部分關於「辣妹」的研究中，分析單位就是電影中的角色。而除了分析單位之外，你在內容分析中也會使用其他單位，像是記錄單位、脈絡單位、以及計算單位等等，不過，這些單位不一定會和你的分析單位完全相同。在這些單位之間僅有一些細微的差異，而且彼此之間也很容易混淆，但是每一種單位都扮演了獨特的角色。在簡單的計畫中，這三種單位都是完全相同的。例如，你或許會去注意關於汽車和貨車的電視廣告的特色（電視廣告是記錄單位），以及在此種廣告之前與之後出現的其他廣告（脈絡單位），並且計算在每個廣告中出現的人物數量以及特徵（人物是計算單位）。

內容分析裡的測量所採用的是**結構式觀察**（structured observation），也就是基於書面規則而進行系統化的、謹慎的觀察。書面規則解釋了如何將觀察加以分門別類，而且，正如同其他的測量一般，所有的類別應該都要能夠彼此互斥而且全容。書面的規則讓其他人可以複製觀察過程，並且也能增加信度。雖然研究者是從初步的編碼規則開始的，但是他們通常都會進行前導研究，並且依據前導研究的結果來改進編碼的方式。編碼系統確認了文本內容的四種特質：頻率、方向、強度、以及空間。在

編碼系統 一組用於內容分析中的指示或是規則，以解釋研究者如何以系統化的方式把文本的象徵性內容轉變成為量化資料。

結構式觀察 一種觀看在一個社會情境中發生的事件的方法，此種方法具有高度的組織性，並且依循著觀察以及記錄的系統性規則。

一個內容分析的研究計畫中，研究者可能會去測量一種或者是全部四種特質（請參閱資訊箱 11.3 我們所測量的事物）。

顯性編碼　一種內容分析的編碼類型，在使用此種編碼方法時，研究者先發展出一份關於字詞、片語、或者是符號的清單，然後在一個溝通媒介中去找出這些字詞、片語、或是符號。

編碼、效度、以及信度

在進行編碼時，研究者必須謹慎小心地審視文本，並以一種非常系統化的方式把文本內容轉化成爲重要的文字、符號、或訊息等等所形成的量數。內容分析編碼的主要類型可分爲兩種：顯性編碼及隱性編碼。

顯性編碼（manifest coding）牽涉到在文本中可見的、表面的內容。例如，你計算了一個片語或詞彙（例如：紅色）在一份書面文本中出現的次數、或一個特定的動作（例如：親吻）在影片中是否出現。顯著編碼系統中有一份你想找出的詞彙或動作的清單。對於書面的文字而言，你可以先把資訊掃瞄成電子檔案的格式，並且利用電腦程式尋找特定的字詞或片語，讓電腦計算這些字詞或是片語出現的次數。而爲了達到這個目的，你首先必須創造一份相關字詞或片語的完整清單。[7]

顯性編碼具有相當高的信度，因爲某個片語或字詞要不是有出現、就是沒出現，這是十分明確且幾乎沒有爭議空間的。不幸的是，顯性編碼並不會考量到字詞或是片語的其他含意。依據脈絡的不同，即使是完全相同的字詞，也可能會有許多不同的意義。由於相同的字詞可能會有多重的意義，因此也限制了顯性編碼的測量效度。

例如，我閱讀了一本*紅*色封面的書，這本書最近很*紅*。不幸的是，這本書的出版商在營收方面卻是滿篇*紅*字（譯註：表示虧損），因此編輯無法處理當一本書*紅*得發紫時所必須面對的堆積如山的、上面綁著*紅*色帶子（red tape）的各種文件（譯註：red tape 在英文中意爲「繁文縟節」，但若以此種方式翻譯，就沒有「紅」這個字，故只能依字面意義直譯）。這本書所敘述的故事內容，是一輛*紅*色的消防車在葉子轉*紅*的季節之後，才會在*紅*燈時停下來。另外也有一群的*紅*色激進份子（Reds）（譯註：此字在英文中是指激進人士，爲了配合作者強調「紅」這個字，故只好加上「紅色的」）帶著*紅*色的旗子來到一間小型的*紅*色校舍前面。而另外一群面色*紅*潤而充滿活力的*紅*脖子人士（redneck）（譯註：這個字所指的是脖子被曬紅的美國中南部的貧苦農民，尤其是指其中那些觀念比較狹隘保守者）則強烈反對他們，這群紅脖子人士喜歡吃*紅*肉，並且尊敬由*紅*、白、藍三色所組成的美國國旗。書中的主角是一位*紅*鼻子的鬥牛士，他是用他的*紅*色斗蓬去鬥*紅*狐狸、而不是公牛。嘴唇*紅*潤的小*紅*帽也出現在這本書中。她因爲在*紅*燈區吃了大量的*紅*色辣椒，所以眼睛發*紅*，而且面*紅*耳赤。她媽媽有著一頭*紅*髮，在知道這件事之後非常生氣，所以把她的屁股打到發*紅*。（譯註：在本段文字中，作者爲了強調 red 這個字而編寫出這段故事，因此故事內容中有許多不甚合理之處，此外，在英文中一些與 red 有關的片語，在中文裡的翻譯不一定會有「紅」這個字，但爲了忠於原文起見，譯者盡量以與「紅」有關的方式譯出，但可能因此而使得中文有一些不通順之處，請讀者不必深究）。

資訊箱 11.3　我們所測量的事物

1. *頻率（frequency）*　頻率所指的是某個事物是否出現，以及如果出現的話，其出現的頻率有多高。例如，在一個星期之中，有多少年長者出現在電視節目當中？這些年長者在所有的人物角色中所佔的百分比率是多少？或者是有他們出現的節目在所有節目當中所佔的百分比率是多少？
2. *方向（direction）*　方向所指的是內容中的訊息在一條連續線上（例如：正面或是負面、支持或是反對）的方向性。例如，我們製作了一份清單，內容是關於在電視節目中的年長者的角色所表現出的行動。我們把這些行動分成三個類別：正面的（例如：友善的、有智慧的、仁慈的、體貼的）、中性的、或者是負面的（例如：喜怒無常的、遲鈍的、自私的、緩慢的、健忘的）。
3. *強度（intensity）*　強度所指的是在某個方向上的訊息的強度或是力量。一個在電視節目中的人物角色可能是活躍的（例如：到處亂跑、很快而且很大聲地在說話）、或是較為文靜的（例如：幾乎是靜止不動地站著，並且輕聲細語地說話）。某個特質，像是健忘等等，可能是輕微的（例如，在離開家裡的時候忘記帶鑰匙，要花比較長的時間才能回想起一個已經有十年沒見面的人的名字）、或是嚴重的（例如，不記得自己的名字、不認得自己的孩子）。
4. *空間（space）*　空間所指的是文本訊息的大小、時間長度、或者是某個訊息所佔的空間多寡。我們可以輕易地去測量一份廣告傳單或是照片的大小或是空間。在測量書面文本中的空間時，我們可以去計算文字、句子、段落、或是在一個頁面裡所佔的空間（例如：多少平方英吋）。對於影片以及聲音文本而言，我們則是去測量所佔的時間長短。例如，某個電視節目中的人物可能出現幾秒鐘、或是出現在一個長達一小時的節目中的每一個畫面。

隱性編碼（latent coding）〔也被稱爲*語意分析*（*semantic analysis*）〕則是去尋找文本內容當中隱藏的、未加以言明的意義。例如，你閱讀了一整個段落，並且要去決定其中是否包含了色情的主題或是浪漫的情緒。你的編碼系統中包含了一些廣泛的規則，以指引你對於文本的詮釋，並且決定特定的主題或是情緒是否有出現在文本當中。在本章一開頭部分關於「辣妹」的研究中所使用的編碼方式就是隱性編碼，但是隱性編碼的信度通常比顯性編碼來得低，因爲隱性編碼的信度取決於編碼者對於語言以及社會意義的知識。[8] 訓練、練習、以及書面的規則都可以增加信度，但是要能夠以前後一致的方式去辨識出主題以及情緒等等之類的事物，仍然是一件相當困難的事。然而，隱性編碼的效度可能會超過顯性編碼，因爲我們在溝通意義時，所使用的是許多含蓄的、未加以言明的方式，而這些方式則是取決於脈絡、而不僅僅是特定的字詞而已。

隱性編碼　一種內容分析的編碼類型，在使用此種編碼方式時，研究者先確認一些主觀的意義，像是主題或是動機等等，然後再以系統性的方式在一個溝通媒介中去找出這些主觀的意義。

你或許會想要同時使用顯性以及隱性編碼來研究文本的內容，因爲若是此兩種方法所得到的結果完全相同的話，將能夠使你的研究結果更有說服

力。然而，若是此兩種方法所得到的結果不同的話，你就應該要重新檢視你的操作性定義以及理論性定義。

在許多研究中，你可能必需把來自於非常大量單位的資訊進行編碼。你可能看了三十本書、數百個小時的電視節目、或者是大約一百個網站（就像在本章一開頭部分所提到的那個研究一樣）的內容。除了親自對資訊進行編碼之外，你或許會雇用助理來協助你進行編碼的工作。如果你雇用了助理的話，你就必須教導這些編碼者關於編碼系統的內容，並且訓練他們填寫一份記錄單。編碼者應該要瞭解各種變項、遵循編碼系統、並且在遇到意義模糊不清時要提出疑問。在編碼工作開始進行之後，當你遇到一些新的特定編碼情境，並且作出相關的決定之後，你必須把這些決定通通都記錄下來，如此一來，你才能保持一致性與連貫性而不會產生混淆。

如果你使用了數個不同的編碼者，你就必須一直去確認編碼者之間的一致性。爲了達到這個目的，你必須要求編碼者去獨立地針對同一份文本進行編碼，然後再確認不同編碼者之間的一致性。你必須去測量**編碼者間信度**（intercoder reliability），這是一種等值信度的類別，而測量的方式是利用一種統計係數，此種係數能夠確認在不同的編碼者之間達到一致性的程度〔請參閱資訊箱 11.4 克里本多夫的阿爾法係數（Krippendorff's alpha）〕。[9] 在報告內容分析研究的結果時，你一定要一併報告這個係數。在本章一開頭的部分所描述的那項研究中，就針對了每個被測量的變項報告了一個編碼者間信度的量數（也就是*克里本多夫的阿爾法係數*）。爲了要得到這個係數，三位編碼者每個人都針對在研究中所使用的所有影片的十分之一獨立進行編碼。該項研究的阿爾法係數範圍大約落在 0.72 到 1.0 之間，而大多數則都是超過 0.80。

如果編碼的過程持續了相當長的一段時間（例如：超過三個月），你應該也要去確認穩定信度，方法是要求每一位編碼者針對之前已經編碼過的樣本文本獨立進行編碼，以確認編碼結果是穩定的、或者是在變化中的。例如，你在四月時已經完成了 6 個小時的電視影集內容的編碼工作，而你在九月時則要求編碼者針對這 6 個小時的內容再次進行編碼，而且不讓編碼者看到原來的編碼結果。如果兩次的結果完全相同的話，那麼你的研究結果就具有穩定信度。但是如果你在兩次的編碼結果中看到相當大的差異，那麼你或許就必須重新訓練編碼者，並且重新進行編碼。

研究者已經研究過許多形式的視覺「文本」，像是照片、繪畫、雕像、建築、衣物、錄影帶、以及電影等等。視覺「文本」比書面文本更加難以分析，因爲視覺文本是透過影像、符號、以及隱喻等等事物，以間接的方式來傳達訊息或者是情緒的內容。除此之外，視覺影像通常會包含了各種不同的、混合在一起的訊息，而且會在多重不同的意義層次上運作。

大多數的人對於主流文化中的關鍵符號都共享相同的意義，但是人們卻有可能以不同的方式來解讀某一個符號或象徵。例如，你應該把國旗「解讀」爲意指愛國精神、對於國家的責任、以及對於傳統的尊重，還是「解讀」爲宰制、濫用權力、以及監督或者是軍事侵略？在 1945 年到 1999 年之間，日

編碼者間信度　在有多位內容編碼者的內容分析中所使用的等值信度，此種信度要求在不同的編碼者之間必須具有高度的一致性。

資訊箱 11.4 克里本多夫的阿爾法係數

克里本多夫的阿爾法係數（Krippendorff's α）是在測量編碼者間一致性或者是評分者間信度時，最被廣為使用而且最為知名的量數。克里本多夫（Klaus Krippendorff）發展出此種編碼者間信度係數的目的，是為了要測量在觀察者、編碼者、裁判者、評分者、以及測量工具之間的一致性。當觀察者之間完全一致時，$\alpha = 1$，也就是表示完美的信度。而如果在觀察者之間因為機率因素而造成任何不一致的情況時，$\alpha = 0$，也就是表示缺乏信度，這就好像是編碼者無法觀察文本或是資訊、因而利用像是擲骰子之類的方式來編造出他們的資料一樣。

此係數的一般公式為：$\alpha = 1 - \frac{D_o}{D_e}$

其中 D_o 所指的是被觀察到的不一致性（observed disagreement），而 D_e 所指的意思是，當對於分析單位的編碼結果只能被歸因於機率、而不是因為這些單位的特質時，研究者所預期會出現的不一致性。

這個公式背後的數學原理以及更為進階的細節，已經超過了本書的程度（請參閱 Hayes and Kreppendorf, 2007，以及 Krippendorf, 2004）。這個公式的資料，來自於兩個或是多個一同接受訓練的編碼者，這些編碼者獨立地對一組共同的分析單位中的每一個變項指定數值。這個係數的細節會依據編碼者的數目以及變項的數值範圍等等而有所變化。這個 α 係數可以應用於許多情境之中：任何數目的編碼者、任何數目的變項類別或是量數、任何的測量尺度（名目尺度、順序尺度、區間尺度、比率尺度）、任何不完整的或是遺失的資料、以及任何大小的樣本。現在已經有數種與統計有關的電腦程式能夠計算這個統計數值。

本很少在公立學校中展示出日本國旗。之後，日本政府官員執行了一項要求展示國旗以及演奏國歌的法律，因而引起了極大的爭議。保守派的政客想要藉由展示國旗方式，以在年輕人心中灌輸更多的愛國精神。然而，許多教師以及其他人卻反對這樣的作法，因爲他們認爲日本國旗會讓人聯想到日本過去的軍事侵略以及對於民主的打壓，除此之外，日本的極右派團體也經常鼓吹要展示國旗，因此也造成許多人的反感。

而在美國，國旗對於不同的社會團體而言，則包含了截然不同的意義。[10] 對於許多非裔美國人而言，國旗象徵了種族隔離、奴隸制度、以及在種族隔離時期白人對於黑人的暴力壓迫。對於許多年紀較大的白人而言，國旗象徵了地區的傳統，以及一種上流的「老南方」生活型態。對於其他人而言，國旗代表了起義造反、個人自由、以及拒斥外在力量強加於個人身上的權威。對於某些身處於美國之外的人而言，國旗純粹只是一個與美國有關的、色彩豐富的時尚宣言。換句話說，當我們把國旗當作一個符號或是象徵時，其實有許多種可能的解讀方式。

爲了研究視覺形像，你必須學習「解讀」（"read"）視覺文本的多重意義，以及學習如何詮釋各種不同的象徵形象。此種「解讀」並不是機械式的（也就是說，某個形象永遠代表了某個意義），而是取決於文化脈絡，因爲形象的意義是以文化爲基礎的。此種解讀也取決於在一個由許多符號所形成的領域內的交互關係。美國國旗

的意義可能會因爲年齡、種族團體、以及地理位置等等因素而有所變化，而且也可能因爲其展示的方式而有不同的意義。在三 K 黨的集會上展示國旗、在密西根大學的美式足球賽中懸掛國旗、與喬治亞州州旗一同懸掛的國旗、或者是印在摩托車騎士夾克背部的國旗，所代表的意義可能都不一樣。在我的家鄉裡，我讀到一篇報紙的文章，該篇文章提到警方正在嚴密監視一個提倡種族仇恨的高中幫派，該幫派會去毆打以及恐嚇威脅非白人的學生。而這個幫派所使用的象徵符號（在帽子以及夾克上），正是美國國旗。美國的國家象徵，像是自由女神像等等，也被用來傳遞一些社會或是政治訊息（請參閱範例 11.3 雜誌封面以及移民）。

社會政治團體會去建構新的象徵符號、或者是去爭取權力以掌控主要的現存象徵符號的意義。例如，有些人想要爲聖誕樹賦予宗教的意義；其他人則想要讓聖誕樹代表一種對於傳統以及家庭價值的慶祝方式，而且不帶任何宗教意涵；還有另外一些人爲了商業的理由而想讓聖誕樹代表一種節慶、歡樂假期的意義。由於象徵符號具有複雜的、多重層次的意義，因此你必須結合關於形象的質化判斷、以及內容分析裡的量化資料，才能作出比較正確的判斷。

如何進行內容分析研究

1. *形成研究問題*　在進行研究時，你必須先從一個主題以及研究問題開始。當研究問題所牽涉到的變項是訊息或是符號時，就很適合進行內容分析。舉例而言，你想要研究當地電視台報導市長選舉的方式，如此一來，你的研究問題可能是每一位候選人被報導的次數或時間是否相同。「報導」這個構念包括了報導的時間長度（出現在電視上的時間）、報導的顯著性、以及某次報導是否偏好某位候選人等等。你可以去對人們進行調查，並且詢問他們對於某次報導的想法，但是，比較好的策略，則是利用內容分析去直接地檢視新聞報導的內容。

2. *決定分析單位*　你必須決定要進行分析的單位。例如，以一場政治上的競選活動而言，你的分析單位有可能是在數個當地的電視台上每天的新聞報導、或者是在所有當地電視台中每天的晚間新聞裡的每一篇新聞報導或是片段。你也可以去計算由候選人或是議題團體所播出的電視廣告次數。除此之外，你也可以去研究那些有候選人參加的電視辯論會或者是訪談節目的內容。

3. *發展出抽樣計畫*　隨機抽樣的技術在內容分析中是非常有用的。首先，你必須先定義母體以及抽樣元素。例如，母體可能是在某一個特定時期內的某些類型文件中的所有文字、所有句子、所有段落、或者是所有文章。相同地，母體也可能包括在某一個特定時期內某些類型的電視節目中的每一場對話、情境、場景、或者是插曲。例如，你可能想要研究在選舉之前的一年之中、以及在選舉過後的一個月之內，在電視新聞節目、電視廣告、辯論或者是訪談節目當中，每位候選人被報導的方式。你必須決定母體是否要包括白天的新聞節目、以及周一到周五晚間新聞中的兩次特別報導。你是否應該要把在一周之內任何一天的白天時間所播放的廣告也算在內？或者是把母體限制在一天之中比較多人收看電視的時段？你的分析單位可能

範例 11.3

雜誌封面以及移民

查維茲（Chavez, 2001）曾經針對主要的美國雜誌的封面進行過一項內容分析研究，這些雜誌所討論的都是美國的移民議題。在檢視了十本雜誌從 1970 年代中期到 1990 年代中期的封面之後，他把這些雜誌的封面分類為傳達下列三種主要訊息的其中一種：肯定的、憂慮的、或者是中立的或平衡的。除了進行分類以及辨識出訊息中的趨勢之外，他也注意到了在封面照片中各種不同人們的組合（也就是不同的種族、性別、年齡、以及衣著）、以及重覆地使用重要的象徵符號（像是自由女神像或者是美國國旗），會以何種方式來傳達一些訊息。查維茲主張，雜誌封面是一種創造文化意義的場所或是地點。在雜誌封面上的視覺影像擁有多重層次的意義，而讀者在閱讀這些影像時，會利用他們的文化知識來建構出特定的意義。以集體的層次來說，這些封面傳達了一種世界觀，並且表達了關於一個國家以及其人民的訊息。例如，在一本雜誌的封面中，編輯們可以把自由女神描繪為既強大而且充滿了同情心（訊息：歡迎新進的移民們）、或是讓自由女神具備了明顯的亞洲人臉部特徵（訊息：來自亞洲的移民扭曲了美國的文化，而且改變了美國的種族組成方式）、或者是讓自由女神高舉一個大型的「停止」標誌（訊息：移民們滾出去）。查維茲（Chavez, p. 44）觀察到，「在雜誌上的影像不但會暗示了當代的『美國』認同，而且在這個過程中，也會有助於建構或是創造出此種認同。」

是一些焦點在於選戰活動以及提及某位候選人的新聞節目片段、一個候選人的圖片或是名字出現在內的電視廣告、或者是邀請一位或是多位候選人作爲來賓的訪談節目。你的母體可能包含了在 13 個月的期間之內，於四家當地電視台中所播放的全部新聞節目片段、電視廣告、以及訪談節目。

在你已經具體指定了目標母體以及你的抽樣元素之後，你必須以更爲詳細的方式來進行規畫，而這就包括了建構抽樣架構，以及決定樣本大小與抽樣比。正如同你已經在第八章中學過的，在製作出一份精確的隨機樣本時，抽樣架構是相當關鍵的一個步驟。爲了針對市長候選人在 13 個月的期間之內的相關報導進行抽樣，你應該列出所有的新聞片段、電視廣告、以及辯論或訪談節目。實際上，你或許會有三個單獨的樣本，其中一個是電視廣告、另外一個是新聞片段、而最後一個則是辯論或訪談節目。由於相同的一個電視廣告可能會出現許多次，你或許會針對電視廣告所有可能出現的時段列出一張清單（抽樣架構），然後針對這些時段進行隨機抽樣。由於在一個新聞節目中可能有零個、一個、或者是數個片段的主題是關於選舉以及候選人的，因此你可能必須搜尋每個電視台的節目表，以取得一張列出所有可能片段的清單，然後再把這份清單當作是你的抽樣架構。如果只有少數的辯論或訪談節目的話，你或許必須把整個母體都包括進來。

當你在進行規劃的時候，你應該要計算必要的工作量。例如，在前導測試當中，你或許會發現，就觀看以及進行編碼工作的時間而言，一個 30 秒到 60 秒的廣告平均大約會花費你 15 分鐘的時間、一個 3 到 5 分鐘的新聞片段大約要花費你 20 分鐘的時間、而一個 30 分鐘的辯論或是訪談節目則可能要花費你 2 個小時的時間。以上的這些時間還不包括進行抽樣或是找出這些廣告、新聞片段、或者是辯論節目時所必須花費的時間。讓我們假定在你的抽樣架構中有 300 個電視廣告、80 個新聞片段、

以及 5 個辯論或訪談節目，而你則是從中抽樣了 100 個電視廣告、40 個新聞片段、以及全部的 5 個辯論或訪談節目。如此一來，在你已經蒐集完所有的影片、並且加以整理組織之後，你在進行編碼工作時所必須花費的時間將會是：（15 分×100 個廣告）＋（20 分鐘×40 個新聞片段）＋（180 分鐘×5 個辯論節目）＝ 2,200 分鐘，或者是將近 37 個小時。你或許必須考慮一下花錢雇用一些助手來當作編碼者。

4. *建構編碼的類別以及記錄單*　你必須確認所有感興趣的變項。這些變項通常來自於文獻回顧中的一些想法、你自己的想法或是理論、或者是對於前導資料所進行的初步分析。你應該爲你自己以及編碼者們（如果你有雇用編碼者的話）建立一個非常清楚明確的編碼系統。顯性或是隱性編碼系統將會精確地描述如何把編碼者所看到或是聽到的事物，轉化爲少數的編碼類別。爲了整理編碼的資料，你應該要製作一張記錄單，在記錄單上有表格或是一個空位來記錄單位的識別碼，並且也有記錄每個變項的編碼資訊的空間（請參閱圖 11.1）。你應該總是以一些前導測試的資料（大約十多個單位）來檢驗你的編碼系統以及記錄單。

5. *編碼以及編碼者間信度檢驗*　最後，如果你使用了多位編碼者的話，你必須去檢驗編碼者間信度，而這通常意謂著先從你全部的樣本中挑選出大約 10%，然後要求每一位編碼者分別獨立地使用編碼系統來針對相同的單位進行編碼。如果有必要的話，你應該把那些不正確的編碼者的資訊丟棄，並且重新進行編碼。

6. *資料蒐集以及分析*　在你已經準備好編碼系統以及記錄單，並且也已經把所有的編碼者都訓練好了之後，你就可以開始去蒐集以及檢查資料了。之後，你可以把資料輸入電腦以進行統計分析、詮釋結果、並且準備撰寫研究報告。

推論

你依據結果所能夠進行的推論，在內容分析中是一個相當重要的議題。內容分析所描述的是文本中所包含的訊息，但是卻無法揭露文本創造者的意圖、或者是文本中的訊息對於接收者所產生的影響或效果。例如，內容分析顯示出童書中包含了性別的刻板印象，然而，這並不必然就意味著童書中的刻板印象會形塑了兒童的信念以及行爲，你必須另外針對兒童的看法進行一個獨立的研究，才能夠證實上述的推論。在本章一開頭部分所描述的研究中，作者們以調查法進行了第二項研究，以瞭解他們的內容分析結果如何影響了青少年電影的觀眾。

現存的統計數據／文件

適當的主題

其實研究者已經可以取得關於社會世界的許多類型資訊，其中有些是以統計文件的形式（像是書籍、報告等等）保存著、或者以集刊的形式被收藏在圖書館或者

圖 11.1　市長選舉內容分析中的電視廣告記錄單範例

廣告編號____　廣告日期____　廣告時間____　編碼者#____

播放廣告的電視台____時間長度____廣告的贊助者____

候選人 A 的影像是否出現？____　候選人 A 的聲音是否出現？____
候選人 B 的影像是否出現？____　候選人 B 的聲音是否出現？____
候選人 C 的影像是否出現？____　候選人 C 的聲音是否出現？____

除了候選人之外，出現在廣告中的其他人物的數目____

年長者的數目____　非年長者的成人數目____　兒童的數目____

性別_____種族 W__ B__ H__ A__ O__

職業_____

勾選廣告中有提到的議題：

1. 稅賦____
2. 學校____
3. 犯罪____
4. 居住問題____
5. 公共運輸____
6. 道路____
7. 社會服務____
8. 公園____
9. 商業活動____
10. 都市發展____
11. 水／空氣品質____
12. 對於未來的願景____
13. 社區____
14. 消防／治安____
15. 分區規劃____
16. 圖書館____
17. 生活品質____
18. 年長者服務____
19. 年輕人服務____
20. 多元文化的議題____
21. 公共衛生____
22. 創造就業____
23. 有效率的政府____
24. 民眾的意見____
25. 與其他政府機構的合作____

候選人 A 所提出的競選承諾_____
候選人 B 所提出的競選承諾_____
候選人 C 所提出的競選承諾_____

_____以上述議題編號_____對候選人 A 提出批評
_____以上述議題編號_____對候選人 B 提出批評
_____以上述議題編號_____對候選人 C 提出批評

是電腦化的記錄中。不論是哪一種形式，你都可以在心中先想好研究問題以及變項，去搜尋此種資訊的集合體，然後以新的方式去重新整理資訊、並且以統計方式來分析這些資訊，以解答一個研究問題。

要具體說明適合使用現存統計數據研究的主題，是相當困難的事，這是因爲這些主題的差異性其實相當地大。你可以去研究一些某個組織已經蒐集過相關資訊的主題，並且把這些資訊公諸大眾。事實上，現存的統計數據研究，並無法完全符合於研究設計的演繹模型。更確切地說，在你找到了一些可以取得的資料之後，你就必須以有創意的方式，將現存的資訊加以重新安排成爲研究問題中的變項。

你應該還記得，最適合進行實驗研究的主題，是那些能夠被當成自變項來加以控制與操弄的主題。最適合進行調查研究的主題，則是那些我們去提出相關的問題以瞭解人們自我回報的態度或是行爲的主題。最適合進行內容分析的主題，則是牽涉到文化傳播或者是溝通中的訊息內容的主題。至於最適合進行現存的統計數據研究的主題，則牽涉到由大型官僚組織所蒐集的資訊。公共或是私人組織會系統性地蒐集許多類型的資訊，以作爲政策決策時的參考之用、或者是當成一種公共服務，但是這些組織幾乎不可能爲了一些與特定研究問題直接相關的目的而去蒐集資料。因此，最適合使用現存的統計數據研究的狀況，是去考驗一些牽涉到社會、經濟、以及政治情況的正式報告中的對於變項的假設。這些正式報告中也包括了對於組織以及組織內成員的描述。一般而言，組織都會在一段很長的時間內去蒐集資訊。例如，你可以利用現存的統計數據，來瞭解在二十年的時間之內，150 個城市中的失業率以及犯罪率之間的關聯性。

在研究長時間以及橫跨不同國家的狀況時，現存的統計數據是非常有價值的。在第一章裡，你已經讀過了麥可貝以及索柏列斯基（McVeigh and Sobolewski, 2007）對於紅色郡以及藍色郡的現存統計數據研究了。在現存統計數據的資料當中，人口普查（請參閱資訊箱 11.5 人口普查）是一種相當有價值的類型（請參閱範例 11.4 現存的人口普查統計數據以及二十世紀早期的歸化入籍狀況）。

社會指標

在 1960 年代時，有許多社會科學家對於政策決定者所能取得的資訊感到不滿意，因此發起了「社會指標運動」（"social indicator's movement"）來測量社會福祉。這群社會科學家想要把關於社會福祉的資訊以及一般被廣泛使用的經濟表現指標〔例如：國民生產毛額（GNP）〕結合起來，以擴展人們對於社會世界的瞭解，同時也能讓制訂政策的官員們取得更多的資訊。這個運動的成員希望，社會生活品質的測量結果，能夠對公共政策的制訂產生影響。[11] 今日，有許多書籍、文章、以及報告都在討論社會指標。而《社會指標研究》（*Social Indicators Research*）這本學術期刊則是致力於社會指標的創造以及評估。自從 1976 年開始，美國人口普查局每隔三年就會出版一份報告，也就是《社會指標》（*Social Indicators*），而聯合國也會去蒐集不同國家裡的許多社會福祉的測量結果。

資訊箱 11.6　人口普查

幾乎每一個國家都會進行人口普查，或者是定期地去計算國家的所有人口。例如，澳洲自從 1881 年以來就一直在進行人口普查，加拿大是從 1871 年開始的，而美國則是從 1790 年就開始進行人口普查至今。大多數的國家每隔 5 年或是 10 年就會進行一次人口普查。除了人口的總數之外，人口普查的官員也會蒐集其他主題的資訊，像是住宅供給情況、種族、宗教信仰、以及教育情況等等。

人口普查是一個高品質的現存統計資料的來源，但是此種資料卻可能會引發爭議。在加拿大，官員們曾經試圖去計算居住在一起的同性伴侶的數目，然而此舉卻引起了關於政府是否應該去記錄社會變遷的公共爭論。在英國於 2001 年進行人口普查時，信奉回教的少數族群相當歡迎被問到關於宗教信仰的問題，因為他們覺得他們一直以來都被官方所忽視。而在美國，關於種族以及民族的調查測量方式已經引起了激烈的爭論，因此，自從 2000 年的人口普查開始，人們已經可以把自己歸類為屬於多重種族的類別。

美國於 2000 年所進行的人口普查也造成了一個嚴重的公共爭議，因為該次人口普查遺漏了數千位民眾，其中大多數人是來自於新近移民以及少數民族的成員們高度集中的低收入地區。另外，也發生了一些重覆計算居住在高收入地區的人們的狀況，因為在這些地區當中，有許多較為富裕的居民擁有第二棟房子。在政客之間也有一些爭議性的辯論，其中有些人主張利用科學抽樣方法以及調整人口普查的進行方式，以終止上述計算錯誤的狀況。然而，事實證明，許多政客們所真正關心的，並不是改善人口普查在科學方面的正確性，而是要保留傳統的人口普查方式來增加他們自己的政治利益、或者是幫助選區內的選民，因為美國政府在劃分區域以及分配公共經費給各地區時所依據的資料，正是人口普查資料。

所謂的**社會指標**（social indicator），所指的是任何與社會福祉有關的量數，而且能夠爲那些制訂政策的官員們提供各種相關的資訊。許多特定的指標都能夠去測量與下列領域有關的福祉：人口、家庭、住宅、社會安全與福利、衛生與營養、公共安全、教育與訓練、工作、收入、文化與休閒、社會流動性、投票、以及對社會與宗教組織的參與情況等等。美國聯邦調查局（FBI）製作的統一犯罪指數（uniform crime index），所指出的就是美國社會中的犯罪數量。社會指標能夠測量社會生活的負面資料，像是嬰兒的死亡率（也就是剛出生的嬰兒在一年之內死亡的比率）以及酗酒問題。社會指標也能夠去測量正面的資料，像是工作滿意度、或者是室內有供水管線的家戶比例等等。社會指標通常會牽涉到一些隱而不顯的價值判斷（例如，哪一些犯罪行爲比較嚴重、或者是構成良好生活品質的條件等等）。

美國的瓦薩爾學院（Vassar College）的社會政策創新研究所（Institute for Innovation in Social Policy）曾經爲美國製作了一個社會福祉指數（Index of Social Well-Being），該指數在各種現存的美國政府統計資料文件中，挑選出十六個社會問題領域（請參閱圖表 11.1）的量數，並且把這些量數結合起來。利用這個指數，你就可以把某個項目在每一年的情況與記錄上最佳的程度進行比較，其

社會指標　一種測量社會福祉的量化指標。

範例　11.4

現存的人口普查統計數據以及二十世紀早期的歸化入籍情況

布羅伊姆拉德（Bloemraad, 2006）曾經研究過現存的統計數據記錄，以檢視美國在二十世紀早期時移民取得公民權或者是歸化入籍的情況。她注意到許多評論家認為目前美國移民取得公民權的機會變低，與他們假定的 80 年到 100 年前歐洲移民快速而統一的歸化入籍情況形成強烈的對比。然而，由於幾乎沒有此種關於早期歸化入籍情況的具體證據，因此其中的一些空缺部分就被許多迷思所填補了。布羅伊姆拉德檢視了美國政府在 1900 年、1910 年、以及 1920 年的人口普查中關於成年男性移民的資料。這些人口普查資料是由美國政府所蒐集的，而且美國政府也會將這些資料公諸大眾，以利研究者進行統計分析。歸化入籍是一種法律程序，能夠使原本不具公民身份的人成為某個國家的公民。此種程序會要求申請者必須在某個國家內居住一段特定的時間長度、沒有任何犯罪記錄、通過語言能力檢定、以及具備某些其他的特質。在 1900 年到 1920 年之間，由於來自於歐洲的大規模移民潮，擁有美國公民權的移民比例，從 67%下降到 49%。布羅伊姆拉德運用了複雜的統計分析技術來調查影響歸化入籍情況的四種解釋：個人的資源以及技能、取得公民權時在法規以及官僚方面所會遇到的障礙、擁有公民權的相對成本以及利益、以及針對移民而發起的政治動員程度。在人口普查資料當中已經有關於居住年數、年齡、讀寫能力、以及英語能力的資料。布羅伊姆拉德發現，不同地理區域的歸化入籍比率有著極大的差異。造成此種差異的關鍵因素，正是當地居民對於新近移民所展現出的歡迎以及溫暖程度。雖然有一些地區性的歷史暗示了此種過程，但是一直到布羅伊姆拉德進行了大規模的全國性研究、並且以統計方式來分析數百萬個個案之前，我們都沒有任何全國性的、可以類推的研究發現。在她的研究中最為引人注目的結果，就是移民的居住地點對於歸化入籍的情況有著最為顯著的影響，而且其影響程度超過了移民的出生地、說英語的能力、或者是讀寫能力。簡而言之，真正決定移民能夠以多快的速度成為公民的關鍵因素，是當地居民對於移民的接納程度，而不是像在過往的研究中或者是政治評論家所強調的那些移民的個人特質。

量表的範圍是從 0 到 100，其中 100 是最高分，而且該指數是從 1970 年開始製作公布的。根據這個指數，美國在 1973 年時，社會福祉達到最高的程度（指數分數爲 77.5），但是自從那個時候開始就每下愈況。整體而言，在 1970 年到 2007 年之間，該指數一路從 66 下滑到 56。當前的社會福祉指數低於近來幾年的分數，而且每一個州的分數差異也非常地大。[12]

找出資料

政府或是國際機構、以及私人來源，是現存統計數據的主要提供者。如果你計畫要進行一項現存統計數據研究，比較明智的作法是先和一位資訊專家討論一下你感興趣的事物－在目前的案例中，你可以去找圖書館員討論，因爲他可以指出可能的資料來源。大多數現存的文件都是「免費的」，換句話說，你可以在公共圖書館裡輕易地取得這些文件，然而，在尋找特定的資訊時所必須花費的時間和努力，卻可能是相當龐大的。進行現存統計數據研究的研究者會花費許多時間在圖書館裡或是在網路上進行搜尋的工作。在他們找到資料之後，他們會把資料存在電腦檔案裡或者是記錄單上，以方便之後進行分析。通常這些資料都已經有電子格式可以讓我們

圖表 11.1 美國的社會健康指數

用來編製社會健康指數的十六種社會指標

嬰兒死亡率
兒童貧窮
虐待兒童
青少年自殺
青少年吸毒情況
高中畢業率
失業
平均薪資
健康保險覆蓋率
老化：年長者之間的貧困狀況
年長者之間的自殺情況
殺人
與酗酒有關的交通死亡事故
食物券涵蓋情況
可負擔的住宅
收入不平等的狀況

2008 美國各州的社會健康分數

排名	社會健康分數
1. 明尼蘇達州	75.5
2. 愛荷華州	71.1
3. 新罕布夏州	67.2
4. 內布拉斯加州	67.0
5. 夏威夷州	63.1
6. 佛特蒙州	62.7
7. 康乃狄克州	61.2
8. 北達科塔州	61.1
9. 猶他州	60.6
10. 紐澤西州	59.9
11. 愛達荷州	59.7
12. 維吉尼亞州	59.7
13. 賓夕法尼亞州	58.6
14. 緬因州	57.4
15. 印第安那州	55.9
16. 堪薩斯州	55.9
17. 德拉威州	55.7
18. 伊利諾州	55.2
19. 威斯康辛州	55.2
20. 馬里蘭州	54.9
21. 南達科塔州	54.4
22. 俄亥俄州	53.8
23. 懷俄明州	53.4
24. 麻薩諸塞州	53.1
25. 華盛頓州	52.2
26. 密蘇里州	51.4
27. 密西根州	48.9
28. 奧勒岡州	47.8
29. 羅德島州	46.8
30. 科羅拉多州	44.6

31. 紐約州	43.9
32. 喬治亞州	43.7
33. 阿拉斯加州	43.6
34. 內華達州	42.6
35. 加州	41.7
36. 西維吉尼亞州	40.8
37. 奧克拉荷馬州	40.1
38. 蒙大拿州	39.4
39. 阿拉巴馬州	38.8
40. 南卡羅萊那州	38.0
41. 德州	37.8
42. 路易斯安那州	37.5
43. 阿肯色州	36.4
44. 肯塔基州	36.2
45. 田納西州	35.5
46. 佛羅里達州	34.3
47. 北卡羅萊那州	33.4
48. 亞歷桑那州	32.8
49. 密西西比州	31.0
50. 新墨西哥州	26.8

資料來源：取自於 http://iisp.vassar.edu/socialhealth08.html. Institute for Innovation in Social Policy. Vassar College. 轉載已經過同意。

取得。例如，研究者現在已經不必去辛苦地記錄在參考資料中的投票記錄了，相反地，研究者可以去使用密西根大學的社會科學資料檔案庫（之後將進行討論）來尋找相關的投票記錄。除此之外，也請參閱資訊箱 11.6 將報紙的報導當作資料來源。

研究者能夠以非常具有創意的方式來運用現存的統計數據（請參閱範例 11.5 現存的統計數據、雙重性別的名字、以及集體行爲）。雖然目前有許多可以取得的資料來源，在此我只打算討論其中的一小部分。在關於美國的統計資訊來源中，其中一個最具有價值的來源，就是《美國統計摘要》（*The Statistical Abstract of the United States*）。自從 1878 年以來，這份摘要每年都出版一次（只有少數的例外情況），而且在所有的圖書館以及網路上都可以取得這份資料。這份資料是從美國政府機構所蒐集的許多報告以及統計資料表中加以挑選並且彙編而成的。由於這份資料中有來自於數百個政府以及私人機構的 1,400 個圖表、表格、以及統計清單，因此，除非你花時間去仔細瀏覽其中許多表格，否則你將會很難去掌握其中所包含的全部內容以及資訊。

大多數的政府都出版類似的統計資訊，像是《澳洲年報》（*Yearbook Australia*）、《加拿大年報》（*Canada Yearbook*）、《紐西蘭官方年報》（*New Zealand Official Yearbook*），以及在英國所出版的《年度統計摘要》（*Annual Abstract of Statistics*）等等。[13] 許多其他的國家也都會出版歷史統計資料的書籍。

光是找出這些政府出版的統計資料文件，本身就是一門藝術，而且某些刊物存在的唯一目的，就是爲了要協助研究者，像是《美國統計索引：美國政府統計刊物

資訊箱 11.6　將報紙的報導當作資料來源

許多社會研究者會將報紙中的報導當作是一個資料來源，不只是去分析這些文章的內容，而且也當作是一種辨識以及計算像是社會抗議之類重大事件的方式。即使報紙並未報導所有的事件（也就是選擇偏誤）、或者是並未涵蓋所報導事件的所有資訊（也就是描述偏誤），報紙仍然可以算是一種公共資訊的重要來源，其重要性幾乎無可取代。除此之外，這些類型的偏誤也有可能因為地理區域或者是歷史時期的不同而有所差異。雖然大型的報紙都會有主題索引，但是這些索引的組織方式對於社會研究的目的而言，可能並不總是那麼有用。特別是在允許言論以及出版自由的國家裡，報紙可以當作一種測量在一段長時間裡的社會事件的方式，尤其是「對於許多歷史以及比較研究設計而言，報紙仍然是抗議事件相關資料的唯一來源」（Earl et al., 2004: 76）。

綜合指引與索引》（*American Statistics Index: A Comprehensive Guide and Index to the Statistical Publications of the U.S. Government*）以及《統計數據來源：美國與國際工業、商業、社會教育、金融、以及其他主題資料之主題式指引》（*Statistics Sources: A Subject Guide to Data on Industrial, Business, Social Education, Financial and Other Topics for the United States and Internationally*）等等。[14] 聯合國以及像是世界銀行之類的國際機構也都有自己出版品，其中有涵蓋了許多不同國家的統計資訊（例如：識字率、農業勞動人口比例、出生率等等）的出版品，例如，《人口統計年報》（*Demographic Yearbook*）、《聯合國教科文組織統計年報》（*UNESCO Statistical Yearbook*）、以及《聯合國統計年報》（*United Nations Statistical Yearbook*）等等。

其他的出版刊物則提供了特殊主題的資料來源。例如，有一些出版品中包含了一些知名人士的社會背景、職業、以及其他的傳記資訊，而這些知名人士都是根據某些標準而被認定爲重要的人物。這些出版品的內容都是來自於這些被認爲重要的人士所自願提供的資訊。另外一個資訊的來源則涵蓋了企業或者是企業的主管。[15] 最後，還有一些出版品則是專門蒐集關於政治、投票、以及政治人物的資訊（關於涵蓋美國的來源出版品，請參閱資訊箱 11.7 提供社會資料的特殊出版品）。

調查資料的次級分析

次級分析是現存統計數據的一種特殊個案，這是以統計的方式來分析原本由某人所蒐集的調查資料，而不是以初級研究（例如，實驗、調查、以及內容分析）的方式來蒐集量化資料。在過去的二十年間，由於有越來越多的資料可以取得，因此也有越來越多的社會科學家進行了次級分析。次級分析相對而言成本較爲低廉；允許研究者橫跨不同的團體、國家、或者時間來進行比較；而且也允許研究者提出一些原本的研究者所沒有考慮到的議題，像是利用青少年的健康調查資料來研究宗教信仰等等（請參閱範例 11.6 次級資料分析：以舊資料來回答新問題）。

範例　11.5

現存的統計數據、雙重性別的名字、以及集體行為

所謂雙重性別的名字（androgynous first name），是一個可能是指女孩或是男孩的名字，而且不會清楚地標示出某個孩子的性別。有些人宣稱，女性主義運動降低了在兒童名字當中的性別標記，以作為該運動在減少性別區隔以及性別不平等的社會影響方面的一部分。其他人則觀察到，在大多數的社會裡，性別仍然是在命名時唯一最具有主導力量的特質。即使當種族團體或是社會階級發明了一些特殊的新名字，性別的區隔仍然被保留了下來。

李柏森等人（Lieberson et al., 2000）檢視了以電腦記錄儲存的現存統計資料，這些資料是來自於伊利諾州在 1916 年到 1989 年之間的一千一百萬名白人兒童的出生證明。他們發現雙重性別的名字是相當罕見的（大約只有 3%），而且有一種稍微傾向於使用雙重性別名字的歷史趨勢，不過該趨勢是在最近幾年才出現的。除此之外，家長為女孩子取雙重性別名字的情況比男孩多，而且在命名時的性別區隔也不穩定（換句話說，某個名字可能會隨著時間流逝而失去其雙重性別的意義）。作者們也注意到，家長為孩子命名的方式，與一種集體行為的模式相當類似，而此種集體行為的模式則已經被發現運作於另外一個研究領域之中：社區的種族隔離。在不同種族之間，更換居住地的情況是不平等的，其中佔有優勢的團體比較不常搬家；較不具力量的團體會遷移到一些被優勢團體所放棄的地區；而且由於新的隔離情況在一段時間之後就會再度出現，因此種族之間的融合情況也是不穩定的。

在進行大模規的調查時，蒐集資料的過程可能會非常昂貴而且難以進行。對於大多數的研究者而言，由於進行使用嚴謹技術的大型全國性調查所需要的成本和時間都高到嚇人，因此他們根本不可能去進行此種調查。幸運的是，大型調查資料的組織、保存、以及傳播方式已經有了長足的進步。時至今日，許多過去的調查的檔案資料都已經開放給研究者去查詢使用了。

位於密西根大學的「大學校際政治及社會研究聯合會」（Inter-University Consortium for Political and Social Research, ICPSR）是全世界社會科學資料最重要的檔案資料庫。其中儲存了超過 17,000 份調查研究以及其相關的資訊，而且研究者只需支付一些費用，就可以取得這些資訊。在美國以及其他國家裡的各種不同的資料中心，也都擁有一些調查資料。[16]

關於美國的調查資料，最被廣泛使用的資料來源是綜合社會調查（General Social Survey, GSS）。自從 1973 年以來，位於芝加哥大學的全國意見調查中心每隔兩年就會進行一次此種社會調查。而在近幾年來，該項調查的涵蓋範圍也包括其他的國家。研究者只需要支付相當低廉的價格，就能取得這些調查資料，進行次級資料分析[17]（請參閱資訊箱 11.8 綜合社會調查）。

限制

雖然次級資料分析以及現存統計數據研究不斷持續成長而且廣受歡迎，此種類型的研究仍然有一些限制。我們不能單純地以爲由於資料是由政府機構或者是研究組織所蒐集的，所以使用此種技術就完全沒有任何的問題。

資訊箱 11.7 提供社會資料的特殊出版品

關於知名人物的出版資訊來源

《美國名人錄》(*Who's Who is American*)是一個廣受歡迎的傳記資料來源，而且從 1908 年開始就一直出版至今。在該書中會列出被收錄者的姓名、出生日期、職業、得過的獎項、出版品、會員資格、教育、擔任過的職位、配偶、以及孩子的名字。除此之外，該書也會針對美國的特定地區（例如，《美東地區名人錄》(*Who's Who in the East*)）、特定職業（例如，《金融與企業名人錄》(*Who's Who in Finance and Industry*)）、以及特定的次團體（例如：女性、猶太人、黑人等等）出版專門的特殊版本。

《美國傳記字典》(*Dictionary of American Biography*）所列出的人物比《名人錄》(*Who's Who*)來得少，但是卻提供更為詳盡的細節。這份出版品是從 1928 年開始出版的，而且會不定時地出版一些增刊或附錄以補充最新的資訊。例如，在附錄七中列出了 572 個人，而且每個人的資料都佔了大約一頁的篇幅，其中包括了關於職業、旅遊經歷、出版品標題、以及與其他知名人士之間的關係等等項目的細節。

《傳記字典索引》(*Biographical Dictionaries Master Index*）是一種索引，裡面列出了在各種《名人錄》出版品以及許多其他傳記資料來源（例如：《曲棍球名人錄》(*Who's Who in Hockey*)）中的姓名。如果研究者已經知道某個人物的姓名的話，這本索引就會告訴研究者可以到哪兒去找到關於此人的傳記資訊。

企業以及公司的資料來源

《鄧恩以及布瑞德史崔特主要工商企業名錄》(*Dun and Bradstreet Principal Industrial Businesses*)是一份指引，裡面收錄了在 135 個國家裡將近 51,000 個企業的資訊，其中包括了銷售額、員工人數、高級主管、以及產品等等的資訊。

《關係企業名錄》(*Who Owns Whom*）共有好幾冊，內容涵蓋了數個國家或是地區（例如，北美地區、英國、愛爾蘭、以及澳洲等等），其中會列出母公司、子公司、以及相關的公司等等資訊。

《標準普爾企業、董事、與經理人名錄》(*Standard and Poor's Register of Corporations, Directors and Executives*）列出了大約 37,000 家美國以及加拿大的公司，其中的資訊包括了企業名稱、產品、經理管理階層、產業、以及銷售額等等。

政治議題的資料來源（美國）

《美國政治年鑑》(*Almanac of American Politics*）是一份雙年刊，其中包括了美國政府官員的相片以及簡短的傳記資料。該份出版品列出了各種委員會的開會記錄、投票記錄、以及其他類似的資訊，以提供給國會中的議員以及在行政部門中的官員當作參考。

《美國選舉：當代美國選舉統計數據手冊》(*America Votes: A Handbook of Contemporary American Election Statistics*）包含了幾乎所有美國地方性以及全國性選舉中各郡的投票資訊與記錄。甚至連各郡的初選結果也都包含在內。

《美國政治重大統計數據》(*Vital Statistics on American Politics*）提供了數十種關於政治行為的表格，其中包括了每位競選國會議員候選人的在進行選戰時的各種支出、他們在初選以及最後選舉中的結果、各種不同的政治組織的意識型態評分、以及各州選民登記規定的摘要等等。

其中的危險之一，是次級資料或者是現存的統計數據可能並不適合你的研究問題。在進行研究之前，你必須考量到資料中的單位（例如：人們或者是組織的類型

範例　11.6

次級資料分析：以舊資料來回答新問題

為了進行次級資料分析，研究者可以使用已經蒐集完畢的調查資料，來處理一些與調查原本的目的不相關的新研究問題。烏埃克爾等人（Uecker et al., 2007）利用來自於一份健康調查的資料來研究宗教。這份資料來自於「美國青少年健康長期研究」（National Longitudinal Study of Adolescent Health），而這項研究原本是一項針對健康以及相關的社會行為、並且以學校為基礎的三階段同樣本多次研究（你已經在第二章當中學過同樣本多次研究（panel study）的相關知識了）。作者們使用了該項調查中在第一個階段所蒐集的資料，對象是 1994 年到 1995 年之間在美國 132 所學校中介於 7 年級到 12 年級的 20,745 學生所組成的隨機樣本；並且也使用了該項調查中第三個階段的資料，也就是在 2001 年到 2002 年之間針對 15,197 名原本的受訪者（當時的年齡約在 18 歲到 25 歲之間）所進行的訪談資料。作者們所真正感興趣的，是要去解釋當青少年從青春期進入到成年期時，他們在宗教參與程度方面下降的情況，而不是健康方面的問題。然而，該項健康調查中也包含了一些與宗教有關的問題，像是宗教參與的程度、宗教在個人生活中的重要性、以及對於大型宗教的感覺等等。作者們利用這三個問題當作是他們的依變項量數。過去的研究已經把年輕人之間宗教參與度下降的情況，解釋為肇因於就讀大學所產生的世俗化效果。在以統計方式分析了資料之後，作者們發現，念大學的青少年和沒有念大學的青少年一樣地虔誠。作者們也發現到，真正使得宗教在年輕人生活中的重要性下降的原因，並不是他們有沒有上大學，而是因為年輕人之間的同居、婚前性行為、以及吸毒或喝酒等等情況所造成的，而且，與人們過去的想法相反的是，高等教育本身對於宗教信仰的影響可以說幾乎是微乎其微。

等等）、蒐集資料的時間與地點、使用的抽樣方法、以及在像是人口普查之類的資料中所涵蓋的特定議題或者是主題。例如，你想要檢視美國境內存在於拉丁美洲人以及白人之間的種族緊張關係，但是你手上僅有的次級資料卻只包含了太平洋西北部地區以及新英格蘭地區而已。在此種情況之下，你應該要重新考慮你的研究問題或者是運用資料的方式，否則就無法運用這些資料來進行研究。

第二個危險之處，則是你必須瞭解真正實質的主題，才能去使用這些資料。由於你可以輕易地取得這些資料，你可能會擁有許多資料，但是對於某個主題卻所知甚少。如此一來，你就有可能作出錯誤的假定、或者是對結果作出錯誤的詮釋。在使用任何資料之前，你都應該先去研究實質主題的細節。例如，如果你使用了關於德國的中學畢業率的資料，但是你卻不太瞭解德國的中學教育系統以及該系統中特殊的學術以及職業的雙軌制，並且假定德國的教育體制和美國的教育體制是完全一樣的，那麼你就有可能在詮釋結果時犯下嚴重的錯誤。

第三個危險則是以過度詳盡的方式來引用統計數據，以給別人一種具有科學嚴謹性的印象，而此種情況有可能會導致**具體性錯置的謬誤**（fallacy of misplaced concreteness）。當某個人引用了過度詳盡的統計數據、並且使細節「超載」（"overloading"），以給人一種具有精確性的錯誤印象時，就是發生了具體性錯置的謬誤。例如，根據現存的統計數據，澳洲的人口數是 19,169,083，但是，比較好的說法是說澳

具體性錯置的謬誤　在一項量化的量數中使用了過多的數字，以試圖使讀者產生資料極為正確的（錯誤）印象。

資訊箱 11.8　綜合社會調查

在社會研究者用來進行次級分析的調查資料當中，綜合社會調查（General Social Survey, GSS）是最有名的。綜合社會調查的任務是「提供及時的、高品質的、具科學相關性的資料給社會科學研究社群使用」（Davis and Smith, 1992: 1）。這份資料以電腦能夠讀取的許多格式來儲存，而且只要花費些許費用就能在許多地方取用這份資料。不論是原始資料庫或者是編碼資料都是不具版權的，因此，使用者可以在不必取得授權的情況下去複製或者是散佈這些資料。你可以在超過 2,000 份研究文章以及書籍中找到運用綜合社會調查所得到的研究結果。

自從 1972 年以來，美國全國意見研究中心（National Opinion Research Center, NORC）幾乎每年都會進行一次綜合社會調查。在每年的典型調查中，會包含由大約 1,500 名美國成年居民所組成的隨機樣本。一組研究者會挑選出一些問題，而其他的個別研究者也可以建議一些問題。該中心每年會重覆一些問題以及主題，有些問題則是每隔四到六年就會重覆出現一次，而在特定年度則會再加入一些其他的主題。例如，在 1988 年時的特殊主題是宗教，而在 1990 年時的特殊主題則是族群關係。

訪談者利用面對面訪談的方式來蒐集資料。全國意見研究中心裡的工作人員則謹慎地挑選訪談者，並且為這些訪談者進行社會科學方法論以及調查訪談方面的訓練。綜合社會調查每年大約會雇用 120 到 140 名訪談者，其中差不多有 95%的訪談者是女性，而且大多數都是中年人。全國意見研究中心也會雇用具有雙語能力以及來自於少數族群的訪談者。訪談者會與受訪者的種族進行配對。訪談一般會進行大約 90 分鐘左右，並且包含了將近 500 個問題。一直以來，回應率大約都在 71%到 79%之間。而未回應的主要原因，則是有些受訪者會拒絕參與調查。

至於國際社會調查計畫（International Social Survey Program）則在其他國家進行類似的調查。一開始時只有德國的全國大調查（German ALLBUS）以及英國的社會態度調查（Social Attitudes Survey）加入，之後則有越來越多國家參與，其中包括了澳洲、奧地利、義大利、匈牙利、愛爾蘭、以色列、荷蘭、瑞士、波蘭、以及其他國家。該計畫的目標是要定期地進行大型的全國綜合調查，並且在合作的國家中以跨國的方式去詢問一些共同的問題。

洲的人口略多於一千九百萬。在針對 2000 年的綜合社會調查進行了次級資料分析之後，你可能計算出離婚人口的比率是 15.65495%，但是，比較好的作法是報告離婚人口大約佔了 15.7%。[18]

分析單位以及變項屬性　在現存的統計數據中的一個常見的問題，就是要找出適合的分析單位。許多統計數據所刊載的都是整體的數據，而不是個別的數據。例如，某份政府文件的一個表格裡所包含的是某個州的整體資訊（例如，失業率、犯罪率），但是研究問題的分析單位卻是個人〔例如：「失業的人是否比較有可能犯下財產犯罪（property crimes）？」〕。在此種情況下，研究者很有可能會犯下生態謬誤。不過，對於次級調查分析而言，這個問題比較不是那麼嚴重，因為我們可以從檔案資料中去取得每個受訪者的原始資料。

另外一個相關的問題，則牽涉到在現存的文件或是調查問題中所使用的變項屬性類別。如果當初蒐集原始資料的組織使用了許多極爲精確的類別時，這就不會有太多問題。然而，當組織是以廣泛的類別來蒐集原始資料、或者是這些類別並不符合目前研究的需求時，就會產生問題了。例如，你對於有亞洲血統的人很感興趣。如果某份文件中的種族類別是白人、黑人、以及其他時，你就會遇上問題，因爲在「其他」這個類別裡包括了亞洲人以及其他血統的人。有些時候，雖然組織是以較爲精細的類別來蒐集資訊，但是在出版時卻只使用了較爲廣泛的類別。因此，你必須更加深入地去探究，才能發現該組織是否蒐集了較爲精細的資訊。

效度　當你的理論定義並不符合蒐集資訊的政府機構或是組織所下的理論定義時，就會產生效度方面的問題。在官方政策以及程序當中，將會詳細說明官方統計數據的定義。例如，你把工作傷害定義爲包括在工作中所發生的輕微割傷、瘀血、以及扭傷，但是在政府報告的官方定義中，卻只有包括必須到醫院去就診的重大傷害。如此一來，你所定義的許多工作傷害將不會被包含在官方統計數據當中。另外一個例子，則是你把失業者定義爲包括那些如果能夠找到好工作就願意去工作的人、那些想要得到全職工作但是卻只能找到兼職工作的人、以及那些已經放棄找工作的人，然而，在官方的定義中，所謂的失業者只包含了那些積極尋找工作（不論是全職或者是兼職）的人而已。因此，官方的統計數據就會排除了一些你定義爲失業的人。在上述的兩個例子裡，你的定義都與官方統計數據裡的定義有著一些差異性。

另外一個會發生效度問題的狀況，則是當你把官方統計數據當作是某個構念的代理時。由於你無法蒐集原始的資料，因此這種作法是必要的。例如，你想要知道有多少人是憎恨罪（hate crimes）的受害者，因此你利用警方的統計數據當作代理，但是這個量數並不是全然有效的。這是因爲許多受害者並未向警方報案說他們自己是憎恨罪的受害者，而且在官方的報告中也不一定總是會揭露所有事件的細節（請參閱資訊箱 11.9 憎恨罪的官方統計數據，以及在精確性方面的緩慢進步）。

或許你想要測量因爲婚前懷孕而「被迫」結婚的情況。你可以利用官方紀錄中的結婚日期以及孩子的出生日期，來估計是否發生了「奉子成婚」的情況。不過，這並不能告訴你懷孕是否就是結婚的動機。某對夫妻可能早就計畫要結婚了，因此婚前懷孕與否是無關緊要的，或者，也有可能這對夫妻在結婚時根本不知道太太已經懷孕了。相同地，某些夫妻在結婚之後沒有孩子的出生紀錄，可能是因爲對於懷孕的錯誤信念而被迫拿掉孩子，或者是雖然有懷孕，但是孩子卻因爲流產而沒有出生記錄。除此之外，某個孩子也有可能是在結婚之後才懷上的，但是後來卻是一個比預產期提前很多時間出生的早產兒。如果你把奉子成婚定義爲孩子是在結婚之後的九個月之內出生的，那麼就有可能發生認定錯誤的情況，因而降低了研究的效度。

發生第三種效度問題的情況，則是來自於你對於資訊的蒐集方式缺乏控制力。出現在官方政府報告中的資訊，是由在官僚機構裡工作的普通人所蒐集的，因此，你必須依賴這些人正確地去蒐集、組織、報告、以及出版資料。系統性誤差可能發生在蒐集原始資訊的過程中（例如，進行民意調查的工作人員刻意避開貧窮的社區

資訊箱 11.9　憎恨罪的官方統計數據，以及在精確性方面的緩慢進步

在各種現存的統計數據類型中，政府對於犯罪的統計數據是經常在社會研究中被使用的許多類型之一。在美國，從 1920 年就開始運作的「統一犯罪報告」（Uniform Crime Reporting, UCR）計畫，是最常被使用的美國國家犯罪統計數據來源。每一個州以及地區的執法機構，都會把大多數主要犯罪行為的統計數據（也就是向警方報案的犯罪行為、以及警方所逮捕到的罪犯），寄送到聯邦調查局（Federal Bureaus of Investigation）。在 1973 年時，美國司法部增加了第二個犯罪資料的來源，也就是「全國犯罪受害調查」（National Crime Victimization Survey, NCVS）。這項調查是一項每年進行一次的調查，而且其中所使用的是 49,000 個家戶所組成的代表性樣本。這項調查的進行方式，是去詢問家戶成員是否有任何年滿 12 歲以上的人曾經是犯罪的受害者。在 1980 年代時，一個新的計畫則補足了統一犯罪報告中的簡單犯罪行為的統計數據。「全國案件報案系統」（National Incident-Based Reporting System, NIBRS）包括了關於犯罪情況（例如，犯罪地點、牽連在內的人、以及犯罪時間等等）的許多細節，而這個系統目前涵蓋了大約 15%的美國人口，但是正在緩慢地擴增當中。

在 1990 年時，美國施行了一條新的憎恨罪法令，其中規定，凡是因為受害者的種族、宗教信仰、國籍、或者是性傾向而對其產生偏見或是歧視者，就算是犯下憎恨罪，而且必須加重刑罰。聯邦調查局會在其年度報告《憎恨罪統計數據》（Hate Crime Statistics）中刊載此種犯罪行為的摘要，而這也是此種犯罪行為的全國資訊的主要來源。回報此種新犯罪行為的精確性，可以說明官方統計數據的一些複雜性。

在執行了美國的憎恨罪法令之後，聯邦調查局訓練了地方性的執法官員應該如何去執行此項新的法令。在 1991 年時，只有 29%的美國人口擁有一個參與此種訓練計畫的執法機構。而到了 1999 年時，這個比率已經上升到了 85%。然而，大多數參與訓練計畫的執法機構，每一年所回報的憎恨罪案件數量都是零。自從 1994 年開始，大約有 85%受過訓練的執法機構回報的憎恨罪案件數量為零。這些案件數量為零的報告，可能並不完全正確。有一項針對憎恨罪案件數量為零的地區所進行的研究，發現其中有 37%的地區其實曾經發生過憎恨罪的案件，但是在報案系統中發生了一些失靈的情況。大型的地區報案系統之間也存在著差異。美國南方各州的一般犯罪率是最高的，但是憎恨罪的案件數量卻是最少的。有些研究認為，各地區對於憎恨罪的態度，會影響了受害者向警方報案的意願。除此之外，並不是所有的地方性警察機構，都會以同樣認真的態度來面對犯下憎恨罪的情況。從 2000 年開始，全國受害調查已經開始增加了憎恨罪的相關問題。

總而言之，在憎恨罪開始執行之後，已經過了十幾年，但是關於此種犯罪的資料仍然相當有限。在不同的地理區域以及警察機構之間的回報正確性也不盡相同。蒐集此種犯罪資料能夠讓我們對於憎恨罪的數量以及類型有著基本的瞭解，但是資料蒐集過程的進步速度非常緩慢。想要認真研究憎恨罪的研究者，必須結合官方的全國報告資料以及其他的資訊來源，因為這些資訊來源包含了較多的細節，而且也把焦點集中在地方性的地理區域上面（請參閱 Nolan, Akiyama, and Berhanu, 2002）。

以及編造資訊，或者是人們在駕照上謊報年齡）、在組織以及報告資訊的過程中（例如，某個警察部門在把犯罪報告歸檔時過於草率，因此遺漏了某些報告）、以及在出版資訊的過程中（例如，在某個表格中發生了排版錯誤），而這些系統性誤差則都會降低測量效度。

這樣的問題，就發生在關於美國永久性失業人數的統計數據上。一位大學的研究者重新檢視了美國勞工統計局（U.S. Bureau of Labor Statistics）用來蒐集資料的方法，並且發現了一個錯誤。關於永久性失業的資料是來自於一項針對 50,000 名受訪者所進行的調查，但是這項調查的未回應率很高，而且政府機構並未針對此種狀況進行調整。在原本的調查結果中，從 1993 年到 1996 年之間的失業人口下降了 7%，然而，經過修正之後的數據顯示，失業人口的比例其實並沒有任何變化。[19]

信度　信度的問題可能會對現存的統計數據研究造成困擾。當官方的定義或者是蒐集資訊的方法隨著時間而發生變化時，就會出現信度的問題。政府每隔一段時間就會修改對於工作傷害、殘障、失業等等的官方定義。即使我們瞭解此種變化，要在長時間中進行一致的測量仍然是不可能的。舉例而言，在 1980 年代早期時，計算美國民眾失業率的方式就曾經有過變化。在 1980 年代之前，政府計算失業率的方式，是把失業的人口總數除以民間企業的員工總數。而修改過後的新計算方式，則是把失業的人口總數，除以民間企業的員工加上在軍隊裡工作的人們的總數。同樣地，當警察部門把他們的記錄電腦化時，民眾報案的犯罪數目似乎增加了，但是這純粹只是因爲保管記錄的方式進步所致。

等值信度可能也會是一個問題。例如，一個關於全國犯罪記錄的量數，必須依賴每個警察部門提供正確的資訊。如果在某個地區的警察部門在記錄資料時漫不經心、馬虎草率，那麼這個回報的量數就會失去了等值信度。相同地，對於警察部門所進行的研究顯示出，對於增加逮捕率的政治壓力，與進行逮捕的數量有著相當密切的關聯性。例如，在某個城市裡的政治壓力可能會增加逮捕的數量（例如，強力打擊犯罪），而在另外一個城市裡的政治壓力則可能會減少逮捕的數量（例如，在選舉過後表現出犯罪率下降的情況，以使官員們比較有面子）。

代表信度在官方政府統計數據當中也可能會是一個問題。例如，美國勞工統計局發現，在使用了性別中立的測量程序之後，女性的失業率稍微增加了 0.6%。一直到 1990 年代中期爲止，訪談者在詢問女性受訪者時，僅僅只有問她們是否一直在做家事或者是做其他的事情而已。研究者把回答「做家事」的女性受訪者歸類爲家庭主婦，而不是當作失業人口來看待，即使這些女性一直在找工作也一樣。然而，一旦訪談者向女性提出與男性相同的問題：「你正在工作中、或者是在做其他事情？」，許多女性會回答她們並非在工作中、而是在做「其他事情」，像是找工作之類的。上述這種情況顯示出，官員們在編製政府統計數據時的方法論細節的重要性。

官方統計數據允許進行國際間的比較，但是各國政府蒐集資料的方法各自不同，而且資料蒐集的品質也良莠不齊。例如，在 1994 年時，各國所報告的失業率如下：美國是 7%、日本是 2.9%、而法國則是 12%。如果這些國家是以完全相同的方式來下

定義以及蒐集資料，也就是把喪失士氣的工作者以及非自願性的兼職工作者的比率都包括在內的話，那麼美國的失業率將會是 9.3%、日本是 9.6%、而法國則是 13.7%。爲了要評估官方政府統計數據的品質，《經濟學人》（*The Economist*）雜誌邀請一個由 20 名統計學家所組成的團隊去評估十三個國家的統計數據品質，評估的基礎包括了免於政治干擾的自由程度、信度、統計方法論、以及主題的涵蓋範圍等等。品質最好的前五名依序是加拿大、澳洲、荷蘭、法國、以及瑞典，而美國則是與英國、德國並列第六名。美國官方統計數據的問題在於高度分權化、統計學家的數量比其他任何一個國家還少、而且蒐集資料的範圍曾經因爲政治的因素而被削減。

在國際間所蒐集的資料可能會具有爭議性。聯合國國際勞工組織（International Labor Organization of the United Nations）曾經報告過，在某些國家中，關於總體經濟活動的官方統計數據是不正確的，因爲這些數據排除了性工業的資料。在某些國家裡（尤其是泰國以及菲律賓），有數百萬名工作者（主要都是年輕的女性）是受到雇用的，而且有數十億美金的收入是來自於賣淫以及性工業。這些活動對於經濟造成很大的衝擊，但是卻沒有出現在任何的官方報告或者是統計數據裡。[20]

遺失資料 對於使用現存統計數據以及文件的研究者而言，另外一個困擾他們的問題，則是遺失資料的問題（我們已經在第七章中討論過這個問題）。有些時候，雖然已經蒐集到資料，但是後來卻遺失了。而更加常見的情況，則是某些資料根本從來就沒有被蒐集過。某些資料之所以遺失了，是因爲在政府機構裡的研究者以及官員決定不要去蒐集這些資訊。那些決定要蒐集什麼資料的人，可能沒有去蒐集一些後來的研究者在處理新問題時所可能需要的資料。政府機構可能會爲了政治的、預算的、或者是其他的理由而開始或是停止蒐集資訊。例如，在 1980 年代早期，美國聯邦政府由於實行了削減支出的措施，因而停止蒐集一些社會研究者認爲相當有價值的資訊。遺失資訊會造成嚴重的問題，尤其是當研究者所研究的事物涵蓋了相當長的一段時間的時候。舉例而言，某個對於美國的罷工事件數目感到興趣並且想要進行研究的研究者，可以取得自 1890 年代以來迄今的資料，但是其中在 1911 年之後，有 5 年的時間之內都沒有任何相關的資料，這是因爲當時聯邦政府並未蒐集此項資料所致。

推論的議題以及理論考驗

當你在根據非反應式資料來推論因果關係或者是考驗一個理論時，你必須要特別地小心謹慎。在使用非反應式以及非侵擾性量數時，我們很難去建立時間順序以及排除一些替代的解釋。在內容分析中，你無法從內容去推論出它對於那些讀到文本的人所會產生的影響，但是可以利用調查研究的相關性邏輯來顯示出變項之間的關聯性。

倫理的考量

在大部分的非反應式研究中，倫理的考量並不是最重要的事，因爲你所研究的人們並未直接參與你的研究。主要的倫理考量，是在使用由其他人所蒐集的資料時所牽涉到的隱私以及保密性。另外一個較大的倫理議題，則是官方的統計數據其實是一種社會以及政治的產物。有一些研究者或是官方機構是依據一些隱而不顯的理論以及價值假定來蒐集資料的。官方所提供的量數或者是統計數據有可能會成爲政治鬥爭的對象，以及一種把政策推向某個特定政治方向的方法。一旦政府機構把某個量數定義爲官方的，該量數就能夠去影響公共政策，並且會導致一些特定的結果，而如果當時官方所使用的是另外一個同樣有效的替代量數的話，則可能會產生完全不同的結果。例如，在 1930 年代的經濟大蕭條中的政治行動主義，就模擬了針對許多社會情況（例如，當住在公共精神病院裡的時候死亡的病患人數）所蒐集的資訊。在當時的政治行動主義之前，許多政府以及民眾並不認爲這些情況的重要性大到值得引起民眾的注意。同樣地，關於不同年齡的非白人學生就讀於美國學校的百分比率的資訊，一直到 1953 年之後才能取得，而對於各種非白人的種族而言，則是自從 1970 年代以後才能取得這項資訊。而在更早之前，這類的資訊在公共政策當中可以說是幾乎完全沒有的。

官方蒐集統計數據的作法，可以促使大眾把注意力轉移到某個議題上面，而社會大眾對於某個議題的關注，則能促使政府去蒐集新的官方統計數據。例如，一直到了政府機構開始持續公布酒醉駕車肇事的交通事故數量之後，酒醉駕車才成爲一個受到關注的公共議題。

政治以及社會的價值觀，也會影響到關於政府機構應該蒐集何種統計數據的決定。大多數官方統計數據的設計以及蒐集過程，都是配合由上而下的行政規畫目的，因此，這些資料可能不符合你的研究目的、或者是那些不認同官僚決策制訂者的思維模式的人的目的。例如，某個政府機構測量了鋼鐵生產量的總噸數、鋪設的道路公里數、以及在一個家戶裡的平均人口數。而關於其他情況的資訊，像是飲用水的品質、通勤上班所必須花費的時間、與工作有關的壓力、以及需要兒童照護措施的兒童數目等等，則沒有被蒐集，因爲政治官員認爲這些都是不重要的資訊。在許多國家裡，官員們都認爲國民生產毛額（GNP）是一個衡量社會進步情況的重要指標，但是 GNP 忽略了社會生活中的非經濟面向（例如，與兒童一起玩耍所花費的時間）以及沒有支領薪水的工作類型（例如：家事）。因此，可以取得的資訊其實所反映出的是政治論辯的結果，以及那些決定蒐集何種統計數據的官員們的價值觀。[21]

結論

在本章裡，你學到了數種類型的非反應式研究技術，這些技術都是在不影響研究對象的情況下，去測量或者是觀察社會生活各種面向的方法。這些技術能夠產生客觀

的、數值的資訊，而你則可以分析這些資訊，以處理一些研究問題。你也可以把這些技術和其他類型的量化或是質化社會研究結合在一起，以處理許多不同的問題。

就像任何其他形式的量化資料一樣，我們也必須去關心測量的議題。雖然我們可以很容易地從一項調查或者是政府文件中取得資訊，但是這非意味著這些資訊可以測量到我們真正感興趣的構念。

你應該要注意到在非反應式研究中的兩個潛在的問題。首先，現存資訊的可得性，會限制了我們所能處理的問題。其次，非反應式變項的效度通常不高，因爲這些變項並未真正地去測量我們所感興趣的構念。雖然現存的統計數據以及次級資料分析是低成本的研究技術，但是研究者對於資料蒐集的過程不但沒有控制力量，而且也缺乏實質的瞭解。此種偏誤的潛在來源，意味著研究者必須特別地戒慎恐懼以及謹慎小心。

在下一章裡，我們會從設計研究計畫以及蒐集資料的階段，進展到分析資料的階段。這些分析技術可以應用於在前面章節裡所提到的量化資料上面。到目前爲止，你已經看到如何從一個主題、進展到研究設計與測量、然後再到蒐集資料的階段。接下來，你將會學到如何去檢視資料，並且看看這些資料能夠告訴你什麼關於假設或者是研究問題的訊息。

問題回顧

1. 內容分析適合應用於哪種類型的研究問題？
2. 文本內容中的哪四種特質應該要被觀察並且記錄於編碼系統當中？
3. 使用現存統計資料的研究者應該要注意什麼樣的信度問題？
4. 次級資料分析的優點以及缺點爲何？
5. 爲什麼內容分析的研究者要使用好幾位不同的編碼者，而在這麼做的時候可能會有什麼問題？
6. 在內容分析中進行推論時有什麼限制？
7. 在內容分析中所使用的分析單位是什麼？
8. 在現存統計數據中的整體性問題是什麼？
9. 在內容分析中的三種效度問題爲何？
10. 研究者在使用現存的統計數據時應該要注意到哪些限制？

注解

1. See Webb and colleagues (1981:7–11).
2. For an inventory of nonreactive measures, see Bouchard (1976) and Webb et al. (1981).
3. See Krippendorff (1980:13). For definitions of content analysis, see Holsti (1968a:597), Krippendorff (1980:21–24), Markoff et al. (1974:5–6), Stone and Weber (1992), and Weber (1985:81, note 1).
4. Weitzman and colleagues (1972) is a classic in this type of research.

5. See Ariés (1977) for an example.

6. Examples of content analysis studies can be found in Berelson (1952), Carney (1972), McDiarmid (1971), Myers and Margavio (1983), Namenwirth (1970), Sepstrup (1981), Stewart (1984), and Stone et al. (1966). Harwood (2000) described elderly people on television. Also see Weber (1983) for a discussion of measurement issues in content analysis.

7. Stone and Weber (1992) and Weber (1984, 1985) review computerized content analysis techniques.

8. See Andren (1981:58–66) and Holsti (1969:94–126) on reliability and latent or semantic analysis.

9. See Krippendorff (1980) for various measures of intercoder reliability. Also see Fiske (1982) for the related issue of convergent validity.

10. For information on the confederate flag issue, see Cooper and Knotts (2006), Holyfield, Moltz, and Bradley (2009), and Newman (2007).

11. On social indicators, see Bauer (1966), Carley (1981), Duncan (1984:233–235), Juster and Land (1981), Land (1992), Rossi and Gilmartin (1980), and Taylor (1980). Also see Ferriss (1988) on using social indicators for planning and social forecasting.

12. See Herbert (2003), Miringoff and Opdycke (2007), and Ravo (1996).

13. Many non-English language yearbooks are also produced; for example, *Statistiches Jahrbuch* for the Federal Republic of Germany, *Annuaire Statistique de la France* for France, Year Book Australia for Australia, and Denmark's *Statiskisk Ti Arsoversigt*. Japan produces an English version of its yearbook called the *Statistical Handbook of Japan*.

14. Guides to government include the *Guide to British Government Publications*, Australian official publications, and Irish official publications. Similar publications exist for most nations. For example, *DOD's Parliamentary Companion for the United Kingdom* and the *Parliamentary Handbook of the Commonwealth of Australia* are both similar to the *Almanac of American Politics*.

15. See Churchill (1983:140–167) and Stewart (1984) for lists of business information sources.

16. Other major U.S. archives of survey data include the National Opinion Research Center, University of Chicago; the Survey Research Center, University of California–Berkeley; the Behavioral Sciences Laboratory, University of Cincinnati; Data and Program Library Service, University of Wisconsin–Madison; the Roper Center, University of Connecticut–Storrs; and the Institute for Research in Social Science, University of North Carolina–Chapel Hill. Also see Kiecolt and Nathan (1985) and Parcel (1992).

17. See Alwin (1988) and Davis and Smith (1992).

18. For a discussion of these issues, see Dale et al. (1988:27–31), Horn (1993:138), Maier (1991), and Parcel (1992).

19. See Stevenson (1996).

20. See *The Economist*, "The Good Statistics Guide" (September 11, 1993); "The Overlooked Housekeeper" (February 5, 1994); and "Fewer Damned Lies?" (March 30, 1996). Also see "U.N. Urges Fiscal Accounting to Include Sex Trade," *New York Times* (August 20, 1998).

21. See Block and Burns (1986), Carr-Hill (1984a), Hindess (1973), Horn (1993), Maier (1991), and Van den Berg and Van der Veer (1985). Discussions by Norris (1981) and Starr (1987) are also helpful.

第12章

量化資料分析

資料處理
單變項的結果分析
雙變項的結果分析
超過兩個變項
推論統計
結論

> 統計也可能被視為一種處理資料的方法。此種定義所強調的觀點，是把統計當成一種與蒐集、組織、以及分析數值事實與觀察結果有關的工具……對於敘述統計而言，最重要的事情就是以一種方便的、可用的、以及可以理解的形式來呈現資訊。
>
> ——Richard Runyon and Audry Haber, *Fundamentals of Behavioral Statistics*, p. 6

如果你閱讀了一份以量化資料爲基礎的研究報告或者是文章，你將很有可能會看到許多圖表、曲線圖、以及填滿了數字的表格。不過，你可別被這些圖表或是表格給嚇到了。這些研究報告或者是文章的作者提供這些圖表、曲線圖、以及表格的目的，是要讓身爲讀者的你能夠快速而簡要地掌握其中的資料。這些圖表以及表格讓你能夠看到研究者所蒐集的證據，並且可以自己去檢視這些證據。當你在蒐集你自己的量化資料時，你也將會運用類似的技術來揭露出隱藏在這些資料中的訊息。你將必須去組織並且操作這些量化資料，以顯現出你對於社會世界所感到興趣的事物。在本章中，你將會學習到組織以及分析量化資料的基礎知識。由於分析量化資料是一個相當複雜的知識領域，因此，本章並不能夠取代社會統計學的課程。本章所討論到的內容，只有包括了在瞭解社會研究時所必須具備的基本統計概念以及資料處理的技術而已。

利用在之前幾章中的技術所蒐集到的資料，都是數字形式的資料。這些數字代表了變項的值，而這些變項所測量的是參與者、受訪者、或者是其他個案的特質。在問卷上、便條紙上、記錄單上、或者是電腦檔案裡的數字都只是原始的形式而已，我們還必須對這些原始資料進行下列的一些處理動作，才能看出這些資料中關於假設的訊息：把這些資料以適合輸入電腦的格式來加以重新組織、以表格或者是曲線圖的方式呈現這些資料以概述其特質、以及去詮釋或者是對結果賦予理論方面的意義。

資料處理

資料編碼

在我們檢視量化資料以考驗假設時，我們必須將這些資料置入於特定的格式當中。你已經在上一章裡閱讀過關於資料編碼的概念了，而在此處，資料編碼所指的意思是以系統性的方式將原始資料重新組織爲一種特定格式，好讓我們可以利用電腦上的統計軟體來輕鬆地分析資料。正如同在內容分析中的資料編碼過程一樣，在把資料從一種格式轉換到另外一種格式時，研究者也會先訂定出一些規則，並且把這些規則應用於資料轉換的過程當中。[1]

如果你已經把資料以數字的方式記錄在組織良好的記錄單上面的話，那麼編碼可能就會像是一種簡單的文書工作而已，但是，如果你想要以一種類似於隱性內容分析的過程，把開放式問題的答案加以編碼成爲數字格式時，編碼的工作就有可能會變得非常困難。在針對開放式的調查資料、或者是其他並不是數字形式的資料進行編碼時，你必須要有一個編碼程序以及一本編碼簿。**編碼程序**（coding procedure）所指的是一組規則，這些規則說明了你如何把某個數字指派給變項的屬性。例如，你把男性編碼爲 1、把女性編碼爲 2，或者是在李克特量表上，你把「非常同意」編碼爲 4、「同意」編碼爲 3，以此類推。你必須針對所有的變項以及遺失的資訊的每一個類別都指定一個編碼。編碼程序將會詳盡地解釋你應該如何把非數字的資訊轉換成爲數字。

編碼簿（codebook）所指的是一份文件（也就是包含了一頁或者是多頁的資訊），在其中描述了編碼程序，並且以特定的格式來描述變項資料的電腦檔案位置。當你在進行資料編碼的工作時，製作一本組織良好而且詳盡的編碼簿是非常重要的，而且你也應該多影印幾份編碼簿。如果你沒有寫下編碼程序的細節、或者是把編碼簿隨手亂放的話，你就可能會失去瞭解資料的關鍵，而且也可能必須將資料重新再編碼一次。

在你動手去蒐集任何資料之前，你應該要先開始去思考編碼程序。舉例而言，有許多的調查研究者會在進行訪談或者是蒐集資料之前，就先針對問卷進行預先編碼（precode）。預先編碼所指的是把編碼類別（例如，1 代表男性、2 代表女性）放入問卷裡，並且把編碼簿中所列出的特質加入問卷當中。[2] 如果你沒有進行預先編碼，那麼你在蒐集完資料之後的第一個步驟，就是要先製作一本編碼簿。你也必須爲每一個個案指定一組識別號碼，以利於你去追蹤記錄所有的個案。接下來，你則必須把每一份問卷上的資訊轉換成爲電腦能夠讀取的格式。

編碼程序　由量化研究者所創造出來的一組規則，以指派數字給特定的變項屬性，通常是為了進行統計分析而作準備，而且會被詳細地記錄在編碼簿裡面。

編碼簿　一份文件，其中描述了對變項進行編碼的程序、以及以電腦能夠讀取的格式描述變項的位置。

輸入資料

大多數被設計來進行數值資料分析的電腦程式，都會要求資料必須以方格的格式來輸入。在方格的格式裡，每一橫列（row）都代表了一個受訪者、參與者、或者是個案。而在電腦的專門用語中，這些資料就被稱爲是**資料記錄**（data records），而每一筆資料記錄都是一個單獨的個案。至於一直行（column）或者是許多直行所代表的則是特定的變項。如此一來，我們就可以由某行與某列交叉所形成的特定位置（例如，第 7 列、第 5 行），回溯到該筆資料的原始來源（例如，第 8 位受訪者在問卷上關於婚姻狀態的題目）。而被指定給某一個變項的一個欄位或者是許多欄位，則被稱爲**資料欄位**（data field），或者是簡稱爲*欄位*（*field*）。

舉例而言，你以一種電腦檔案的格式對三位受訪者的調查資料進行編碼，就像是在圖 12.1 中所呈現的資料檔案的開頭部分一樣。人們無法輕易地解讀此種格式的資料，而且如果沒有編碼簿的話，這些資料可以說是毫無價值的。這個資料檔案把三位受訪者對於 50 個調查問題的答案濃縮成爲三個橫列。許多研究計畫的原始資料看起來都類似於圖 12.1，但是卻可能有超過 1,000 個橫列，而且可能會有超過 100 行的資料。例如，一份對 250 名受訪者進行 15 分鐘的電話調查所得到的資料，可以產生一個 250 列乘以 240 行的資料方格。

依據在圖 12.1 裡的編碼簿的內容說明，前兩個數字是識別碼，因此，在圖中的範例資料是來自於第一位（01）、第二位（02）、以及第三位（03）受訪者的資料。請注意，爲了減少在 l（英文字母）和 01（數字）之間的混淆情況，我們在數字前面會加上零來表示這是數字而不是字母。除此之外，第一行總是代表十位數，而第二行則是代表個位數。編碼簿告訴我們，在第 5 行中包含了「性別」這個變項，個案 1 和 2 都是男性，而個案 3 則是女性。而在第 4 行中的資訊則告訴我們，卡洛斯訪談了個案 1 和個案 2，而蘇菲亞則是訪談了個案 3。

在把原始的量化資料輸入電腦時，共有以下的四種方法：

1. *編碼單*（*code sheet*）　先蒐集資訊，然後把資訊從原本的來源轉換成爲方格的格式（編碼單）。接著再一行一行地把編碼單上的資訊輸入電腦裡。
2. *直接輸入法*〔包括電腦輔助電話訪談（CATI）〕　在蒐集資訊時，訪談者坐在電腦鍵盤前面（或者是類似的記錄裝置），然後一邊傾聽或是觀察資訊，並且將資訊輸入電腦；或者，也可以去要求受訪者自行把資訊輸入電腦中。在使用**直接輸入法**（direct-entry method）時，電腦必須先被設定好能夠接受所輸入的資訊。
3. *光學掃瞄*（*optical scan*）　先蒐集資訊，然後藉由畫記正確的「點」的方式，把資訊輸入光學掃

資料記錄　在電腦格式的資料中的單位，其中包含了某個個案的所有變項。

資料欄位　在以電腦格式儲存的資料中的一個或是多個欄位，代表了某個特定變項的資訊位置。

直接輸入法　在不使用條碼或是光學掃瞄單的情況下，以打字的方式直接把資料輸入電腦的過程。

圖 12.1 三個個案的編碼資料以及編碼簿

調查問卷節錄

受訪者編號_____ 訪談者姓名_____

受訪者的性別：___ 男性 ___女性

1. 第一個問題與美國總統有關。關於下列的敘述，請問你覺得「非常同意」、「同意」、「不同意」、「非常不同意」、或是「無意見」？

美國總統目前的工作表現相當稱職。

_____ 非常同意 _____ 同意 _____ 不同意 _____ 非常不同意 _____無意見

2. 你的年紀多大？___歲

編碼資料節錄

```
00000000011111111112222222222333333333444 ……(十位數)
12345678901234567890123456789012345678901 ……(個位數)
01 212736302 182738274 10239 18.82 3947461……
02 213334821 124988154 21242 18.21 3984123……
03 420123982 113727263 12345 17.36 1487645……
```

前三個個案的原始資料，從第 1 行到第 42 行。

編碼簿節錄

行	*變項名稱*	*描述*
1-2	編號	受訪者的編號
3	空白	
4	訪談者	蒐集資料的訪談者： 1 = 蘇珊 2 = 卡洛斯 3 = 黃恩 4 = 蘇菲亞 5 = 克蕾倫絲
5	性別	訪談者回報受訪者的性別 1 = 男性；2 = 女性
6	總統的表現	美國總統目前的工作表現相當稱職 1 = 非常同意。 2 = 同意 3 = 無意見 4 = 不同意 5 = 非常不同意 空白= 遺失資訊

瞄單裡（或者也可以要求受訪者或參與者自行把資訊填入表單裡）。接下來，利用光學掃瞄機或者是讀卡機，來把資訊傳輸到電腦中。

4. *條碼*（*bar code*） 先蒐集資訊，並且把資訊轉換成爲以不同寬度代表不同特定數值的條碼，然後再利用條碼閱讀機把資訊傳輸到電腦中。

清理資料

在對資料進行編碼時，正確性是極度重要的（請參閱範例 12.1 資料處理的範例）。如果你在進行編碼或是把資料輸入電腦時犯下錯誤，將會威脅到量數的效度，並且產生具有誤導性的結果。如果你擁有完美的樣本、完美的測量結果、而且在蒐集資料時也沒有犯下任何錯誤，但是在編碼的過程中、或是在把資料輸入電腦的過程中卻犯下錯誤的話，你就有可能因此而毀了整個研究計畫。

在非常小心謹慎地進行編碼工作之後，你還必須去確認編碼的正確性，或者是去「清理」（"clean"）資料。一般而言，你可以隨機地去抽取整體資料的 10%到 15%來當作樣本，並且進行第二次的編碼。如果你在重新編碼的樣本中沒有發現任何錯誤的話，你就可以繼續進行接下來的步驟。但是如果你發現了錯誤的話，你就必須重新檢查所有的編碼資料。

在把資料輸入電腦之後，你可以利用下列兩種方法來確認編碼的正確性。**可能編碼式清理**（possible code cleaning）〔又稱爲*可能性檢查*（*wild code checking*）〕牽涉到去檢查所有變項的類別，以找出不可能出現的編碼。例如，在受訪者的性別編碼中，1 代表男性，2 代表女性。如果在某個個案的性別變項欄位中出現了 4 這個數字，那麼這就代表出現了編碼錯誤。另外一種方法是**列聯式清理**（contingency cleaning）〔又稱爲*一致性檢查*（*consistency checking*）〕，此種方法牽涉到把兩個變項交叉分類（cross-classifying），以找出在邏輯上不可能出現的組合方式。例如，你針對學歷和職業這兩個變項進行交叉分類，如果你發現某個受訪者根本連國中二年級的課程都沒讀完，但是在記錄上的資料卻顯示是一個醫生的話，你可能就必須去找出其中的編碼錯誤了。

當資料被輸入電腦之後，你可以利用一些方法來修正資料，但是你所使用的類別，無法比在蒐集資料時所使用的類別更加精細。例如，你或許會把比率尺度的收入資料分成五個順序尺度的類別，而且你也可以合併變項的類別，並且結合來自於數個不同指標的資訊，以創造出一個新的指數變項。

可能編碼式清理　一種利用電腦來清理資料的方式，作法是去搜尋一些不可能有個案存在的回答或者是答案的類別。

列聯式清理　一種利用電腦來清理資料的方式，在使用此種方式時，研究者會去檢視兩個變項類別的組合方式，以找出在邏輯上不可能出現的個案。

單變項的結果分析

次數分配

統計（*statistics*）一詞可以用來指稱一組蒐集而來的數字（例如：居住在某個城市裡的居民數目），以及應用數學的一個分支，而我們則利用此種應用數學來操弄

範例 12.1

資料處理的範例

沒有任何方法可以取代親自處理資料所能帶來的好處。以下是一個資料準備的範例，這個範例是來自於一項我和我的學生一同進行的研究。我所任教的大學調查了大約三分之一的學生，以瞭解學生們對於校園性騷擾的想法以及經驗。一組研究團隊抽取了一份隨機樣本，並且製作以及分發了一份自填式的問卷。受訪者把答案劃記在一張光學掃瞄單上，這份掃瞄單與他們平時考選擇題時所使用的答案卡相當類似。在分發了超過 3,000 張的光學掃瞄單之後，研究就開始進行了。

在回收了掃瞄單之後，我們以目視的方式快速地瀏覽了每一張掃瞄單，以找出明顯的錯誤。雖然在掃瞄單上已經清楚地說明，受訪者應該要用鉛筆把每一個圓圈乾乾淨淨地塗滿、塗黑，但是我們仍然發現有大約 200 名受訪者使用鋼筆來作答，而且還有另外 200 名受訪者非常粗心大意、或者是顏色塗得太淺。我們先把這些掃瞄單清理過，然後再用鉛筆重新畫記一遍。我們也發現了 25 張已經不堪使用的掃瞄單，這些掃瞄單若不是面目全非、受到損傷，就是答案太不完整（例如，在 70 個題目中，只回答了前 2 題而已）。

接下來，我們用電腦去讀取那些仍然可以使用的光學掃瞄單。我們利用電腦來產生每一個變項屬性的發生數目或者是次數。在看著次數分配時，我們發現了數種不同類型的錯誤。有些受訪者在回答一些只有一個答案的問題時，竟然畫記了兩個答案。有些受訪者則是畫記了一些不可能的答案編碼（例如，把性別的數字畫記為 4，但是合理的答案只有兩個，1 代表男性，2 代表女性）。而有些受訪者對於每一個題目的答案都是完全相同的，此種情況暗示了他們根本沒有認真在看待這項調查。對於每個有錯誤的個案，我們就會找出原本的光學掃瞄單，看看我們是不是還能夠找回一些資訊。如果我們無法找回任何資訊，我們會把該個案重新分類為未回應，或者是把某個回答重新編碼為遺失的資訊。

在這份問卷上有兩組條件式問題。對於這兩組問題，任何在第一個問題中回答「否」的受訪者，會被要求跳過接下來的五個題目。我們為每個問題製作了一張表格，並且仔細地進行檢查，看看對第一個問題回答「否」的所有受訪者是否都有跳過接下來的五個問題。我們發現在大約 35 個個案中，雖然受訪者對第一個問題的答案是「否」，但是仍然接著回答接下來的五個問題。我們只好找出原本的光學掃瞄單，並且試著去搞清楚該名受訪者的真正意圖是什麼。在大多數此類的個案中，受訪者的答案似乎確實真的是「否」，但是卻沒有注意到那些告訴他們要跳過接下來五個題目的指示。

最後，我們檢視了每個變項屬性的次數，以看看是否有任何的意義。我們很驚訝地發現到，在關於種族血統的問題方面，有大約 600 名受訪者所標記的是「美國原住民」（"Native American"），除此之外，在這些把自己標記為美國原住民的受訪者當中，有超過一半以上的人是大一新生。在確認過學校的官方記錄之後，我們發現在本校所有註冊入學的學生當中，只有大約 20 個人真的是美國原住民或者是美洲印第安人，而且有超過 90%的學生是白人或者是非西班牙裔的白種人。至於把自己標記為黑人、非洲裔美國人、或者是墨西哥裔美國人的受訪者比例，則與官方記錄相符。我們得到的結論是，有一些白人受訪者對於「美國原住民」這個詞彙不熟悉，所以不知道這個詞彙所指的是「美洲印第安人」（"American Indian"）。顯而易見的是，他們本應該把自己標記為「白人」的，但是卻錯誤地把自己標記為「美國原住民」。因為我們原本預期在樣本中應該只有 7 名美國原住民，所以我們把所有回答「美國原住民」的受訪者，全部重新編碼為「白人」，而這代表了我們把在樣本中的美國原住民重新分類為白種人。在完成這個工作之後，我們就準備好要去進行分析資料的工作了。

以及總結一群數字的特質。社會科學家會使用上述兩種類型的統計，但是在這兒，我們的焦點則是在於第二種類型的統計，也就是一些操弄以及總結數字的方法，而這些數字則是代表了來自於某個研究計畫裡的資料。

敘述統計（descriptive statistics）（也可譯爲描述統計）所描述的是數值資料。我們可以利用所牽涉到的變項數目來把數值資料加以分類成爲單變項（univariate）、雙變項（bivariate）、或者是多變項（multivariate）等等（也就是一個、兩個、或者是多個變項）。單變項統計所描述的是單獨一個變項（*uni*-這個字首的意思是指「一個」，而-*variate* 這個字根所指的意思是「變項」）。在描述單一變項的數值資料時，最簡單的方式就是利用**次數分配**（frequency distribution）。你可以把次數分配應用於名目尺度、順序尺度、區間尺度、以及比率尺度的資料上。例如，我手邊有一份關於 400 名受訪者的資料，並且想要以簡單明瞭的方式來製作關於性別資訊的摘要。要達成此種目的最容易的方法，就是去計算原始資料，或者是製作比例的次數分配（請參閱圖 12.2）。我也可以利用圖表的形式來呈現出相同的資訊。

一些常見的統計圖表包括了直方圖、長條圖、以及圓形圖。長條圖主要是使用在間斷變項上，圖形中的長條可以是垂直方向、也可以是水平方向，而且在長條之間留有一些小空間。雖然**直方圖**（histogram）這個術語不太精確，但是直方圖通常是一組用來表示區間尺度或是比率尺度資料的筆直長條圖。[3]

對於區間尺度或者是比率尺度的資料而言，我們通常會把資訊分爲數個類別。這些歸類在一起的類別必須要彼此互斥。我們也可以在一個**次數分配多邊圖**（frequency polygon）中去繪製出區間尺度或者是比率尺度的資料，其中縱軸所代表的是個案的數目或者是次數，而橫軸所代表的則是變數的值或是分數。當我們把所有的點連接起來之後，就會得到一個多邊形。

敘述統計　一種廣泛類型的簡單統計，研究者使用此種統計來描述資料中的基本模式。

次數分配　一種表格，其中會顯示出個案在某個變項的類別中的分佈情況，換句話說，也就是在每一個類別裡的個案數目或者是百分比例。

直方圖　一種以圖形來表示單變項的次數分配或是比例的方式，通常是以垂直的線條來表示數量或者是比例。

次數分配多邊圖　一種由連結的點所形成的圖形，用以顯示在某個變項的每一個類別當中的個案數目。

集中趨勢量數　一種統計量數的種類，此種量數可以把關於某個變項的次數分配的資訊總結成為一個單一的數字。

眾數　一種單一變項的集中趨勢量數，所指出的最常出現或者是最常見的分數。

集中趨勢量數

一般而言，我們通常會想要以一個單獨的數字來總結關於某個變項的資訊。爲了要總結某個變項的資訊，我們會使用三個**集中趨勢量數**（measures of central tendency）（也就是次數分配中央部分的量數：平均數、中位數、以及眾數）。許多人會把這三種量數稱作*平均數值*（*averages*），不過，這種說法雖然也是指同樣的事物，但卻比較不是那麼精確或是清楚。

眾數（mode）是最容易使用的集中趨勢量數，

圖 12.2　單變項統計量數的範例

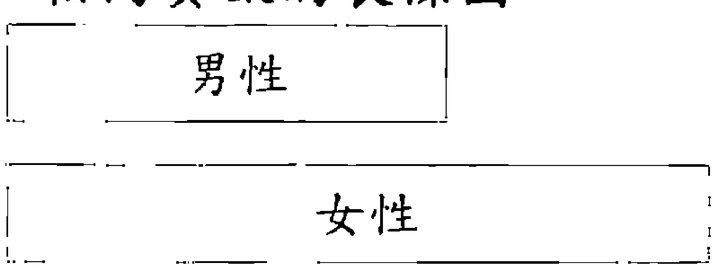
原始計數的次數分配

性別	*次數*
男性	100
女性	300
總數	400

百分比例次數分配

性別	*次數*
男性	25%
女性	75%
總數	100%

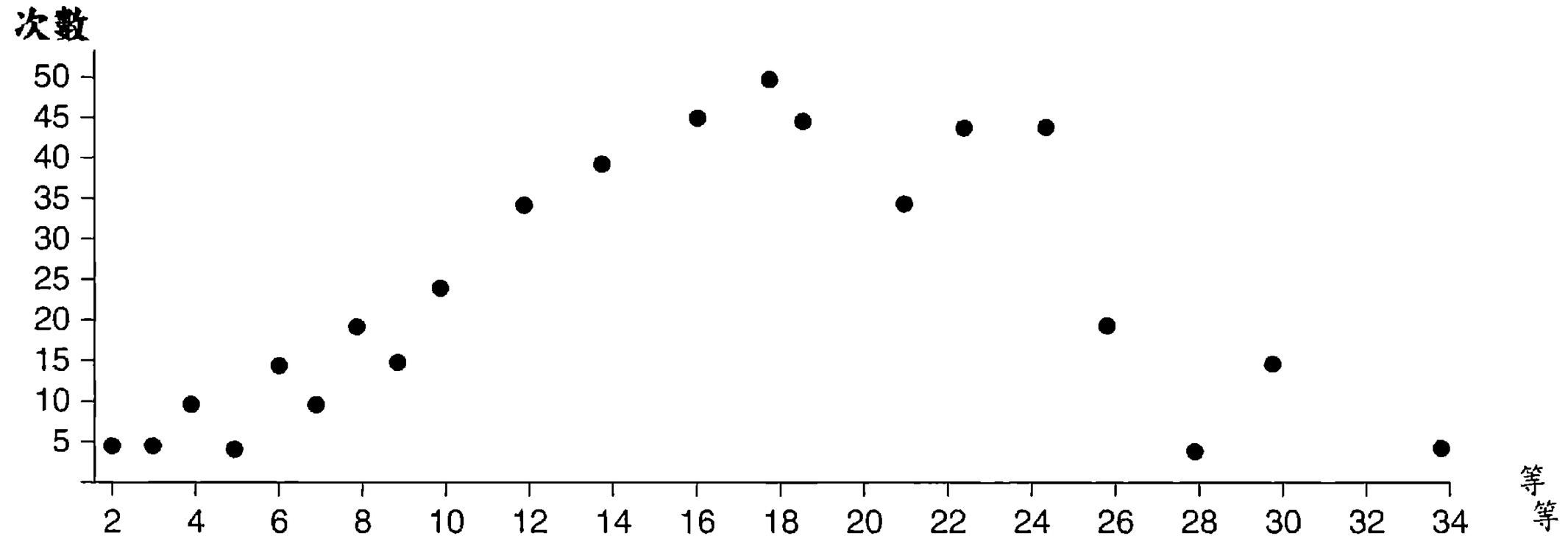
相同資訊的長條圖



歸類資料次數分配的範例

第一份工作的年收入	*次數（N）*
低於$5,000	25
$5,000 到$9,999	50
$10,000 到$15,999	100
$16,000 到$19,999	150
$20,000 到$29,999	50
$30,000 以上	25
總數	400

次數分配多邊圖的範例



而且我們可以把眾數運用在名目尺度、順序尺度、區間尺度、以及比率尺度的資料中。眾數所指的就是最常見或者是最常出現的數字。例如，以下有一列數字串：6、5、7、10、9、5、3、5，而在這串數字當中的眾數就是 5。一個分配可以有超過一個以上的眾數，例如，在以下這列數字串裡：5、6、1、2、5、7、4、7，其中的眾數就是 5 和 7。如果數字串太長的話，只要利用次數分配就能輕易地看出眾數，因

爲你只要去找出次數最多的分數就可以了。在次數分配中，一定會有至少一個個案的分數等同於眾數。

中位數（median）所指的就是中點，也就是第 50 個百分位數值、或者是高於以及低於這個點的個案數目剛好各佔一半。我們可以在順序尺度、區間尺度、以及比率尺度的資料中使用中位數（但是不能用於名目尺度的資料裡）。我們可以用「目測」的方式看出眾數，但是要計算出中位數時，就必須花點時間才行。計算中位數最簡便的方法，就是先把所有的分數由低到高排列出來，然後再從頭開始計數到中間的數值。如果分數的數量是奇數的話，計數的過程就會相當簡單。例如，假設有 7 個人正在等公車，這些人的年齡分別是 12、17、20、27、30、55、80，如此一來，中位數就是 27。請注意，中位數並不會輕易地改變。在上述的例子裡，如果 55 歲和 80 歲這兩個人都搭上公車離開了，而後來則有另外兩個 31 歲的人加入了等公車的行列，那麼中位數仍然會維持不變。如果分數的數目是偶數的話，計算中位數的過程就會有一點點複雜。例如，有 6 個人正在公車站牌旁邊等公車，他們的年齡分別是 17、20、26、30、50、70，如此一來，中位數就會是介於 26 到 30 的中間值。在此種情況中，計算中位數的方式就是把兩個位於中間部分的分數相加，然後再除以 2（26 + 30 = 56，56/2 = 28），因此，這些等公車的人們的年齡中位數就是 28，雖然在這些人當中其實沒有任何人的年齡是 28 歲。另外也請注意到，由於每個人的年齡都是不同的數字，因此在上述的 6 個年齡的數字串中並沒有眾數。

中位數　一種單一變項的集中趨勢量數，所指出的是一個中間點或分數，也就是高於或低於這個點或分數的個案數目，剛好各佔一半。

平均數　一種單一變項的集中趨勢量數，所指的是算術平均數，也就是把所有的分數加總、並且除以分數的數目之後所得到的數值。

常態分配　一種表示個案分布情況的鐘形次數分配多邊圖，圖形中央部分為最高點，而在中心點的兩邊則具有完全相同的曲線斜率。常態分配是許多自然發生現象的分配型態，而且是許多統計理論的基礎。

平均數（mean）〔又稱爲*算術平均數*（*arithmetic average*）〕可以說是最被廣泛使用的集中趨勢量數，不過，我們只能在區間尺度以及比率尺度的資料中使用平均數。[4] 在計算平均數時，我們必須先把所有的分數加總起來，然後再除以分數的數目。例如，在前面提過的例子裡，那些等公車的人們的年齡平均數計算過程如下：17 + 20 + 26 + 30 + 50 + 70 = 213；213/6 = 35.5。請注意，在這些人當中，其實沒有任何一個人的年齡是 35.5 歲，而且平均數也不等同於中位數。

極端值（非常大或是非常小的數值）若有任何變化，可能會對平均數造成很大的影響。例如，在上述的例子裡，假設 50 歲和 70 歲的那兩個人離開了，後來則是兩個 31 歲的人加入的話，則年齡分布的情況會變成下列這個樣子：17、20、26、30、31、31。如此一來，中位數仍然維持不變，也就是 28。至於平均數則是 17 + 20 + 26 + 30 + 31 + 31 = 155；155/6 = 25.8。因此，當少數的極端值被移除時，平均數就會產生很大的落差。

如果次數分配形成**常態分配**（normal distribution）或者是鐘形曲線（bell-shaped curve），上述的這三種集中趨勢量數就會完全相等。然而，如果次

數分配所形成的是**偏態分配**（skewed distribution）（也就是說，有較多的個案落在較高的或者是較低的分數那一端），那麼這三種量數就不會相等。如果大多數的個案分數都比較低、而且有少數很高的極端值的話，平均數將會是最高的、中位數次之、而眾數則會是最低的。相反地，如果大多數的個案分數都比較高、而且有少數很低的極端值的話，那麼平均數將會是最低的、中位數次之、而眾數將會是最高的。一般而言，雖然平均數通常會用於大多數其他的統計量數當中，但是中位數卻是最適合用於偏態分配當中的（請參閱圖 12.3）。

變異量數

集中趨勢量數是以單一數字來總結關於某個分配的資訊，然而，這些量數只有提供了關於分配的中心區域的資訊。對於一個分配而言，還有另外一項重要的特質，也就是圍繞著中心區域的*離散程度*（*spread*）、*分散情況*（*dispersion*）、或者是*變異性*（*variability*）。某兩個分配可能會有完全相同的集中趨勢量數，但是在中心區域附近的離散程度卻不相同。例如，有 7 個人正站在某個酒吧前面的公車站牌旁邊等公車，他們的年齡分別是 25、26、27、30、33、34、35 歲。在這個數列當中的中位數以及平均數都是 30。在另外一個位於一家冰淇淋店前面的公車站牌旁邊，也有另外的 7 個人在等公車，他們年齡的中位數以及平均數與前一群人是完全相同的，但是他們的年齡卻分別是 5、10、20、30、40、50、55 歲。在冰淇淋店前面的這一群人的年齡，與中心區域的離散程度較高，換句話說，就是他們的年齡分配中的變異性較大。

*變異性*具有重要的社會意義。舉例而言，在 X 城市中，家庭年收入的中位數以及平均數是 37,600 美元，而且變異程度爲零。所謂*零變異*（*zero variation*）指的是在該城市中的每個家庭年收入剛剛好都是 37,600 美元，不多也不少。而在 Y 城市中的家庭年收入中位數以及平均數都與 X 城市完全相同，但是其中有 96%的家庭，其年收入僅有 14,000 美元，而另外 4%的家庭，其年收入則高達 350,000 美元。在此種情況下，X 城市中的家庭年收入是完全均等的，而在 Y 城市中則存在著極嚴重的不平等狀況。如果我們不知道在這兩個城市裡的家庭年收入的變異性，我們就會錯失了非常重要的資訊。

偏態分配　一種表示單一變項類別中的個案分布情況的圖形，但此種圖形並不是常態分配，換句話說，其形狀並不是鐘形。在偏態分配中，兩個端點的個案數目並不是相等的，在其中某一個端點的個案數目會多於另一端。

全距　一種單一變項離散程度的量數，能夠指出最高分數以及最低分數。

我們可以利用以下三種方式，測量變異程度：全距、百分位數、以及標準差。**全距**（range）是其中最爲簡單的方式，且是由最大和最小的分數所構成。例如，在前面提過的例子裡，那些站在酒吧前面的公車站牌旁邊等公車的人們，其年齡的全距是從 25 到 35，或是 35－25 ＝ 10。如果那個 35 歲的人搭上公車離開了，並且由一個 60 歲的人所取代，那麼全距就會變成 60－25 = 45。由於全距只能告訴我們最高值與最低值，所以全距會有一些限制。例如，以下兩組數字的全距都是 35：30、30、30、30、

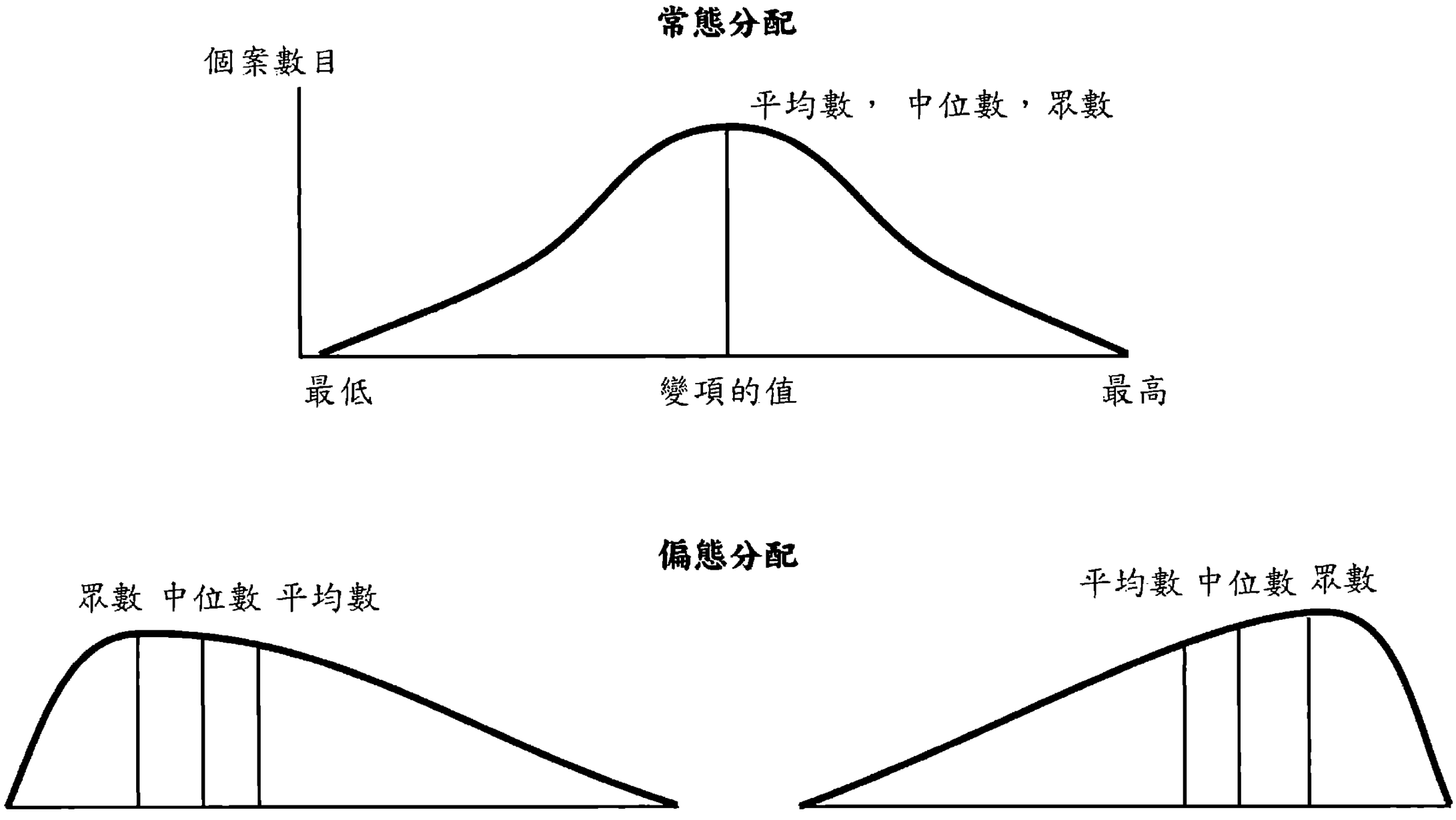


圖 12.3　集中趨勢量數

30、65 以及 20、45、46、48、50、55。

百分位數（percentiles）能夠告訴我們某個分數在分配中的特定位置。有一個你已經學習過的百份位數，也就是中位數，意思是指位於第 50 個百分位數的分數。有時候第 25 個以及第 75 個百分位數，或者是第 10 個以及第 90 個百分位數也會被用來描述一個分配。例如，位於第 25 個百分位數的分數，就意味著在該分配中有 25% 的個案的分數等於或者是低於該分數。計算百分位數時所依循的邏輯，與計算中位數時的邏輯完全相同。如果你有 100 個人而且想要找出第 25 個百分位數，你首先將所有的分數依序由低到高排好，然後再由最低分往上數，直到你數到第 25 個分數爲止。如果個案的總數並不是剛好 100 的話，你只要把分配依照百分比例的基準來進行調整就可以了。

標準差（standard deviation）是最難計算的變異量數，而且標準差也是最爲全面而且最被廣爲使用的量數。你在順序尺度、區間尺度、以及比率尺度的資料中都可以使用全距以及百分位數，但是你只能在區間尺度或者是比率尺度的資料中使用標準差。標準差以平均數爲基礎，並且能夠告訴我們在所有的分數與平均數之間的「平均距離」爲何。如果個案數量較多的話，人們幾乎已經不會親自動手去計算標準差，因爲只要利用電腦，在短短幾秒之內就可以計算出標準差。

請看一下在圖 12.4 中所顯示出的標準差計算過程。如果你把每一個分數與平均數之間的距離差

百分位數　一種單一變項離散程度的量數，能夠指出等於或是低於某個點或是分數的個案數量的百分比例。

標準差　一種單一變項離散程度的量數，能夠指出所有分數與平均數之間的平均距離。

圖 12.4 標準差

計算標準差的步驟

1. 計算出平均數。
2. 把每一個分數減去平均數。
3. 把每個分數減法平均數之後的差予以平方。
4. 把每個分數與平均數之間的平方差加總，以得到平方和。
5. 把平方和除以個案總數，以求得變異數。
6. 將變異數開根號以得到平方根，而這個平方根就是標準差。

計算標準差的範例

【共有 8 位受訪者，變項為接受學校教育的年數】

分數	*分數－平均數*	*（分數－平均數）*
15	15 − 12.5 = 2.5	6.25
12	12 − 12.5 = −0.5	.25
12	12 − 12.5 = −0.5	.25
10	10 − 12.5 = −2.5	6.25
16	16 − 12.5 = 3.5	12.25
18	18 − 12.5 = 5.5	30.25
8	8 − 12.5 = 4.5	20.25
9	9 − 12.5 = −3.5	12.25

平均數 = 15＋12＋12＋10＋16＋18＋8＋9 = 100，100/8 = 12.5

平方和 = 6.25＋.25＋.25＋6.25＋12.25＋30.25＋20.25＋12.25 = 88

變異數 = 平方和／個案總數 = 88/8 = 11

標準差 = 變異數的平方根 = $\sqrt{11}$ = 3.317 年

以下是計算標準差的公式：

符號：

x=每一個個案的分數　　Σ = Sigma（希臘字母），代表總和

$\bar{x}$=平均數　　N = 個案的總數

標準差的公式：[a]

$$\text{標準差} = \sqrt{\frac{\sum(x-\bar{x})^2}{N-1}}$$

[a] 依據研究者所使用的是母體的資料、或者是以樣本所估計得到的母體母數之間的不同，這個公式會有些微的差異。

異（也就是用個別的分數減去平均數）加總起來的話，你所得到的結果會是零，這是因爲平均數與所有分數之間的平均距離都是相同的。另外也請注意到，與平均數距離最遠的分數，對於平方和以及標準差所造成的影響也是最大的。

標準差本身的用途其實是有限的，主要還是用在進行比較的目的上。例如，假設現在有 A、B、C 三個班級，對 A 班裡的學生而言，其家長接受學校教育年數的標準差是 3.317，B 班爲 0.812，至於 C 班則是 6.239。在此種情況中，這些標準差

的數字所告訴研究者的訊息是，B 班學生家長的教育程度的相似度較高，而 C 班學生家長的教育程度的差異性則比較大。事實上，在 B 班中，以某個「普通的」家長接受學校教育的年數而言，該數字高於或是低於該班所有家長的平均數僅僅不到一年，因此 B 班學生的家長同質性很高。然而，在 C 班裡，某個「普通的」家長接受學校教育的年數可能高於或低於該班的平均數超過六年以上，因此，C 班家長的異質性就相當高。

Z 分數　在一個分數的次數分配中，某個分數的標準化位置，其表示方式是以該分數高於或者是低於平均數多少個標準差來標示出該分數所在的相對位置。

我們可以利用標準差以及平均數來計算得到 **Z 分數**（z-scores），而 Z 分數則讓你可以去對兩個或者是更多個分配或群組進行比較。Z 分數也被稱作*標準化分數*（*standardized score*），是以距離平均數多少個標準差的方式，來表示某個點或是分數在一個次數分配上的位置，換句話說，Z 分數所指的是某個分數在一個次數分配中的相對位置，而不是絕對值（請參閱資訊箱 12.1 計算 Z 分數）。Z 分數可以告訴我們許多訊息，例如，凱蒂（Katy）是 A 公司裡的銷售經理，年收入爲 70,000 美元，而麥克（Mike）則是 B 公司的銷售經理，年收入爲 60,000 美元。雖然在凱蒂和麥克的年收入之間有著 10,000 美元的絕對差距，但是相對於在同一家公司中的其他員工而言，他們這兩位經理所得到的薪水其實在比例上是相同的。換句話說，凱蒂和麥可的收入都高於他們個別公司中三分之二的員工。

以下還有另外一個如何使用 Z 分數的例子。漢斯（Hans）與海蒂（Heidi）是雙胞胎姐弟，但是漢斯的身高比海蒂還要矮。與其他同年齡的女孩相較之下，海蒂的身高剛好是平均值，因此她的 Z 分數爲零。相同地，在與他同年齡的男孩之中，漢斯的身高也剛好是平均值，所以他的 Z 分數也是零。因此，在個別的比較群體中，這對雙胞胎姐弟的 Z 分數都是相同的，所以他們的相對身高是完全相同的。

只要利用平均數以及標準差，就可以輕易地計算出 Z 分數。例如，某位雇主面試了來自於國王學院以及皇后學院的學生。這位雇主知道這兩間學院相當類似，而且都是以 4.0 分的量表來評定學生的成績。然而，在國王學院裡的學業平均成績的平均數是 2.62，標準差是.50，而在皇后學院裡的學業平均成績的平均數是 3.24，標準差則是.40。這位雇主認爲皇后學院的分數有灌水的嫌疑。來自於國王學院的蘇捷特（Suzette）的學業平均成績是 3.62，而來自於皇后學院的喬治（Jorge）的學業平均成績則是 3.64。這兩位學生所修過的課程完全相同，而這位雇主則想要依據這兩所學院的評分方式來對這兩名應徵者的分數進行調整（也就是計算出標準化的分數）。她先以每位學生的分數減去個別學院中的成績平均數，然後再除以個別學院中的成績標準差，以計算出他們個別的 Z 分數。例如，蘇捷特的 Z 分數計算方式如下：3.62−2.62 = 1.00/.50 = 2，而喬治的 Z 分數則是：3.64−3.24 = .40/.40 = 1。因此，這位雇主就可以知道，蘇捷特的成績比國王學院學生的學業成績平均數高了兩個標準差，而喬治的成績只比皇后學院學生的平均數高了一個標準差而已。也就是說，雖然蘇捷特的絕對學業成績低於喬治的成績，但是相對於他們各自學校裡的學生而言，蘇捷特的成績則是遠高於喬治的成績。

資訊箱 12.1 計算 Z 分數

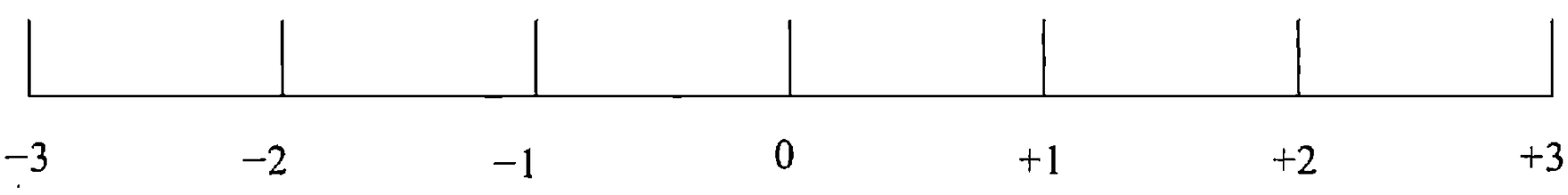
就我個人而言，我不是很喜歡下列這個計算 Z 分數的公式：

Z 分數 =（分數－平均數）／標準差，或者，以符號來表示的話，公式如下：

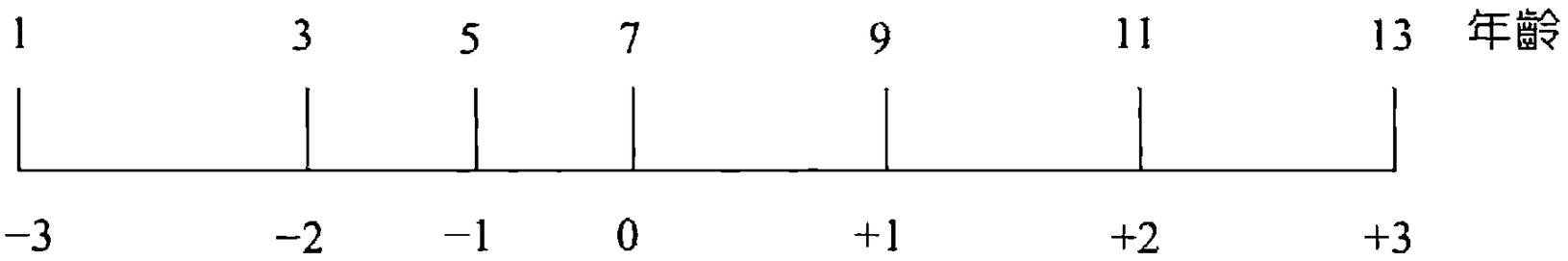
$$z=\frac{x-\bar{x}}{\delta}$$

其中的 x= 分數，$\bar{x}$= 平均數，δ= 標準差

我通常會使用一個簡單的概念圖表，這個圖表不但同樣可以計算出 Z 分數，而且還能顯示出 Z 分數的真正意義。例如，讓我們來思考一下以下這項關於學童年齡的資料，其中的平均數是 7 年，而標準差則為 2 年。那麼，我該如何去計算出米奎爾（Miguel）這個 5 歲學童的 Z 分數？或者，如果我知道亞修達（Yashohda）的 Z 分數是+2，那麼我該如何去得知她的真正年齡？首先，我先畫出一條從−3 到+3 的直線，而且把 0 標示在正中央的位置。然後，我接著把平均數的數值放在零這個位置，因為平均數的 Z 分數正好就是零，而 Z 分數所測量的正是某個分數高於或是低於平均數的距離。我的 Z 分數只有標示到 3 而已，因為在大多數的情況中，幾乎所有的個案都會落在離平均數三個標準差的範圍之內。我所畫出來的這條線看起來就像下列的這個樣子：



現在，我把平均數的數值標記上去，並且沿著這條線去加上或者是減去標準差。當平均數是 7 年、而且標準差為 2 年的時候，則高於平均數一個標準差的位置（1）的數值就是 7+2，也就是 9 年。而在 Z 分數為−2 的位置上，我則標上 3 年，這是因為這個位置比平均數低了兩個標準差（也就是 4 年），所以該位置的數值就是 7−4 = 3。這麼一來，我所畫出來的線現在看起來就會變成下列這個樣子：



現在，我們可以很輕易地看出來，米奎爾的年齡是 5 歲，所以他的 Z 分數是-1，而亞修達的 Z 分數是+2，所以她的年齡就是 11 歲。利用這張圖表，我可以輕鬆地把 Z 分數轉換為年齡、或者是把年齡轉換為 Z 分數。就算 Z 分數不是整數，而是像−1.5 這種數字的話，我也可以直接對應我的圖表而得知其年齡為 4 歲。同樣地，如果年齡是 12 歲的話，那麼 Z 分數就是+2.5。

雙變項的結果分析

雙變項關係

單變項統計（univariate statistics）是以個別的方式來描述單一變項，而**雙變項統計**（bivariate statistics）則具有更高的價值，這是因爲此種統計分析方式可以讓我們同時考慮兩個變項，並且去描述這兩個變項之間的關係。即使是簡單的假設也都必須要有至少兩個變項。雙變項統計分析可以顯示出兩個變項之間的**統計關係**（statistical relationship），也就是那些傾向於同時出現的事物。例如，某條河流的水已經遭到污染，而喝了這條河流的水的人們則都會生病，因此，在上述的這兩件事之間就存在著某種關係。更明確地說，是在下列這兩個變項之間存在著統計關係：其中一個變項是河水遭受到污染，而另外一個變項則是喝了河水的人們的健康情形。

統計關係是依據下列兩個概念而來的：共變以及統計獨立。**共變**（covariation）所指的是某兩項事物會一起出現，或者是彼此之間具有關聯性。*共變*（*covary*）所指的就是一起改變，換句話說，在某一個變項中具有某個數值的個案，很有可能在另外一個變項中也會有某個數值。例如，那些在收入變項中具有較高數值的人們，很有可能在預期壽命這個變項上也會有較高的數值。同理可推，那些收入較低的人，他們的預期壽命也會較低。此種情況通常會以比較簡短的方式來敘述，也就是只有簡單地說收入與預期壽命之間具有關聯性，或者是這兩者之間有共變的關係。我們也可以說，在知道了某人的收入之後，我們就可以去預測這個人可能的預期壽命；或者，我們也可以說預期壽命是由收入來決定的。

統計獨立（statistical independence）則是剛好與共變相反，因爲統計獨立所指的是在變項之間沒有任何的關聯性或是關係。如果兩個變項彼此獨立，當某個個案在某一個變項上出現了特定的數值時，在另外一個變項上並不會因此而出現任何特殊的數值。例如，麗塔（Rita）想要知道兄弟姐妹的數量是否與預期壽命有關。如果這兩個變項是彼此獨立的，那麼擁有許多兄弟姐妹的人、和身爲獨生子或是獨生女的人的預期壽命就會是完全相同的。換句話說，就算麗塔知道某個人有多少個兄弟姐妹，她也無法去預測或是得知這個人的預期壽命爲何。

單變項統計　只有處理單一變項的統計量數。

雙變項統計　只有牽涉到兩個變項的統計量數。

統計關係　利用基本的應用數學為基礎，來表達兩個或者是多個變項之間是否有互相影響的情況，也就是說，在變項之間是否有關聯性或是獨立性。

共變　一種指兩個變項會一起產生變化的概念，在此種情況下，在先得知了某個變項的數值之後，就能得知在另一個變項中的可能數值的資訊。

統計獨立　在兩個變項之間沒有任何統計關係的狀況，換句話說，就算研究者知道了某個變項的值，也無法得知關於另外一個變項的值的任何資訊。也就是在這兩個變項沒有任何關聯性。

散佈圖　一種展現出兩個變項之間的統計關係的圖形，而繪製的方式是去畫出每一個個案在兩個變項上的數值。

我們通常會以因果關係或者是預期共變的方式來陳述假設，而如果我們使用了虛無假設，意思就是變項之間是各自獨立的。虛無假設通常會被使用在正式的假設考驗當中，而且也經常會使用於推論統計中（稍後將針對推論統計進行討論）。

我們可以運用數種技術來決定在兩個變項之間是否有任何關係存在。其中三種基本的技術如下：散佈圖，也就是一個標示出變項之間關係的圖形；百分比表；以及關聯量數，也就是一種利用單一數字來表達出共變量的統計量數（例如：相關係數）。另外，關於如何以圖表來表示資料，請參閱圖表 12.1。

散佈圖

散佈圖的定義　所謂的**散佈圖**（scattergram，或者也被稱爲 *scatterplot*），指的是一種你在其中畫出每一個個案或是觀察結果的圖形。散佈圖中的每一條軸線都代表了某個變項的值。散佈圖通常使用於以區間尺度或者是比率尺度來進行測量的變項，但是很少用於順序尺度的變項。除此之外，如果其中任何一個變項是名目變項的話，則無法使用散佈圖。關於哪一個變項（自變項或是依變項）應該放在橫軸或者是縱軸，其實並沒有一定的規則，但是自變項（以字母 *X* 來代表）通常會放在橫軸，而依變項（以字母 *Y* 來表示）則是放在縱軸。每個變項的最低值都應該位於圖形的左下角，而最高值則是應該在縱軸的最高點或者是橫軸的最右邊。

建構散佈圖　先從兩個變項的全距開始。畫出橫軸與縱軸，把每個變項的值標記上去，並且在每一條軸線上寫下數字（使用方格紙會比較方便）。接下來把變項名稱標記在軸線上，並且在頂端寫下標題。現在，你已經可以開始輸入資料了。針對每一個個案，你先找出每一個變項的值，並且在圖上與兩個值對應的位置畫上一個點。例如，假設你想要製作一個關於學校教育年數以及子女數目的散佈圖。你首先看看第一個個案，並且找出其學校教育年數（例如：12）以及子女數目（例如：3）。然後你在散佈圖上去找出「學校教育年數」變項的值是 12、以及「子女數目」變項的值爲 3 的交叉位置，並且在這個位置畫上一個點來標示出這個個案。你接著必須對每個個案重覆這個過程，一直到所有的個案都被標記在圖形中爲止。

在圖 12.5 中所顯示出的散佈圖，是依據 33 位女性的資料所繪製的，而在圖形中顯示，在女性接受教育的年數以及她們所生育的子女數目之間呈現出負相關的關係。

散佈圖可以展現出雙變項關係中的三個面向：形式、方向、以及關聯的強度。

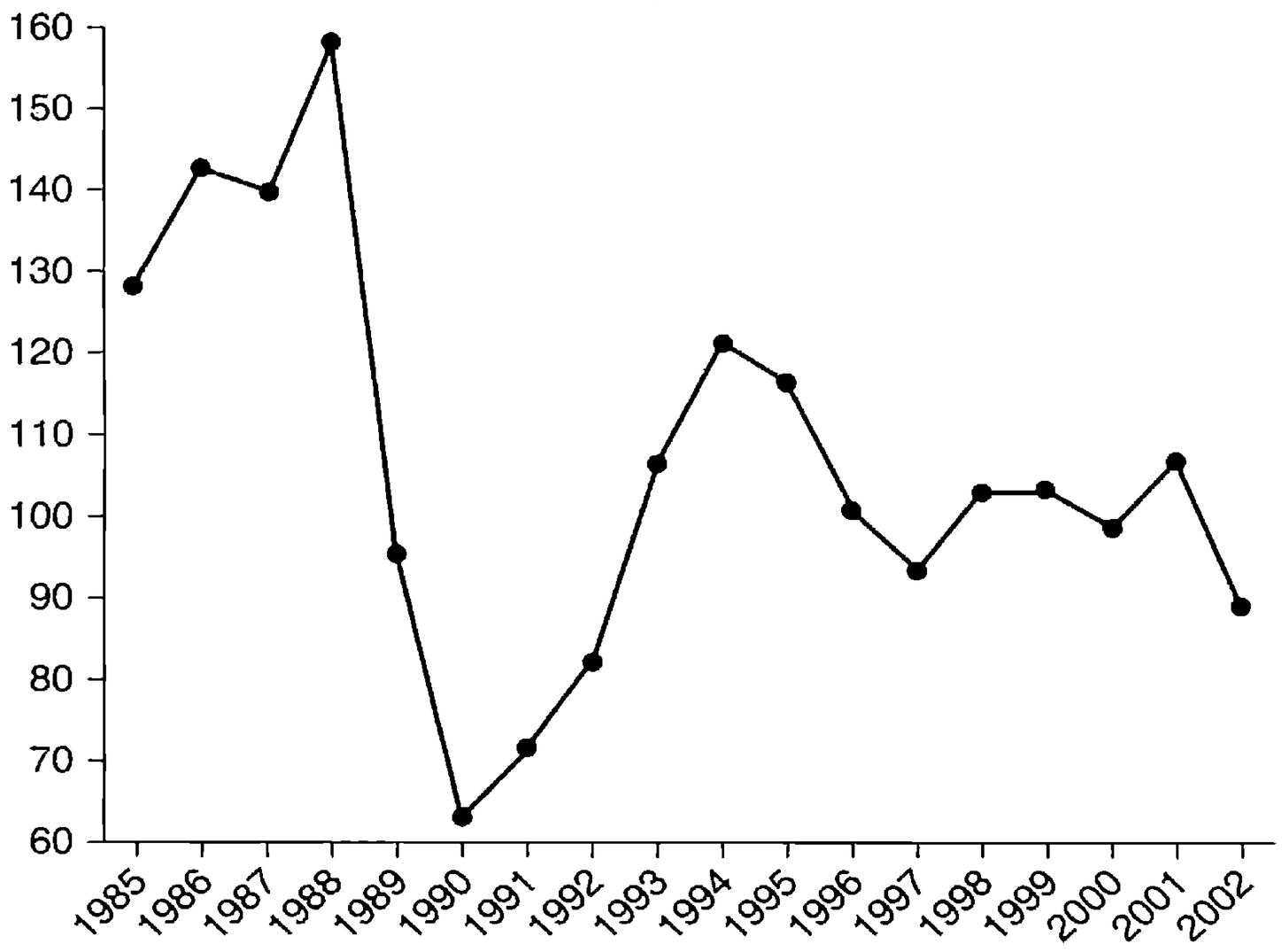
1. *形式（form）*　兩個變項之間的關係可能會有三種形式：彼此獨立的、線性的、以及曲線的。獨立或沒有關係的情況是最容易看出來的，因爲在此種情況下，散佈圖中的點看起來就像是隨機散落在圖上，其分布的情形可說毫無任何模

圖表 12.1　畫出正確的圖形

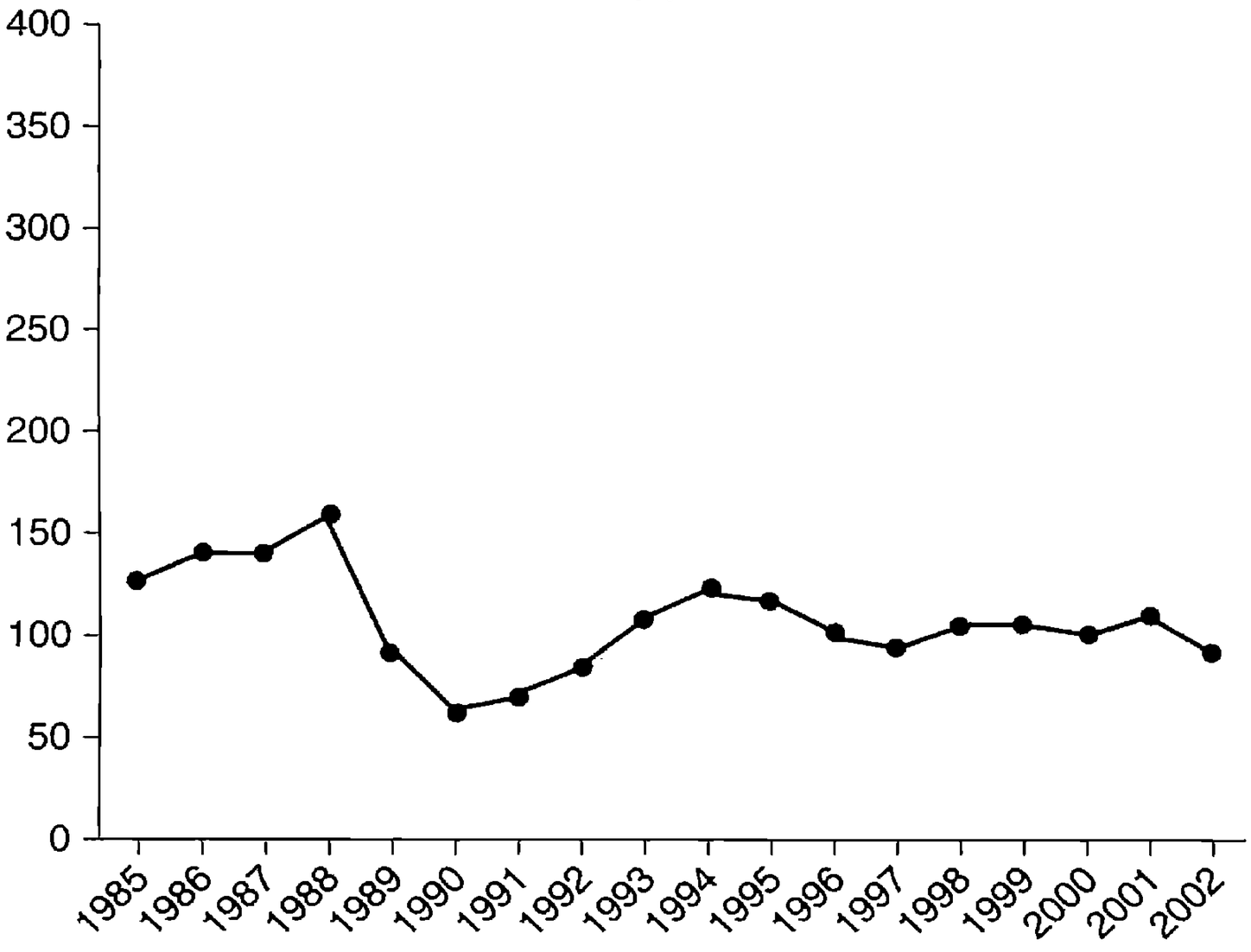
在圖 A 中的圖形顯示出劇烈的變化。在 1990 年時出現了陡峭的下滑趨勢，但是之後卻又急速上升，並且呈現出不穩定的狀態。而在圖 B 中的圖形則看起來比較穩定。從 1989 年到 1990 年之間的下滑狀態看起來比較平順，而且其他年份則幾乎都是維持在相當平穩的情況。這兩個圖都是依據完全相同的資料所繪製出來的，也就是從 1985 年到 2002 年之間的美國企業倒閉率。其中的 X 軸（在底部）所標示的年份是完全相同的。

至於 Y 軸的刻度，在圖 A 中是從 60 到 160，而在圖 B 中則是 0 到 400。圖 A 中的圖形之所以看起來變化比較劇烈，其實是 Y 軸的刻度所導致的。因此，當你在閱讀圖形時，一定要小心地去確認各軸的刻度，這是因為有些人會為了特定的目的而刻意選擇某種刻度，以使資料所展現出來的圖形看起來比較沒有變化、或者是呈現出戲劇性的變化。

圖A



圖B



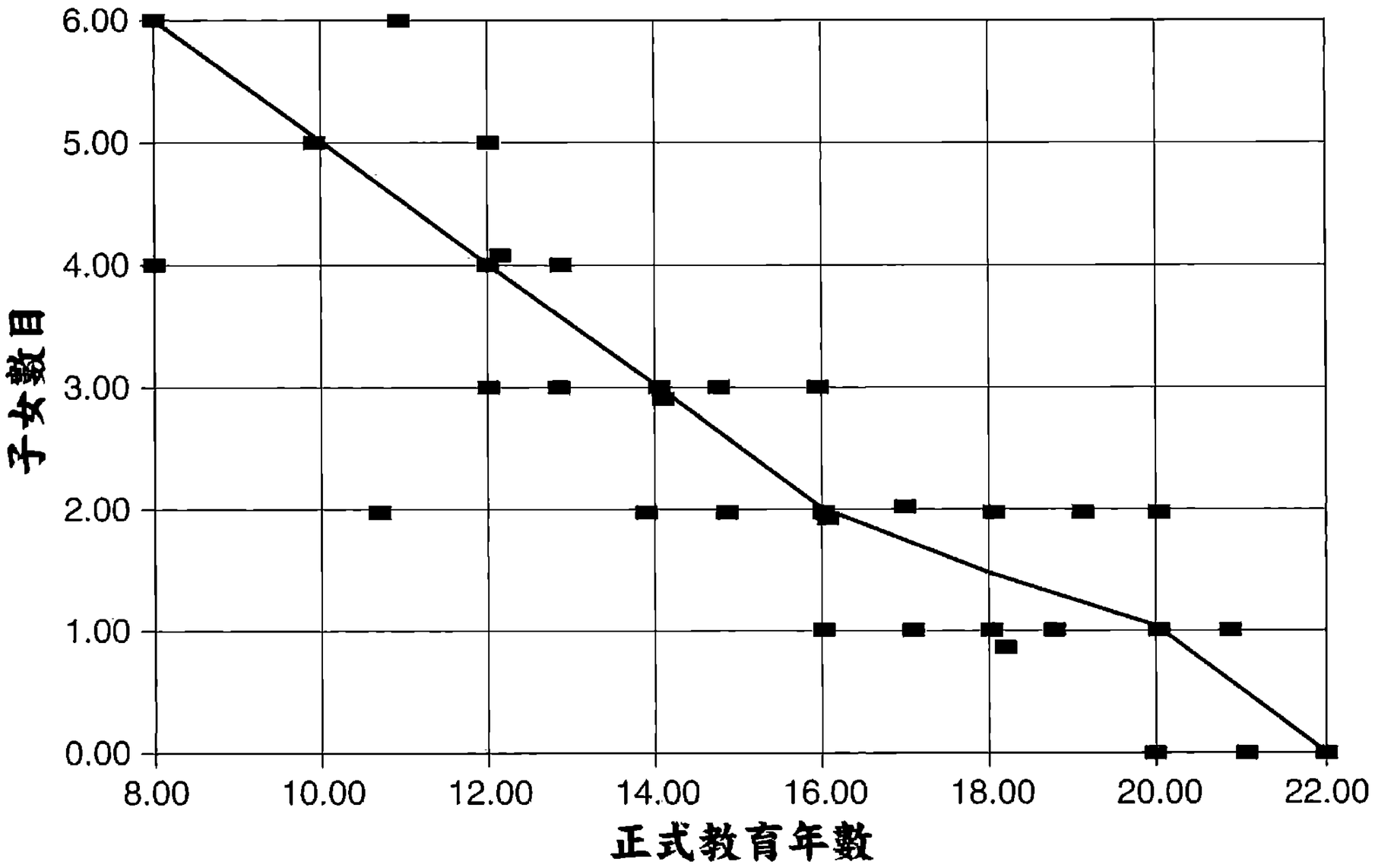


圖 12.5　散佈圖的範例：33 位女性的教育年數與子女數目之間的關係

式可言；或者，圖中的點也有可能形成一條與橫軸或縱軸完全平行的直線。而**線性關係**（linear relationship）所指的則是在由個案所畫出的點當中，你可以看出一條從一個角落延伸到另一個角落的直線。至於**曲線關係**（curvilinear relationship），則指在一群個案的中心位置，形成了一個 U 形的曲線，這個 U 形曲線的開口部分可能朝上、也可能朝下；或也可能形成一種像是 S 形的曲線。

2. *方向*（*direction*）　線性關係的方向可能是正向的、也有可能是負向的。正向關係的圖形看起來就像是一條從左下角延伸到右上角的斜線。在 X 變項上較高的值通常會與 Y 變項上較高的值一起出現，而且反之亦然。在收入與預期壽命的例子中，就顯現出此種正向的線性關係。至於負向關係看起來就像是一條從左上角延伸到右下角的斜線，而這表示在某一個變項上較高的值，會與在另一個變項上較低的值一起出現。例如，教育程度比較高的人，曾經被逮捕的次數就會比較少。如果我們看看一個針對一組男性的資料所畫成的散佈圖，其中的 X 軸代表學校教育的年數，而 Y 軸代表被逮捕的次數，那麼我們會看到，大多數曾經被逮捕許多次的個案（或者是男性），都集中在右下角的區域，這是因爲這些個案中大多數人接受學校教育的年數都比較少。而大多數幾乎未曾被逮捕過的個案則是會集中在左上角的區域，這是因爲在這些個案當中，有大部分的人都接受過較多的學校教育。此種根據

線性關係　兩個變項之間的關聯性，此種關聯性可能是正向的或是負向的，而且會出現在變項的所有水準中。當以散佈圖來畫出此種關聯性時，其圖形會形成一條沒有彎曲的直線。

曲線關係　在兩個變項之中的一種關聯性，當其中一個變項的值增加時，另外一個變項的值會顯示出一種特定的變化模式，像是先減少、然後增加、最後又減少。曲線關係並不是一種線性關係。

變項之間的關係所形成的虛構直線，其斜度可能很平緩、也可能很陡峭。較爲進階的統計學，會提供關於這條直線斜率的精確數值量數。

3. *關聯的強度（precision）*　雙變項關係的關聯性程度會有所差異，而所謂關聯的強度，指的是在散佈圖中各個點的離散程度。當圖形中的各個點緊密包圍著代表雙變項之間關係的那條直線時，就是具有較高的關聯強度。而當圖形中的各個點鬆散地分佈在直線的周圍時，其關聯強度就比較低。我們可以用「目測」的方式來看出關聯強度較高的關係，或者是使用進階的統計方法來測量某個關係的強度，就好比是在單變項統計中的標準差一樣。

雙變項表格

我們會在許多情境下去使用雙變項的**列聯表**（contingency table），此種表格以一種更爲濃縮的方式來展現出與散佈圖完全相同的資訊。列聯表勝過散佈圖的優勢之一，就是其中的資料可以使用任何一種測量尺度，不過，在使用區間尺度以及比率尺度時，資料必須先加以歸類分組才行。

雙變項的列聯表是依據**交叉列表**（cross-tabulation）（也就是同時列出兩個或是多個變項的表格）而來的。因爲在某個變項的每一個類別中的個案，都會再被分配到第二個（或者是額外）變項的每一個類別中，所以才被稱爲是「列聯」的。列聯表會把個案同時分配到數個變項的類別當中，並且顯示出在某個變項類別中的個案如何受到其他變項的類別所影響。

建構百分比例表　由個案的次數所組成的列聯表，其用處是有限的，因爲要利用個案的次數來看出型態或者是變項之間的關係，其實是相當困難的。利用把資料「標準化」、或者是把資料轉換成爲百分比的方式，即使不同個案的次數差異很大，我們也能夠輕易地看出變項之間的型態或者是關係。建構一個百分比例表並不困難，而且有好幾種方法可以使百分比例表看起來更具有專業的形象。我們首先來回顧一下親手建構一個表格的步驟，如果你是利用電腦來製作表格的話，這些相同的原則也同樣能夠適用。我們先從原始資料開始著手（請參閱在範例 12.2 原始資料以及次數分配，其中的資料是來自於一項虛構的調查）。

如果你親手建構一個表格的話，你可能會發現有一個在原始資料以及畫出表格之間的步驟是相當有用的。這個步驟就是去建立一個複合次數分配（compound frequency distribution, CFD），此種次數分配與一般的次數分配很類似，差別在於此種次數分配所計算的是兩個變項類別之間的每一種組合方式。例如，你想要去瞭解下列兩個變項之間的關係：其中之一是年齡，而另外一個則是對於可以喝酒的法定年齡的態度。年齡是一個比率變項，所以你必須對這個變項進行分組歸類，也就是把這個比率尺

列聯表　一種兩個或是多個變項的交叉列表的摘要格式，可以顯示出變項間的雙變項量化資料，其表現形式是某個變項的所有類別在橫列或是直行中的百分比例。

交叉列表　把兩個變項的資料放入一個列聯表中的過程，其目的在於顯示出在變項類別交叉位置的百分比或者是個案數目。

範例 12.2

原始資料以及次數分配

原始資料範例

個案	*年齡*	*性別*	*學校教育年數*	*態度*	*政黨，等等……*
01	21	女	14	1	民主黨
02	36	男	8	1	共和黨
03	77	女	12	2	共和黨
04	41	女	20	2	無黨派
05	29	男	22	3	民主社會黨
06	45	女	12	3	民主黨
07	19	男	13	2	遺失資訊
08	64	男	12	3	民主黨
09	53	女	10	3	民主黨
10	44	男	21	1	保守黨

（態度分數：1＝ 同意，2＝ 沒意見，3＝ 不同意）

兩個次數分配：年齡以及對於修改法定飲酒年齡的態度

年齡群組	*個案數目*	*態度*	*個案數目*
未滿 30 歲	26		
30—45 歲	30	同意	38
46—60 歲	35	沒意見	26
61 歲以上	15	不同意	40
遺失資訊	3	遺失資訊	5
總數	109	總數	109

複合次數分配：年齡群組以及對於修改法定飲酒年齡的態度

年齡	*態度*	*個案數目*
未滿 30 歲	同意	20
未滿 30 歲	沒意見	3
未滿 30 歲	不同意	3
30—45 歲	同意	10
30—45 歲	沒意見	10
30—45 歲	不同意	5
46—60 歲	同意	4
46—60 歲	沒意見	10
46—60 歲	不同意	21
61 歲以上	同意	3
61 歲以上	沒意見	2
61 歲以上	不同意	10
	小計	101
任一變項的遺失資訊		8
總數		109

度的變項當成順序變項來看待。而在百分比例表中，我們會把比率尺度以及區間尺度的資料加以分組歸類，以把此種資料轉換成爲順序尺度的資料。否則的話，我們光是一個變項可能就會有高達 50 個類別，而所製作出來的百分比例表可能根本就無法閱讀。

複合次數分配中包括了全部類別的所有組合方式。在上述的例子中，年齡有 4 個類別，而態度有 3 個類別，因此我們總共會得到 3×4 = 12 種組合。製作出複合次數分配的步驟如下：

1. 決定變項類別之間所有可能的組合方式。
2. 每當某個個案符合某種類別的組合方式時，就在該組合方式旁邊標上一個記號。
3. 把每個類別組合方式旁邊的記號數目加總起來。

如果沒有任何遺失資訊的問題，就把類別的數目相加（例如，把所有「同意」的數目相加、或者是把所有「年滿 61 歲以上」的數目相加）。不過，在上述的例子裡，其實有出現了遺失資訊的問題。在複合次數分配中，把四個「同意」類別的個案數目加總之後的總數是 37（20 + 10 + 4 + 3），但是在單變項次數分配裡的個案總數卻是 38，這兩個總數之間出現了不一致的情況，這是因爲在這 38 個個案中，有一個個案的年齡資訊遺失了。

下一個步驟就是利用把橫列和直行標上名稱的方式，來設定表格中的各個部分（請參閱圖 12.6）。自變項通常會被放在直行，但是有些人並不一定會依循此種約定俗成的作法。接下來，在複合次數分配中的每一個數字會被放在表格的細格（cell）當中，而所要放入的細格必須對應於變項類別的組合方式。例如，如同在圖 12.6 中的情形，複合次數分配顯示出，未滿三十歲、而且抱持同意態度的個案數目是 20 個（在表格中左上角的那個格子）。

圖 12.6 是原始計數表、或者又稱爲次數表，表格中的細格都包含了個案的總數。此種表格很容易製作，但是卻非常難以解讀，因爲橫列和直行的總數可能會不一樣。我們所真正感興趣的，是在和其他的細格進行比較之後，每一個細格中數字的相對大小。

我們可以利用三種方式來把原始計數表轉換成爲百分比例表：橫列的百分比例、直行的百分比例、以及總數的百分比例。前兩種經常被用來顯示出變項間的關係，至於總數的百分比例則幾乎從來都沒有人使用過，而且也無法輕易地揭露出變項之間的關係。

那麼，到底應該要以橫列、還是直行來計算百分比例才是最好的作法呢？其實任何一種都有可能是適當的。以下是製作百分比例表的技巧：在計算直行的百分比例時，你應該去計算每一個細格相對於直行總數的百分比，而這也包括了整個直行或者是**欄外值**（marginal），而欄外值所指的是某一橫列變項或者是直行變項的總數。例如，請先看一下在圖 12.6 中的直行欄外值。第一行的總數是 26（也就是總共有 26 個人的年齡未滿 30 歲），而在該直行的第一個細格內的數目則是 20（也就是年齡未滿 30 歲、而且抱持同意態度的人有 20 個），因此，該細格的百分

欄外值　在列聯表中列出總數的橫列或者是直行。

圖 12.6 年齡群組以及對於修改法定飲酒年齡的態度：原始計數表

原始計數表（a）

	年齡群組（b）				
態度（b）	未滿 30 歲	30－45 歲	46－60 歲	61 歲以上	**總數（c）**
同意	20	10	4	3	37
沒意見	3(d)	10	10	2	25
不同意	3	5	21	10	39
總數（c）	26	25	35	15	101

遺失的個案（f）= 8

（e）

表格中的各個部分

（a）給每個表格一個*標題*，在標題中說明變項名稱以及提供背景資訊。

（b）標記橫列變項以及直行變項的名稱，並對每一個變項類別給予一個名稱。

（c）這個部分包含了橫列以及直行的總數，而這些總數被稱為**欄外值**。欄外值等同於該變項的單變項次數分配。

（d）對應於每個變項類別的交叉位置的每一個數字或是格子就是**表格的細格**（cell of a table）。

（e）在標記的變項類別以及總數中的數字，就是**表格的主體**（body of a table）。

（f）如果有任何遺失的資訊（也就是發生了受訪者拒絕回答、提早結束訪談、或者是說「我不知道」等等情況的個案），把遺失資訊的個案的數目標記在表格附近，以便於計算出原始的個案總數。

比就是 20/26 = 0.769，也就是 76.9%。或者，以在橫列欄外值中的第一個數字而言（該數字爲 37），其百分比爲 37/101 = 0.366 = 36.6%。而這告訴你的是，在所有的個案中，有 36.6%的人是同意的。除了因爲四捨五入而造成的一些誤差之外，所有的百分比相加之後的總數應該等於 100%。

在計算橫列的百分比例時，計算的過程也相當類似，也就是去計算出每一個細格相對於橫列總數的百分比例。例如，利用其數字爲 20 的同樣一個細格，你現在要去計算的是這個細格中的數字 20 在橫列總數 37 中所佔的百分比例，因此，其百分比例爲 20/37 = 0.541 = 54.1%。除非欄外值完全相同，否則同樣一個細格在橫列或是直行中的百分比例是不同的。

橫列以及直行的百分比例可以讓你處理不同的問題。橫列的百分比例表所回答的是下列這個問題：在那些想要降低法定飲酒年齡的受訪者當中，每一個年齡群組所佔的百分比是多少？根據橫列的百分比例表，在表示同意的受訪者當中，有 54.1%的人是屬於未滿 30 歲這個年齡群組的。至於直行的百分比例表所處理的則是下列這個問題：在每一個年齡群組當中，抱持著不同的態度的人各自佔了多少的百分比例？根據直行的百分比例表，在未滿 30 歲的受訪者當中，有高達 76.9%的人抱持著同意的態度。從橫列的百分比裡，你得知在抱持同意態度的受訪者當中，有稍微高於一半的人是未滿 30 歲的。而從直行的百分比裡，你得知在未滿 30 歲的受訪者當中，

有超過 75%的人抱持著同意的態度。第一種百分比所告訴你的，是關於抱持特定態度的人們的資訊。而第二種百分比所告訴你的，則是在特定年齡群組中的人們的資訊，並且讓你可以去進行比較。

你的假設通常會告訴你要去看橫列百分比例表或者是直行百分比例表。在一開始的時候，你可能會想要把兩種百分比例都計算出來，並且練習去詮釋每一種百分比例所傳達的訊息。例如，你的假設是一個人的年齡會影響到他或她對於法定飲酒年齡的態度，而且你有興趣想要知道的，是哪一個年齡層的人們最支持以及最不支持降低法定飲酒年齡。如此一來，這就暗示著你應該要去看直行的百分比例表，因爲你想要去比較不同年齡群組之間的態度。然而，如果你感到興趣的，是去描述在每一個抱持不同態度的群組裡，其中的人們在年齡方面的組成情況，那麼橫列的百分比就是比較適合的。或許你想要去買下關於該項議題的電視廣告，而且你想要知道什麼樣的年齡群組會去收看這些電視廣告。正如同鄒塞（Zeisel, 1985: 34）所提過的，每當在交叉列表上的某個因素可以被視爲是另一個因素的起因時，若要得到最能夠提供有用訊息的百分比，就是去計算該原因因素方向的百分比。因此，如果年齡是你的原因變項，那麼你就應該去計算出橫列的百分比例表。

不幸的是，關於應該把自變項以及依變項放在百分比例表中的橫列或者是直行的位置、或者是應該列出橫列或直行的百分比，目前並沒有所謂的「業界標準」（“industry standard”）。大多數的研究者習慣上會把自變項放在直行，而且列出直行的百分比，但是也有不少人會把自變項放在橫列，並且列出橫列的百分比。

閱讀百分比例表　一旦你瞭解了如何製作百分比例表的過程，你將會發現，要閱讀以及理解百分比例表所傳達的訊息，就會變得容易多了。在閱讀百分比例表時，首先要去看標題、變項的名稱、以及任何的背景資訊。接著，再去看看計算百分比例時所使用的方向，也就是橫列或者是直行。請注意到在表 12.1 中的標題與名稱都是一樣的，之所以會出現這種情況，是因爲在表中所使用的變項是完全相同的。如果標題中包括了把資料轉換成百分比例的方式，那麼這個表格會更加容易理解，不過，在表 12.1 中並沒有這麼做。有時候你會看到一些較爲簡略的表格，其中省略了總數 100%或者是欄外值的部分，不過，此種作法只會徒然增加閱讀者的困擾而已。因此，當你在製作百分比例表時，最好還是把所有的部分都包含在內，並且使用清楚明白的標題。

當你在閱讀百分比例表時，你可以先看看在表中計算百分比例的方向爲何，然後在相對的方向進行比較。雖然此種作法聽起來讓人覺得很困惑，但是在實際去做的時候卻很簡單。基本的原則是，如果在表格中的百分比例是由上到下的方向來計算的話（也就是以直行來計算），那麼你就應該針對橫列來進行比較，同理可推，如果表格中的百分比例是由左到右的方向來計算的話（也就是以橫列來計算），那麼你就應該針對直行來進行比較。

舉例而言，像是在表 12.1 中的橫列百分比例表裡，你應該要去對直行或者是年齡群組進行比較。如此一來，你就可以看出，在抱持同意態度的受訪者當中，有大多數的人都是屬於最年輕的群組。而在各個年齡群組中的比例，則告訴我們，隨著

年齡逐漸增加，同意的比例也會跟著下降。大多數沒意見的人，都是屬於中年的群組，而抱持不同意態度的人，年紀則比較大一些，尤其是在 46 到 60 歲這個年齡群組中。而當你在閱讀表 12.1 裡的直行百分比例表時，你則應該要針對橫列來進行比較。如此一來，你就可以看出，在最年輕的年齡群組當中，絕大多數的人都是抱持著同意的態度，而且這也是唯一一個大多數人抱持同意態度的年齡群組。除此之外，相較於年齡最大的其他兩個年齡群組，在最年輕的這個年齡群組中，只有 11.5%的人是抱持不同意態度的。

想要從百分比例表中看出某項關係的存在，必須要多加練習才行。如果在某個百分比例表中沒有任何關係存在的話，不論是在橫列或者是直行中，每一個細格中的百分比例看起來幾乎都是相等的。若是存在著線性關係的話，在百分比例表上斜角方向的細格就擁有比較大的百分比例。如果存在著曲線關係的話，最大的百分比例會在細格中形成某個特定的型態。譬如說，百分比例最大的細格或許會出現在右上角、中間底部、以及左上角的細格裡。當表格大小適中（表格裡約有 9 到 16 個細格）、表格中的細格裡有一定數目的個案（建議至少要有 5 個個案）、而且其中的關係強度很高而且很清楚時，是最容易從百分比例中看出某項關係的存在的。

閱讀散佈圖的原則也能夠幫助你看出在百分比例表中的某項關係。想像一下把一張散佈圖分割成爲大小相等的 12 個區塊。再想像一下，如果把百分比例表重疊在散佈圖上的話，那麼在散佈圖上每個區塊中的個案數目，就會對應於百分比例表上

表 12.1　年齡群組以及對於修改法定飲酒年齡的態度：百分比例表

直行百分比例表

	年齡群組				
態度	*未滿 30 歲*	*30－45 歲*	*46－60 歲*	*61 歲以上*	**總數**
同意	76.9%	40.0%	11.4%	20.0%	36.6%
沒意見	11.5	40.0	28.6	13.3	24.8
不同意	11.5	20.0	60.0	66.7	38.6
總數	99.9%	100%	100%	100%	100%
（*N*）	（26）*	（25）*	（35）*	（15）*	（101）*

遺失的個案=8

橫列百分比例表

	年齡群組					
態度	*未滿 30 歲*	*30－45 歲*	*46－60 歲*	*61 歲以上*	**總數**	**(N)**
同意	54.1%	27%	10.8%	8.1%	100.0%	(37) *
沒意見	12.0	40.0	40.0	8.0	100.0	(25) *
不同意	7.7	12.8	53.8	25.6	99.9	(39) *
總數	25.7%	24.8%	34.7%	14.9%	100.1%	(101) *

遺失的個案=8

*在百分比例表當中，你應該要在總數 100%的旁邊區域，以括弧的方式提供計算百分比時所使用的個案總數或者是 *N*。如此一來，讀者們才能夠在百分比例表以及原始計數表之間來回進行比對查核。

每一個細格中的個案數目。你可以把百分比例表想成是散佈圖的濃縮形式。在散佈圖中的雙變項關係，會對應到百分比例表中在斜角線上的細格。因此，看出強烈關係的簡易方法之一，就是在橫列百分比例表中把每一個橫列中數字最大的百分比例圈起來、或者是在直行百分比例表中把每一個直行最大的百分比例圈起來，並且去看看是否會連成一條直線。

此種「圈出數字最大的細格」的規則相當有用，但是在此必須提醒你一件重要的事。在百分比例表中的類別必須是順序尺度或者是區間尺度，而且排列的順序要與散佈圖中的順序完全相同。在散佈圖中，數值最低的變項通常會被放在左下角的位置，如果在百分比例表中的類別順序與散佈圖中的順序不同的話，那麼「圈出數字最大的細格」的規則就會失去作用。

例如，表 12.2A 像是正相關，而表 12.2B 來則像是負相關。這兩個表所使用的是完全相同的資料，而且也都是橫列的百分比例表。然而，真正的關係其實是負相關。請你再仔細地看一下這兩個表，你會發現在表 12.2B 中的年齡類別排列順序才是和散佈圖中的排列順序完全相同的。當你有所疑慮時，你應該回頭去看看正相關和負相關之間的基本差異。正相關所指的是當某一個變項的值增加時，另外一個變項的值也會跟著增加。至於負相關所指的是當某一個變項的值增加時，另外一個變項的值則是會跟著減少。

表 12.2A　年齡與學校教育年數

	學校教育年數				
年齡	*0-11*	*12*	*13-14*	*16+*	總數
未滿 30 歲	5%	25	30	40	100
30－45 歲	15	25	40	20	100
46－60 歲	35	45	12	8	100
61 歲以上	45	35	15	5	100

表 12.2B　年齡與學校教育年數

	學校教育年數				
年齡	*0-11*	*12*	*13-14*	*16+*	總數
61 歲以上	45%	35	15	5	100
46－60 歲	35	45	12	8	100
30－45 歲	15	25	40	20	100
未滿 30 歲	5	25	30	40	100

消減誤差比例　在許多統計量數中所使用的邏輯，該邏輯所測量的是兩個變項之間的關聯強度。若是關聯強度很高，當研究者在使用自變項的資訊來預測依變項時，就能減少大多數的錯誤或是誤差。

無百分比例的雙變項表格　另外一種雙變項表格則會將資訊加以濃縮－也就是一種集中趨勢的量數（通常是平均數）。當某一個變項是名目尺度或者是順序尺度、而另一個變項則是以區間尺度或者是比率尺度來測量時，你就可以使用此種表格。區間尺度或者是比率尺度變項的平均數（或者是類似的量數）會出現在名目或者是順序尺度變項的每一個類別裡。不過，你不能從複合次數分配中去建構集中趨勢量數。相反地，你應該把個案分配到順序或是名目變項的類別中，然後從原始資料中去計算出每個變項類別中的所有個案的平均數。表 12.3 中列出了在每一個態度類別中，人們的年齡平均數。這些結果顯示出，那些抱持不同意態度的人們的平均年齡，遠遠高於那些抱持同意態度或是沒意見的人們的平均年齡。

關聯量數

關聯量數（*measure of association*）是指一個單一的數字，該數字能夠表達出某項關係的強度，而且通常也能表達出該關係的方向。關聯量數把關於某個雙變項關係的資訊濃縮成爲一個單一的數字。關聯量數的種類很多，至於何時應該使用哪一種正確的關聯量數，則是取決於資料的測量尺度，以及特定的研究目的。許多關聯量數都是以希臘字母來表示的，像是 λ（lambda）、γ（gamma）、τ（tau）、χ（chi）、以及 ρ（rho）等等字母都是被廣爲使用的量數。此處的重點在於如何去詮釋這些量數，而不是在討論如何計算出這些量數的方法與過程。如果你想要徹底瞭解每一個量數的話，你將必須修完至少一門統計學的相關課程。有一些關聯量數，像是 γ，是用於以順序尺度測量的資料上的（請參閱資訊箱 12.2 伽瑪值 γ）。至於其他的量數，像是相關係數（correlation coefficient）等等，則是針對以比率尺度所測量的資料（請參閱資訊箱 12.3 相關）。

大多數在此處所討論的基本量數，都是依循著一種**消減誤差比例**（proportionate reduction in error）的邏輯。此種邏輯所詢問的問題是，我們對於某一個變項所具有的知識，能夠減低多少我們在猜測另外一個變項的數值時所犯下的錯誤。*獨立*意味著我們對於某個變項的知識無法減少在另外一個變項上發生錯誤的機會。如果變項

表 12.3　對於修正法定飲酒年齡的態度與受訪者平均年齡

對於飲酒年齡的態度	平均年齡	個案數目（N）
同意	26.2	(37)
沒意見	44.5	(25)
不同意	61.9	(39)

遺失個案數目 ＝8

資訊箱 12.2　伽瑪值（γ）

γ 值（gamma）（中文音譯為「伽瑪」）是一個相對較為簡單的統計量數，主要是用於測量在兩個順序尺度變項之間的關聯強度。在使用這個雙變項量數時，你必須詳細指定在假設中何者是自變項、何者又是依變項。這個量數可以用來說明其他關聯量數的基本邏輯。

γ 值讓你可以依據某一個變項的等級，來預測另外一個變項的等級。在本質上，γ 值所回答的是下列的這個問題：如果你已經知道我在變項 1 上面的等級，那麼當你預測我在變項 2 上面的等級時，你的預測的正確性會有多高？例如，如果你知道我在數學這一科的分數等級，那麼當你預測我在文學這一科的分數時，正確性會有多高？完全預測或是 γ 值的最高可能分數是 +1 或者是−1，至於到底是+1 或−1，則是取決於兩個變項彼此的等級究竟是完全相同（正向）、或者是剛好相反（負向）。如果在兩個變項之間的關係是完全統計獨立的話，那麼 γ 值就會等於零。計算 γ 值的公式，所使用的是在交叉列表主體的細格中的資料。

讓我們看看一個簡單的例子，這個例子所使用的資料，是來自於 2008 年的美國成人的全國性樣本（綜合社會調查）的真實資料。總共有 672 人被詢問了關於他們的快樂以及健康狀態的問題。許多健康照護專業人員以及社會科學家都曾經提到，情緒方面的快樂感受與較為健康的狀態之間有一些關聯，因此，我們所要考驗的假設，就是快樂的人身體也會比較健康。

在我們先看一看原始計數表或是次數表之後，我們可以從欄外值看到，大多數的人都相當快樂，而且有更多人說他們目前的健康狀況相當良好。

	若把所有的狀況都考慮在內，你認為自己最近的心情如何？			
你認為自己的健康狀況是？	非常快樂	相當快樂	不太快樂	總數
極佳	63A	100D	19G	182
良好	93B	190E	53H	336
還可以或是不好	27C	77F	50I	154
總數	183	327	122	672

γ 值所依據的是「配對觀察」（“paired observation”）的概念（也就是說，以在自變項以及依變項上的相對等級來對觀察結果進行比較）。一致性（相同順序）配對觀察值所顯示的就是正相關，換句話說，當某個配對中的成員在自變項中的等級越高時，在依變項上的等級也會越高。不一致性（反向順序）配對觀察值則顯示出負相關。也就是說，當個配對中的成員在自變項上的等級越高時，在依變項上的等級就會越低。

γ 值的計算公式如下：

$$\gamma = [(P - Q)/(P + Q)]$$

其中 P = 一致性（concordant）配對，而 Q = 不一致性（discordant）配對。

γ 值的範圍是從−1.0 到 0 再到+1.0，而且也是一個消減誤差比例的統計量數。如果 γ = 0，就代表由自變項所提供的額外資訊，對於預測沒有任何的幫助。γ 值越高，在預測依變項時的正確性就越高。γ 值可能是正值、也可能是負值，而正負值所代表的是兩個變項之間關聯的方向。當一致性配對的數目較多時，γ 值將會是正數，而當不一致性配對的數目較多時，γ

值就將會是負數。

γ 值是針對在自變項以及依變項上具有一致性的細格（也就是相同的等級）、以及那些不一致的細格（也就是相反的等級）來進行比較，並且忽略那些等級相同的細格（也就是那些自變項以及依變項的等級相同的細格）。在前面所顯示出的表格中總共有 9 個細格，首先，讓我們一起來找出所有「一致性」配對的細格（為了讓你清楚地知道我們正在談論的是哪一個細格，所以在每一個細格中都加上了英文字母）。

在左上角的細格 A 與細格 E 是一致的，這是因為他們位於從左上到右下的斜角線，而在假設中已經預測了此種情況（也就是說，非常快樂的人會擁有極佳的健康狀況，而相當快樂的人則會擁有良好的健康狀況，以此類推）。其他的一致性配對還包括了 E 和 I、B 和 F，以及 D 和 H，其原因和上述的情況相同。而相反方向的不一致性配對，則包括了 G 和 E、E 和 C、D 和 B、以及 H 和 F。我們把在每一個配對當中的個案數目相乘，因此，一致性配對的計算公式為(A × (E + F + H + I)) + (D × (H + I)) + (B × (F + I)) + (E × I)，而不一致性配對的計算公式則為(G × (B + E + C + F) + (D ×(B + C)) + (H× (C+ F)) + (E ×C)。把每一個細格中的個案數目代入公式中，則會變成(63 × (190 + 77 + 53 + 50)) + (100 ×(53 + 50)) + (93 × (77 + 50)) + (190× 50) = 23310 + 10300 + 11811 + 9500 = 54,921 個一致性配對。此外，不一致性配對的數目則為(19 × (93 + 190 + 27 + 77)) + (100 × (93 + 27)) + (57× (77 + 27)) + (190× 27) = 7353 + 5512 + 12000 + 5130 = 29,995 個不一致性配對。把這兩個總數代入計算 γ 值的公式，則結果為：(54921−29995)/(54921 + 29995) = 0.2935。不過，我們通常可以利用電腦來完成計算的過程。而 γ 值為.2935 時，暗示了薄弱的正向關係，或者是健康情況與快樂程度之間的關聯程度不太強。

以下是對於各種 γ 值的詮釋（+表示正相關，−表示負相關）：

γ 值	意義
0.00 到 0.24	沒有關係
0.25 到 0.49	薄弱的關係（正向或負向）
0.50 到 0.74	中等強度的關係（正向或負向）
0.75 到 1.00	強烈的關係（正向或負向）

資訊箱 12.3 相關

對於大多數的人來說，計算相關係數（ρ）的公式看起來相當可怕。如果是親自動手去計算相關係數的話，尤其是當資料有好幾個位數時，可能會是一件非常耗時耗力的辛苦工作。因此，現在研究者們都是交給電腦去計算相關係數。然而，當研究者依賴電腦來進行計算工作時，所可能會產生的問題之一，就是研究者或許會不瞭解相關係數的真正意義。以下就用一個簡短的、簡化過的例子來說明如何計算相關係數的過程。

相關係數的用途在於顯示出兩個變項「一起出現」或者是共變的程度有多高。理想上，變項應該具有比率的測量尺度（有些人會使用區間尺度的變項）。在計算相關係數時，我們首先要把變項的每一個分數都轉換成為 Z 分數，這個步驟可以依據平均數和標準差來把變項加以「標準化」。接下來，我們要把每一個個案的兩個 Z 分數一起相乘，這個步驟可以告訴我們，

每一個個案中的兩個變項一起發生變化的程度，換句話說，對於那些在兩個變項的 Z 分數都很高的個案而言，其分數將會變得更高；而對於那些在兩個變項的 Z 分數都很低的個案而言，其分數則會變得更低。最後，我們則要把相乘過的 Z 分數的總和，除以個案的總數。如此一來，就會得到一種已經標準化之後的「平均」共變類型。簡而言之，相關係數就是把 Z 分數的平方值加總、並且除以個案總數之後所得到的結果。相關係數的數值範圍一定會介於+1.0 到−1.0 之間，並且是以一個單一的數字來總結在散佈圖上關於某個關係的資訊。

讓我們來看看在紅酒的年齡以及五種小瓶裝紅酒的價格之間的相關程度。首先，任何一個自認為心臟夠強、或者是沒有數學符號恐懼症的人，都可以看一看下列這個計算相關係數時最常被使用的公式之一：

$$(\Sigma\,[z\text{−score}_1][z\text{ −score}_2])/N$$

其中的 Σ 代表總和，z-score1 代表第一個變項的 Z 分數（請參閱資訊箱 12.1），z-score2 代表第二個變項的 Z 分數，而 N 則是代表個案總數。

以下是如何在不直接使用這個公式的情況下去計算出相關係數的方法：

			（差）		平方差		Z 分數		Z 分數
酒的種類	**年齡**	**價格**	*年齡*	*價格*	*年齡*	*價格*	*年齡*	*價格*	*乘積*
A	2	$10	−2	−5	4	25	−1.43	−0.70	1.00
B	3	5	−1	−10	1	100	−0.71	−1.41	1.00
C	5	20	+1	+5	1	25	0.71	+0.70	0.50
D	6	25	+2	+10	4	100	+1.43	+1.41	2.00
E	4	15	0	0	0	0	0.00	0.00	0.00
總數	20	$75			10	250			4.50

平均數：　年齡 ＝4；價格 ＝$15
變異數：　年齡 ＝10/5 ＝2；價格 ＝250/5 ＝50
標準差：　年齡 ＝2 的平方根 ＝1.4；價格 ＝50 的平方根 ＝7.1
相關係數：　4.5/5 ＝.90

步驟 1：計算出每一個變項的平均數以及標準差。（計算標準差的時候，首先用每一個分數減去其平均數，接著把相減之後的差予以平方，然後把平方差加總，最後則把平方差的總和除以個案總數，以得到變異數。然後再取變異數的平方根，即可得到標準差）

步驟 2：把變項的每一個分數都轉換為 Z 分數。（只要把每一個分數減去平均數，然後再除以其標準差，即可得到 Z 分數）

步驟 3：把每一個個案的兩個 Z 分數相乘。

步驟 4：把所有 Z 分數的乘積加總，並且把總和除以個案數目，即可得到相關係數。

之間是彼此獨立的話，那麼關聯量數就會等於零。

如果變項以及依變項之間有著強烈的關聯或者是關係，那麼當我們依據自變項的知識來預測依變項時，就比較不會犯下錯誤或是誤差，也就是說，此時消減誤差

的比例就會很大。當我們作出大量的正確猜測時，就暗示了如果在兩個變項之間存在著關聯的話，關聯量數就是一個不爲零的數字。表 12.4 描述了五個經常被使用的雙變項關聯量數，請注意到，大多數關聯量數的範圍都是從−1 到+1，其中負數表示負向關係，而正數則代表了正向關係。當一個量數是 1.0 時，就代表了消減誤差比例是 100%，也就是完全的預測。

超過兩個變項

統計控制

在瞭解資料時，證明兩個變項之間的關聯性，是相當重要的第一步。然而，就算證明了兩個變項之間具有關聯性，這並不代表你有充分的理由能夠說正是自變項影響了依變項。除了時間順序以及關聯性之外，我們還必須排除一些替代的解釋，因爲這些替代的解釋可能會使得我們所假設的關係變成是虛假的（spurious）。實驗研究者在排除替代解釋時所使用的方法，是去選擇一個特定的研究設計，該設計能夠完全控制對於結果的可能替代性解釋（也就是那些會威脅到內在效度的因素）。

在非實驗研究當中，我們可以利用統計的方式，以控制變項（稍後將進行討論）來控制替代性的解釋。我們以多變項的表格以及統計量數來檢視控制變項，而這些表格以及量數能夠幫助我們決定某個雙變項之間的關係是否是虛假的。我們也可以展現出多個自變項對於某一個依變項所產生的相對效果大小。

所謂的**控制變項**（control variable）是指代表了對某個雙變項關係的替代性解釋的第三個（或是第四個、第五個）變項。之所以被稱作是「控制」的，是因爲此種變項會針對在一個假設中除了主要的自變項與依變項之外的其他變項的效果進行調整、或者是將這些變項的效果列入考慮。例如，在你的雙變項表格中顯示出，身高較高的青少年比身高較矮的青少年更喜歡打棒球。但是在身高以及對於棒球的喜好之間的雙變項關係，卻有可能是一種虛假關係。爲什麼呢？這是因爲你猜想真正的原因可能是青少男的身高比青少女還要高，而且你也猜想男性比女性更加喜歡棒球。爲了測試這個關係是否真的是由身高所造成的，你必須控制性別這個變項。藉由控制性別變項的方式，你可以在統計方面移除性別所造成的影響。一旦你這麼做了，你就可以去看看在身高以及對棒球的喜好之間的雙變項關係是否仍然成立，或者在身高以及對棒球的喜好之間的關聯性是否真的是由性別所造成的。

控制變項　一種「第三個」變項，能夠顯示出某個雙變項關係是否為替代性解釋所造成的，此種變項發生的時間順序可能是在其他變項之前或是之間。

你可以去看看某個雙變項的關係在控制變項的類別裡是否依然存在，並以此種方式來「控制」（“control for”）第三個變項。例如，在你控制了性別變項之後，在身高以及對棒球的態度之間的關係仍然成立，這就表示身高較高的男性和女性，都比

表 12.4　五種關聯量數

λ（lambda）主要用於名目尺度的資料中。此種量數是根據消減誤差的模式而來的，而且其範圍介於 0（獨立）與 1.0（完全預測或者最強烈的可能關係）之間。

γ（gamma）主要用於順序尺度的資料中。此種量數是根據對變項類別的配對進行比較，並且看看某個個案在每一個不同變項上的等級是否相同。γ 值的範圍從−1.0 到+1.0，而 γ 值為 0 時則代表沒有任何關聯。

τ（tau）也是用於順序尺度的資料中。此種量數所根據的是與 γ 值不同的方法，而且能夠處理在使用 γ 值時所可能出現的一些問題。事實上，有好幾個統計量數都叫作 τ（這是一個相當受歡迎的希臘字母），而在此處所討論的是肯達爾的 τ 係數（Kendall's tau），其數值範圍是從−1.0 到+1.0，而數值為 0 時代表沒有任何關聯。

ρ（rho）也被稱為皮爾遜積差相關係數（Pearson's product moment correlation coefficient）（這是以知名的統計學家皮爾遜（Karl Pearson）來命名的，而其基礎則是來自於一種計算積差的統計程序）。這是一種最被廣為使用的相關性量數，當人們使用了相關（correlation）一詞，而且沒有再作更進一步的說明時，他們所指的就是這個相關性的統計量數。此種量數只能使用於以區間尺度或者是比率尺度所測量的資料中。ρ 也被應用於變項的平均數以及標準差，而且能夠告訴我們，不同的個案在散佈圖上與表示關係的直線（或是迴歸線）的距離有多遠。ρ 的範圍是從−1.0 到+1.0 之間，而若是 ρ 等於 0 時，則表示沒有相關。如果把 ρ 的值予以平方，有時也稱為 R^2，則平方之後的數值在消減誤差比例方面具有特殊的意義。R^2 能夠告訴我們，在某一個變項（例如：依變項）當中，有多少比例是能夠由另外一個變項（例如：自變項）來加以說明或是解釋的。ρ 只能夠去測量線性關係，無法去測量非線性的或者是曲線的關係。舉例而言，當 ρ 等於 0 時，所代表的意義可能是沒有任何關係、也有可能是曲線關係（請參閱資訊箱 12.3）。

χ^2（chi-squared）有二種不同的用途。其中之一是在敘述統計中被用來當作是一種關聯量數，就像上面所列出的其他量數一樣；或者，也可以被運用於推論統計當中。作為一種關聯量數，χ^2 可以被用於名目尺度或者是順序尺度的資料上。此種量數的上限是無窮大的，而下限則是 0，意思是指沒有任何關聯性可言（請參閱資訊箱 12.3）。

五種關聯量數摘要

量數	*希臘符號*	*資料尺度*	*高關聯性*	*獨立*
Lambda	λ	名目	1.0	0
Gamma	γ	順序	+1.0, −1.0	0
Tau(Kendall's)	τ	順序	+1.0, −1.0	0
Rho	ρ	區間，比率	+1.0, −1.0	0
Chi-square	χ^2	名目，順序	無限大	0

身高較矮的男性和女性更喜歡棒球。換句話說，也就是控制變項是沒有效果的。當出現了此種情況時，原本的雙變項之間的關係就不是虛假的，而且控制變項（有可能存在的替代性解釋）則是不具任何效果的。

如果在你控制了性別變項之後，原本的雙變項關係變弱或者甚至是消失的話，那又該怎麼辦呢？當出現了此種情況的時候，就代表了身高較高的男性沒有比身高較矮的男性更喜歡棒球，而且身高較高的女性也沒有比身高較矮的女性更喜歡棒

球。因此，這也暗示了原本的雙變項關係是虛假的，而且真正造成對於棒球態度產生差異的因素，其實是第三個變項（在上述的例子中是指性別），而不是身高。

在許多進階的統計技術當中，統計控制都是相當重要的觀念。像是相關係數之類的關聯量數只能暗示某種關係的存在，因此，一直到你考量了控制變項的問題之前，原本的雙變項關係都有可能是虛假關係。這就是爲什麼研究者在直到他們已經考量了所有的控制變項之前，都會以非常小心謹慎的態度來詮釋雙變項關係的緣故。

在你引進了控制變項之後，你就能夠看出自變項的**淨效果**（net effect），也就是在排除掉控制變項的影響之後，自變項對於依變項所能產生的真正效果。我們接著要來看看兩種引入控制變項的方法：三變項百分比例表以及多元迴歸分析。

百分比例表的詳析模型

建構三變項表格　爲了達到建立因果關係的必要條件，我們必須去進行「控制」、或者是看看某個替代性解釋是否會使得某個因果關係變爲不成立。如果某個替代性解釋能夠說明某個關係的話，那麼原本的雙變項關係或許會是一種虛假關係。我們會把替代性解釋加以操作化成爲第三個變項或者是控制變項。

藉由在三變項表格中以統計方法引入控制變項的方式，你就可以去針對第三個變項進行考量。三變項表格與雙變項表格之間只有些微的差異之處。在某種意義上，這兩種表格都是由多個雙變項表格所組成的。三變項表格的內容，是由自變項與依變項分別針對控制變項的每一個類別所製作的個別雙變項表格所組成的。依據你的自變項以及依變項所製作的多個表格（每一個表格都對應到控制變項的每一個類別），就是三變項表格的**局部表格**（partials）。這些表格會以局部的方式來列出控制變項的效果。局部表格的數目取決於控制變項中的類別數目。局部表格看起來就像是雙變項表格一樣，但是局部表格所使用的是個案的次集合。只有那些在控制變項上有特定數值的個案，才會出現在局部表格當中。因此，你可以把所有的局部表格結合起來，以回復到原本沒有加入控制變項時的雙變項表格。

三變項表格具有三種限制。首先，如果控制變項的類別超過四個以上的話，在詮釋三變項表格時就會變得相當困難。第二，控制變項可能屬於任何一種測量尺度，但是你必須把區間尺度或者是比率尺度的控制變項加以歸類分組才行（換句話說，就是把區間尺度或是比率尺度的變項轉換成爲順序尺度）。最後，個案的整體數目是一個具有限制作用的因素，因爲個案會被分配到局部表格的細格中。在局部表格中的個案總數，等於在雙變項表格中的細格數乘以控制變項類別數目的乘積。例如，假設某個控制變項有三個類別，而在雙變項表格中則有 12 個細格，如此一來，則在局部表格中總共將會有 3×12 =36 個細格。一般建議在每個細格中平均要有 5 個個案，因此，至少就應該要有 5×36 = 180 個個案。

淨效果　在利用統計的方式把可能產生影響的控制變項的效果移除之後，某一個變項（通常是自變項）對於另一個變項（通常是依變項）所產生的效果。

局部表格　在三個變項所組成的列聯表中，自變項以及依變項針對控制變項的每一個類別所產生的表格。

就像在建構一個雙變項表格時的過程一樣，在建構三變項表格時也是從複合次數分配開始著手，不過在此處所使用的不是兩個變項、而是三個變項的複合次數分配。在表 12.5 中所顯示的，是在表 12.1 的雙變項關係中加入了「性別」作爲控制變項所形成的三變項表格的範例。

正如同雙變項表格一樣，在複合次數分配中的每一種組合方式，都代表了在最終表格（在此是指局部表格）中的一個細格。每一個局部表格都有在原本的雙變項表格中的變項。對於三個變項的情況而言，建構出三個雙變項表格在邏輯上是有可能的。在表 12.5 的範例中的組合方式爲：（1）性別與態度、（2）年齡群組與態度、以及（3）性別與年齡群組。局部表格是以原本的雙變項關係爲基礎來建構的。在每一個局部表格中的自變項都是年齡群組，依變項是態度、而性別則是控制變項。因此，三變項表格就是由一對局部表格所組成的，其中每一個局部表格都顯示出某個特定性別的年齡/態度關係。

你的理論以及對於社會世界的理解，都會顯示出在原本雙變項關係中的假設、以及哪些變項或許會是替代性解釋（也就是控制變項）。

正如同在雙變項表格中一樣，複合次數分配能夠提供細格（此處是指局部表格中的細格）的原始計算次數。接下來，你利用與在雙變項表格中完全相同的方法，把這些原始次數轉換成爲百分比例（也就是把細格中的數目除以橫列或者是直行的總數）。舉例而言，在女性的局部表格中，左上角細格中的數字是 10，因此，該細格的橫列百分比就是 10/17 = 58%。

詳析典範（elaboration paradigm）是一種在閱讀三變項百分比例表時所使用的系統。[5] 此種典範描述了在你加入一個控制變項之後所可能出現的五種型態，這些型態所描述的是局部表格與原本雙變項表格之間該如何進行比較，或者是在你加入了控制變項之後，原本的雙變項關係會發生什麼變化（請參閱範例 12.3 詳析典範摘要）。在此處提出的型態範例所顯示出的是比較明顯的個案，但是當差異並不是那麼明顯時，你就可能必須使用到進階的統計技術。

在這五種型態當中，**複製型態**（replication pattern）是最容易理解的。當某個局部表格中的關係，複製或者是呈現出在加入控制變項前就已經存在於雙變項表格中的相同關係時，就是出現了複製型態，而且這也意味著控制變項是沒有任何效果的。**標明型態**（specification pattern）則是次一個最容易理解的型態。當某個局部表格複製了原本的雙變項關係，但是其他的局部表格並未出現相同的情況時，就是出現了標明型態。例如，你在交通事故以及大學畢業生之間發現了強烈的（負向）雙變項關係。你控制了性別變項，並且發現到此種關係只有對男性才成立（換句話說，此種強烈的負向關係

詳析典範　在增加了控制變項之後、並且對雙變項列聯表以及局部表格進行比較時，用來描述在不同表格之間的明顯型態的一種系統。

複製型態　在詳析典範中的一種型態，在此種型態中，局部表格所顯示出來的關係，與單獨只有自變項與依變項的雙變項列聯表中的關係完全相同。

標明型態　在詳析典範中的一種型態，在此種型態中，雙變項列聯表所顯現的關係只有出現在某個局部表格中，而在另外一個局部表格中則沒有出現。

表 12.5　三變項分析中的複合次數分配表以及表格

三變項表格的複合次數分配

男性			女性		
年齡	*態度*	*個案數目*	*年齡*	*態度*	*個案數目*
未滿 30 歲	同意	10	未滿 30 歲	同意	10
未滿 30 歲	沒意見	1	未滿 30 歲	沒意見	2
未滿 30 歲	不同意	2	未滿 30 歲	不同意	1
30－45 歲	同意	5	30－45 歲	同意	5
30－45 歲	沒意見	5	30－45 歲	沒意見	5
30－45 歲	不同意	2	30－45 歲	不同意	3
46－60 歲	同意	2	46－60 歲	同意	2
46－60 歲	沒意見	5	46－60 歲	沒意見	5
46－60 歲	不同意	11	46－60 歲	不同意	10
61 歲以上	同意	3	61 歲以上	同意	0
61 歲以上	沒意見	0	61 歲以上	沒意見	2
61 歲以上	不同意	5	61 歲以上	不同意	5
	小計	51	小計		50
遺失任何一個變項的資訊		4	遺失任何一個變項的資訊		4
男性個案總數		55	女性個案總數		54

男性的局部表格

	年齡群組				
態度	未滿 30 歲	30－45 歲	46－60 歲	61 歲以上	**總數**
同意	10	5	2	3	20
沒意見	1	5	5	0	11
不同意	2	2	11	5	20
總數	13	12	18	8	51

遺失的個案=4

女性的局部表格

	年齡群組				
態度	未滿 30 歲	30－45 歲	46－60 歲	61 歲以上	**總數**
同意	10	5	2	0	17
沒意見	2	5	5	2	14
不同意	1	3	10	5	19
總數	13	13	17	7	50

遺失的個案=4

範例　12.3

詳析典範摘要

型態名稱	*在比較局部表格以及原本的雙變項表格時所看到的型態*
複製	在兩個局部表格中的關係，與在雙變項表格中完全相同。
標明	雙變項關係只有出現在一個局部表格當中。
闡明	雙變項關係在局部表格中大幅地弱化或者甚至是消失（控制變項是中介變項）。
解釋	雙變項關係在局部表格中大幅地弱化或者甚至是消失（控制變項的發生時間早於自變項）。
抑制變項	雙變項之間不存在任何關係，關係只有出現在局部表格中。

詳析型態的範例

複製（百分比例表）

雙變項表格

	低	*高*
低	85%	15%
高	15%	85%

局部表格

	控制=低		*控制=高*	
	低	*高*	*低*	*高*
低	84%	16%	86%	14%
高	16%	84%	14%	86%

闡明或是解釋（百分比例表）

雙變項表格

	低	*高*
低	85%	15%
高	15%	85%

局部表格

	控制=低		*控制=高*	
	低	*高*	*低*	*高*
低	45%	55%	55%	45%
高	55%	45%	45%	55%

標明（百分比例表）

雙變項表格

	低	*高*
低	85%	15%
高	15%	15%

局部表格

	控制=低		*控制=高*	
	低	*高*	*低*	*高*
低	95%	5%	50%	50%
高	5%	95%	50%	50%

抑制變項（百分比例表）

雙變項表格

	低	*高*
低	54%	46%
高	46%	54%

局部表格

	控制=低		*控制=高*	
	低	*高*	*低*	*高*
低	84%	16%	14%	86%
高	16%	84%	86%	14%

只有出現在男性的局部表格中，並沒有出現在女性的局部表格中）。此種型態之所以被稱作是*標明*的，是因爲你必須去標示並且詳細說明在控制變項中出現了原本雙變項關係的類別爲何。

控制變項在闡明型態以及解釋型態中都具有強大的效果。在這兩種型態中，原本在雙變項表格中出現的關係，在局部表格中卻消失或者是大幅地減弱了。換句話說，你原本在雙變項表格中看出在自變項與依變項之間存在著某種關係，但是在局部表格當中，此種關係卻消失了，而且原來的變項似乎變成彼此獨立了。你無法單單只藉由看著表格就能夠去辨別出這兩種型態的差異。這兩種型態之間的差異之處，取決於控制變項在所有變項的因果順序中的位置。理論上，控制變項可能出現在下列兩個位置的其中一個：在原本的自變項以及依變項之間（也就是說，控制變項其實是中介變項）、或者是在原本的自變項之前。

闡明型態（interpretation pattern）所描述的情況，就是當控制變項的位置是介於原本的自變項以及依變項之間的時候。例如，你檢視了在宗教背景以及對於墮胎態度之間的關係，並且把政治意識型態當作是一個控制變項。你在推理之後認爲，一個人的宗教背景會影響到他或她目前的政治意識型態以及對於墮胎的態度。你據此提出了一個理論，認爲政治意識型態在邏輯上應該會早於對某個特定議題的態度（例如，對於墮胎的態度）。也就是說，宗教背景導致了政治意識型態，而政治意識型態則影響到對於墮胎的態度。此處的控制變項是一個中介變項，而且這個中介變項能夠幫助你闡明完整因果關係的意義。

解釋型態（explanation pattern）看起來與詮釋型態完全相同，其中的差異在於控制變項發生的時間順序。在解釋型態裡，控制變項發生的時間早於在原本的雙變項關係中的自變項。例如，原本的關係是存在於宗教背景以及對於墮胎的態度之間的，但是現在則加入性別當作控制變項。由於一個人的性別是在出生時就已經決定的，因此性別這個變項的發生時間就會早於宗教背景。解釋型態會改變研究者解釋結果的方式，因爲此種型態暗示了一開始的雙變項關係其實是虛假的（請參閱第六章當中關於虛假關係的討論）。

當原本的雙變項表格暗示兩個變項之間彼此獨立，但是在一個或是兩個局部表格中卻出現了某種關係時，就是產生了**抑制變項型態**（suppressor variable pattern）。例如，在雙變項表格當中，宗教背景和對於墮胎的態度是彼此獨立的。然而，在你引入了國家的宗教這個控制變項之後，你卻在局部表格中看到宗教背景與對於墮胎的態度產生了關聯。也就是說，控制變項抑制了真正的關係，而真正的關係則是出現在局部表格中。

闡明型態　在詳析典範中的一種型態，在此種型態中，雙變項列聯表顯示出了某個關係，但是在局部表格中卻顯示出沒有任何關係，而且在因果解釋中，控制變項是中介變項。

解釋型態　在詳析典範中的一種型態，在此種型態中，雙變項列聯表顯示出了某種關係，但是在局部表格中卻顯示出沒有任何關係，而且控制變項發生的時間早於自變項。

抑制變項型態　當雙變項表格顯示變項之間是彼此獨立的，但是在一個或是兩個局部表格中卻顯示出某種關係時，就是出現了抑制變項型態。

多元回歸分析

多元回歸分析是一種廣受歡迎的統計技術，不過此種技術的計算過程已經超出了本書的程度。雖然你只要利用適當的統計軟體就能夠迅速地計算出多元回歸，你仍然必須要具有統計學的背景，才能避免在計算以及詮釋過程中犯下錯誤。多元回歸必須要使用區間或者是比率尺度的資料。

多元回歸最大的優點之一，就是能夠同時針對數個控制變項（也就是替代性解釋）來進行調整。在使用百分比例表時，你幾乎無法同時使用超過一個以上的控制變項。除此之外，由於多元回歸已經被廣泛地使用，因此你很有可能會在閱讀研究報告或是文章時遇到多元回歸。多元回歸的結果可以告訴讀者兩件事。首先，多元回歸可以告訴讀者，自變項以及控制變項對於依變項的預測能力有多高。其中有一個統計量數，也就是 R^2，可以告訴我們某一組變項在「解釋」依變項時的能力有多好。此處所謂的*解釋*（*explain*），意思是指當我們根據關於自變項的資訊來預測依變項的分數時，犯下的錯誤會比較少。一個有數個變項的良好模型，也許能夠說明、或者是解釋依變項裡大部分的分數。例如，當 R^2 等於 0.50 時，意思是指一旦我們知道了自變項以及依變項時，在預測依變項時的正確率可以提高 50%，而且如果你在不知道自變項以及控制變項的情況下去預測依變項的話，你有一半的機會將會犯下錯誤。

其次，多元回歸的結果，可以告訴我們每一個變項對於依變項所造成的效果的方向以及大小，而且所產生的效果是以數值來精確地測量的。當數值越高時，則某個變項在預測依變項時的效果越強。而效果前面的符號（正號或是負號）則能夠告訴你該變項對於依變項所造成的效果的方向。例如，你可以看到五個自變項或是控制變項如何同時影響依變項，而且是在所有變項都控制了彼此效果的情況下。當我們在考驗一些關於多個自變項影響了單一依變項的理論時，此種能力是特別有用的（請參閱第三章中關於因果關係圖的範例）。

我們利用一個標準化的回歸係數、或者是希臘字母 β，來測量來自變項或者是控制變項對於依變項的效果或影響。此種回歸係數與相關係數類似，而且其範圍是從 0 到+0.99 或是－0.99，其中 0 表示沒有效果。我們可以利用統計考驗來決定一個係數的統計顯著性（將在本章稍後部分進行討論）。兩個變項之間的 β 係數就等同於相關係數。

我們利用 β 回歸係數來決定控制變項是否具有任何效果。例如，在 X 和 Y 之間的雙變項相關是 0.75。接著，我們以統計的方式加入了四個控制變項。如果 β 值仍然維持在 0.75，那麼這四個控制變項就都沒有任何效果。然而，如果 X 與 Y 的 β 值變小了（例如，下降到 0.20），就表示控制變項對依變項具有效果。

讓我們考量一個回歸分析的範例，其中的年齡、收入、參與宗教活動、教育、以及居住地區都是自變項，而依變項則是在一個政治意識型態指數上所得到的分數。多元回歸的結果顯示出，收入以及參與宗教活動具有大型的效果，教育與居住地區具有小型的效果，而年齡則是沒有任何效果。把所有的自變項者加在一起時，

推論統計　一種以隨機抽樣為基礎的應用數學的分支，此種統計允許研究者作出精確的陳述，以說明他們有多高的信心程度去認為樣本中的量數與母體母數是完全相同的。

在預測一個人的政治意識型態方面的正確率爲 38%（請參閱範例 12.4 多元回歸結果的範例）。[6] 這個範例顯示出，高收入、經常參與宗教活動、以及居住在南方地區等等因素，都與偏向保守派的意見有正相關，而教育程度較高則與偏向自由派的意見具有關聯性。收入所具有的效果，是居住在南方地區所產生的效果的兩倍之多。

圖表 12.2 總結了敘述統計的類型以及技術。接下來，我們要把注意力轉向推論統計。

推論統計

推論統計的目的

到目前爲止，在本章中所討論到的統計概念，都是屬於敘述統計的範圍。但是我們想要做到的，通常不只是想要描述或是敘述而已，我們還想要去考驗假設，去發現樣本的結果在母體中是否仍然成立，以及決定結果（例如，在兩個群組之間的平均分數）是否大到足以指出某個關係確實存在、而不是單獨由於機會因素所造成的。**推論統計**（inferential statistics）能夠利用機率理論以正式的方式來考驗假設，允許從樣本對母體進行推論，並且測試敘述的結果是否是由隨機因素所造成的、或者是由一個真正的關係所造成的。以下的部分將會解釋推論統計的基本觀念，但是不會深入地去探討推論統計的任何細節，這是因爲推論統計這個領域比敘述統計更爲複雜，而且也必須要有統計學的背景知識才能理解。

統計推計所依賴的是來自於機率抽樣的原則，而所謂的機率抽樣，意思是指我

範例　12.4

多元回歸結果的範例

依變項是政治意識型態指數（分數越高則態度越偏自由派）

自變項	*標準化的回歸係數*
居住地區=南方	−.19
年齡	.01
收入	−.44
接受教育的年數	.23
參與宗教活動	−.39
	$R^2 = .38$

圖表 12.2　敘述統計主要類型摘要

技術類型	統計技術	目的
單變項	次數分配、集中趨勢量數、標準差、Z 分數	描述單一變項
雙變項	相關、百分比例表、卡方	描述兩個變項之間的關係或者是關聯性
多變項	詳析典範、多元回歸	描述在數個變項之間的關係，或者是去檢視數個自變項如何對一個依變項產生影響。

們利用一個隨機的過程（例如，利用亂數表、隨機電腦過程等等），從整個母體中挑選出一些個案來當作樣本。當我們在談論我們具有多大的信心可以從樣本的結果去推論母體時，推論統計就是一種精確的表達方式。

如果你曾經讀過或者是聽過「統計顯著性」或者是「在 0.05 的水準上具有顯著性」的結果，那麼你就已經接觸過推論統計了。我們會利用推論統計來進行各種不同的統計考驗（例如，*t* 考驗或是 *F* 考驗）。我們會在正式的假設考驗中使用統計顯著性，而在決定是否要接受或者是拒斥虛無假設時，這是一種精確的表達方式。[7]

統計顯著性

具有統計顯著性的結果（*statistically significant results*）這個詞彙的意思是指這些結果不太可能是由於機會因素所造成的。**統計顯著性**（statistical significance）所指的是在樣本中發現某個不存在於母體中的關係的可能性。由於機率樣本牽涉到隨機過程，因此樣本結果永遠都有可能會與母體母數有所差異。我們想要去估計出樣本結果有多大的機率是由於真正的母體母數、或者是由於隨機抽樣的機會因素所造成的。利用某種來自於數學以及特定統計考驗的機率理論，我們就能夠去分辨出某些結果（例如：關聯性、兩個平均數之間的差異、回歸係數等等）究竟是由於隨機抽樣過程中的隨機誤差所造成的、或者是顯示出了真正發生在社會世界中的一些效果。

統計顯著性能夠告訴我們的只是一種可能性，也就是說，統計顯著性無法以絕對肯定的方式來證明任何事物。統計顯著性所敘述的，是某些特定結果出現的機率有多高或多低。除此之外，統計顯著性也不同於實際的、實質的、或者是理論的顯著性。某些結果可能在統計上具有顯著性，但是在理論上卻是無意義的、或者是微不足道的。例如，某兩個變項可能純粹因爲巧合而在統計上具有顯著的關聯，但是在彼此之間其實並沒有任何邏輯性的連結（例如，指甲的長度與說法語的能力）。

統計顯著性　一種可能性，所指的是在樣本結果中的發現或是統計關係有多大的可能性是由於隨機因素、而非是因為存在於整個母體中的實際關係所造成的。

顯著水準

我們通常會以顯著水準來表達統計顯著性（例如，某個考驗在某個特定的水準上具有統計顯

統計顯著水準　一組數字。當研究者在測量某個統計關係是由隨機因素所造成、而不是由變項之間真正存在的關係所造成的機率有多高時，可以利用這組數字來簡單地表達出此種機率。

著性），而不是直接提供特定的機率。在談到某些結果有多大的機率是由機會因素所造成的時候（也就是說，出現在樣本中的某個關係，根本不存在於母體中），**統計顯著水準**（level of statistical significance）（通常是.05、.01、或者是.001）是一種簡便的表達方式。當我們說某些結果的顯著水準是.05 的時候，我們所指的意思如下：

- 此種結果若是由機會因素所造成的話，其發生機率只有 5%。
- 有大約 95%的機率，樣本結果並不是單獨由於機會因素所導致的，而是正確地反映出母體的情況。
- 單獨由機會因素而導致此種結果的機率是.05，或者是 5%。
- 我們可以有 95%的信心，認爲結果是由母體中的真正關係所導致的、而不是由機會因素所造成的。

上述的這些說法，其實只是以不同的方式來陳述完全相同的事情而已。此種情況聽起來有點像是在我們討論抽樣那一章裡，對於抽樣分配以及中央極限定理的相關討論一樣。其實這並不是巧合。這是因爲兩者都是以機率理論爲基礎的，而我們則是利用機率理論來把樣本資料以及母體連結在一起。機率理論讓我們能夠去預測，當使用隨機過程時，在經過了一段長時間以及許多事件之後，會發生什麼狀況。因爲我們只有一個樣本，而且我們想要推論到母體，因此機率理論可以幫助我們去估計我們的特定樣本能夠代表整個母體的機率有多高。除非我們得到整個母體的資訊，否則我們永遠無法確定地知道真實的情況爲何，但是機率理論則允許我們去說明我們所具有的信心程度，換句話說，也就是樣本中的結果有多大的機率會與母體中的實際情況有所差異。

例如，某個樣本顯示出大學男生與女生在念書時間方面有一些差異。那麼，此種結果究竟是不是實際上在母體中根本沒有任何差異、只是因爲選擇了異常的樣本所導致的？或者，此種結果確實反映了大學男生與女生之間的真正差異？（請參閱範例 12.5 卡方檢定）

第一類型與第二類型錯誤

統計顯著性的邏輯所依賴基礎，是機會因素是否可能產生某種結果。你或許想要問，爲什麼要使用.05 的顯著水準呢？當我們使用這個水準時，意思是說由隨機性造成某種結果的機率只有 5%而已。那麼，爲什麼我們不使用一個更加確定的標準，像是 1000 分之 1 的隨機機率呢？這麼一來的話，因爲隨機性、而非因爲某個真正的關係而造成某種結果的可能性就會更小了。

關於此種想法，目前有兩個答案。比較簡單的答案，是科學社群已經約定俗成地同意把.05 當成是應用於大部分目的時的基本原則。對於結果具有 95%的信心，

範例　12.5

卡方檢定

卡方（chi-square, x^2）有兩種使用的方式，而此種情況則造成一些混淆以及困惑。在當作一種敘述統計量數（descriptive statistic）時，卡方能夠告訴我們在兩個變項之間的關聯強度；而在當作一種推論統計量數（inferential statistic）時，卡方則能告訴我們，我們所發現的任何關聯，有多高的機率是單獨由機會因素所造成的。在檢視以名目尺度或者是順序尺度所進行測量的變項時，卡方是一種被廣泛使用、而且相當強而有力的工具。在決定雙變項百分比例表中是否有任何關聯存在時，相較於單純用「目測」的方式，卡方是比較精確的表達方式。

在邏輯上，我們首先要決定在某個表格中的「期望值」（"expected value"），不過，我們只能單獨依據欄外值的資訊來決定期望值。你應該還記得，欄外值其實就是每一個變項單獨的次數分配。我們可以把期望值想像成在沒有檢視表格整體資訊的情況下，我們所擁有的「最佳猜測」（"best guess"）。接著，我們則要去檢視資料，看看資料與「期望值」之間的差異性有多大。如果在表格中的資料與期望值完全相同、或者是非常接近期望值的話，那麼在兩個變項之間就沒有任何關聯，也就是這兩個變項是彼此獨立的。換句話說，獨立意味著在表格中「所發生的狀況」，正是我們單獨依據欄外值的資訊時所預期會看到的狀況。當變項之間彼此獨立時，卡方等於零；而當變項之間的關聯越強時，卡方的值也會隨著增加。如果表格中的資料與期望值的差異非常大時，那麼我們就可以得知，在表格中「所發生的狀況」已經超越了我們單獨依據欄外值的資訊所預期會看到的狀況（換句話話，就是在兩個變項之間具有關聯）。請參閱下列關於身高與成績之間關聯的範例。

原始或是觀察的資料表

在研究方法課程的成績

學生的身高	C	B	A	總數
高	30	10	10	50
中等	10	30	10	50
矮	30	20	50	100
總數	70	60	70	200

期望值表

期望值 ＝（直行總數×橫列總數）／（整體總數）。例如，在原始資料表中左上角細格的期望值為：（70 × 50）／200 = 17.5

在研究方法課程的成績

學生的身高	C	B	A	總數
高	17.5	15.0	17.5	50.0
中等	17.5	15.0	17.5	50.0
矮	35.0	30.0	35.0	100.0
總數	70.0	60.0	70.0	200.0

差異表

差異值=（觀察值－期望值）。例如，同樣以表格中的左上角細格為例，其差異值為（30−17.5）= 12.5

在研究方法課程的成績

學生的身高	C	B	A	總數
高	12.5	−5.0	−7.5	0.0
中等	−7.5	15.0	−7.5	0.0
矮	−5.0	−10.0	15.0	0.0
總數	0.0	0.0	0.0	0.0

卡方 =先把每個差異值予以平方，然後再除以每個細格中的期望值，最後把所有的數值加總起來即可得到卡方值。以表格中左上角的細格為例，其計算過程為：先將 12.5 予以平方，得到 156.25，然後再除以 17.5，即得到 8.93。接著重覆這個過程以求得每一個細格的值，最後，計算卡方值為：

卡方 = 第一橫列（8.93 + 1.67 + 3.21）+
第二橫列（3.21 + 15 + 3.21）+
第三橫列（.71 + 3.33 + 6.43）= 45.7

因為卡方值並不是零，所以在資料中的變項並不是彼此獨立的，也就是變項之間具有關聯。卡方係數無法告訴我們關聯的方向（例如，負向關聯）。就推論統計量數而言，我們必須使用卡方表或者是電腦程式來評估關聯的強度（也就是說，去瞭解此種大型的卡方值單獨由機會因素所造成的機率有多高）。在不去深入去探討卡方表所有細節的情況之下，此種關聯是相當罕見的，也就是說，純粹因為機會因素而發生此種關聯的機率只有 1,000 分之 1。對於有 9 個細格的表格而言，在.001 的水準上，卡方值為 45.7 是達到統計顯著性的。

已經是研究者在解釋社會世界時可以被廣爲接受的標準。第二個、也是比較複雜的答案，則是牽涉到在第一類型錯誤與第二類型錯誤之間的權衡取捨。當我們認爲某個關係存在、但是事實上此種關係並不存在時，就是犯下了**第一類型錯誤**（Type I error）。而當我們認爲某個關係不存在，但是事實上此種關係卻是確實存在時，就是犯下了**第二類型錯誤**（Type II error），而這也意味著我們錯誤地接受了虛無假設（請參閱表 12.6）。當然，我們都想要同時避免犯下這兩種錯誤，也就是說，只有當某種關係確實存在時，我們才說該項關係存在，或者是只有當確實沒有任何關係存在時，才說某種關係不存在。然而，我們在此卻會遇上一個進退爲難的處境：當犯下某一類型錯誤的可能性減少時，犯下另一類型錯誤的可能性卻會隨之增加。

一開始的時候，你可能會覺得第一類型錯誤和第二類型錯誤的觀念有點難懂，但是同樣的邏輯困境也會出現在研究情境之外。例如，陪審團可能會錯誤地認爲某個被告是有罪的、但是事實上這名被告卻是無辜的，或者，陪審團也有可能會錯誤地認爲某個人是無罪的、但是事實上這個人卻是有罪的。陪審團並不想犯下上述的任何一種錯誤，也就是陪審團不想把無辜的人送進監獄、或是讓有罪的人逍遙法外，但是陪審團仍然必須運用有限的資訊來作出判斷。同樣地，製藥公司也必須決定是否要販售一種新藥。該公司可能錯誤地宣稱此種新藥沒有任何副作用，但是事實上此種新藥卻有造成失明的副作用。或者，該公司可能因爲害怕有嚴重的副作用而決定不要販售某種藥品，但是事實上根本沒有任何副作用。這家製藥公司也同樣不想犯下任何一種錯誤。如果這家公司犯下第一種錯誤，將有可能面臨法律訴訟而且傷害到使用藥物的人們。而第二種錯誤則會阻止這家公司去販售一種可能具有療效並且帶來龐大利益的新藥。

第一類型錯誤　當研究者認為某個關係存在、但是事實上卻是不存在時，就是犯下了此種錯誤，也就是錯誤地拒絕了虛無假設。

第二類型錯誤　當研究者認為某個關係不存在、但是事實上卻存在時，就是犯下了些種錯誤，也就是錯誤地接受了虛無假設。

表 12.6　第一類型以及第二類型錯誤

研究者的陳述	在社會世界中的真實情況	
	沒有關係	*具有因果關係*
沒有關係	沒有錯誤	第二類型錯誤
具有因果關係	第一類型錯誤	沒有錯誤

接下來，讓我們把統計顯著性以及二種類型錯誤的觀念結合起來：如果你過度謹慎，並且設定了一個非常高的顯著水準，你就有可能會犯下其中某一類型的錯誤。舉例而言，假設你使用了.0001 的顯著水準。也就是說，只有當某種結果出現的機率罕見到每 10,000 次中只出現一次時，你才會把此種結果歸因爲機會因素所造成的。這麼高的顯著水準，意味著你極有可能會錯誤地認爲某種結果是由機會因素造成的，但是事實上卻並非如此。你可能會錯誤地接受虛無假設，但是事實上變項之間確實存在著某種關係（也就是犯下了第二類型的錯誤）。對比之下，如果你是個願意承擔風險的研究者，並且設定了一個像是.10 之類的低顯著水準，那麼就代表了你所得到結果純粹由機會因素所造成的機率是 10%。如此一來，你就有可能會錯誤地認爲變項之間存在著某種因果關係，但是事實上真正造成此種結果的，卻是一些隨機因素（例如，隨機抽樣誤差）。換句話說，你有可能會錯誤地拒絕虛無假設（也就是犯下了第一類型錯誤）。總而言之，.05 的顯著水準其實是在第一類型錯誤以及第二類型錯誤之間作出折衷取捨之後的妥協結果。

這個部分的內容已經大致描述了推論統計的基本原理以及觀念。此種統計技術是相當精確的，而且是以抽樣誤差、樣本大小、以及中央極限定理之間的關係作爲其基礎。推論統計的強大之處，就在於它能夠讓我們以明確的確定程度，來陳述特定的樣本結果在母體中爲真的可能性。例如，你進行了統計考驗並且得知當顯著水準爲.05 時，某種關係具有統計上的顯著性。因此，你可以說樣本結果很有可能不是由機會因素所造成的。更確切地說，也就是某個真實關係存在於社會世界的機率高達 95%。推論統計的考驗雖然相當有用，但是仍然有其限制，像是資料必須來自於隨機樣本、而且在進行考驗時也只有考量到抽樣誤差而已等等。非抽樣的誤差（例如，一個不佳的抽樣架構、或者是一個設計不良的量數）並未被考量在內。此外，也不要天真地以爲此種考驗能夠提供簡單的、最終的解答。請參閱資訊箱 12.4 電腦上的統計程式裡的相關討論。

結論

本章討論了如何組織量化資料以作好進行分析的準備工作，以及如何分析這些資料（把資料組織成爲圖表的格式，或者是利用統計量數來摘要這些資料中的資訊）。我們利用統計分析來考驗假設並且回答研究問題。你已經看到資料應該要如何

資訊箱 12.4　電腦上的統計程式

幾乎每一個必須計算許多統計量數的社會研究者，都會利用電腦程式來完成計算的工作。我們可以利用一些像是 Excel 之類的基本試算表（spread-sheet）程式來計算一些統計量數。不幸的是，試算表主要是以會計以及簿計的目的來設計的，因此，雖然此種程式會包含一些統計功能，但是這些統計功能其實既不強大、能力也很有限。目前有許多電腦程式是專門設計來計算一般的統計量數的。對於初學者而言，可能會覺得這些統計程式令人感到眼花撩亂，因為隨著電腦科技的進展，這些產品也在迅速地演變當中。而在十幾、二十年前的時候，人們必須懂得電腦語言、或者是能夠寫一些基本的程式，才能夠利用電腦來計算統計量數。

在近幾年來，對於使用者來說，統計軟體已經變得越來越友善、也越來越容易使用。目前在社會科學中最受歡迎的統計程式是 Minitab、Microcase，以及 Statistical Package for the Social Sciences（SPSS）。其他的統計軟體還包括了 Statistical Analysis System（SAS）、BMPD（已經被 SPSS 公司所收購）、由 StratSoft 所推出的 STATISTICA、以及 Strata。其中有許多在一開始時都是為了研究目的而設計出來的簡易的、低成本的統計程式。時至今日，其中有許多軟體都是由私人公司所設計與販售的，而且這些公司也很有興趣把一些精密複雜的軟體產品販售給許多不同的企業以及政府使用者。

在社會科學中最被廣為使用的統計程式是 SPSS。這套軟體的優點在於，社會科學家已經廣泛地使用這套軟體超過三十年之久、這套軟體包含了許多操弄量化資料的方法、而且其中也包含了大多數的統計量數。而這套軟體的缺點，則是因為其中有太多選項以及複雜的統計技術，所以必須要花上很長的時間才能夠學會如何使用這套軟體。除此之外，購買這套軟體的費用很高，除非是購買一種較為便宜的、「精簡」過的學生版本，其中還包含了一本教材以及練習簿。

隨著電腦科技使得運用統計程式變得更為容易，另外一種風險也正在增加當中，換句話說，有些人也許會操作這些程式，但是卻不瞭解某些統計量數的真正意義、或者是這些程式的運作方式。這些人可能會在不知不覺之間就違反了某種統計程序所要求的基本假定，或是以不適當的方式使用這些統計技術，並且產生一些毫無意義、但是看起來在技術上非常複雜的統計結果。

進行編碼，然後利用單變項或者是雙變項統計技術來進行分析。由於雙變項之間的關係可能是虛假的，所以通常必須要運用到控制變項以及多變項的分析技術。除此之外，你也學到了一些關於推論統計的基本觀念與原理。

對於剛開始進行研究的研究者而言，如果他們所得到的結果不支持某個假設時，有時候他們會認爲有可能是因爲他們在過程中犯下了一些錯誤所致。事實上，拒絕一個假設並沒有什麼不對。科學研究的目標，就是要產生能夠真正反映社會世界的知識，而不是去捍衛一些敝帚自珍的想法或者是假設。一個假設只是依據有限的知識所作出的理論性猜測，因此，假設必須要經過檢驗。品質絕佳的研究可能會發現某個假設是錯誤的，而品質不佳的研究則可能會去支持某個假設。良好的研究

取決於高品質的方法論，而不是取決於支持某個特定的假設。

良好的研究意味著要去減少當我們從資料對社會世界作出真正的推論時所可能遇到的錯誤或者是障礙。在研究過程當中，錯誤可能會發生在許多階段當中，並且會影響到研究結果。也就是說，在研究設計、測量、資料蒐集、編碼、計算統計量數以及建構表格、或者是詮釋結果的階段中，都有可能會出現錯誤。即使你能夠毫無錯誤地設計研究、進行測量、蒐集資料、進行編碼、以及進行計算，你仍然必須完成研究過程中的另外一個步驟：對表格、圖形、以及統計量數提出詮釋，並且回答下面這個問題：這些事物的意義是什麼？爲事實、圖形、表格、或者是統計量數賦予意義的唯一方式，就是去善加運用理論、洞見、以及理解力。

單單只有資料、表格、以及電腦所計算得到的結果，並不能夠去回答研究問題。事實並不會爲自己說話。身爲研究者，你必須回到你的理論上（也就是各種概念、概念之間的關係、假定、以及理論性定義等等），並且爲結果賦予意義。不要把你自己限制在一開始時的觀念當中，在研究中仍然有讓你發揮創造力的空間，而且當你試圖去理解結果的真正意義時，新的觀念有可能會從中出現。重要的是當你在設計以及進行研究時，你必須要小心謹慎，如此一來，你才能夠有足夠的信心認爲你的結果反映了社會世界中的某個真正的關係，而不必去擔心這些結果是否來自於某種錯誤、或者是研究過程本身的問題。

在結束量化研究的討論之前，我們必須再提出最後一個議題。有越來越多的記者、政治人物、以及其他人物使用統計結果來提出某個論點或者是支持某種主張。然而，此種作法並未增加公共論辯中的正確性或者是清楚性。相反地，此種作法反而更常造成大眾的混淆。因此，我們更應該去瞭解統計能做什麼事、以及不能做什麼事。有一句老生常談是，你能夠利用統計來證明任何事，但是這種觀念是錯誤的。然而，有些人卻會刻意地使用或者是誤用統計來假裝他們能夠證明任何事情。可能是出於無知、或者是刻意的欺騙，有些人會利用統計來愚弄他人。保護你自己不會被統計數據誤導的最佳方法，並不是刻意地去忽視這些數據、或者是避而不談統計數據，相反地，你應該要去瞭解研究過程以及統計技術、去思考你所聽到的事物、並且適時地提出問題。

我們接下來要討論的對象是質化研究。質化研究的邏輯和目的與過去幾章中所談論的量化的、實證主義的取向截然不同。質化研究比較不關心數字、假設、以及因果關係，而是比較關心文字、規範與價值、以及意義。

問題回顧

1. 編碼簿是什麼？又該如何運用於研究之中？
2. 研究者如何清理資料以及檢查他們的編碼？
3. 請描述研究者使用光學掃瞄單的方法。
4. 研究者可以用何種方式展現出次數分配的資訊？

5. 請描述平均數、中位數、以及眾數三者之間的差異。
6. 從散佈圖中可以看出一項關係的哪三種特徵？
7. 何謂共變？我們應該如何使用共變？
8. 研究者何時能夠從散佈圖類推到百分比例表，以發現在變項之間的關係？
9. 請討論應用於三變項分析中的控制的概念。
10. 當某位研究者說「在.001 的水準上具有統計顯著性」時，這是什麼意思？在此種情況中，該位研究者比較可能犯下的是第一類型錯誤、還是第二類型錯誤？

注解

1. Practical advice on coding and handling quantitative data comes from survey research. See discussions in Babbie (1998:366–372), Backstrom and Hursh-Cesar (1981:309–400), Fowler (1984:127–133), Sonquist and Dunkelberg (1977:210–215), and Warwick and Lininger (1975:234–291).
2. Note that coding gender as 1 = Male, 2 = Female, or as 0 = Male, 1 = Female, or reversing the gender for numbers is arbitrary. The only reason one uses numbers instead of letters (e.g., M and F) is that many computer programs work best with all numbers. Sometimes coding data as a zero can create confusion, so the number 1 is usually the lowest value.
3. For discussions of many different ways to display quantitative data, see Fox (1992), Henry (1995), Tufte (1983, 1991), and Zeisel (1985:14–33).
4. Other statistics measure special types of means for ordinal data and for other special situations, which are beyond the level of discussion in this book.
5. On the elaboration paradigm and its history, see Babbie (1998:400–409) and Rosenberg (1968).
6. Beginning students and people outside the social sciences are sometimes surprised at the low (10 to 50 percent) predictive accuracy in multiple regression results. There are three responses to this. First, a 10 to 50 percent reduction in errors is really not bad compared to purely random guessing. Second, positivist social science is still developing. Although the levels of accuracy may not be as high as those of the physical sciences, they are much higher than for any explanation of the social world possible 10 or 20 years ago. Finally, the theoretically important issue in most multiple regression models is less the accuracy of overall prediction than the effects of specific variables. Most hypotheses involve the effects of specific independent variables on dependent variables.
7. In formal hypothesis testing, we test the null hypothesis and usually want to reject the null because rejection of the null indirectly supports the alternative hypothesis to the null, the one we deduce from theory as a tentative explanation. The null hypothesis was discussed in Chapter 6.

第13章

田野／實地研究

瞭解田野／實地研究
田野／實地研究訪談
資料品質
田野／實地研究中的倫理兩難困境
焦點團體研究
結論

田野研究的研究對象，是在日常生活的自然過程中行動的人們。田野工作者大膽地進入其他人的世界中，以得知關於下列這些事物的第一手資料：這些人的生活方式、他們談話以及行為的方式、以及什麼事物吸引他們或者是使他們感到苦惱...除此之外，田野研究也被視為是一種研究方法，而實行此種研究方法的研究者則是試圖去瞭解，他們所觀察到的活動對於從事這些活動的人們具有什麼意義。

——Robert Emerson, *Contemporary Field Research*, p. 1

*性別*是一種認同與表現，而且我們透過日常的互動而再製以及重新創造此種認同與表現。*婚姻*是一種根據性別而形成的關係，而婚禮則是一種具有明確規範的儀式化事件，其目的在於強化傳統的男性特質與女性特質。同樣地，新娘婚前派對（bridal shower）（譯註：在西方傳統中，當新娘要出嫁之前，新娘的媽媽、伴娘、或是好友會幫新娘舉辦一場婚前派對，以表達祝福之意）也是有性別差異的。這個活動被稱為「新娘」派對、而不是「結婚」派對，就意味著這是一種屬於女性的儀式。而男性在婚前與此相對應的儀式，則是告別單身派對。在過去十年裡，一種新的社會形式、也就是混合式的或者是兩性共同舉辦的婚前派對，則漸漸蔚為風潮。蒙特慕洛（Montemurro, 2005）曾經研究過混合式以及傳統的新娘婚前派對。她利用滾雪球抽樣的技巧，對 51 名女性進行了深度訪談。在進行訪談之前的五年之內，這些女性曾經擔任過新娘婚前派對裡的貴賓，或者是規畫、主持、或是參加過超過 280 場以上的新娘婚前派對，但是，蒙特慕洛把焦點集中在訪談之前一年內的 148 場新娘婚前派對。除此之外，她也以參與觀察者的身份參加了五場新娘婚前派對，其中有三場是傳統的新娘婚前派對（參加者全部都是女性），而另外二場則是混合式的婚前派對。她記錄了參與派對的來賓資訊（年齡、性別、以及與新娘的關係）、在派對當中依序發生了什麼事情、來賓送給新娘的禮物、以及來賓們的行為與感受。在傳統的新娘婚前派對裡，男性都只是配角或者是不出現，而且會聚集在家中除了舉行派對之外的其他地方，或者甚至是完全不在家裡。此種情況暗示了新娘婚前派對是一種專屬於女性的空間，而且許多女性則提到她們在傳統的新娘婚前派對中覺得很無聊。蒙特慕洛發現了三種類型的混合式派對：只有未婚夫出席、夫妻一起

出席、以及以新郎為中心（也就是「新郎婚前派對」）的派對。混合性別的婚前派對的舉行時間（週末的夜晚），通常會與傳統婚前派對的舉行時間不同，而且也比傳統的婚前派對更加隨興。在混合式的婚前派對中比較有可能會供應酒精飲料，而且不會把打開禮物當作是最重要的、或者是唯一的活動。除此之外，相較於傳統的婚前派對，在混合式婚前派對中的禮物比較多樣化，而且也比較不會過於強調女性特質。在混合式婚前派對中，雖然性別角色仍然相當清楚明確，但是卻比較傾向於平等。在只有女性能參加的婚前派對裡，保留了正式的形式並且期望女性「表現出」女性特質，而在混合式的婚前派對裡，則傾向於大肆揮霍，並且也偏向於展示出社會地位、而不是性別的轉變。

從這一章開始，我們將把焦點從量化研究轉移到質化研究，並且要針對田野研究以及焦點團體研究進行討論。*田野研究*（*field research*）包含了許多特定的技術，但是研究者通常會進行直接觀察或者是參與小型的社會情境，而這些社會情境大多數時候都是研究者本身的文化。在第一章中關於芝加哥地區幫派的研究（Venkatesh, 2008）、以及在第二章中關於紐約市區街頭小販的研究（Duneier, 1999）中，你已經讀過關於田野研究的內容了。正如同本章一開始關於新娘婚前派對的研究所顯示的，田野研究的對象並不僅僅是都市裡的貧苦人們而已。

許多人喜歡進行田野研究，因爲此種研究牽涉到與人們「混在一起」（“hanging out”）。田野研究沒有冰冷的數學公式、或者是複雜的統計量數，而且也沒有抽象的演繹假設。相反地，田野研究包含了在自然的社會情境中去與「真實的人們」進行直接的、面對面的社會互動。因此，對於喜歡觀察人群的人，田野研究是相當具有吸引力的。田野研究的報告也可能會是相當迷人的，揭露出陌生的社會世界裡的各種情況，像是天體營、街友或是職業賭徒、街頭幫派、警察小組、急診室、以及藝術家常常聚會的地方等等。有些田野研究報告就像是小說一樣，讀起來令人欲罷不能，甚至就像是恐怖小說或是神秘小說一樣地充滿了令人感到刺激興奮的情節。

田野研究要求研究者直接與被研究的人們進行對話、並且去觀察他們。透過經年累月的個人互動，你就能瞭解這些人以及他們的生命史、嗜好、習慣、希望、恐懼、以及夢想。遇見新的人群以及發現新的社會世界，會是一件很有趣的事。然而，田野研究也有可能會是相當困難、充滿緊張關係、耗費時間、在情緒上令人精疲力盡的，而且有時候甚至會遭遇到身體方面的危險。

瞭解田野／實地研究

當我們想要去得知、瞭解、或者是描述一群彼此互動的人群時，田野研究是最適合的研究方法。田野研究幫助我們回答像是下列的這些研究問題：人們如何在社會世界裡進行Y這件事？或者是X這個社會世界看起來是什麼樣子的？我們可以運用田野研究來發現一些利用其他研究方法（例如：調查、實驗）無法接觸或處理的

社會世界面向，像是研究街頭幫派或者是新娘婚前派對等等。

大多數田野研究的焦點都是在於一個特定的地點或者是情境，而其範圍則是從一個小型的團體（二十或是三十個人）到整個社區都有可能。剛開始進行研究的田野研究者，應該先從一個相對較爲小型的團體開始，其中的成員會在一個固定的場所（例如：在某個街頭角落、教會、酒吧、美容院、棒球場等等）定期地與彼此進行互動。有些研究者也會使用一些沒有固定形式的社會經驗，此種社會經驗雖然沒有固定的場所，但是密集的訪談以及觀察是我們能夠接觸這些經驗的唯一方式，例如，像是曾經被攻擊以及搶劫的人們的感受，或者是因爲先生自殺身亡而成爲寡婦者的感受。[1]

爲了要使術語具有一致性，我將會把在田野情境中被研究的人們稱作是*成員*（*members*），這些人在田野情境中是內部者或是本地人，而且屬於一個外來的田野研究者想要瞭解的團體、次文化、或者是社會情境。

田野研究者已經探索了各式各樣不同的社會情境、次文化、以及社會生活的面向[2]（請參閱圖 13.1）。我的學生曾經成功進行過短期的、小型的田野研究的場域，包括了美容院、托育中心、麵包店、賭博遊樂場、保齡球館、教會、咖啡店、洗衣店、警察派出所、療養院、脫衣酒吧、刺青店、以及健身房等等。

田野／實地研究簡史

我們可以追溯田野研究到那些旅行到遙遠外地的人們所寫的報告。[3] 自從十三世紀以來，歐洲的探險家以及傳教士已經寫下關於他們所遇到的奇特文化以及人群的描述。其他人則是透過閱讀這些描述來瞭解異國的文化。到了十八世紀以及十九世紀時，隨著歐洲國家向外大舉擴張，這些到外地的旅行者變得更加具有文化水準，因此，此種關於奇特地區與人群的報告的數量與品質也就隨著提升了。

學術性質的田野研究則是在十九世紀晚期與人類學一同開始發展。最初的人類學家僅僅閱讀了由探險家、政府官員、或者是傳教士所撰寫的報告，因此他們並未與他們所研究的人們有過任何直接的接觸。許多旅遊報告的焦點都是在於異國風情，並且充斥著種族歧視以及種族優越感的論調。這些旅行者幾乎都不會說當地的語言，因此都相當依賴翻譯員。一直到 1890 年代時，歐洲的人類學家才開始旅行到遙遠的異地去瞭解以及學習其他的文化。

英國的社會人類學家馬林諾斯基（Bronislaw Malinoski, 1844-1942），是第一位長期與一群人生活在一起並且撰寫關於資料蒐集過程的研究者。在 1920 年代時，他把密集的田野工作當作是一種新的方法，並且主張把直接觀察與當地人的陳述，和觀察者的推論區分開來。他認爲要發展出對於某個社群或是文化的深度理解的最佳方式，就是由研究者直接與當地的人們進行互動並且生活在一起，以及學習當地人們的風俗習慣、信仰、與社會過程。

很快地，研究者們就把田野研究的技術應用於研究他們自己的社會上面。在 1890 年代時，布思（Charles Booth）以及韋布（Beatrice Webb）同時運用了調查研究

圖 13.1 田野／實地研究的場所與主題的範例

小型的情境

一架飛機上的乘客
酒吧或是小酒館
受暴婦女庇護所
攝影社
洗衣店
社會運動組織
社會福利辦公室
電視台
候診室

社區情境

退休社區
小型城鎮
都市中的種族社區
勞動階級社區

兒童活動

遊樂場
棒球小聯盟
學校中的青少年
國中女生團體
夏令營

職業

空服員
藝術家
雞尾酒女侍
捕狗人
挨家挨戶拜訪的業務員
工廠工人
賭徒
醫學院的學生
脫衣舞孃
警察
餐廳主廚
社工人員
計程車司機

偏差以及犯罪行為

身體／生殖器官穿洞以及刺青
狂熱崇拜
藥頭以及吸毒者
嬉皮
天體營
神秘團體
娼妓
街頭幫派、機車幫派
街友、街友避難所

醫學情境以及與醫學相關的事件

死亡
急診室
加護病房
懷孕與墮胎
阿茲海默症照護者的支持團體

以及田野研究的技術來研究倫敦的貧苦民眾。他們直接去觀察在自然情境裡的人們，並且使用了一種歸納式的資料蒐集方法。而田野研究技術中的參與觀察則可能是在 1890 年時源自於德國。戈爾（Paul Gohre）在工廠裡充當學徒並且在工廠裡生活了三個月，而且每天晚上都在家裡記下詳盡的筆記，以研究工廠裡的生活。他在之後所出版的作品影響了許多在大學裡的學者，其中包括了社會學家韋伯（Max Weber）。

至於在美國的田野研究，則是可以追溯回到芝加哥大學的社會系，也就是廣爲人知的*芝加哥社會學派*（*Chicago School of sociology*）。芝加哥社會學派對於田野研究的影響可以分爲兩個階段。第一個階段是從 1910 年代到 1930 年代，在這個階段裡，研究者運用了以個案研究或者是生活史方法爲基礎的各種研究方法，其中包括了直接觀察、非正式訪談、以及閱讀文件或是官方記錄等等方法。在 1916 年時，帕克（Robert E. Park, 1864-1944）草擬了一份對芝加哥市進行社會調查的研究計畫。由於受到他身爲報社記者的背景所影響，他力勸其他研究者離開圖書館並且藉由在

街角、酒吧、以及高級旅館大廳中去進行直接觀察以及傾聽對話的方式來「弄髒他們的手」（"get their hands dirty"）（譯註：此片語的意思是實際去從事某些活動，而不是只有空談而已）。一些早期的研究，像是《遊民》（*The Hobo*）（Anderson, 1923）、《傑克洛勒：一個犯罪少年自身的故事》（*The Jack Roller*）（Shaw, 1930）以及《幫派》（*The Gang*）（Thrasher, 1927）等等，則是確立了早期的芝加哥社會學派的風格，也就是對於街頭生活所進行的敘述性研究，而且幾乎沒有任何分析。

早期的田野研究混合了新聞報導以及人類學的技術。新聞報導的技術要求研究者必須深入探究隱藏在外表以及行爲背後的真相、運用告密者、注意各種衝突、並且揭露出「真正發生」的事情。人類學的技術則告訴我們要長時間地與一小群人相處在一起、進行詳盡的觀察、然後寫出一份報告來描述團體成員之間進行互動以及觀看世界的方式。

芝加哥學派的第二個階段是從 1940 年代到 1960 年代，在這個階段中，藉由擴展人類學技術的方式，學者們將參與觀察發展成爲一種獨特的技術，以去研究或是探討研究者本身的社會。就在此時，有三條原則漸漸地浮現出來：（1）在人們所處的自然情境中、或者是*原來的情境*（*in situ*）中去研究他們；（2）以長時間重覆進行直接互動的方式去研究人們，以及（3）深入理解成員們對於社會世界的觀點，並以此理解爲基礎而發展出廣泛的理論性洞見。

在第二次世界大戰之後，田野研究面臨了來自於調查研究以及量化研究的激烈競爭。一直到 1970 年代以及 1980 年代爲止，田野研究在所有的社會研究中所佔的比例逐年下滑。田野研究者開始從認知心理學、文化人類學、民俗學、以及語言學等領域中借用以及調整一些觀念與技術。田野研究者也重新審視了社會科學的認識論根源以及哲學假定（請參閱第四章的內容），以詳盡地闡述質化方法。除此之外，這些研究者也更加關心研究技術的議題，而且也以更具系統性的方式，來詳細地說明在研究社會生活時，田野研究是一種獨特的科學方法。

時至今日，田野研究者會去直接觀察在自然情境中的成員，並且與這些成員進行互動，以獲取「內部」的觀點。許多此種類型的田野研究者支持對於社會生活的行動主義或是社會建構主義的觀點。換句話說，他們不再把人們視爲一種社會力量藉以運作的中立媒介、或者是把社會生活當作是某個「就在那兒」等著被測量的事物，相反地，他們認爲人們會透過其日常的互動而不斷地創造並且定義社會生活。田野研究者假定人們在過濾其經驗時，是透過持續的、流動的、以及主觀的現實感受而達成的，而且此種現實感受會形塑了我們看待事件的方式以及對於事件所做出的行動或反應。此種關於社會生活的假定，暗示了我們必須關注的是日常生活中面對面的協調、討論、以及協商等等社會過程，因爲人們正是透過這些過程來建構以及修正社會意義。在進行田野研究時，研究者同時要去描述社會世界、而且也是其中的一個行動者。當研究者是某個社會情境中的一部分時，進行田野研究就不只是一種被動的、中立的資料蒐集活動而已。更確切地說，研究過程本身已經變成一種自我覺察的真實生活社會經驗了。

民族誌以及俗民方法論

民族誌以及俗民方法論是田野研究的兩個延伸領域，而且都是以社會建構主義爲基礎而發展出來的。**民族誌**（ethnography）來自於文化人類學。Ethno-的意思是指人們或是民俗，而 graphy 則是指撰寫或者是描述某個事物。民族誌就是對於人們以及／或是其文化的描述。我們會經常作出推論－也就是說，超越了那些被明確地說出來的事物、或者是顯而易見的事物——並且去探討真正的意義或者是間接暗示的事物。人們會在特定的社會脈絡下透過外在的行爲（例如：言論以及行動）來展示出他們的文化（也就是他們所思考、考慮、或是相信的事物），然而，我們無法單獨從明顯的、外在展現出來的行爲當中捕捉到完整的社會意義。因此，藉由使用民族誌的方式，我們不但會去描述人們的生活以及行爲，而且也會去試圖推論出行爲的意義（也就是隱藏在行爲背後的想法或者是信念）。民族誌的主要目的，就是要超越我們能夠輕易地從外部觀察到的事物，進而去發現我們所觀察的人們內在的真正感受以及想法。例如，某人邀請你參加一場「新娘婚前派對」。根據你的文化知識，你或許會猜想這將會是一個非正式的派對，而且你應該爲即將出嫁的新娘準備一份禮物。文化知識包括了象徵符號、歌曲、俗諺、事實、行爲方式、以及物品（例如，行動電話、漢堡等等），而我們則是藉由看電視、傾聽父母親以及朋友所說的話、以及觀察他人等等方式來學習文化。

文化知識（*cultural knowledge*）包括了*顯性知識*（*explicit knowledge*）（也就是那些我們所知道以及談論的事物）以及*隱性知識*（*tacit knowledge*）（也就是那些我們隱約知道、但是很少直接公開承認的事物）。例如，顯性知識包括了社會事件（例如，一場新娘婚前派對）。大多數的人都能夠描述在某個社會事件中所發生的事情。而隱性知識則包括了一些未說出的文化規範，像是關於什麼禮物才是恰當的、以及送禮的方式之類的規範。人們或許根本不會去思考這些規範，而且如果這些規範不明確的話，人們可能會爲了應該如何適當地運用這些規範而感到焦慮不安。當有人違反這些規範時，他們會覺得不舒服，但是卻很難明確具體地指出造成此種不舒服感覺的來源。民族誌學者會去描述成員們所使用的顯性以及隱性文化知識，而且他們會使用詳盡的描述以及謹慎的分析，去分解並且重新結合這些事件。

民族誌　一種田野研究，此種研究強調的是以內部者的觀點，提出對於某個不同文化極為詳盡的描述，以增進其他人對於該文化的瞭解。

深厚描述　一種質化資料，在此種資料中，研究者試圖以極為詳盡的描述方式來捕捉某個社會情境的所有細節，並且對該情境以及其中的人們的內在生活傳達出一種熟悉而親密的感受。

人類學家吉爾茲（Clifford Geertz, 1926-）曾經提過，在民族誌中的一個極度重要的部分就是**深厚描述**（thick description）（又可譯爲深描、厚描、厚實描述、稠密描述等等）。[5] 這是一種對於具體細節的豐富的、極爲詳盡的描述（與提出摘要或概論、或者是使用標準變項的作法剛好相反）。針對一個僅有 3 分鐘的事件所作出的深厚描述，其篇幅可能會長達好幾頁。深厚描述能夠精確地捕捉真正發生的

事情，並且把事件的經過放置於一個更大的脈絡之中。除此之外，深厚描述也允許多重的詮釋或者是觀點，並且提供了更廣闊的社會文化脈絡，因此讓讀者能夠推論出更深層的文化意義。

俗民方法論（ethnomethodology）是於 1960 年代開始發展的一種獨特的方法，其研究對象是常識知識。[6] 爲了要研究常識，俗民方法論學者會去觀察在自然情境裡持續進行的社會互動中，常識被創造出來以及被使用的過程。俗民方法論是一種田野研究的極端形式，其基礎來自於現象學哲學以及社會建構主義取向，並且融合了理論、哲學、以及方法。在梅漢以及伍德（Mehan and Wood, 1975: 3,5）的著作中，我們可以看到他們對於俗民方法論的描述：

> 俗民方法論不是一種研究發現的主體、不是一種方法、不是一個理論、也不是一種世界觀。我把俗民方法論視為是一種生活形式……俗民方法論試圖要展現出某種層次的實在，此種實在的層次超越了社會學的層次……俗民方法論與社會學之間的差異，就像是社會學和心理學之間的差異一樣。

俗民方法論牽涉到對於微觀情境進行專門化的、極爲詳盡的分析（例如，將簡短對話或是社會互動影片的內容謄寫成逐字稿）。與芝加哥學派的田野研究相較之下，俗民方法論比較關注於方法本身，而且也把研究發現視爲不但來自於我們運用於研究的特定方法、而且也來自於我們所研究的社會生活。

俗民方法論的核心假定之一，就是社會意義是脆弱的以及流動的，而且不是固定、穩定、或是堅牢的。我們會在一個持續的過程中不斷地創造以及再創造意義。因此，俗民方法論學者會密切地分析我們所說的話，包括我們的停頓以及言談的脈絡。他們假定人們是利用隱性的社會文化規則來「完成」（"accomplish"）對於常識的理解。俗民方法論學者想要揭露出一些未被言明的隱性規則，雖然我們會依循這些規則來行事、但是我們卻沒有明確地意識到這些規則的存在。俗民方法論學者認爲我們會應用我們的文化知識、以及利用特定社會脈絡中的線索，而不斷地針對日常生活的事件進行詮釋（也就是去理解這些事件、或是爲這些事件賦予意義）。

利用仔細檢視日常社會互動的方式，俗民方法論學者試圖去發現建構社會實在以及常識的規則。他們想要詳細地記錄我們如何應用微觀層次的社會規則，以及「飛快地」（"on the fly"）創造出新的規則。例如，一位實證主義的量化研究者認爲標準化的測驗或者是正式的調查訪談能夠產生關於某個人的客觀事實，而俗民方法論學者則認爲這些測驗或是訪談所展現出的，是這個人接收隱性線索以及應用常識文化知識的能力。

破壞性實驗（breaching experiment）是俗民方法論學者經常使用的技術之一，這是一種揭露出以及展現出簡單隱性規則的力量的方法，而我們則是依賴這些隱性

俗民方法論　一種社會科學取向，此種取向結合了哲學、社會理論、以及方法，其目的在於研究常識知識，以及調查在小型規模的情境中的社會互動，以揭露出人們用來建構以及維持其日常生活社會實在的一些規則。

破壞性實驗　一種研究技術，田野研究者利用此種技術而去刻意地違反社會規則以及行為模式，以揭露出關於社會意義以及關係的面向。

規則在日常生活中去創造出一種現實感受。在此種「實驗」中，俗民方法論學者會刻意地違反某種隱性的社會規範。此種破壞規則的作法，通常會引來強烈的社會反應（例如，人們會變得焦慮而且困惑、緊張地笑、或者是表達出惱怒或氣憤的感受）。此種反應不但證實了某種隱性規則的存在，而且也證明了此種隱性規則是日常社會生活運作時不可或缺的特質。除此之外，此種破壞規則的作法也顯現出社會實在的脆弱性。在一個知名的破壞性實驗中，葛芬柯（Harold Garfinkel, 1917-）派他的學生到附近的商店去，並且告訴他們故意把其他的顧客「誤認」爲是店員。剛開始的時候，這些顧客覺得很困惑，並且結結巴巴地解釋自己並不是店員，但是當這些學生依然堅持把他們當作店員時，許多困惑的顧客只好心不甘情不願地接受這個情境的新定義，並且笨拙地試圖去扮演店員的角色。其他顧客則是「勃然大怒」（"blew up"）並且「沉不住氣」（"lost their cool"），也因此違反了與其他顧客維持禮貌而且客氣的互動關係這一項更大的社會規範。此種社會破壞性實驗具體地顯示出我們有多麼依賴隱性知識來使社會生活持續運作下去（例如，分辨出店員與顧客的不同）。電影導演也曾經使用類似的社會情境以達到喜劇性的效果。他們在電影中安排了來自於不同文化的人，由於這些人並未共享關於適當行爲的共同隱性規則，因此他們在不知不覺之中就違反了許多社會規範。[7] 此種情境之所以會讓人產生幽默滑稽的感受，是因爲當一個正常的成人違反了某項共同的日常隱性規範時，就會打亂了日常社會實在的運作，並且產生一種我們可以透過笑聲來紓解的緊張關係。如果一個年紀很小的兒童、或者是一個認知能力受損的人違反了隱性規範的話，那麼幾乎不會有人認爲很有趣，反而可能會覺得「可愛」或是「令人難過」。心理健康的實務工作者則是把個人辨識以及應用日常隱性文化知識的能力，當作是分辨個人心理能力健全與否的一種指標。

田野／實地研究的邏輯

田野研究比較偏向於進行實際的社會研究、而比較不像是一種特定的研究技術。不過，田野研究者會利用一些範圍相當廣泛的特定技術。[8] 正如同夏茲曼以及史特勞斯（Schatzman and Strauss, 1973: 14）所說的：「田野研究比較像是由許多活動所組成的傘狀概念，在這個概念之下，任何技術都可以被用來獲得想要得到的知識、以及應用於思考此種資訊的過程。」田野研究者是富有機智的、具有天份的人，而且具有在田野情境中獨立進行思考的能力與巧思。田野研究也會牽涉到拼湊（bricolage）的技巧（我們已經在第六章中討論過拼湊的概念了），此種技巧不僅僅是把資訊的不同片段結合在一起而已，而是要把研究者所研究的事物連結到這些事物出現的脈絡、把研究者和所研究的人們連結在一起、以及把意義和經驗整合起來（Kincheloe, 2005）。

自然主義　一種原則，依據此種原則，研究者應該要去檢視的，是在自然的、日常的、持續的社會情境中所發生的事件。

田野研究者所依賴的基礎是**自然主義**（naturalism）的原則，此種原則可以應用於各種現象的研究，像是海洋、動物、植物等等。自然主義

告訴我們要在自然的情境中去觀察一般事件，而不是在人爲的、設計過的、或者是由研究者所創造的情境中去觀察這些事件。我們瞭解事物的最佳方式，就是去捕捉那些發生於真正現實狀況中的事件，因此，我們必須離開像是辦公室、實驗室、或者是教室之類可以預測的、安全的情境，親自到「田野」之中去進行我們的研究。

田野研究的另外一項原則，則是在持續進行的社會生活當中包含了許多的觀點，而且人們會在自然的社會情境當中使用這些觀點。因此，爲了要瞭解社會生活，我們必須把所有的觀點都包括在內。田野研究者試圖去進入各種不同成員的「腦海」中、或者是意義系統中，然而再轉換回到一種外部者或者是研究的觀點。正如同馬能（Van Maanen, 1982: 139）所說的：「田野工作意味著涉入以及分離，既是忠誠也是背叛，既是開放也是秘密，而最有可能的狀況則是愛恨交織。」你應該要能夠順利而且快速地在不同的觀點之間進行轉換，並且同時從多重的觀點來看待各種事件。雖然小型的研究團隊比較有效，但是通常都是由個別的研究者單獨進行田野研究。因此，田野研究者就必須同時做許多事情，而且還要維持高度的注意力（請參閱資訊箱 13.1 田野研究者在做些什麼事？）。

因爲你在對社會生活進行研究時，也會直接地參與了「真實的」社會生活，因此，與大多數的量化研究不同的是，個人的人格特質在田野研究中具有高度的關聯性與重要性。魏克斯（Wax, 1979: 509）曾經說過：

> 在正式的以及量化的方法當中，個人的特質通常是不會被注意到的。在電子化的資料處理過程中，研究主持者或者計畫者的年齡、性別、或者是種族都是完全被忽視的。但是，在田野工作中，這些個人特質的基本面向就會被突顯出來，而且會強烈地影響到田野研究的過程。

資訊箱 13.1　田野／實地研究者在做些什麼事？

田野研究者所做的是下列這些事情：

1. 觀察在自然情境中所發生的普通事件以及日常活動，同時也要觀察一些不尋常的事件。
2. 直接涉入研究對象的生活，並且在田野情境中去親自體驗日常社會生活的過程。
3. 獲得圈內人的觀點，同時維持著分析的觀點或者是保持身為圈外人的距離。
4. 隨著情境的需求而以彈性的方式來運用各種不同的技術以及社交技巧。
5. 以大量的書面筆記以及圖表、地圖、或者是照片的形式來提出資料，以提供極為詳盡的描述。
6. 在事件的社會脈絡之下，以整體性（也就是作為一個完整的單位、而不是以片段的方式）而且獨立的方式來看待事件。
7. 去理解田野情境中成員的感受，並且發展出同理心，而非只是記錄下「冰冷的」客觀事實。
8. 留心注意文化中顯性的（公認的、有意識的、被說出來的）以及隱性的（比較沒有被公開承認的、內隱的、未被說出來的）面向。
9. 在不強加外部觀點的情況下，去觀察持續進行的社會過程。
10. 處理高度的個人壓力、不確定性、倫理道德的兩難困境、以及模稜兩可的狀況。

此種直接參與田野情境的作法，可能會造成情緒方面的衝擊。田野研究可能會既有趣又令人興奮，但是也有可能會擾亂你的個人生活、人身安全、或者是心理健康。與其他類型的社會研究相較之下，田野研究更有可能會因爲進行了研究而重新建立友誼、家庭生活、自我認同、或者是個人的價值觀：

> 進行田野工作的代價非常高，但不是在金錢方面（田野工作的金錢成本比大多數其他種類的研究還要來得低），而是在生理以及心理的努力方面。田野工作是非常艱苦的工作，同時過著兩種生活是非常累人的。（Bogdan and Taylor, 1975: vi）

進行田野研究必須花費許多時間，在進行田野研究時，研究者可能必須花上數百或是上千個小時在直接觀察上面，並且可能必須花上數個月或者甚至是數年的時間與成員進行互動，其中可能必須幾乎每天都要去到田野現場當中。如同范恩（Fine, 1996: 244）在他針對四家餐館的廚房所進行的研究中所提到的：「我試著在每一週裡都出現在現場六天……我試著交錯安排我的觀察時間……在每一家餐館裡，我先花一個月的時間進行觀察，然後去訪談所有的全職廚師。我總共進行了三十次的深度訪談，而每次訪談的時間都持續大約一到三個小時左右。」

進行田野／實地研究的步驟

與量化研究相較之下，進行田野研究的過程比較有彈性、而且也是比較沒有結構的。因此，你在進行田野工作之前，必須先妥爲安排以及作好準備。以下所列出的計畫步驟，只能充當大略的指引或者是參考的準則而已（請參閱資訊箱 13.2 田野研究的步驟）。我們可以把整個過程區分爲六個部分：準備、選擇以及進入田野場域、田野策略、田野中的關係、蒐集資料、以及離開。

步驟一：準備進入田野　在準備進入田野的階段，總共有四個面向：學習具有彈性、準備、去除焦點、以及具有自我意識並且瞭解你自己。

具有彈性　在進行田野研究時，靈活敏捷是一個優點。在田野研究中，你不必去依循那些明確安排好的、事先設定的、固定不變的步驟。你不必去應用某一組方法、或者是去測試一些明確的假設，而是依據各種技術在特定情境中提供珍貴資訊的價值性，來挑選出最合適的研究技術。在一開始的時候，你應該預期到自己將會幾乎無法掌控資料、而且也可能沒有什麼焦點。你應該要能夠視情況需要而轉換方向並且依循一些線索、學習辨識機會並且把握機會、而且能夠迅速地針對不斷變動的社會情境作出調整。你不應該把自己困在一開始的錯誤想法當中，相反地，你應該學習打開心胸去發現新的想法。如果你想要對田野中社會生活的某一部分找出最有意義或是最有價值的問題，通常你必須要有耐心、付出時間、保持敏感、並且進行反思。

資訊箱 13.2　田野／實地研究的步驟

1. 作好自我準備、閱讀文獻、並且去除焦點。
2. 選擇田野場域並且取得進入的權利。
3. 進入田野並且與成員建立社會關係。
4. 扮演一個社會角色、學習其中的訣竅、並且與成員相處。
5. 觀看、傾聽、並且蒐集質化資料。
   - 開始分析資料，並且產生以及評估可行的假設。
   - 把焦點集中於田野情境的特定面向，並且使用理論抽樣。
   - 對成員中的資訊提供者進行田野訪談。
6. 脫離並且離開該田野場域。
   - 完成分析並且撰寫研究報告。

*註：*每個步驟所需要的時間並沒有固定的比例。若要大約估計近似值的話，江克爾（Junker, 1960: 12）的建議是，在進入田野之後，研究者應該預期花費大約六分之一的時間進行觀察、三分之一的時間記錄資料、三分之一的時間分析資料、以及六分之一的時間報告結果。請同時參閱丹辛（Denzin, 1989: 176）所提出的田野研究的八個步驟。

讓自己作好準備　人性以及個人因素可能會對任何的研究計畫造成影響，但是在田野研究中，這些因素卻都是極爲重要的成分。田野計畫通常開始於一些偶然發生的事件或者是個人的興趣，像是工作、嗜好、或者是身爲病患或是行動份子等等。[9] 爲了要進行田野研究，你必須磨鍊你仔細觀察以及傾聽、短期記憶、以及規律書寫的技巧。在你進入到田野場域之前，你應該先練習觀察不同情境的一般細節，並且在之後把這些細節寫下來。只要不斷地練習，就能提升對於細節的關注能力以及短期記憶。同樣地，每天寫日記或者是個人日誌，也是練習撰寫田野筆記的好方法。就如同所有的社會研究一樣，閱讀學術文獻將會幫助你瞭解概念、潛在的危險或陷阱、蒐集資料的方法、以及解決衝突的技巧。新手田野研究者在開始進行研究之前，應該要先閱讀大量的田野研究報告。除此之外，你可能也會發現到，日記、小說、電影、新聞報導、以及自傳等等都是珍貴的工具，可以幫助你更加熟悉田野情境，並且使你在情緒上爲進入田野而作好準備。

去除焦點　在開始進行田野研究之前，你必須先清除心中所有先入爲主的想法，並且採取一種更爲寬廣的觀點，而不是把焦點狹隘地集中在特定事物上面。然而，在你進入了田野情境之後，你就可以開始聚焦在你所想要研究的事物上面。**去除焦點**（defocusing）的意思是指有意識地以全新的觀點來看待事件、高度警覺並且充滿好奇心、而且不會受到假設以及先前的判斷所影響。去除焦點有兩種類型。[10] 第一種類型是廣泛地注意所有的事物，以親眼見證範圍廣泛的情境、人群、以及場域，也就是在決定把什麼事物包括進來以及排

去除焦點　一種在田野研究早期階段中所使用的技術，研究者利用此種技術來移除自己過去所抱持的假定以及先入為主的想法，以使自己對於田野中的事件抱持著更加開放的態度。

除在外之前，先對整體情境有一些瞭解與感覺。第二種類型則是去超越狹隘的研究者角色，並且不要把你自己完全地限制在研究者的角色裡。正如同道格拉斯（Douglas, 1976: 122）曾經提過的，將你的經驗擴展超越於一個嚴格的專業角色之外，是一件重要的事情。你應該要超越你自己的「舒適區」（"comfort zone"），以盡可能地去體驗田野中的一切，但是不能違背你身爲研究者的主要義務與承諾。

自我覺察　一個良好的田野研究者，也是一個高度自覺的人。身爲田野研究者，你必須瞭解你自己，並且能夠反思你的個人經驗。你或許會預期在田野中將會感受到焦慮、自我懷疑、挫折、以及不確定性等等感覺。特別是在一開始的時候，你可能會覺得你所蒐集的資料是錯誤的，而且可能會遭受到情緒混亂、孤立、以及困惑的狀況。你也可能會覺得自己處於雙重邊緣化的狀態：你在田野情境中是一個圈外人，而且也遠離了朋友、家人、以及其他的研究者。[11] 你的個性、個人經歷、以及文化經驗，在田野研究中都是非常重要而且具有相關性的。因此，你有必要去瞭解你自己的限制、個人的承諾、以及內在的衝突（請參閱接下來關於壓力的部分）。正如同艾莉亞索芙（Eliasoph）在研究一間鄉村西部酒吧時所發現的，自我覺察是非常重要的（請參閱範例 13.1 在一間鄉村西部酒吧所進行的田野研究）。

田野工作可能會對你的認同以及觀點產生強烈的衝擊。許多研究者都曾經報告說自己已經因爲田野研究的經驗而發生很大的轉變。有些人採取了新的價值觀、興趣、以及道德觀，或者是改變了宗教信仰或是政治意識型態。[12] 麥可德莫特（McDermott, 2006: 161）利用在亞特蘭大以及波士頓地區的便利商店裡打工的方式，研究了黑人與白人之間的種族關係。她在研究報告中曾經提到：「當我完成了在便利商店裡的工作時，我覺得自己已經變成完全不一樣的人。」在對職業賭博進行了密集的田野研究之後，早野（Hayano, 1982: 148）也提到了類似的感受：

> 到了此時，我覺得坐在牌桌旁邊比坐在教職員會議或者是在自己的課堂中還要更加自在。我大部分的社交生活都是以玩撲克牌為主，而且我通常會感受到一種放棄大學教授這個職位的渴望，尤其是在大贏一場之後，這樣一來，我才能花更多的時間來玩撲克牌。

步驟二：選擇一個田野場域並且進入現場　大多數的田野研究都是發生在一個特定的情境之中的。在進行一項研究的早期階段裡，你必須要選擇一個場域、與守門員交涉、取得進入的權利、扮演一個社會角色、採取某種程度的涉入狀態、並且與成員建立和諧信任的關係。

選擇研究場域　我們通常會談論到在一個情境或是**田野場域**（field site）中去進行田野研究，但是這個詞彙卻是具有誤導性質的。一個場域是指事件或者是活動發生的脈絡，也是一種由社會所定義的領域，而且其界線是具有彈性、並且會不斷變動的。研究者所感興趣的個案、活動、或者是群體可能會橫跨數個實體的場域。舉例而言，一支大學橄欖球隊的隊員可能會在球場上、在更衣室裡、在宿

田野場域　研究者在其中進行田野研究的一個自然場所。

範例　13.1

在一間鄉村西部酒吧所進行的田野／實地研究

艾莉亞索芙（Eliasoph, 1998）曾經對一個加州社區裡的數個團體進行過田野研究，以瞭解美國人如何避免表達政治立場，而她進行研究的地點之一，是一個社交俱樂部。艾莉亞索芙把她自己描述為是一個「城市的、住在美國東西兩岸的、戴著眼鏡的猶太裔博士候選人，而其祖先當中曾經有過共產黨人、無神論者、自由派、書呆子、理論學家、以及喜好爭辯者」（p. 270）。而社交俱樂部的世界對她而言是非常陌生的。這個社交俱樂部稱為西爾夫拉多俱樂部（Silverado Club），其重心是圍繞著一個酒吧，而且是以播放鄉村以及西部音樂為主。艾莉亞索芙如此描述這個俱樂部：

> *西爾夫拉多俱樂部位於一個廣闊的、充滿車痕的停車場一角，這片土地曾經是一片濕地，而現在則是讓卡車司機中途休息以及用餐的地方。此地距離阿瑪歌（Amargo）（一個城鎮的名字）的核子戰艦基地大約一英里半。在鋪了柏油的平坦道路上、以及加油站旁的地面上，冒出了一些香蒲花，偶爾會有一些小溪流澆灌這些花兒。停車場裡擠滿了巨大的四輪驅動車輛，此種情景使得我那部小小的本田汽車看起來就像是一個玩具一樣……在沒有窗戶的西爾夫拉多俱樂部裡，剛開始進入時會覺得一片漆黑，幾乎什麼都看不見，不久之後，漸漸就可以看到在演奏臺的後面釘著一面巨大的國旗，而且還有在酒吧裡常見的那種霓虹燈啤酒招牌與鏡子，男人們戴著牛仔帽、並且穿著牛仔襯衫和牛仔褲，女人們則是燙了捲髮、穿著有像荷葉一般的蕾絲的裙子或丹寧裙子、或者是穿著牛仔褲，而且在她們的皮帶上還以閃亮的小飾品繡上她們自己的名字。（1998: 92）*

艾莉亞索芙介紹她自己是一個學生。在她為期兩年的研究中，她曾經忍受過煙霧瀰漫的房間、而且也買過價格昂貴的啤酒以及罐裝水；她也曾經參加過一場婚禮以及許多舞蹈課程，而且參與了無數次的談話以及聽過許多不登大雅之堂的黃色笑話或是種族歧視的笑話。她仔細傾聽、提出問題、用心觀察、並且在化粧室裡偷偷記下筆記。每當她花了兩個小時與俱樂部成員相處、並且回到家裡之後，她所面對的是一群大學同學，這些同學對於她所研究的世界幾乎一無所知。對於這些大學生而言，充滿詼諧的對話是最重要的事，而且要盡量避免無聊。對比之下，俱樂部裡的成員使用的非言語溝通方式多過於言語的溝通方式，而且就算覺得無聊、或者只是坐在那邊無所事事，都是無所謂的。這個研究迫使艾莉亞索芙去重新檢視她自己的觀點以及品味，而在進行研究之前，她一直都把這些觀點與品味視為理所當然而從未去深入思索過。

舍裡、在訓練營中、以及在彼此的住處裡進行互動，因此，這支隊伍的田野場域就包括了上述五個地點。選擇田野場域是一個重要的決定，而且你應該詳細記錄下選擇場域的過程。

你的研究問題將會在選擇場域方面給予你一些指引。當你在選擇一個進行田野研究的場域時，有三個因素是具有關聯性的：資料的豐富性、陌生的程度、以及適合度。[13] 有一些場域比其他的場域更有可能提供豐富的資料。那些呈現了社會關係網絡、各種不同類型活動、以及在長時間裡各種不同事件的場域，將能夠提供更爲豐富以及更加有趣的資料。對於田野研究的初學者而言，選擇一個不熟悉的情境通常會比較容易一些，這是因爲在一個新的場域中，比較容易看出文化事件以及社會關係。博格丹與泰勒（Bogdan and Taylor, 1975: 28）曾經提過：「我們會建議研究者去選擇一些其中的成員是陌生人的情境、以及研究者不具有特別的專業知識或專門

知識的情境。」同時，新手田野研究者可能會因爲面對一個全新的社會情境而感到手足無措或者是畏縮膽怯。當你在「挑選」可能的田野場域時，你應該要考慮到一些實際的議題，像是你的時間以及技巧、在場域中的人們之間的嚴重衝突、你的個人特質以及感受、以及進入場域中各個部分的可能性。

你與生俱來的特質也有可能限制了你進入某些場域的可能性。例如，一名美國的黑人研究者，就別指望能夠去研究三 K 黨或者是新納粹主義份子，不過，話雖如此，仍然有一些研究者成功地跨越了這些與生俱來的界線。[14] 有時候「圈內人」以及「圈外人」也可能一起共同合作。例如，圈外人道格拉斯（Douglas）就與一名圈內人成員弗拉那干（Flanagan）一起進行一項關於天體營的研究，而且也曾經有一名白人研究者與一名黑人共同研究一項黑人的住宅計畫。[15]

進入一個場域的權利也可能會是一個議題。各種不同場域的性質就像在一條連續線上一樣，其中一端是開放的、公開的地區（例如：公共餐廳、機場候機室等等），而在另一端則是封閉的、私人的環境（例如，私人公司、俱樂部、在某個人家裡所舉辦的活動等等）。你或許會發現自己在某個場域中不受歡迎、或者是不被允許進入某個場域，或者在進入某個場域時會遇上一些法律或是政治方面的障礙。在某些機構（例如：公立學校、醫院、監獄）中的法律與規定會限制了你進入場域的可能性。除此之外，機構審查委員會也有可能因爲道德方面的理由而限制田野研究的進行（我們已經在第五章中討論過相關的主題了）。

田野研究通常是一種個案研究，但是選擇一個田野場域與專注於一個研究個案這兩者之間並不完全相同。一個田野場域是一個社會空間或是地點，而各種活動則發生於該空間或地點之內。而一個個案則是某種類型的社會關係或是活動。一個個案可能會延伸超越了某個場域的界線，並且連結到其他的社會情境。你可以先選擇一個場域，然後在其中找到一些個案來詳加審視。

與守門員交涉 大多數的田野場域都會有**守門員**（gatekeepers），這些人具有一些正式的或是非正式的職權，可以掌控進入某個場域的權利。[16] 守門員可能是一個在角落裡的惡棍、醫院裡的管理人員、或者是某個企業的老闆。非正式的公共區域（例如：人行道、公共等候室等等）幾乎不會有守門員，但如果是正式組織的話，你就必須先向管理機構或是當局請求取得進入的許可。不論有沒有正式的頭銜，守門員就是一個領導者，而且在場域中的其他成員都會遵守他或她的指示。除此之外，你可能必須花上一些時間才能發現誰是守門員（請參閱範例 13.2 守門員與進入的權利）。你應該預期必須得去與守門員溝通協商，並且進行協議以取得進入的權利。守門員可能根本不在乎概念距離或者是倫理平衡的需要，因此，你必須設定不能退讓的底線，以保護研究的完整性。如果在一開始的時候有太多的限制，你通常可以在之後重啓協商的大門，而且隨著彼此的信任程度逐漸增加，此時守門員有可能會忘記了他們最初的要求。親自去拜訪守門員在倫理上和政治上都是精明有效的作法。大多數守門員都不會在乎研究的發現，除非這些發現可能會提供別人當作批評他們的證據。

守門員 一個掌控進入某個場域權利的人，此人可能具有正式或非正式的角色。

當你進入新的層次或者是田野場域的其他地區時，與守門員交涉就會變成一個一再發生的議題。除此之外，守門員也有可能會影響了研究的方向。「即使是最友善、最合作的守門員或是贊助者，都有可能會影響到研究的進行與發展。在某種程度上，民族誌學者會被引導，而出現了在友誼與敵對的現存網絡、領域、以及其他類似的界線中與守門員一致的情況」（Hammersley and Atkinson, 1983:73）。在某些場域裡，守門員的同意反而可能會造成一種污名，因而阻止了其他成員的合作。例如，如果囚犯知道了研究者已經獲得典獄長的同意來進行研究的話，他們可能就不會願意與研究者共同合作了。

進入場域以及取得進入的權利　當研究者在進入並且取得進入田野場域的權利時，必須要有常識的判斷力以及社交技巧。田野場域通常會有不同的層次以及地區，而且進入每個層次或地區時都會遇到一些問題。進入場域的過程比較像是在一層層

範例　13.2

守門員以及進入的權利

范卡德胥（Venkatesh, 2008）曾經針對芝加哥低收入住宅計畫中一個稱為黑人國王（the Black Kings）的販毒幫派進行過研究，而在這項研究中，他在取得進入的權利時遇上了困難。他詳細地描述了如何取得進入權利的過程。當時他幸運地偶然遇上同情他的幫派老大，這名老大叫作 J.T.，而且在幫派活動以及住宅計畫當中都是最重要的守門員。范卡德胥本身是來自於中產階級的加州郊區，而且具有南亞的血統。當時他天真地帶著一大疊的調查問題去研究住宅計畫，根本還沒準備好要去面對在這個充斥著破敗高樓住宅區中的赤貧問題、腐敗問題、以及生活中各種赤裸裸的現實狀況。在他進入一棟建築物不久之後，有幾個兇惡的年輕人就在一個黑暗、骯髒、充滿尿騷味的樓梯間堵住他。這些人把他誤認為是一個墨西哥人（而且是這個幫派的死對頭「拉丁國王」（Latin Kings）中的一個成員），因此似乎打算要好好地教訓他一頓，幸好此時 J.T.剛好出現並阻止了他們。正如同范卡德胥（Venkatesh, 2008: 17-19）在研究中所報告的：

> *J.T.瞪了其中一個年輕人一眼，然後轉向我，他說：「你不是芝加哥人。你應該不要隨便在這個地區亂走。在這裡是會出事的。」接著，J.T.就問了我一堆問題…我整晚幾乎都坐在冰冷的台階上，努力避免露出身上所帶的任何金屬碎片。我也很想睡覺，但是我實在太緊張了。*
>
> 隔天下午，范卡德胥帶了六手啤酒（譯註：一手啤酒為六罐）回到了昨天的那個地方。
>
> *「要不要啤酒？」我邊說邊丟了一罐啤酒給他。「你說過，如果我想知道這裡的人過的是什麼樣的生活，我就應該多跟他們混在一起。」J.T.沒有回答我。其他一些人則是不可置信地笑了出來，其中一個人說：「我早就告訴過你，他瘋了！」「這傢伙竟然以為他可以跟我們混在一起！我仍然覺得他是拉丁國王那掛的人。」終於，J.T.開口說話了。「好吧，既然這位兄弟想要跟我們混在一起」，他平靜地說：「那就讓他加入吧。」（p. 23）*

在試圖取得進入權利的過程當中，范卡德胥走過許多冤枉路、也犯了很多錯誤，他面臨過嚴重的身體危險，克服了不確定感以及恐懼，但是也擁有一些不可思議的好運，尤其是在守門員方面。

地剝去洋蔥的外皮，而不像是單純地打開一扇門那麼簡單。除此之外，當你在交涉協商並且取得可以進入場域的承諾之後，這些承諾不一定是穩定的。你必須要有備案的計畫、或者在之後可能必須再回來重新進行協商。由於研究的具體焦點可能會到研究的後期階段才會出現、而且也可能會發生改變，因此，比較明智的作法，是不要被守門員所訂下的一些細節給困住了。

我們可以把接近以及進入田野場域的過程想像成是一個**進入的階梯**（access ladder）（請參閱圖 13.2）。你從最底部的那一階開始，在這個階段裡，要進入場域是很容易的，而且你的身分是一個天真的圈外人，尋找著可見的、公開的資訊。下一階則要求取得更多進入的權利，當你開始認真地在場域中進行觀察時，就會進入這個階段。在這個階段裡，你是一個被動的觀察者，不去質疑成員們所說的事物，但是你正在慢慢地深入當地的社會生活。隨著你待在田野中的時間越來越長，你會再往上移動一階。此時你會觀察以及聽到一些具有潛在敏感性的事物，而且你也會開始去澄清你所看到或是聽到的事物。在進入的階梯中要達到這個階段是相當不容易的。到了最後，你可能會試著去形塑互動過程，以使此種過程揭露出特定的資訊，而且你或許會要求去看看一些具有高度敏感性的資料。很少人能夠在進入的階梯中達到這麼高的階段，因爲要到達此階段必須要建立起極爲深厚的信任關係才能辦得到。[17]

扮演一個社會角色　你在日常生活中會扮演許多社會角色，像是女兒／兒子、學生、顧客、運動迷等等，而且也會去維持與其他人之間的社會關係。其中有些角色是你自己選擇的，而其他的角色則是爲你而量身打造的。幾乎沒有人可以選擇要不要去扮演兒子或是女兒的角色。有些角色是正式的（例如：銀行行員、警察首長），而其他的角色則是非正式的（像是調情者、年長的公民、夥伴等等）。你可以轉換角色、扮演多重角色、也可以用特別的方式來扮演某一個角色。

進入的階梯　田野研究者在一開始的時候或許只能夠看到以及得知一些公開的、不具爭議性的事件，但是隨著研究者付出越來越多的時間以及努力，研究者就能獲得一些更為隱密的、私密的、以及具有爭議性的資訊。

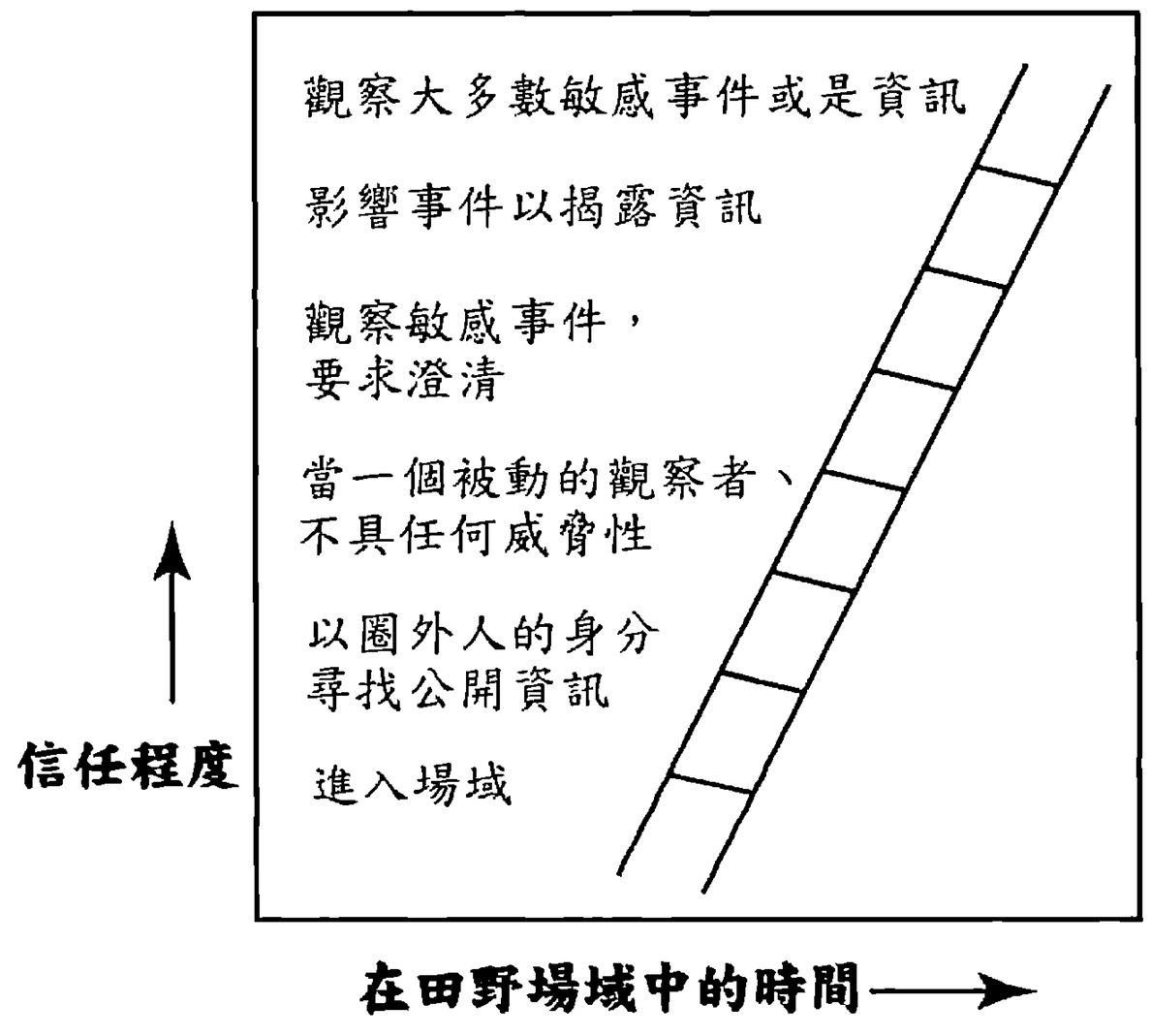


圖 13.2　進入的階梯

你在田野中扮演了兩種類型的角色：在場域中的社會角色（例如：顧客、病患、員工等等），以及你的田野研究者角色（將在下一個部分進行討論）。哈靈頓（Harrington, 2003）曾經提過，田野研究者的成功與否，取決於他或她如何有技巧地去協商象徵性互動的過程，像是自我的呈現以及扮演某些社會角色等等。她觀察到（p. 609）：

> 進入田野場域的研究者所遇到的不僅是參與者而已，而且還有參與者用來理解這個世界的、先前就已經存在的類別－參與者會把這些類別應用於研究者身上，以理解研究者出現在田野場域中的意義，並且想出應該如何與研究者進行互動……研究者被定義的方式，必須要能強化或者不威脅到參與者的團體認同。

你必須對一些預先存在的社會角色進行協商，這些社會角色是在早期的田野場域互動過程中，成員們所指派給你的角色。你被指派的角色以及你在這個角色當中的表現，會影響到進入場域的容易性以及深入程度，而且也會影響到你在發展社會信任以及取得合作機會方面的成功機率。有一些現存的角色能夠比其他角色提供更多的進入機會。這些角色能夠給予你去進行觀察以及與所有成員互動的能力、到處移動的自由、以及一種平衡研究者與成員之間的要求的方式。有些時候，你或許能夠引入一個新的角色，或者是修改一個現存的角色。例如，范恩（Fine, 1987）就創造了一個「成人朋友」的角色，而當他在研究前青春期的男孩時，他就以不帶成人權威的方式來扮演這個角色，如此一來，他就能夠去觀察這些男孩文化的某些部分以及行爲，而在其他情況下，成人是無法接觸到這些文化部分或是行爲的。隨著你在田野中的時間越來越長，你或許會扮演數個不同的角色。

你與生俱來的特質以及身體外觀也可能會限制了社會角色。雖然你可以改變外觀的某些部分，像是衣著或是髮型等等，但是卻無法改變與生俱來的特質，像是年齡、種族、性別、以及吸引力等等。然而，在取得進入的權利時，這些因素卻可能具有相當的重要性，而且也可能會限制了你能扮演的角色。例如，葛妮（Gurney, 1985）就報告說，在一個男性具有壓倒性優勢的環境中，身爲一名女性研究者，必須要進行額外的協商以及忍受許多「麻煩」（“hassles”）。然而，她的性別提供了深刻的理解，並且創造出了一些男性研究者無法達到的情境。

因爲許多角色都是依據性別來分類的，因此性別就成爲一項重要的考量。當田野環境被認爲是危險的或是粗魯下流的、而且是由男性所主導掌控的環境（例如，警察工作、消防隊員）時，女性研究者就會遇上許多困難之處。場域中的成員可能會刻意迴避女性研究者，而且女性研究者可能會落入一些具有限制性的性別刻板印象中（例如：「小甜心」、「中看不中用的花瓶」、「長舌婦」等等）。在一些由男性主導的例行性以及行政管理的場域中（例如，法院、大型公司），男性研究者則可能會遇上較多的問題，而在由女性所主導掌控的領域中，男性研究者也有可能不被接受。在同時包含了男性與女性的場域中，男性與女性研究者或許就都能夠進入場域並且被成員所接受。[18]

幾乎任何角色都會限制了進入田野場域中某些部分的權利。例如，在酒吧裡的酒保角色，就會限制了對於顧客私密行爲的知識，而且當顧客在其他地點聚會時，

也不能出現在現場。當你在選擇角色（或者是被分派角色）時，你應該要謹慎小心，但是應該也要瞭解到，所有的角色都有其各自的利弊與得失。

大多數的社會情境都會包括了派系、非正式的團體、科層體制、以及對抗的行爲。某個角色可能會幫助你獲得某個派系的接納或是被某個派系排除在外、讓你被當作一個有權威的人或是地位卑微的人來對待、或者是某些成員的朋友或是敵人。你必須注意，一旦選擇了你所要扮演的角色，你可能會招來盟友或是敵人，而這些人則可能會協助或是限制了研究。

在某些環境中（例如：警察工作、暴力犯罪幫派）可能會有危險以及高風險的面向，而這些面向可能會影響了社會角色。你應該要先瞭解你在安全方面的風險、評估這些風險、然後再決定你是否願意去進行研究。有些觀察者主張，田野研究者應該要勇於承擔在某個情境中的危險以及風險，才能真正地瞭解該情境以及其中的成員。例如，魏斯特馬蘭德（Westmarland, 2000）就主張，田野研究者若想真正瞭解警察的觀點與想法，唯一的方式就是穿上防彈背心和警察一同衝到暴力犯罪現場、並且與警察們一同躲避子彈。承擔風險意味著有些研究者可能會遇到「千鈞一髮」的緊急狀況、或者是會因此而受傷。

除了身體上的傷害之外，你也可能會面臨到法律或者是財務方面的風險，或者是因爲你在田野中的行動而對你專業或是個人的名譽造成損害。對某些環境進行研究（例如：精神病院、創傷中心、戰爭地區），可能會造成情緒或心理上的不適，並且傷害到研究者內在的幸福感。曾經研究過高風險環境（像是都市貧民區內的販毒者）的田野研究者，則提供了一些如何保護自身安全的建議（請參閱資訊箱 13.3 在不安全的環境中保護自己的安全）。

決定涉入的程度　我們可以依據研究者對於成員的涉入程度多寡，在一條連續線上排列出研究者所扮演的角色。在其中的一個極端是超然的外來觀察者，而在另外一個極端則是親密涉入的內部參與者。數位作者已經發展出一些討論研究者角色的系統（請參閱圖表 13.1）。

你的涉入程度會因爲下列幾個因素而所有變化：與成員的溝通協商、田野環境的細節、你個人的自在程度、以及你在田野場域中所扮演的社會角色。隨著你在田野中的時間變長，你或許會從一個圈外人逐漸變成圈內人。每一個層次都有其優點與缺點，而不同的田野研究者則主張不同的涉入程度。例如，有些人會批評阿德勒夫婦（Adlers, 1978）的完全成員角色是過度涉入，並且會喪失研究者的觀點。其他人則主張此種方式是瞭解成員社會世界的唯一方法。

在連續線中處於圈外人那一端的角色，會減少獲得成員接納所需要的時間、不太重視建立極爲融洽的關係、而且有時候能夠幫助成員打開眼界。這些角色有利於研究者採取超然的立場，並且保護研究者的自我認同。梅瑞本（Rueben May）在研究崔娜酒吧（Trena's bar）的十八個月期間內，就採取了此種角色，每週只有去酒吧三到四次而已。他回報說（2001: 174）：「我身爲民族誌學者的目標，就是詳細記錄崔娜酒吧那些熟客的日常生活型態，而且同時要盡可能保持不去干擾他們的狀

態……我把大部分的時間都花在傾聽顧客們的對話，並且詳細記錄那些顧客們認爲重要的主題。」雖然如此一來，與當地人同化（go native）的風險會比較小（請參閱之後對於這個主題的討論），但是身爲圈外人，你就比較不可能去完整而深入地掌握圈內人的經驗，而且也比較有可能會作出錯誤的詮釋。

許多人排斥外在觀察者角色，並且主張要瞭解成員的唯一方法，就是去加入他們並且參與田野環境。何利（Holy）觀察到：

> 研究者並不是為了要觀察研究對象才去參與他們的生活，而是在完全參與研究對象的生活時去進行觀察……也就是透過與研究對象共同生活的方式來進行觀察……研究者在積極參與研究對象的社會生活的過程中，漸漸地與他們共享相同的意義……在此種意義上，研究意味著去參與和適應所研究的文化。

而在連續線的圈內人那一端的角色，則有助於培養同理心，並且分享某位成員的生活經驗。此種角色可以幫助你去完整地體驗一個成員的私密社會世界。然而，與成員之間毫無距離、對成員有太多同情心、或者是過度涉入於成員的生活等等，都是有風險的。讀者可能會質疑你的報告、你在蒐集資料時會遇上更多困難、對於你自己的衝擊可能會更劇烈、而且你可能會缺乏在進行嚴肅的資料分析時所必須要保持的社會距離。[19]

資訊箱 13.3　在不安全的環境中保護自己的安全

1. 第一印象是很重要的，你應該採取適合該環境的個人風格以及行為舉止。
2. 學習瞭解「街頭生活」並且融入其中，不要讓自己的穿著以及行動看起來太像一個圈外人。
3. 向他人解釋你自己、你的身分、以及為什麼你會來到這個環境中。
4. 粗略觀察一下實體環境，看看有沒有明顯的危險徵兆（例如，看起來就快陷落的地板、或是看起來就快塌下來的天花板）。
5. 保持警覺性並且準備好快速地回應具有潛在危險性的情況（偏執狂、性騷擾、搶劫、偷竊、槍擊、警察突襲、以及警方的逮捕行動）。
6. 找到一個「保護者」（也就是在環境中具有權力或是影響力的人，你可以和此人建立堅強的信任關係，而且此人會提供你言語或是肢體方面的保護）。
7. 發展出一種確定的、自信的心態，而且行為舉止不要看起來像是一個受害者。過度害怕的行為可能會引來他人的攻擊。
8. 培養出「第六感」，並且利用細心謹慎以及常識來處理變化的情況。預先藏好一些錢以備不時之需。
9. 與一群人發展出「安全區域」，你可以信任這些人、與這些人可以自在地相處、而且這些人也能夠接受你。
10. 如果感覺不舒服，馬上離開該環境，之後再另外找時間回來。對於女性研究者而言，遭遇到性侵害或是強暴的威脅通常是一種真實而切身的問題，而且也應該被認真地看待。

*資料來源：*改編自 Bourgeois (1996), Lee-Treweek and Linkogle (2000), and Williams and Dunlap (1992).

圖表 13.1　在田野／實地中的涉入程度

江克爾（Junker, 1960）（同時也請參閱 Denzin, 1989; Gold, 1969; Roy, 1970）描述了四種研究者角色：

1. *完全的觀察者*　研究者彷彿身處於一個單向的鏡子後面，或者是採取一種「隱形的角色」，像是一個偷聽的守衛一樣。
2. *作為參與者的觀察者*　成員們從一開始就知道研究者的存在，但是與研究者之間僅有有限的接觸而已。
3. *作為觀察者的參與者*　研究者是參與者們公開而且親密的朋友。
4. *完全的參與者*　研究者的行為就像一個成員一樣，而且可以分享圈內人的秘密資訊。

甘斯（Gans, 1982）提供了一個類似的系統，但是把中間的兩個類別合併為研究者參與者。他所強調的是在每一個層次中的依附／情感涉入、或者是疏離。

阿德勒與阿德勒（Adler and Adler, 1987）則建議了三種角色：

1. *外圍成員*　研究者保持著自己與被研究的成員之間的距離，或者是依據自己的信念、或者是對成員活動的不自在程度而設下一些限制。
2. *積極成員*　研究者扮演了成員的角色，經歷一種典型的新成員入會儀式，並且以成員的身分參與活動，與其他成員之間維持高度的信任關係，並且會偶爾地從田野中離開。
3. *完全成員*　研究者轉變成為一個完全涉入參與的成員，並且經歷了與他人完全相同的情緒。研究者與成員「同化」（"goes native"），並且發現要離開田野並且返回研究者的角色是非常困難的、或者是完全不可能的。

建立融洽的關係　當你一進入田野的時候，你應該就要開始建立融洽的關係。在一開始時，這只是單純地意味著與成員在田野裡共處，而且要建立融洽的關係也需要時間、韌性、以及開放的態度。爲了建立融洽的關係，你應該要建立友善的關係、共享同樣的語言、並且學著和成員們一起歡笑以及哭泣。當你做了這些事，就是在理解成員方面向前進展了一步，而且也能漸漸地超越理解而進展到同理心的地步，也就是說，以另外一個人的觀點來看待以及感受各種事件。

建立融洽的關係並非總是一件容易的事。社會世界並非全然是處於和諧的狀態，而且也不必然會有溫暖、友善的人們。某個環境可能會引起恐懼、緊張、以及衝突。成員可能會是令人不愉快的、不值得信任的、或者是不誠實的，而且他們甚至有可能會做出一些妨礙你或者是令你討厭的事情。你應該要作好準備去面對許多不同的事件以及關係。然而，你可能會發現到，要深入於一個情境當中、或者是與成員培養真正親近的關係，是一件不可能做到的事。當你遇到一些不可能取得成員的合作或是同情心的情境時，你就必須採取不同的技術。[20] 除此之外，你可以依照表面的意義來接受你所聽見或是看見的事物，但是不要輕易地受騙上當。如同夏茲曼以及史特勞斯（Schatzman and Strauss, 1973: 69）所說的：「研究者同時既要相信『所有的事物』、而且也不相信『任何事物』」。

步驟三：應用策略　一旦進入了田野場域，你很快地就必須要應用一些不同的策略：協商、將研究正常化、決定揭露多少資訊、抽樣以及聚焦、利用陌生的態度、注意社會崩潰、以及處理壓力。

協商　在整個田野工作的過程當中，你都可能必須要進行協商以及形成新的社會關係。[21] 隨著你獲得更多進入的權利、建立信任、取得資訊、以及容忍成員的抗拒或是敵意，你將必須一直與成員溝通協商，直到你建立起一種穩定的關係爲止。你應該預期你在田野中可能必須去一再地進行協商，並且不斷地解釋你所做的事情。對於那些被邊緣化的人、從事非法或是不正當活動的人、以及屬於菁英階層的人，你都必須進行更多密集的協商，才能增加你與他們進行接觸的機會。例如，爲了接近行爲偏差者的次文化，田野研究者已經利用過來自於他們自己私生活中的熟人、拜訪了社會福利機構或是執法機構、刊登廣告以徵求自願者、提供某種服務（例如，諮商）以交換進入的權利、或者是去到行爲偏差者經常出入的地點並且加入一個團體。哈潑（Harper, 1982）取得進入權利的方式，是在身無分文的情況下住進一個貧民區裡的慈善機構裡，並且與一些瞭解街頭生活的街友們交朋友。芭特（Bart, 1987）則主張，她身爲女性主義行動主義者的背景、以及非專業的行爲舉止，在進入一個非法的女性墮胎診所時是不可或缺的因素。麥可德莫特（McDermott, 2006: 160）說：

> 「由於我是在南卡羅萊那州長大的，而且之前曾經在那兒當過便利商店的店員，所以才能夠融入於亞特蘭大這個場域。我也因此而能夠以當地的腔調說話、輕鬆地理解當地的腔調以及我們都被要求穿制服的事實……這意味著我也能在服裝方面與所有其他人都完全相同。」[22]

在發展出社會關係之後，你或許會維持這些關係數個月或者是數年之久。而在接觸菁英人士以及專業人士時，成功與否通常是取決於你的運氣或者是個人的人脈。[23] 霍芙蔓（Hoffmann, 1980）運用了她的家族人脈，並且在請求訪談的信件中附上了一些推薦函，因而成功地接觸了一個在公司董事會當中的富裕人士。丹濟格（Danziger, 1979）則是因爲她的父親本身就是一個醫生的緣故，因而能夠進入並接觸到醫生們的活動。強森（Johnson, 1975）在進入一個社工機構時，則是因爲提到該機構裡的某個員工是他太太的朋友，才能順利得到進入的權利。

將研究正常化　一名田野研究者不只是在觀察以及研究田野中的成員而已，他或她自己本身也在被成員們所觀察以及研究。「當田野工作者在進行一項對於他人的研究時，其他人也在進行一項對於該名田野工作者的研究」（Van Maanen, 1982: 110）。這位與外界隔絕的研究者其實並不是獨自一人在進行著田野工作，在田野環境中的每一個人都在幫助創造田野工作（Wax, 1979: 363）。在明顯的田野研究中，成員們一開始對於研究者出現在身邊都會覺得不自在。大多數成員對於田野研究並不熟悉，而且根本無法分辨社會學家、心理學家、諮商者、以及社工人員之間的差異。因此，他們可能會把你當作是一個外界的批評家或危險的間諜、或者是一個救世主或無所不知的專家。

當你採取了一個明顯的角色時，你必須**將社會研究正常化**（normalize social research），換句話說，就是幫助成員們把社會研究從某種未知的、具有威脅性

將社會研究正常化　田野研究中的一種技術，此種技術試圖讓被研究的人們對於研究過程覺得比較自在，而且也幫助他們接受研究者的存在。

的事物，重新定義爲某種正常的、可以預測的事物。你可以利用一些方法來幫助成員們做到這件事，像是提出你自己的傳記、每次解釋一點田野研究的概念、表現出不帶任何威脅性的樣子、或者是接受在環境裡一些較不嚴重的偏差行爲（像是輕微地違反官方規定等等）。[24] 例如，強森（Johnson, 1975: 99-104）曾經進行過一項關於社會工作者的研究，當進行研究時，在這些社會工作者們瞭解到他能夠接受他們一些比較不是那麼嚴重的違規行爲（像是提早一點下班，然後去游泳）、以及在他說他覺得其他人也會這麼做之後，他才被他們所接受。而在麥可德莫特（McDermott, 2006）逮到了順手牽羊的人、並且同意單獨在便利商店裡上夜班之後，其他的同事才接受了她，因爲上述的事情證明了她很有韌性，而且她也不害怕去做一些「普通的、吃力不討好的工作」。

另外一個將研究正常化的方法，就是以成員們能夠理解的方式來解釋研究。正如同范恩與葛雷斯納（Fine and Glassner, 1979）、勒梅斯特斯（LeMasters, 1975）、以及范卡德胥（Venkatesh, 2008）所發現的，有些時候，成員們對於自己的事情被寫成書籍會覺得很興奮，而此種興奮之情對你是有幫助的。當勒梅斯特斯在研究威斯康辛州的一家社區小酒館時，他在五年的研究期間之內成爲該酒館的常客，一週之內有好幾個夜晚都會待在那兒。他（1975: 7）提到了他如何向其他成員解釋他正在做的事：

> 一開始我所扮演的是顧客的角色，也就是一個喜歡喝啤酒和打打撞球的人而已。但是，因為我待在酒館中的時間很長，所以有人開始對我提出問題，而此種情況也讓我很難繼續再演下去。我後來才知道，有些酒館裡的熟客甚至已經認為我是由州政府酒類監督委員會所派來的臥底人員……後來，當又有人詢問我為何會出現在這家酒館裡的時候，我就採取了下列的說法：我是一名社會學家，而社會學家必須要具備美國社會各種不同面向的知識，才能當一個稱職的教師；我發現在這家「綠洲」酒館內的人們有助於我瞭解藍領人士對於美國社會的感受；除此之外，我對於經常與白領人士來往已經覺得厭煩，而在這個酒館裡所遇到的人事物都讓我覺得耳目一新等等。以上這些話都是我的肺腑之言。

決定揭露多少資訊 你必須決定要揭露多少關於你自己以及研究計畫的資訊。揭露出你自己的私生活、嗜好、興趣、以及背景等等，能夠幫助你建立信任以及親密的關係，但是你同時也會失去一些隱私，並且也必須確保焦點仍然是在於田野中的事件上面。

你揭露資訊的不同程度可以用一條連續線來表示，其中一端是完全隱密的研究，也就是在田野當中沒有人察覺到你正在進行一項研究，而另外相反的一端，則是每個人都知道研究計畫的所有細節（請參閱第五章裡關於隱密研究的討論內容）。揭露資訊的程度以及時機，取決於你自己本身的判斷、以及在情境中的具體情況。隨著時間經過，當你覺得更加放心、也更有把握時，你或許也會願意逐漸揭露更多的資訊。

除非有非常好的理由，否則的話，最好的作法還是揭露研究計畫給守門員以及其他人知道。即使如此，雖然你或許會揭露出你作爲研究者的身分，但是你可能會裝作是一個在表面上看起來順從的、不會造成傷害的、以及對一些不具威脅性的議

題感到興趣的人（請參閱稍後部分討論作爲一個可被接受的無能者的部分）。麥可德莫特（McDermott, 2006）編造了一個封面故事，告訴人們她想要研究的是經濟結構調整對於勞工所造成的影響，但是並沒有揭露出她真正感興趣的是人們的種族態度。她（p. 36）在研究報告中說：「如果我從一開始的時候就明白說出我真正的研究意圖的話，就非常有可能會影響到資料的效度……」當她完成研究並且要離開田野場域時，她才向那些與她合作的人們進行簡報，並且揭露出她研究的真正目的。

在你選擇了一個田野場域並且獲得進入的權利之後，你必須瞭解其中的規則、與成員們發展出融洽的關係、在環境中扮演一個角色、並且維持社會關係。在正式面對這些問題之前，你應該問問你自己下列這些問題：我要如何呈現我自己？對於我自己而言，身爲「測量工具」的意義是什麼？我要如何去採取一種「陌生的態度」？

人們會以明顯的以及暗示的方式來對他人呈現自己。我們會透過我們的身體外觀、我們所說的話、以及我們的行動方式來展示出我們的特質或身分，也就是我們是什麼類型的人、或者是我們想要成爲什麼樣的人。對於自我的呈現方式，會傳遞出一種象徵性訊息，此種訊息可能是「我是一個認真、勤奮的學生」、「我是個很酷的人」、或者是「我是一個叛逆而且喜歡在派對中狂歡的人」等等。當然，人們也可能會擁有多個自我，而在呈現出這些自我時，可能會因爲場合的不同而有所變化。

對於在田野中呈現出自我，你應該要非常謹慎小心。例如，你應該要以什麼樣的穿著打扮出現在田野中？最佳的指引方式，就是去尊重你自己以及在田野中的成員。不要穿得過份講究，以免冒犯他人或者是讓他人覺得你很突兀。你也不一定總是要去模仿研究對象的穿著方式。一名研究街友的教授不一定必須在衣著或是行動上都要像個街友一樣，這名教授只要以非正式的方式來打扮以及行動就已經很足夠了。同樣地，當你在研究企業主管或者是高層官員時，通常就必須穿著比較正式的服裝、以及表現出專業的風範。[25]

自我的呈現可能會在某種程度上影響到田野關係。然而，誠實通常都是最好的策略。這是因爲要表現出高超的騙人外表、或者是以一種和你平時的樣子相差極大的方式來呈現你自己，其實是一件相當困難的事。

例如，表現出真實的自己、並且揭露出她自己身爲猶太女性的個人背景，就幫助梅兒霍芙（Myerhoff, 1989）成功地進入一個猶太人老年安養之家裡的年長居民所組成的田野場域，並且建立了融洽的關係。與此同時，由於她在田野中所進行的互動，她對於自己身分認同的理解以及覺察也因而發生了變化。史妲克（Stack, 1989）一開始所扮演的是圈外人的角色，也就是一個研究低收入黑人工業社區的白人女性。到了最後，成員們不但接受了她，還和她建立了像家人一樣的關係。她被暱稱爲「白色的卡洛琳」，就是一種接受以及親密的暗示。她表現出許多小小的善意行爲，像是開車載送人們到醫院或是福利辦公室、購物、以及拜訪生病的兒童等等。她能做到這些事，是來自於她與其他人的互動方式－她很開放，而且樂於分享個人的感受。雖然安德森（Anderson, 1989）本身是一個黑人男性，但是在一個黑人酒吧裡，他卻發現社會階級會形成障礙。他進行田野研究的環境，是在芝加哥南部一個窮困的黑人社區裡的一家小酒吧以及販賣烈酒的商店。安德森與成員們發展出一種信任

的社會關係，而且他所結交的一個叫作何爾曼（Herman）的圈內人也「幫助」他很多。何爾曼是一個機智、隨和的人，不但熟知街頭的一切狀況，而且在該場域中擁有廣泛的人脈。安德森之所以成功，是因爲「我採取了一種低調的、不張揚的角色……不去破壞在此種環境中對於社會秩序的共識性定義」（Anderson, 1989: 19）。

聚焦以及抽樣 一旦你置身於田野之中，你首先應該對整體狀況有大致的理解。因爲唯有如此，你才能逐漸地聚焦於一些特定具體的問題或者是議題（請參閱圖 13.3）。[26] 只有在你親自體驗過田野的情境之後，你才能夠去決定特定的研究問題以及發展出暫時性的「假設」。一開始，每一件事可能看起來都像是具有關聯性的，不過，之後你就可以選擇性地把注意力聚焦在特定的問題以及主題上面。

田野研究中的抽樣方式不同於調查研究中的抽樣方式，雖然有時候在這兩種研究中都會使用到滾雪球抽樣的技術（請參閱第八章）。[27] 在本章一開頭的部分曾經提到一項關於新娘婚前派對的研究，在該項研究中就使用了滾雪球抽樣的技術。在田野研究中，我們通常會使用理論抽樣，此種抽樣方式是由發展理論以及抽樣次數、情境、事件種類、地點、人們的類型、或者是感興趣的脈絡來提供指引的。

麥可德莫特（McDermott, 2006）所抽取的樣本是分別在波士頓以及亞特蘭大這兩個地區中的某個混合種族社區中的勞動階級，因爲她想在理論上去比較北部以及南部城市中的情況。爲了要完全理解田野場域如何維持相同情況、或者是發生改變的過程，你可以在不同的時間點去觀察在某個環境中所發生的事件：像是在一天當中的不同時間點、一週裡的每一天、以及所有季節。當你在進行抽樣時，最好的方式通常是讓時間有重疊的部分（例如，把抽樣時間訂爲從早上 7 點到 9 點、從早上 8 點到 10 點、以及從早上 9 點到 11 點等等）。

你應該要在不同的地點進行抽樣，因爲雖然只在一個地點抽樣或許能夠比較深入，但是卻只能提供一個狹隘的觀點。坐在或是站在不同的地點，可以幫助你更加瞭解整個場域。例如，在學校教師之間的同儕對同儕的行爲，通常會發生在教師休息室、以及當教師們聚集在當地的酒吧或是咖啡廳裡的時候、或者是在一個暫時當作教師開會場地的教室裡。除此之外，研究者也會追隨成員走過的路徑，以造訪在田野中各種不同的地點。

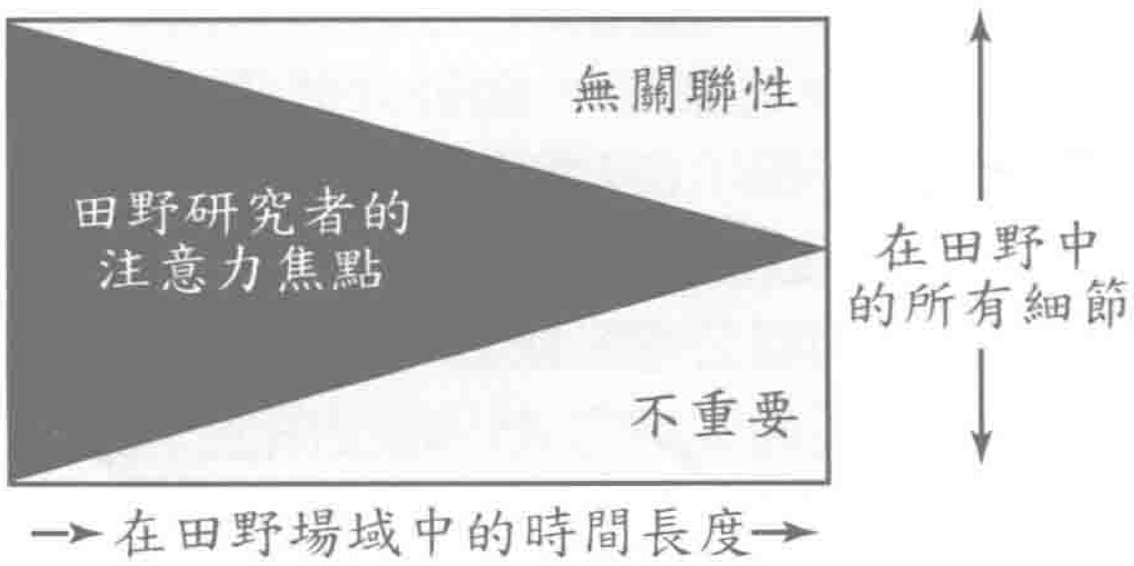


圖 13.3 在田野／實地研究中聚焦

當我們在對人們進行抽樣時所利用的方式，是把注意力或是互動過程集中於不同類型的人們身上（例如：老住戶以及新來者、年長者以及年輕人、男性以及女性、領導者以及追隨者等等）。當你辨識出各種不同類型的人們、或者是擁有對立觀點的人們時，你或許必須要試著去與所有類型的人們進行互動、並且去瞭解他們。

例如，你可能抽樣了三個種類的田野事件：例行性的、特殊的、以及意料之外的。雖然例行性的事件（例如：每天開店做生意）每天都會發生，但是你不應該單單只因爲這些事件是例行公事，就認爲這些事件無關緊要。特殊的事件（例如：年度的辦公室派對）都會事先就公布以及進行規畫。這些事件把注意力集中在成員身上，並且揭露出以其他方式無法看到的社會生活面向。意料之外的事件則是指那些當研究者在場時剛好碰巧發生的事件（例如，當店經理生病請假而無法監督員工時，員工們處於一種無人監督的狀況）。在此種情況下，你可以看到一些不尋常的、未經規畫的、或者是相當罕見的事件。

採取陌生的態度　要認出我們已經非常熟悉的事物，是一件困難的事情。我們所處的日常世界裡充滿了無數的細節。如果我們一直注意每一件事情，我們就會遭受到嚴重資訊超載的情況。因此，我們會下意識地忽略周遭大多數的事物，並且會利用慣性思考的模式來設法應付生活中的一切。不幸的是，如此一來，我們就無法看到熟悉事物的獨特之處，而且也會假定其他人所經歷的現實情況也和我們完全一樣。我們傾向於把我們自己的生活方式當作是自然的、或者是正常的。此種對於熟悉事物的「盲目狀態」，使得在熟悉的環境裡進行田野研究變得相當困難。事實上，「當某人對於自己的文化極爲熟悉時，此種情況所產生的盲目狀態可能和洞見一樣多」（McCracken, 1988: 12）。藉由研究其他文化或者是次文化的方式，在關於什麼才是重要的事物、以及如何完成任務等等方面，你會遇到非常不同的假定或是想法。這種文化之間的直接比較、或者是文化衝擊，會使你更加容易看到文化的基本元素，並且促進自我發現。

田野研究者們會採取一種**陌生的態度**（attitude of strangeness）以獲得上述這些益處，而這意味著你應該要去質疑以及注意到平常的細節、或者是透過陌生人的眼睛來觀察一些平常而普通的事件。此種方法可以幫助你克服在觀察普通的細節時所會產生的沉悶無聊感。除此之外，此種方法也有助於揭露出成員們未能有意識察覺到的一些情境中的面向。

人們幾乎無法去認出他們視爲理所當然的風俗習慣。例如，當有人送給我們一份禮物時，我們會說謝謝並且稱讚這份禮物有多麼棒。相對之下，在許多文化的送禮習俗當中，還包括了一種抱怨禮物不適合的習慣。陌生的態度有助於揭露出隱而不顯的文化－例如，送禮的人會期待聽到「謝謝你」或者是「這份禮物很棒」之類的話，而且如果沒聽到這些話就會覺得失望難過。你應該同時採取陌生人以及圈內人的觀點。陌生人會認爲所有的事件都是特殊的社會

陌生的態度　一種田野技術，在使用此種技術時，研究者在心理上會進行調整，因此會以好像是第一次、或者是以圈外人的身分來「看待」田野中的事件。

過程，而對於圈內人而言，這些事件則似乎是再自然不過的事了。戴維斯（Davis, 1973）把此種狀況稱爲火星人與皈依者：火星人認爲每件事都是陌生的、並且會去質疑所有的假定，但是皈依者則會接受一切事物，並且想要成爲信徒。你不但需要這兩種觀點，而且也需要一種在這兩種觀點之間迅速進行轉換的能力。[28]

陌生的態度也鼓勵你去重新思考你的社會世界。沉浸在一個不同的環境中，有助於打破老舊的思考習慣以及行動習慣。當你遭遇到不熟悉的事物時，不論是不同的文化、或者是透過陌生人的眼光來看待熟悉的事物，你都會發現到，反思以及內省變得更加容易、而且也更加強烈了。

注意社會崩潰　當兩種文化傳統或者是社會假定無法互相配合協調時，就會發生**社會崩潰**（social breakdown）。社會崩潰將會突顯出社會意義，這是因爲在崩潰的過程當中，隱藏的例行性期望以及假定都會變得明顯可見。例如，我去到一家餐廳並且找了個位置坐下來，然後我就等著服務生出現來爲我提供點餐的服務。但是在過了十分鐘之後，仍然沒有任何人爲我提供服務，因此我忍不住生氣了。我環視四周，並且注意到我根本沒看到任何一位服務生。相反地，我看到顧客們都端著自己的餐點從一個門口進來，因此我終於發現自己有一些誤會了。我內心原本的期待是這家餐廳會提供餐桌服務，但是事實上在這家餐廳裡，顧客們必須自己走到櫃檯去點餐，然後自己把食物拿到桌上。一旦我認出了在此種脈絡下要應用何種規則，我就能夠解決此種社會崩潰的狀況了。

社會崩潰會產生尷尬的情況，因爲文化意義的錯誤配對通常會使某個人看起來很愚蠢、無知、或者是不明事理。例如，你被邀請去參加一場派對，而派對開始的時間是晚上八點。你穿上你平常的服裝，也就是一條舊牛仔褲以及一件皺皺的毛衣，並且像平時一樣，在參加八點開始的派對時，你總是在八點三十分抵達。當你抵達派對現場時，你發現門是開著的，然後你走了進去。你很驚訝地看到其他所有人都穿著正式的服裝，並且在享用一頓正式的晚餐，而且主人早在三十分鐘之前就準時開始上菜了。所有的人都盯著你看，讓你覺得自己好像走錯地方了。你的文化期待（這不過就是一個非正式的學生派對，裡面會有吵雜的音樂、舞蹈、啤酒、以及非正式的服裝等等）並未符合這個情境（這是一個正式的晚餐宴會，在此種宴會中，人們預期要以正式的方式飲食、進行禮貌的交談、以及行動）。此種衝突的情況會使得那些「大家都知道」或者是想當然爾的、未說出來的社會規則變得明顯可見。

社會崩潰可能是無法預料的，或者，你也可以刻意地創造出社會崩潰來測試假設。就如同俗民方法論學者所採用的破壞性實驗一樣，你或許可以刻意去違反一些社會規則，以暴露出某些隱性規則的存在、以及這些規則的重要性。你可以去觀察未經規畫的社會崩潰、或者是製造出小型的社會崩潰，然後去觀察成員的反應，以正確地指出隱性的社會期待。

社會崩潰　在田野場域中，社會規則以及行為模式未能如預期般運作的狀況，此種狀況會揭露出許多關於社會意義以及關係的資訊。

處理壓力　田野工作可能是具有高度回報的、令人興奮的、而且充滿成就感的，但是也可能是困難的：

> 在人類為自己所設計出來的、較為惹人不快的活動裡，田野工作必然能夠名列其中之一。至少可以說，田野工作通常是不方便的、有時候在生理上是不舒服的、經常令人尷尬的、以及在某種程度上總是充滿緊張的（Shaffir et al., 1980: 3）。

新手研究者在田野中會面臨尷尬、經歷不舒服、以及被細節淹沒的窘境。例如，當她在研究二次世界大戰期間美國對於日裔美國人的再安置營（relocation camps）時，備受尊敬的田野研究者魏克斯（Wax, 1971）就報告說她必須忍受高達華氏 120 度的高溫所引起的不適感、骯髒以及破爛的生活環境、腹瀉、以及蚊子等等。她覺得自己孤立無援、常常哭泣、而且因爲強迫性飲食而增胖了 30 磅。在田野中待了幾個月之後，她覺得自己完全失敗，因爲她不但被成員所懷疑，而且還與再安置營裡的管理階層發生爭執。

要維持一種「邊緣性」的地位是充滿壓力的，而要當一個不是完全涉入的圈外人也是相當困難的，尤其是當你在研究一些充滿強烈感受的情境時（例如：政治競選活動、宗教皈依過程等等）。田野工作中的寂寞以及孤立的感覺，可能會結合想要發展出融洽關係以及同理心的渴望，因而造成過度的參與。你可能會**與當地人同化**（go native）並且放棄專業研究者的角色，以成爲你所研究的團體中的完全成員。或者，你會因爲在成員放下戒心時得知了許多私人的細節而覺得有罪惡感，以及過度認同成員。[29] 正如同范卡德胥（Venkatesh, 2008: 176-177）在他對於芝加哥幫派的研究中所提到的：

> 我開始覺得自己有一點點精神分裂的症狀，彷彿我在住宅計畫中是某一種人——有時候我發現自己甚至連說話方式都變得不一樣了——而在倫敦的海德公園裡則是另外一種人。漸漸地，我發現自己對於整個社會科學領域感到越來越憤怒……我覺得其他的學者好像都是活在一個與世隔絕的泡泡裡……我沒有和我的女朋友、室友、以及朋友們分享我的挫折感——事實上，他們之中大多數人都是很支持我的——相反地，我把自己的經驗藏在心裡面……當我真的嘗試著去談論我所進行的田野工作時，我有一種進退兩難的感受。事實上，我有時候會表現得好像在為幫派份子以及他們的暴力行為辯護一樣、或者是好像在把這些住宅計畫中的狀況浪漫化了……我變得越來越安靜、越來越孤獨。我的研究所同學、甚至是一些教授們，都認為我是一個難以親近的人。於是，開始有一些關於我的流言蜚語，說我是一個企圖心太強、過於冷漠的人。但是我覺得我似乎也只能對這些虛構的流言充耳不聞了。

在田野研究中，某種程度的情緒壓力是無法避免的。與其刻意去壓抑情緒的反應，你應該要對情緒反應保持敏感。有些方法可以幫助你處理在田野環境中的情緒壓力，其中包括了書寫個人的日記、

與當地人同化　當田野研究者過度參與研究對象、失去了所有距離或是客觀性，並且完全加入研究對象時，就是與當地人同化了。

情緒日誌、把內心的感受寫下來，或者是找一些在田野場域之外的、你可以信任的朋友，然後向他們傾吐自己的壓力。[30]

步驟四：維持在田野中的關係　當你努力維持在田野中的關係時，你必須使用到許多社交的策略以及技巧。

調整以及適應　隨著時間經過，你會發展出並且修正社會關係。某些一開始對你冷漠的成員，之後可能會對你比較友善，或者，有些人可能一開始裝作很友善的樣子，但是之後他們的恐懼以及懷疑就會浮現出來了。你處於一種非常微妙的地位。在計畫的早期階段，當你還未完全察覺到關於田野場域的一切狀況時，你應該不要太過急著去建立起親密的關係，因爲情況在之後可能會有所變化。但是，如果你真的和某些人建立起親密的友誼，他們可能會變成你的盟友，爲你在田野中提供保護，並且幫助你獲得接近其他事物的權利。

你必須密切注意你的行動或是外表會如何影響到田野中的成員。例如，一位在外表上極富吸引力的研究者，在與異性成員互動時，可能會遇到迷戀、調情、以及嫉妒等等狀況。因此，這位研究者就必須注意到這些田野中的關係，並且學習如何處理這些問題。[31]

除了發展出社會關係之外，你也必須能夠結束某些關係、或者是從這些關係中抽身退出。你可能必須結束與某位成員的關係，才能與其他成員形成新的關係、或者是去探索田野環境中的其他面向。正如同結束任何友誼關係時一樣，不論是研究者或是成員，在結束一段關係時都會承受情緒方面的痛苦。你必須在社會敏感性以及研究目標之間找到一個平衡點。

運用魅力以及培養信任　你必須運用社交技巧以及個人魅力來建立起融洽的關係。信任、友善的感覺、以及受到歡迎，都能促進你與成員之間的溝通，並且能夠幫助你瞭解其他人的內心感受。而要做到這件事，其實並沒有什麼神奇而快速的方法。表現出對他人真誠的關心以及興趣、坦白誠實、以及分享感覺等等，都是良好的策略，但是這些策略並不是完全不會失誤的，而且也必須取決於特定的環境以及成員。你的行爲舉止應該要總是不帶有任何威脅性，而且如果可能以及恰當的話，你應該要表現出既溫暖又友善的態度。

許多因素都會影響到信任以及融洽的關係，其中包括了：你呈現你自己的方式；你在田野中的角色；以及一些會鼓勵信任、限制信任、或者是使獲得信任變成不可能的事件等等。信任並非是一蹴可幾的，換句話說，信任是一種過程，是隨著時間流逝、經過許多微小的社會細節（例如，分享個人經驗、述說故事、手勢、暗示、臉部表情等等）才能慢慢建立起來的。信任是一直在不斷地重新創造之中的，而且一旦建立起了信任，要失去信任會比在一開始時建立信任還要更加容易。雖然建立信任關係是很重要的，但是這並不能保證成員就會因此而向你透露所有的資訊。除此之外，信任也可能被限制於某些特定的領域。例如，你可能在財務方面與成員建立起信任關係，但是卻無法在揭露親密的約會行爲方面去建立信任關係。在研究的

每個領域中，你可能都必須要去重新建立信任關係，而且，一旦建立了信任關係，也必須要不斷去小心謹慎地維持才行。

出局者　在田野研究中，拒絕與研究者合作或是參與研究的人們。

有些成員可能不願意開放或者是與你合作。所謂的**出局者**（freeze-outs），指的是那些表達出不合作態度、或者是明顯不願意參與研究的成員。你或許永遠也無法取得所有人的合作，或者，也許只有在長期的堅持之後，才能發展出稍微溫暖的關係。

融洽的關係可幫助你理解成員，但是理解只是爲了達到更深入知識的先決條件而已，其本身並非目的。當你遇到一個新的或是不尋常的社會意義系統時，你在一開始時會有一些迷惑之處，但是隨著你克服了這些迷惑，你也會緩慢地在田野中發展出對於成員的理解。一旦你能夠理解成員的觀點，下一步就是學習如何以成員的觀點來進行思考以及行動。這就是同理心、或者是去採取（至少是暫時地）其他人的觀點。同理心不必然意味著同情、同意、或者是贊成，而只是指與其他人擁有相同的感受。[32] 融洽的關係有助於發展出理解，以及最終發展出同理心，而發展出同理心則有助於營造更爲融洽的關係。在《梅崗城故事》（*To Kill a Mockingbird*）這本小說中，就提到了融洽的關係以及同理心理解之間的聯結：

> 「首先，」（阿提克斯）說：「史考特，如果你能夠學會一個簡單的技巧，你就可以和所有種類的人都相處得更好。在你學會從某個人的觀點去考慮事情之前，你永遠無法真正地瞭解這個人。」
>
> 「您的意思是？」
>
> 「——也就是你要去瞭解這個人的真正想法、並且以他的觀點來行動才行。」（Lee, 1960: 34）

提供舉手之勞　在田野中也會發展出交換關係，在此種關係中，人們會彼此交換一些小型的象徵或是舉手之勞，其中包括順從以及尊敬等等。[33] 你或許可以藉由在一些小地方提供幫助的方式，來獲得成員們的接納。當你不容易接觸到一些敏感的議題時，交換或許能夠有所助益。你可以先施予小惠，但是不要求成員們給予你任何回報。當你與這些成員再次見面並且分享經驗時，他們就會想起你曾經幫忙過他們，並且爲了要報答你而願意提供你接近某些事物的權利。例如，當范恩（Fine, 1987: 242）以他「成人朋友」的角色提供了許多舉手之勞時（例如：開車載那些男孩去看電影），他就因此而得知了許多資訊。他（1996: x）也報告說，當他待在餐廳廚房裡的時候，曾經洗過馬鈴薯、清理過豆莢、而且也做了許多雜務。

避免衝突　在田野中可能會爆發出爭執、衝突、以及意見不合的狀況，或者，你所研究的不同團體，彼此之間的立場可能是對立的。在這樣的情況裡，你將會感受到選邊站的壓力，而且成員們也會去測試看看你是否值得信任。在此種場合當中，你通常應該要盡可能地保持中立，並且小心謹慎地遊走於對立的雙方之間，這是因爲

一旦你加入了某一邊，你與另外一邊的接觸關係就會被切斷。[34] 除此之外，你也將只能從一個觀點來觀察整個情境而已。然而，有些研究者〔例如：馬能（Van Maanen, 1982: 115）〕則主張，真正的中立只是一種幻覺而已。當你與成員們交往日漸密切，並且被捲入了由各種關係與承諾所交織而成的網絡時，要去完全避免衝突其實是不太可能的。

表現出感興趣的樣子　我們在田野中會試圖去維持一種**表面上的興趣**（appearance of interest）。一位經驗豐富的研究者，即使他或她其實並不是真正地感到興趣，也會在言語上或是行動上（例如，利用臉部表情、選擇喝咖啡、組織一場派對等等）表現出對田野事件感到有興趣以及願意投入的樣子。之所以要這麼做，是因爲如果成員們認爲你看起來好像覺得很無聊或者是心不在焉的樣子，就有可能會傷害到你和他們之間的關係。當你對於田野場域中的事件表現出不感興趣的樣子時，你就是在傳遞某種訊息給成員們，也就是你覺得他們是乏味、無趣的人，而且你根本不想待在那兒－這麼一來，你根本就別奢望要與他們建立信任、親密、以及強大的社會聯結了。暫時裝作對某件事感興趣的樣子，是我們在日常生活中常用的小小欺騙手段，而且也是以禮待人這個更爲廣泛的社會規範的一部分。[35]

當然，選擇性的忽視（也就是不去直視、或者是表現出沒有注意到的樣子）也是一種禮貌的表現。如果某個人犯了社會方面的錯誤（例如，不小心使用了不正確的字眼、或者是放屁等等），忽視此種錯誤就是一種有禮貌的表現。選擇性的忽視在田野中也能發揮作用，如果你警覺性夠高的話，選擇性忽視會讓你有機會去不經意地偷聽到別人的對話、或者是觀察到一些不打算公開的事件。

成為可被接受的無能者　身爲研究者，你在田野中的目的是去學習，而不是去當一個專家。依據情境的不同，你應該要當一個友善但是天真的圈外人，也就是一個**可被接受的無能者**（acceptable incompetent），換句話說，就是一個有興趣去瞭解田野中的社會生活、但是在田野環境中只有部分能力（也就是只有部分技術以及知識）的人，而且成員們因爲認爲此人不具威脅性而願意接受他或她，並且視爲一個需要被引導或者是被教導的人。[36]

在一開始的時候，你對於田野環境或是當地文化可能一無所知。你可能會被當作是一個容易受騙上當的傻瓜，並且因爲你在田野中的笨拙行爲而成爲他人的笑柄。即使是當你已經瞭解了田野環境之後，你也可以裝作仍有不懂的地方，好讓成員們願意提供一些資訊。當然，你也有可能會做得太過火，並且表現得過於無知，以致於人們根本不把你當成一回事。

表面上的興趣　一種田野研究者在田野場域中用來維持關係的技巧，在使用此種技巧時，即使研究者實際上對於研究對象的活動沒有興趣，也會假裝對這些活動很有興趣、並且也感到興奮的樣子。

可被接受的無能者　這個詞所指的是一名假裝技巧或是知識不足的田野研究者，而其目的則是在於藉此去瞭解更多關於田野場域的知識。

步驟五：蒐集以及記錄資料　這個部分所討論的是如何去獲得良好的質化田野資料。田野資料所指

的是你所經歷過的、記得的、以及記錄在田野筆記中的資料。

吸收以及體驗　研究者本身就是田野資料的測量工具。正如同羅夫蘭等人（Lofland et al., 2006: 3）所觀察到的：「在使他或她自己去親身經歷他人的生活、以及與其他人一起過日子與感受這些生活時，研究者會成爲進行研究時最主要的工具或者是媒介。」這句話有兩個意義。首先，這句話是指你必須去承受一些壓力，也就是你必須對於在田野中所發生的事保持警覺以及敏感度，以及在記錄資料方面要有條不紊。第二，這句話的意思是田野工作會對你個人造成一些影響。田野工作牽涉到社會關係以及個人感受。你在其中會有自己的主觀洞見以及感受，或者是「來自於經驗的資料」（"experiential data"）。[37] 個人的、主觀的經驗是田野資料的一部分。這些資料本身就極具價值，而且在詮釋田野中的事件時也具有其重要性。你不必刻意去保持客觀以及忽視個人的反應，你對於田野事件的感受也是資料的一部分。例如，卡爾普（Karp, 1973, 1980）曾經針對販賣色情書刊的書店進行過一項研究，而他對於緊張狀況的個人感受，就是資料中極爲重要的一部分。他本身在田野中所感受到的不舒服感，揭露出了在該情境中的一些動態關係。除此之外，根據克萊門以及寇普（Kleinman and Copp, 1993: 19）的說法：「如果我們避而不寫我們的反應，我們就無法去檢視這些反應。若是不把我們的主觀性加入其中的話，我們就無法達到完全沉浸於其中的狀態。」

田野研究可能會提高對於個人感受的察覺與體認。例如，當你身處於一個天體營之中時，你可能才會完全察覺到你個人對於裸體的感受，或者是當你身處於一個其他人經常向你「借用」許多物品的環境中時，你才會完全察覺到你對於個人財物的感受。這麼一來的話，你的驚訝、憤怒、或者是質疑，就會變成進行反思以及獲得深刻理解的機會。[38]

觀察以及傾聽　你在田野中所做的事，有很大一部分是去密切注意、觀察、並且仔細傾聽。你必須運用你所有的感官，用心地去留意你所看到、聽到、聞到、嚐到、或者是碰觸到的事物。你應該要成爲一個能夠吸收所有資訊來源的工具。你應該要仔細審視外在的實體環境，以瞭解該環境的氣氛。例如，地板、牆壁、天花板是什麼顏色？某個房間有多大？窗戶和門在什麼地方？傢俱是以什麼方式來擺放的、傢俱的狀況如何（例如，新的、又舊又破的、骯髒的、或者是乾淨的）？照明設備是什麼類型的？有沒有任何標誌、繪畫、或是植物？在裡面聽到的聲音或是聞到的味道是什麼？

爲什麼你要大費周章地去特別留意這些細節呢？你或許已經注意到，商店以及餐廳通常會利用照明、顏色、以及播放特定的音樂來營造出某種氣氛。或許你也已經知道，販售中古車輛的業務員，會在車子裡噴灑一些新車特有的味道，或者是購物中心也會刻意讓空氣中瀰漫著剛烤好的麵包香氣。而這些微妙而難以捉摸的信號都會對人們的行爲造成一些微妙的影響。

在田野研究中，進行觀察通常是一件鉅細靡遺而且冗長乏味的工作。你必須要有耐心、以及一種專注於日常生活中緩慢細節的能力。西弗爾曼（Silverman, 1993:

30）曾經提過：「如果你到電影院去看某個動作場面（像是飛車追逐、持槍搶劫等等），那麼你就不太可能會發現到，要成爲一個良好的觀察者其實是一件容易的事。」在田野研究中，動機是來自於對細節的深層好奇心，而不是來自於那些快速閃過的畫面。藉由仔細傾聽以及觀察的方式，良好的田野研究者對於一些能夠揭露出「到底發生什麼事」的細節感到好奇。你要記住，我們正是透過平凡的、瑣碎的、日常的小事，來傳達出社會生活的核心。大多數人都會忽略了持續流動的細節，但是你必須要學習如何去觀察這些細節。

除了實體的環境之外，你也應該要去觀察人們以及他們的行動，並且去注意每個人可以觀察到的外表特徵，像是年齡、性別、種族、以及身材等等。依據對方是 18 歲、40 歲、或者是 70 歲、男性或是女性、白人或者是非白人、矮小虛弱或者是高大壯碩等等不同的狀況，人們也會以不同的方式來進行社交互動。當你在留意這些特徵時，你應該把自己也包括進去。例如，陌生的態度能夠增加你對於某個團體的種族組成方式的敏感度。如果某位研究者因爲自己是一個白人，就忽視了在一個多元種族社會中某個白人團體的種族組成方式的話，這名研究者在種族方面就是感覺遲鈍的。同樣地，「在研究過程中，當參與者的性別被忽視了的時候，就是發生了性別感覺遲鈍的現象」（Eichler, 1988: 51）。

你應該把這些細節都記錄下來，因爲這些細節可能會揭露出某些具有重要性或是意義的事物。你寧可因爲過度小心而把所有細節都記錄下來，也不要忽略了任何一個具有潛在重要性的細節。例如，「當那名高大、壯碩的 19 歲白人男性，穿著一件破舊的 T 恤以及骯髒的牛仔褲衝進了這間很亮的房間時，那名矮小肥胖而且膚色較淡，年紀約六十幾歲、而且穿著正式服裝的黑人女性正優雅地坐進一張破舊的椅子」這句話所傳達的資訊，就遠遠多於只有簡略地說「有一個進入房間，而另一個人坐了下來」。

你應該要留意身體外觀的一些面向，像是整潔乾淨、衣著、以及髮型樣式等等，因爲這些面向會傳達出一些可能影響社會互動的訊息。人們花費大量的時間與金錢去挑選衣服、梳理髮型、用化粧品打扮自己、刮鬍子、燙衣服、以及使用體香劑或是香水等等。而這些事物都是他們自我呈現的一部分。即使是那些不打扮、不刮鬍子、或者是不用體香劑的人，也是以某種方式在呈現他們自己，並且透過外表來傳達出某種象徵性的訊息。沒有任何人的衣著或是外觀是「正常、標準的」，如果你作出了此種陳述，那就暗示了你對於社會信號還不夠敏感。

人們所做的事也具有意義與重要性。你應該要去注意人們坐著或是站著的地點、他們走路的步調、以及他們的非言語溝通方式。人們會透過非言語的溝通方式來表達社會資訊、感覺、以及態度，這些非言語的溝通方式包括了手勢、臉部表情、以及站著或是坐著的方式（以僵硬的姿勢站著、或者是沒精打采地坐著等等）。人們會利用他們在團體中安排位置的方式、以及透過眼神的接觸，來表達出彼此之間的關係。當你注意到某些人們彼此站在一起靠得很近、看起來很放鬆、而且彼此之間也有眼神接觸時，你或許可以從中讀出社會溝通的訊息。

你也可以去留意事件發生的情境脈絡，像是「什麼人出現在現場？」「誰剛剛抵

達或是離開現場？」以及「那個房間是不是又熱又悶？」等等。此種細節或許能夠幫助你爲事件賦予意義，並且瞭解某個事件發生的原因。如果你沒有去注意細節，這些細節就會遺失，而你也就無法完整地瞭解某個事件了。

在田野研究中，意外發現的事物以及偶然的相遇都是相當重要的。在許多時候，你必須一直到後來才會知道你正在觀察的事物所具有的重要性或是關聯性。此種情況具有兩個暗示。第一，隨時進行敏銳的觀察並且記下詳細優秀的筆記是相當重要的，即使是在似乎沒發生什麼事情的時候亦然。第二，隨著時間經過，定期回顧過去的事件並且學習欣賞等待的時間，也是相當重要的。大多數田野研究者都說他們花了很多時間在等待。而剛入門的田野研究者經常會因爲他們似乎浪費了大量時間而感到挫折，因爲他們似乎總是在等待其他人出現、或者是等待事件發生。新手研究者必須要學會的是，等待的時間是田野工作中必要的一個部分，而且也可能是相當具有價值的。

你必須去瞭解田野環境的節奏、學習根據人們的行程表來安排活動、並且去觀察事件是如何在人們自己的時間流動中發生的。除此之外，等待的時間並不總是等同於浪費掉的時間。等待的時間其實是一個進行反思、觀察細節、發展社會關係、建立融洽關係、以及讓田野環境中的人們習慣你在場的好機會。等待的時間也能夠展現出你的投入以及認真，因爲堅忍不拔也是必須加以培養的重要特質。你可能會不耐煩地想要趕快進入田野場域、結束研究、然後繼續過你的「真實生活」，但是，對於在田野場域中的人們而言，這就是他們的「真實生活」。你應該調整個人的需求以去配合田野場域的要求。

一名優秀的田野研究者會去仔細地傾聽人們說話的內容、說話的方式或者是暗示的意義，以及人們在說話時所使用的慣用語、腔調、以及不正確的文法等等。例如，人們通常會使用一些慣用語，像是「你知道的」、「當然」、或是「等等之類的」。你應該要去瞭解這些慣用語背後的真正意義。你可以試著去傾聽所有的對話，但是當許多對話同時發生、或者是當你正在偷聽時，要清楚地傾聽所有的對話就會比較困難。幸運的是，重要的事件以及主題通常會被重覆提起。

當人們彼此進行互動一段時間之後，就會發展出共同的象徵符號以及術語。換句話說，人們會創造出新的詞彙、或者是爲一般的詞彙賦予新的意義。新的詞彙會從特定事件、假定、或者是關係當中發展出來。能夠瞭解並且使用此種語言，就表示了擁有某個獨特次文化中的成員身分。你應該去學習此種專門化的語言，或者又稱爲**行話**（argot）。[39]

你應該要先從一個前提開始，也就是說，在你的世界中所使用的詞彙以及符號，在你的研究對象的世界中可能會有不同的意義。除了你早已熟悉的詞彙之外，你必須去理解這些新的詞彙以及它們在脈絡中的用法（Bogdan and Taylor, 1975: 53）。

你應該要去辨識出行話如何契合於社會關係或是意義之中。行話能夠提供你一些線索，讓你知道成員們認爲什麼事物才是重要的，以及他們以何種方式來看待這個世界。例如，道格拉斯（Douglas,

行話　由一個次文化或者是一群定期互動的成員所使用的特殊語言或是術語。

1976: 125）在關於天體營的研究中發現了「虎視耽耽」（"vultching"）這個詞，而這個詞是天體營的成員們用來稱呼那些在沙灘上圍繞在有吸引力的裸體女性身邊的男性。

卡托維契以及戴蒙德（Katovish and Diamond, 1986）曾經針對販售渡假小屋的公司的銷售手法進行過一項研究，在進行這項研究的時候，他們在六個月的期間之內進行了觀察以及非正式的訪談，而且在這段期間之內，其中一名研究者被該公司正式聘用爲員工，而另外一名研究者則是充當受訓者的角色。他們把銷售中心當作是一個舞台來加以分析，而在這個舞台上會針對潛在的買家呈現出一系列的事件，除此之外，他們也分析了員工們之間所使用的行話。例如，當財務經理進到銷售中心、而且當業務員與潛在的買家在討論房價時「洩漏」（"drops"）了一些資訊，就可能會有「降價」（"drops"）的機會。此種經過設計的事件是爲了要刺激買氣。常見的手法是故意透露一些消息，像是某個之前買了二十個單位的大公司剛剛決定只要買十五個單位就好了，因此突然之間就有五個單位可以用特別的價格來購買；或者是之前的一個客戶因爲銀行不願意核發貸款，所以他原本買下的單位可以用折價的方式出讓；或者是說只有少數特殊的成員才有資格用優惠價格成交等等。

田野研究者通常必須在田野的行話以及外在世界之間來來回回地進行轉譯。史培德利（Spradley, 1970: 80）在引述一名「城市流浪者」（"urban nomad"）所說的話時，提供了一個例子。他說「如果一個人沒和警方打過交道，就不能算是個流浪漢。」這句話的意思是：如果一個人不曾因爲在公共場所喝醉被逮捕、並且在看守所裡待過一夜的話（也就是「和警方打交道」），就不會被認爲是該次文化（也就是「流浪漢」）的真正成員。在你待在田野中一段時間之後，你或許會對使用這些特殊的行話覺得很自在，但是太快就學著撂行話是不明智的舉動，而且也可能會有被當成傻瓜的風險。

記錄資料 在田野研究中，資訊超載是常見的事，而且也會讓個人的能力備受考驗，不論這個人在記錄資料方面有多麼熟練都一樣。大多數的田野研究資料都是以筆記的形式來呈現的。完整的田野筆記包括了地圖、圖表、照片、訪談、錄音帶、錄影帶、備忘錄、田野中的物品、在田野中所記下的隨手筆記、以及在田野環境之外的其他地方所詳細寫下的筆記等等。你應該預期會寫滿許多筆記本、或者是在電腦中佔據許多儲存的空間。你花在寫筆記上面的時間，可能會多於你待在田野中的時間。有些研究者光是三個小時的觀察資料，就可以填滿四十頁不留空行的筆記。經過練習之後，你在田野中每待一個小時，就應該能夠產生數頁的筆記資料。

撰寫筆記通常是一種無趣、冗長乏味的工作，而且需要自律能力。筆記中包含了你利用記憶而寫下的廣泛描述性細節。艾默森以及同事（Emerson and colleagues, 1995）認爲優秀的田野筆記既是一種心態與習慣、也是一種活動，並且也說（p. 40）：「也許比民族誌學者花了多少時間待在田野裡這個問題更加重要的，是撰寫田野筆記的時機問題……在離開田野環境後*立刻*就寫下田野筆記，能夠提供更爲鮮明、更加詳盡的回憶……」（強調的部分是作者原本就加上去的）。如果可能的話，你應該總是在當天的想法以及興奮之情開始消退之前、而且在還沒有向別人重述事件的時

候，就趕緊寫下筆記。在最短時間之內把鮮明的記憶寫進筆記裡，通常能夠引起一種情緒的抒發，並且刺激具有洞察力的反思。在開始寫田野筆記時，你在田野場域中每待一個小時，就應該撥出大約半個小時的時間來撰寫田野筆記。

你必須把筆記保管好，而且也必須把筆記整理好，因爲你將會不斷地回去反覆閱讀之前寫過的筆記。一旦寫下來了，這些筆記就是你的私人資料，而且具有極高的價值。你必須小心地保存這些筆記，並且維持筆記的機密性。成員們有保持匿名的權利，而且大多數研究者在筆記中都會使用假名。由於敵對的另一派人馬、敲詐勒索者、或者是執法官員等等都可能會對田野筆記有興趣，因此有一些研究者會以密碼的方式來撰寫田野筆記。

你的心理狀態、注意的程度、以及在田野中的狀況，都會影響到記筆記的工作。在撰寫筆記之前，最好是先從在田野中待上相對較短的一到三個小時開始。強森（Johnson, 1975: 187）曾經說過：

> 觀察記錄的質與量，會因為田野工作者精力充沛或是精疲力盡的感覺、對於特定事件的反應、與其他人的關係、酒精飲料的飲用量、以及個別觀察的次數等等因素而有所不同。

田野筆記的類型　田野研究者會以許多種方式來記下田野筆記。[40] 這裡所推薦的一些方式（請參閱資訊箱 13.4 關於記田野筆記的建議）就只是我們的建議而已。完整的田野筆記有數種不同的類型，而在這兒則討論了五種主要的類型（請參閱圖 13.4）以及補充的類型。通常最好的作法是把某一段觀察時期的所有筆記都保存在一起，並且把各種不同的筆記放在個別的頁次上，以方便你辨別不同類型的筆記。有些人會把進行推論的筆記和直接觀察的筆記放在一起，但是利用像是括弧或是有顏色的墨水之類明顯而見的方式來區別出這兩者。筆記的數量也會因爲類別不同而有極大的差異。例如，在田野中待了六個小時可能會產生一頁隨手筆記、四十頁直接觀察的筆記、五頁的研究者推論，以及總共兩頁關於方法論、理論、以及個人的筆記。

1. *隨手筆記*　在田野中幾乎不可能好好地寫筆記。即使是一個大家都已經認識的觀察者，如果在公共場合中低頭猛寫筆記的話，看起來也會讓人覺得非常奇怪。更加重要的是，當你低下頭寫筆記時，你就無法看到或是聽到正在發生的事情。當你在寫筆記時，你會把注意力從田野中移到你的筆記上，但是你原本應該要把注意力放在進行田野觀察上面的。具體特定的環境會決定你是否能夠在田野中記筆記。你或許能夠寫筆記，而成員們或許也預期你會這麼做。或者，你也可能必須偷偷摸摸地記筆記（例如，到洗手間裡記筆記）。正如同麥可德莫特（McDermott, 2006: 88）在田野場域中進行了一場重要的互動之後所提到的：「我很快地編了個理由到洗手間去，然後在那兒急速而潦草地記下筆記……」。

當你在田野中的時候，你所寫下的是**隨手筆記**（jotted notes），此種筆記是一些簡短而且能夠引

隨手筆記　當研究者在田野中時，以不引人注目的方式在任何方便取得的物品上所寫下的田野筆記，其目的在於之後能夠「喚起回憶」。

起記憶的記錄，像是一些詞彙、慣用語、或不經意的素描等等。也許是隨手潦草地記在一些方便取得的物品上面（例如：餐巾紙、火柴盒等等）。之後你將會把這些隨手筆記融入你直接觀察的筆記中，但是絕對不要以隨手筆記來取代直接觀察的筆記。

資訊箱 13.4　關於記田野／實地筆記的建議

1. 在每一段待在田野中的時間結束之後，立刻記下筆記，而且在還沒有把觀察記錄下來之前，不要與他人談話。
2. 在撰寫每一次拜訪田野的記錄時，都從新的一頁開始記錄，並且記下日期與時間。
3. 把隨手筆記當作是一種暫時記憶的輔助方式，記下一些關鍵字或詞彙，或者是最早說或最晚說的事。
4. 在頁邊留下大片空白，以便於你可以很容易隨時增加筆記內容。如果你在後來想起了某些事，就回去加到筆記裡面。
5. 作好筆記的分類，並且把各種不同的筆記分開保存好，這樣之後要回去檢視時會比較容易。
6. 依照發生的順序把事件記錄下來，並且標註事件持續的時間長短（例如，等待 15 分鐘、騎車 1 小時）。
7. 筆記應該盡可能具體、完整、以及清楚。
8. 盡可能清楚地分段，並且運用引號。最好的情況是能夠正確地回憶成員們所說過的話，並且用雙引號標註出來，至於解釋意義的部分則使用單引號。
9. 記錄下一些在當時似乎無關緊要的對話或者是例行公事，這些對話或例行公事在之後可能會具有重要性。
10. 「跟著你的感覺走」，並且迅速地寫下筆記，不必擔心拼字錯誤或者是一些「天馬行空的想法」。你可以假定沒有任何其他的人會看到你的筆記，放心地把所有想法都寫下來，但是記得要使用假名。
11. 絕對不要以錄音帶完全地代替田野筆記。
12. 將圖表或者是田野環境的地圖加進筆記裡，並且大略描述你自己在觀察期間內的行動以及其他人的行動。
13. 在筆記中加入你自己的話語以及行為。並且在另外一個部分記錄自己的情緒感受以及個人想法。
14. 避免使用評價性的總結字眼。不要寫下像是「那個水槽看起好噁心」之類的記錄，而是要說「那個水槽已經生銹了，而且看起來好像已經有很長一段時間沒有清理過了。水槽裡的食物殘渣及骯髒的碗盤看起來好像已經堆在裡面好幾天了。」
15. 定期地重新閱讀筆記，並且把在重新閱讀的過程中所產生的想法記錄下來。
16. 一定要製作一份或者是多份備份，把這些備份的筆記放在一個上鎖的安全地點，並且把這些備份筆記放置在不同的地點，以免碰到火災、水災、或者是竊盜等事件時，你的心血可能會毀於一旦。

直接觀察	推論	分析	個人日誌
10月4日，星期天。凱伊咖啡館（Kay's Kafe），下午3點。一位年約40多歲、高大而且過重的白人男性進來了。他穿著一套破舊的棕色西裝。他單獨一個人，坐在第二桌。凱伊走了過去，問說：「要點些什麼？」那個人說：「先給我一杯黑咖啡。」凱伊離開後，他點了一支煙，並且看著菜單。下午3點15分，凱伊打開收音機。	凱伊今天原本看起來很親切，還哼著歌。後來她卻變得嚴肅而且充滿戒心。我覺得她只要一緊張，就會打開收音機。	自從發生過搶案之後，女性都會害怕那些獨自一人出現的男性。	外面正在下著雨，我覺得跟凱伊在一起很自在、很舒服，但是今天蠻無聊的。

圖 13.4　田野／實地筆記的類型

2. *直接觀察的筆記*　田野資料的基本來源就是**直接觀察的筆記**（direct observation notes）。你在離開田野之後，要立刻寫下直接觀察的筆記，而且你在之後隨時可以回去增加內容。你應該針對每一次進入田野的日期、時間、以及地點的順序依次排列筆記。直接觀察的筆記能夠以非常具體的、特定的詞彙來描述你在田野中所聽到以及看到的人事物。在可能的範圍之內，直接觀察的筆記應該能夠精確地記錄特定的詞彙、慣用語、或行動。

直接觀察的筆記　田野研究筆記的一種類型，此種筆記試圖把研究者在田野中所聽見或看見的事物的所有細節，通通都記錄下來，也允許在之後進行多重的詮釋。

你的記憶力會隨著練習而不斷改善，而你很快地就會記得在田野中所真正說過的話。口頭的陳述應該要用雙引號的方式寫下，以和釋義的部分區隔開。對話的一些附加特徵（非言語的溝通、道具、語調、速度、音量、手勢等等）也應該被記錄下來。記錄下實際說過的話，不要加以修飾或改寫，並且把不合文法的說法、俚語、以及口誤（例如，你應該如實地寫下「啊，我要肥家了，小莎」，而不是「我要回家了，莎莉」）都完整地記錄下來。

在筆記中應該寫入具體的細節，而不只是摘要而已。例如，不要只有寫下「我們談論了關於運動的事情」，而要寫下「安東尼和山姆、傑森起了爭執。他說芝加哥小熊隊會贏得下星期的比賽，因爲小熊隊剛換到了一個叫做奇阿佩塔（Chiappetta）的新游擊手。他還說小熊隊比紐約大都會隊強多了，因爲他認爲大都會隊的內野手比較差。他引述了上個星期的一場比賽，在那場比賽中，小熊隊以 8 比 3 贏了波士頓紅襪隊。」你應該記下誰在現場、發生了什麼事、發生地點、發生時間、以及在什麼情況下發生。新手研究者可能會因爲覺得「沒有發生什麼重要的事」，所以沒有記下任何筆記。而經驗豐富的研究者知道，在「沒發生什麼事」時的所發生的事件，其實可以揭露出許多訊息。例如，即使是在看似微不足道的對話當中，成員們也可能表達了自己的感覺，並且把經驗歸類爲各種不同的民俗類別。

澤魯巴費爾（Zerubavel, 1981）提出了一個有用的方式來思考在田野中的時間，他檢視了社會生活的節奏，並且主張社會活動之間的協調是以時間的組織爲基礎的。

在你的直接觀察筆記中，你或許會試著去注意並且記錄的時間模式有以下四種：（1）順序結構（sequential structure）：什麼事是最早發生的、接著發生的、然後再接著發生的等等，也就是事件發生的順序（依序排列，或者是以之前與之後來標記）；（2）持續時間（duration）：社會事件的時間長短（過於冗長或是過於簡短）；（3）時間位置（temporal locations）：一天之中、一週之中、一個月之中、以及一年之中某個時間點的社會意義（太早或是太晚）；以及（4）重覆發生（reoccurrence）：某些事件重覆出現，或者是與社會規範有關聯的時間循環（經常發生或是還不足夠）。

3. *推論筆記* 你應該要仔細去傾聽成員所說的話，以能與他們「感同身受」（"climb into their skin"）或者是「設身處地」（"walk in their shoes"）來思考事物，[41] 而這牽涉到一個三步驟的過程：先傾聽而不去加以分析歸類；把你所聽到的話，和其他時候所聽到的、以及其他人所說過的話來進行比較；以及運用你自己的詮釋來進行推論、或者是理解這些資訊所代表的意義。在一般普通的互動過程中，我們會同時進行這三個步驟，並且很快地就跳到我們自己的推論。而在田野研究中，你必須學習先觀察以及傾聽，不要急著去作出推論、或者是強加某種詮釋。而這些不帶任何推論的觀察結果，必須記錄在直接觀察的筆記裡。

你可以在另外一個單獨的地方把你的推論記錄下來，而這對於直接觀察是很關鍵的。我們永遠也無法直接看見社會關係、情緒、或者是意義，我們只能看見特定的身體行動以及聽見人們所說的話語，然後運用我們的背景文化知識、來自於脈絡的線索、以及已經做過或是說過的事物，來爲這些動作或話語賦予意義。例如，雖然我們無法直接看見愛或是憤怒，但是我們可以看見並且聽見特定的行動（臉紅脖子粗、聲音很大、大動作的手勢、低俗污穢的言語等等），並且從這些事物當中去作出推論（這個人正在生氣）。

我們經常會根據我們所看到以及聽到的事物來推論出社會意義，不過，我們的推論不一定總是正確的。例如，我的侄女來拜訪我，而且陪我到一家商店去買風箏。在收銀機旁邊的店員微笑著問她是不是那天就要和她的「爸爸」（店員看著我）一起去放風箏。顯而易見的，這位店員觀察了我們之間的互動，然後推論出我們的關係是爸爸和女兒，而不是伯父和侄女。這位店員的確看到並且聽到一個成年男性和一個女孩，但是她所推論出來的社會意義卻是不正確的。你應該把推論的意義和直接觀察記錄分開，因爲行動的意義並不一定總是不證自明的。有些時候，人們其實會試圖去欺騙別人。例如，有一對彼此沒有關係的男女一同住進了一家汽車旅館，而他們所登記的名字卻是史密斯夫婦。更爲常見的情況是，社會行爲是模稜兩可的、或者是可能有多重意義的。例如，我看見一對年約三十歲的白人男性和女性下了車，並且一同進入一家餐廳。他們坐在同一張桌子旁、點了餐、刻意壓低了聲音而且表情嚴肅地在談話，有時候甚至還會靠近對方去聽對方所說的話。當他們起身離開時，那名面帶悲傷神色的女性似乎就快要哭出來了，而那名男性則是短暫地抱了她一下。之後，他們就一起離開餐廳了。我所看見的這兩個人之間到底是什麼樣的關係

呢？是一對即將分手的情侶？或兩個朋友正在討論他們的共同朋友？或是因爲發現了彼此的配偶有婚外情而在商量應該怎麼辦？還是一對父親剛剛過世的兄妹？**推論獨立**（separation of inference）允許我們在重新閱讀直接的觀察筆記時，可以產生多重的意義。如果你不把推論的意義獨立區隔開來，你就會漏失其他可能的意義。特喬拉（Tjora, 2006: 433）曾經說過，你應該「把你『知道』已經發生的事物」、以及你『認爲』已經發生的事物通通都記錄下來……而不是把這兩者與實際的觀察結果混在一起。」

4. *分析備忘錄*　當我們在田野裡的時候，會作出許多關於如何進行研究的決定。我們會計畫一些行動（例如：進行一次訪談、觀察某個特定的活動等等），而其他的行動似乎是臨時發生的。大部分的田野研究者都會寫下分析筆記，以留下一份關於計畫、策略、倫理以及程序方面的決定、以及對於策略的自我評判的記錄。

在田野研究中，理論會在蒐集資料期間以及在回顧田野筆記時浮現出來。理論筆記是一種連續的敘述，描述了你在爲田野事件賦予意義時所作出的嘗試。也就是說，你就像是在這些筆記中「自言自語」（"think out loud"）一樣，而在這些筆記裡，你可能會提出不同想法之間的連結、發展出假設、提出猜測、以及發展出新的概念。

分析備忘錄（analytic memos）包括了方法論的策略以及關於理論的筆記，其中包含了你的想法、對於理論的系統性討論、關於你所作過的決定的記錄。當你在仍然留在田野中時，你可以利用分析備忘錄來詳盡闡述以及擴展一些想法，並且可以在重新閱讀以及思考這些備忘錄時，去修正或者是發展出更加複雜的理論。

5. *個人筆記*　正如同前面已經討論過的，個人的感受以及情緒反應會變成資料的一部分，而且也會影響到你在田野中所看到或是聽到的事物。你應該在筆記中保留一個像是個人日記的部分，並且在裡面記錄個人生活事件以及感受（例如：「我今天有點緊繃，我猜想這會不會是因爲我昨天才剛和別人吵過架所造成的……」，或是「今天這種陰鬱、多雲的天氣，讓我頭痛了一整天」等等）。

個人筆記提供了一種應付壓力的方式，因爲這些筆記是關於個人反應的資料來源，而且當你之後在重新閱讀田野筆記時，也能幫助你評估你所寫下的直接觀察筆記或者是推論筆記。例如，如果你在觀察期間內的心情非常好，就有可能會以正面的方式來解讀你所觀察到的事物。

6. *訪談筆記*　如果你進行了田野訪談（稍後將會進一步討論），你應該把訪談筆記分開保存好。[42] 除了記錄下問題以及答案之外，你也應該製作一個**封面頁**（face sheet），封面頁會被安排放在筆記的最前面，其中包括了像是日期、訪談地點、受訪者的特徵、以及訪談內容等等的資訊。在你重新閱讀筆記時，封面頁可以幫助你瞭解這些筆記。

推論獨立　一種過程，在此種過程中，田野研究者在撰寫直接觀察的筆記時，會把所觀察到的事物、與研究者所推論或是認為發生的事物彼此獨立區隔開來。

分析備忘錄　當質化研究者在檢視資料的細節，並且發展出更為抽象的想法、主題、或者是假設時，所記下的筆記。

封面頁　在訪談筆記或是田野筆記最開頭部分的那一頁，上面記載了日期、觀察地點、訪談、以及情境脈絡等等資訊。

7. *地圖、圖表、以及人工製品*　你可能會想要繪製地圖，並且畫出田野場域特徵的圖表或是圖畫。[43] 這些地圖或圖表可以達到兩個目的：幫助你組織在田野中的事件，以及幫助你傳達田野場域給其他人知道。例如，當一位研究者在觀察一個有 15 張高腳凳的酒吧時，這位研究者可能會畫出 15 個圓圈、並且標上號碼，以簡化記錄的過程（例如：「陽介（Yosuke）進了酒吧並且坐在第 12 張高腳凳上，而費比（Phoebe）則已經坐在第 10 張高腳凳上了。）

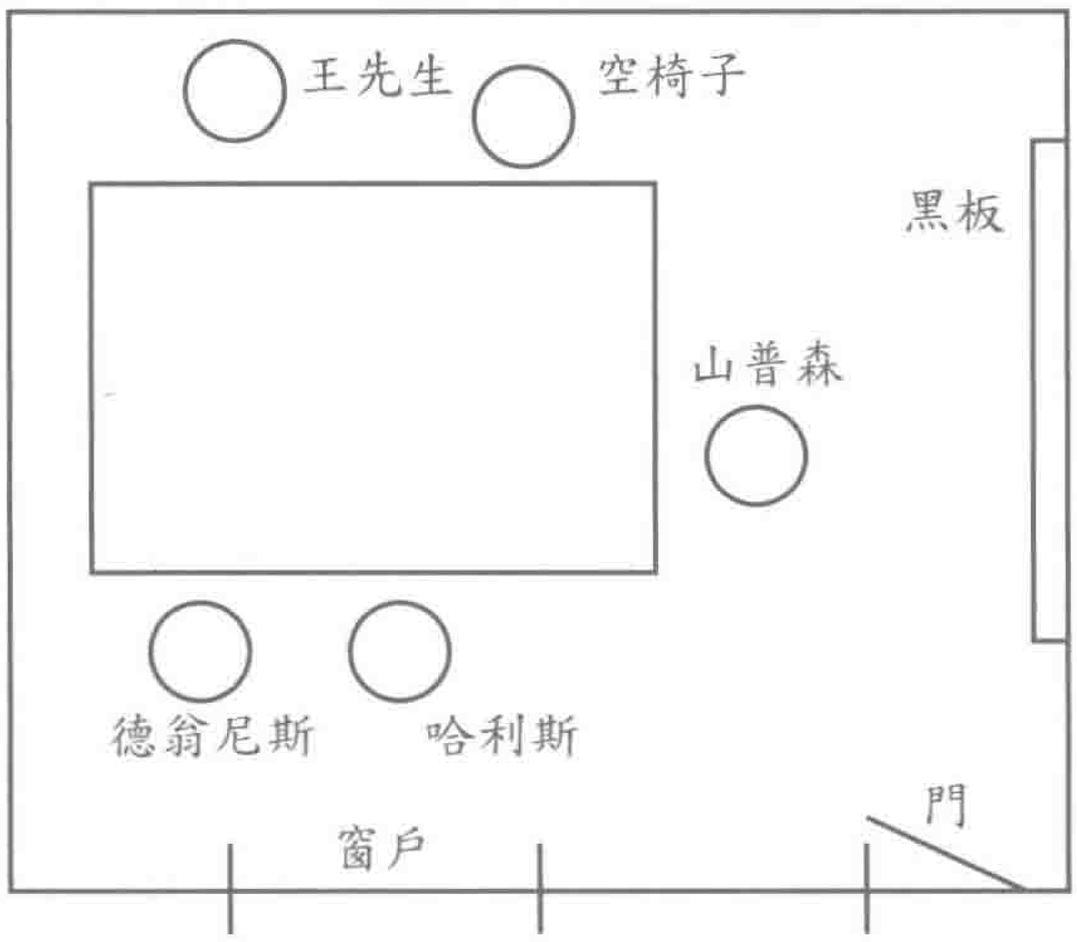
有三種類型的地圖對你會所有幫助：空間的、社會的、以及時間的。第一種地圖可以幫助你定位資料，後兩者則是資料分析的初步形式。*空間地圖*（*spatial map*）可以標示出人群、設備、以及其他事物的實體位置與空間，以顯示出活動發生的地點（請參閱圖 13.5a）。*社會地圖*（*social map*）則顯示出人們的數量或是種類，以及人們之間在權力、影響力、友誼、以及勞力分工等等方面的安排情況（圖 13.5b）。至於*時間地圖*（*temporal map*）則是顯示出人們、物品、服務、以及溝通或是行程方面的來來往往（圖 13.5c）。

除了這些我們製作來分析田野場域的地圖之外，許多田野研究者也會蒐集人工製品，也就是從田野場域中取得的物品。這些可以充當實體證據的物品（例如，一本小冊子、菜單、咖啡杯、T 恤、節目表或是參加者名單、派對帽等等），能夠幫助你回想起在田野場域中的情景。你可以利用這些物品來幫助你喚起回憶、闡述一個主題、或者是象徵或代表某個活動或事件。

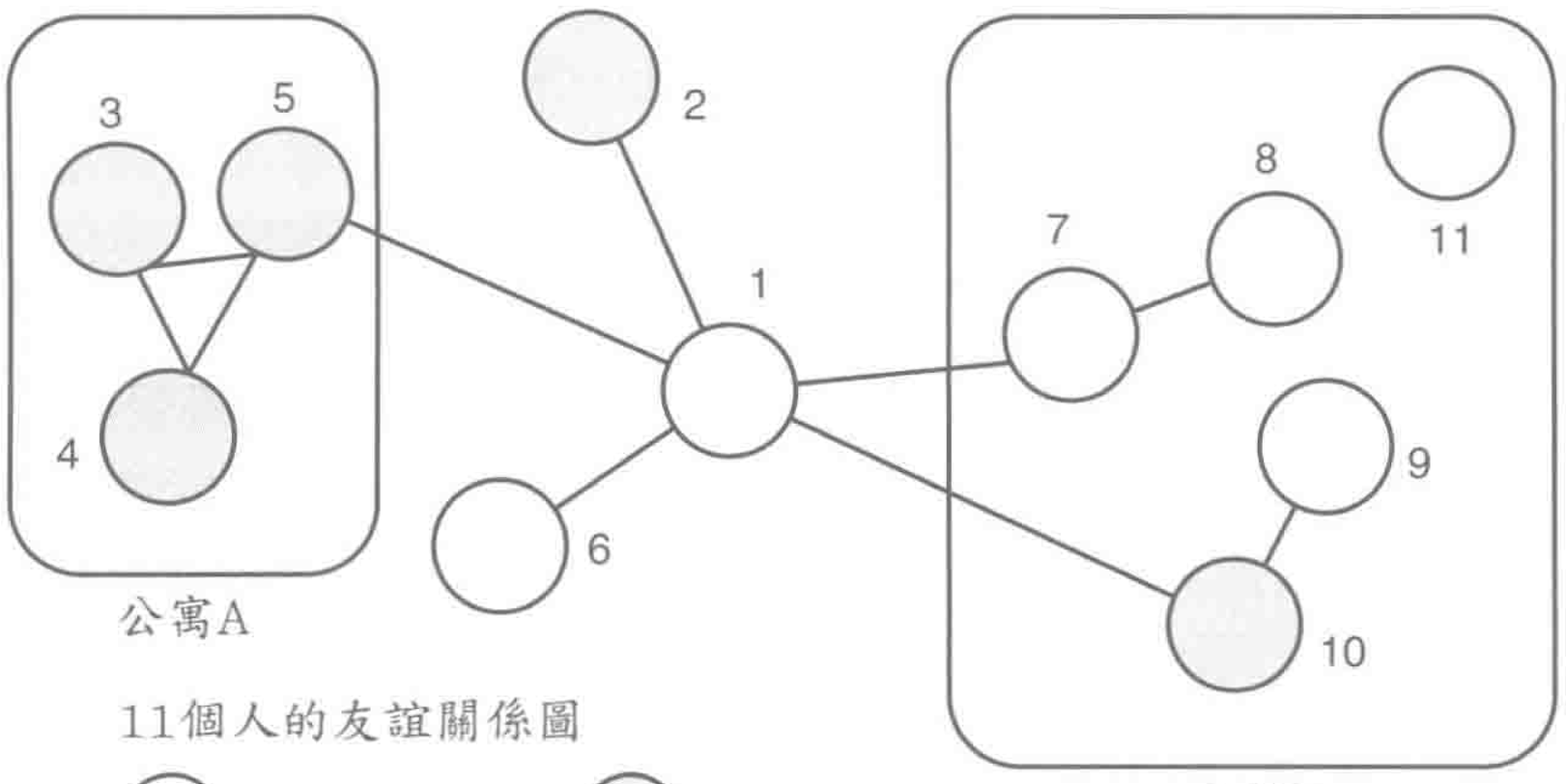
8. *機器記錄的資料*　在田野研究中，照片、錄音機、以及錄影帶等等，都是有用的補充物品，但是絕對不能夠以這些物品來取代田野筆記、或者是取代你親自到田野中的重要性。你無法把這些物品都帶進所有的田野場域中，而且只有在建立起融洽的關係之後，你才能使用這些物品或是器材。錄音帶和錄影帶幾乎可以相當忠實地記錄下所發生過的事情，並且是一種可以讓他人檢視的永久記錄。錄音帶和錄影帶可以幫助你回想事件，並且觀察哪些事件並未發生、或者是沒有回應的狀況，而這些情況都是很容易就會被遺漏掉的。然而，這些器材或是物品也會造成干擾，而且會增加一種被監視的感覺。依賴這些器材的研究者必須去處理一些相關的問題（例如，確定電池已經充飽電、擁有足夠數量的空白錄音帶等等）。除此之外，重聽錄音帶或者是重看錄影帶可能會相當耗費時間，例如，你可能必須花費 100 個小時，才能聽完你在田野中所錄的 50 個小時的錄音內容。如果你想把把錄音帶的內容謄寫出來，通常必須花費極高的代價，而且內容也不一定正確。而且這些謄寫出來的內容不一定能夠傳達出微妙的脈絡意義或者是一些含糊不清的話語。[44]

步驟六：離開田野場域　在田野中的工作可能持續數個星期到十幾年之久。[45] 不論是哪一種情況，到了某個時間點的時候，田野工作仍然必須劃下句點。有些研究者認爲，當發展理論的過程中止了、或者是達到一個結束點時，田野工作也就會自然而然地結束了。其他人則認爲田野工作可能會永無止境地繼續下去，因此就必須要堅決地作出結束一切的決定。

a. 空間地圖



b. 社會地圖



c. 時間地圖

巴茲（Buzz）酒吧一週當中每天的狀況

	星期一	星期二	星期三	星期四	星期五	星期六
開門 **10:00**	老酒鬼	老酒鬼	老酒鬼	老酒鬼	蹺班或是早退的人	去釣魚
5:00	足球賽的觀眾	鄰居以及打橋牌的人	壘球隊（男性之夜）	年輕客人	震耳欲聾的音樂以及各種不同的客人	孤獨的人以及沒有約會的人
打烊 **1:00**						

圖 13.5　在田野／實地研究中所使用的地圖類型

經驗豐富的田野研究者會預期必須經過一個脫離的過程，才能夠離開田野。依據涉入的強度以及待在田野中的時間長度，這個脫離的過程對於研究者以及成員而

言，有可能會造成很大的影響，而且也會造成情緒上的痛苦。當你在離開田野時，你可能會在情緒上感受到放棄親密友誼的痛苦。由於個人以及情緒上的糾纏，你也許會發現要放棄這一切是很困難的。如果你涉入的強度很高、時間很長，而且田野場域和你自己本身的文化之間也有很大的差異性時，你可能需要花上好幾個月的時間來調適，才能在自己本來的文化環境中感到自在。

一旦你決定要離開了——可能是因爲研究計畫已經到達一個自然的終點，而且已經沒有什麼新的發現了，或者是因爲外在的因素迫使研究計畫必須結束（例如：某個工作結束了、守門員要求你離開等等）——就必須選擇一個離開的方法。你可以快速地離開（就是從此再也不回來了），或者是慢慢地從田野中抽離，也就是花上好幾個星期的時間逐漸減少涉入的程度。你也必須決定要如何把這個消息告訴成員們，以及要在多早之前就先提到這件事。

離開的過程取決於特定的田野環境以及你所發展出來的關係。一般而言，你應該至少要提前一些時間讓成員們知道你即將要離開了。如果你曾經和成員們有過協議或是承諾，那麼你就應該一一實現，如此一來才能毫無牽掛地離開。有些時候，某種儀式或是典禮，像是歡送會或者是與所有人握手等等，能夠幫助你和成員道別並且結束彼此的關係。不過，女性主義研究者則主張在離開之後仍然應該要與成員們繼續維持關係。

你的離開將會對成員們造成影響，有些成員可能會覺得受到傷害或者是被拒絕，因爲一段親密的社會關係就這樣結束了。他們的反應可能會是試著再把你拉回到田野裡，並且讓你更加融入成員之中，或者，他們也有可能會因此而生氣並且充滿怨恨。他們也可能因爲意識到你真的只是一個圈外人，因此變得冷淡而且冷漠。不論如何，一直到脫離以及離開的過程結束之前，田野工作都不算是真正結束了。（請參閱重點摘要 13.1 田野研究過程概要）

田野／實地研究訪談

到目前爲止，你已經讀到了田野研究者如何進行觀察以及記筆記的方式了，除此之外，田野研究者也會去訪談成員，但是田野訪談和調查研究的訪談是不一樣的，而接下來的這個部分所要介紹的，就是田野訪談。

在田野研究中所進行的是非結構性的、非指導性的、以及深度的訪談，這和正式的調查研究訪談在許多方面都是不一樣的（請參閱表 13.1）。[46] 田野訪談包括了提出問題、傾聽、表示興趣、以及記錄下受訪者所說的話等等。

田野訪談是由研究者以及一個或者是多個成員共同合作之下的產物，在此種訪談中，成員們是積極主動的參與者，成員們的洞見、感覺、以及合作，都是討論過程的關鍵部分，而此種討論過程則會揭露出主觀的意義。「訪談者的出現以及涉入的形式——他或她去傾聽、參與、鼓勵、打斷、離題、開始一個話題、以及結束的方式——都是在構成受訪者整體敘述時所不可或缺的部分」（Mishler, 1986: 82）。

重點摘要　13.1

田野／實地研究過程概要

步驟一：準備進入田野

- 要有彈性
- 讓自己作好準備
- 去除焦點
- 自我覺察

步驟二：選擇一個田野場域並且進入現場

- 選擇研究場域
- 與守門員交涉
- 扮演一個社會角色
- 決定涉入的程度
- 建立融洽的關係

步驟三：應用策略

- 協商
- 將研究正常化
- 決定揭露多少資訊
- 聚焦以及抽樣
- 採取陌生的態度
- 注意社會崩潰
- 處理壓力

步驟四：維持在田野中的關係

- 調整以及適應
- 運用魅力以及培養信任
- 提供舉手之勞
- 避免衝突
- 表現出感興趣的樣子
- 成為可被接受的無能者

步驟五：蒐集以及記錄資料

- 吸收以及體驗
- 觀察以及傾聽
- 記錄資料
- 田野筆記的類型

1. 隨手筆記
2. 直接觀察的筆記
3. 推論筆記
4. 分析備忘錄
5. 個人筆記
6. 訪談筆記
7. 地圖、圖表、以及人工製品
8. 機器記錄的資料

步驟六：離開田野場域

田野研究訪談有許多種不同的名稱，像是*非結構性訪談*、*深度訪談*、*民族誌訪談*、*開放式訪談*、*非正式訪談*、以及*長時間訪談*等等。一般而言，這些訪談都牽涉到一個人或者是多個人在場、發生在田野之中、而且是非正式以及非指導性的（也就是說成員可以把訪談引導到各種不同的方向）。[47]

田野訪談涉及到相互分享經驗。你或許會分享你的背景，以建立信任關係並且鼓勵資訊提供者開誠佈公，但是你不能強迫他人提供答案或者是使用引導式問題。你應該去鼓勵並且引導一種相互發現的過程。在她對於年輕人次文化所進行的研究當中，威爾金絲（Wilkins, 2008: 21）提到，當她意外地懷孕並且以單親媽媽的身分繼續進行研究時，這件事「以重要而且通常無法預期的方式改變了我的社會位置」，其中包括了讓她擁有更多的機會去接觸一些有獨特想法的人（也就是一些裝扮成波多黎各人的白人青少年）。

在田野訪談中，在成員們表達自己的想法或是意見時，是以他們平時說話、思考、以及組織現實的方式來表達的。因此，你應該要以原本的自然形式來記錄成員們所說的笑話以及敘述的故事，而不要把這些笑話或故事重新包裝成一種標準的形式。你應該聚焦在成員的觀點以及經驗上面。爲了要能夠貼近成員的經驗，你應該

以具體的例子或是情境來提出問題——例如，與其問「你爲什麼離職？」，你應該這麼問：「你能不能告訴我有哪些事情導致你在六月離職呢？」

田野訪談是在一段長時間內逐步進行的。你應該先從建立融洽的關係開始，並且避免去碰觸到任何評價性的、或者是高度敏感的主題。一直到你建立親密的關係之前，都要避免去刺探成員內心的感覺，而且，即使是在已經建立了親密的關係之後，你也應該考量到成員們對你仍然可能會有一些疑慮。在見面數次之後，你或許能夠更加深入地探問敏感的議題，並且針對較不敏感的議題尋求澄清。在之後的訪談裡，你或許可以回到之前問過的主題並且確認過去的答案，而方法則是以不帶判斷的語調來重述之前的答案，並且請成員們再次確認一遍——例如：「在我們上次的談話過程中，你曾經說過，自從老闆減少你的薪水之後，你就開始從店裡面拿走東西，是這樣嗎？」

田野訪談是一種「言談事件」（"speech event"），比較接近於一種朋友之間的對話，而比較不像是在調查研究訪談中那種刺激／反應的模型（請參閱第十章）。你對於朋友之間的對話應該相當熟悉，而此種對話具有本身的非正式規則，以及下列的

表 13.1　調查訪談與田野／實地研究訪談

典型的調查訪談	典型的田野訪談
1. 具有清楚明確的開始與結束。	1. 訪談的開始與結束都是不明確的。訪談可以在過了一段時間之後繼續進行。
2. 以相同的順序對所有的受訪者提出相同的標準問題。	2. 訪談的問題以及提出問題的順序，都是依據特定的人們以及情境而量身打造的。
3. 訪談者一直保持中立。	3. 訪談者表現出對於回答感到興趣的樣子，並且鼓勵成員進一步說明或解釋。
4. 訪談者提出問題，而受訪者則回答問題。	4. 就像是一場朋友之間的對話，但是訪談者所提出的問題比較多。
5. 幾乎總是只有一名受訪者單獨接受訪談。	5. 可以發生在團體的情境中、或者是有其他人在場的區域中，但是也可能會有所變化。
6. 具有專業的語調以及清晰明確的焦點，任何離題的狀況都會被忽略。	6. 訪談中穿插了笑話、題外話、故事、趣事、逸事等等，而且都會被記錄下來。
7. 比較常使用封閉式問題，而且不常使用探問的技巧。	7. 開放式問題是很常見的，而且也會經常使用探問的技巧。
8. 由訪談者一人單獨掌控訪談的步調和方向。	8. 訪談者與成員一起掌控訪談的步調以及方向。
9. 訪談發生時的社會脈絡會被忽略，而且被假定為不會造成任何影響。	9. 訪談的社會脈絡會被記錄下來，而且在詮釋回應的意義時會被視為是很重要的一部分。
10. 訪談者試圖把溝通型態放入標準的架構當中。	10. 訪談者會調整自己，以去適應成員的規範以及語言的用法。

*資料來源：*改編自 Briggs (1986), Denzin (1989), Douglas (1985), Mishler (1986), Spradley (1979a)。

這些元素：（1）打招呼（「嗨！很高興再次看到你」）；（2）沒有明確的目標或目的（我們不會說：「現在讓我們來討論一下，我們上個禮拜所做過的事情吧！」）；（3）避免明確的重複（我們不會說：「你能不能澄清一下你剛剛說過關於……的事？」）；（4）提出問題（「你有看到昨天那場比賽嗎？」）；（5）表現興趣（「真的嗎？真希望我當時也在場！」）；（6）表示不知情（「沒有，我錯過了，發生了什麼事？」）；（7）輪流說話，以使彼此之間的對話達到平衡（不會有某一個人一直在問問題、而另一個人則是一直在回答問題的情況）；（8）省略語（我們會說「我錯過了那場賽馬，但是我會去看另一場賽車」，而不是說「我錯過了肯塔基州那場賽馬，但是我會去看印第安那州的 500 哩賽車」）；（9）當沒有人說話時，暫停或是短暫的沉默是可以被接受的；（10）結束（我們不會說「讓我們結束這場對話吧！」，相反地，我們在真正離開之前，會有一些口頭上的暗示，像是「我現在得回去工作了，明天見！」）。

田野訪談與朋友之間的對話是不一樣的，因爲田野訪談具有明確的目的：去瞭解成員以及環境。你會在田野訪談中加入一些解釋或者是要求，而這種作法與一般朋友之間的對話是不同的。例如，你可能會說「我想要問你關於……」或者是「你能不能看一下這個，確認一下我寫的對不對？」田野訪談是比較沒那麼平衡的，大部分的問題都是由你所提出來的，而且你會表現出較多不知情以及感興趣的狀況。除此之外，田野訪談中也包含了重覆，因此你也許經常會要求成員去詳細說明一些你不太清楚的省略說法。[48]

田野訪談者也會特別去注意**標記**（markers），也就是「在田野訪談中，受訪者不經意而且短暫地提及某個重要事件或者是感覺狀態」（Weiss, 1994: 77）。例如，在與一名 45 歲的醫生進行訪談的過程中，當這名醫生在描述他在某一門高中課程方面所遇到的困難時，不經意地提到：「正是大約在那個時候，我的妹妹在一場車禍當中受到重傷。」這名醫生之前從未提過他的妹妹或者是那場車禍的事，而當他偶然間提到這些事的時候，就暗示了在事發當時，這是一個很重要事件。你應該注意到這個標記，而且或許在之後可以問說：「在稍早之前，你曾經提到過，你的妹妹曾在一場車禍中受到重傷。你能不能再多談談這起事件呢？」最重要的是，你必須仔細傾聽，不要頻繁地打斷受訪者的話、不要一直幫成員完成句子或是任意地接話、不要提供聯想（例如：「喔，那就好像是……一樣！」）、不要在成員開始回答之後還一直堅持要先念完你的問題、不要爭奪訪談過程的控制權、也不要一直固守著某種想法而去忽略了新的線索。[49] 也許你將會得知一些意料之外的資訊，像是在上述的例子中，由於妹妹曾經發生過車禍，因此激起了那名醫生對於醫學的興趣，而且這起車禍對於他選擇醫生這個職業也具有極大的重要性。

生命史

生命史、生命故事、或者是傳記式訪談，是田野訪談的一種特殊類型，此種訪談與口述歷史（請參閱第十四章）有一些重疊之處。[50] 過去所發生的

標記　在田野訪談中，受訪者不經意而且短暫地提到某個人事物，而此種情況實際上暗示了一個非常重要的事件或者是感覺。

生命史訪談　一種開放式訪談，其中的受訪者會去描述他或是她的整個人生。此種訪談是口述歷史的一種次類型。

故事有多重的目的，而且可能會形塑了訪談的形式。在**生命史訪談**（life history interview）中，我們會去進行訪談並且蒐集關於某個特定個人生命的文件資料。這個特定的個人被稱作資訊提供者（informant），而且通常是年紀較大的人。「生命故事的概念被用來標明回顧的資訊本身，而且不具有生命史這個詞通常所暗示的確定性證據」（Tagg, 1985: 163）。我們會提出開放式問題，以瞭解這個人如何理解他或是她自己的過去。故事的正確性其實不如故事本身來得重要。我們知道資訊提供者或許會重新建構或者是把現在的詮釋加到過去的故事之中，也就是他或她可能會「重寫」他或她的故事。此種訪談的主要目的，就是去瞭解資訊提供者如何去看待／回憶自己的過去，而不是去得到某種客觀的事實（請參閱資訊箱 13.5 生命史訪談）。

當我們在詢問資訊提供者在不同日期以及不同的生命階段中所發生的事情時，我們有時候會使用生命故事方格（life story grid），此種方格的內容，可能是在資訊提供者生命中十個不同的階段裡所發生的各種不同類別的事件，像是遷移、職業、教育、或者是家庭事件等等。我們可以利用人工製品（例如：老照片）來補充訪談資訊，並且在訪談中呈現這些人工製品以刺激討論或者是勾起回憶。「作爲一種實徵的活動，生命史的撰寫因爲信件、文件、訪談等等事物而變得更加豐富」（Smith, 1994: 290）。

麥可克瑞肯（McCracken, 1988: 20）舉了一個例子，說明物品如何協助他進行的訪談，並且幫助他去瞭解受訪者看待事物的方式。當他在一位 75 歲的老婦人家中客廳對她進行訪談時，麥可克瑞肯一開始認爲這個客廳裡堆滿了許多雜亂的物品。然而，在請這位老婦人解釋每一件物品的意義之後，顯而易見地，她認爲每一件物品都是一個紀念品或者是能夠勾起回憶的東西。這個客廳就像是一間博物館，裡面藏滿了她生命中重要事件的紀念物。而當麥可克瑞肯以這種新的方式去看待這些物品之後，他才開始不再把這些傢俱以及物品當作是無生命的東西，而是充滿了意義的物品。

有些時候，我們會發現一個關於某個人的現存資料庫，而在其他時候，我們則是必須去找出相關文件，並且創造出一個資料庫。找到此種文件資料可能會是一個相當艱鉅費力的工作，而且之後還得去審閱、分類、以及組織這些資訊。而訪談資料以及文件資料則會一起形成生命故事的基礎。

田野／實地訪談中所提出的問題類型

在進行田野訪談時，我們會提出三種類型的問題：描述性的、結構性的、以及對照性的問題。我通常會同時提出三種問題，但是在研究過程的不同階段中，某一種問題會比其他兩種問題更常出現（請參閱圖 13.6）。在早期的階段裡，你主要提出描述性的問題，並且逐漸增加結構性的問題，一直到了開始進行分析的中期階段

資訊箱 13.5　生命史訪談

一般而言，生命史或者是生命故事訪談通常會包含了兩次到十次的開放式訪談，每次訪談約為 60 分鐘到 90 分鐘左右，而且通常都會錄音。這些訪談可以達到數個目的。首先，這些訪談可以協助受訪的資訊提供者去重新建構他或是她的生命記憶。把一個人的生命事件當作一個敘事故事來加以重新描述並且回憶，能夠產生治療性的效用，並且把個人的智慧傳遞給下一代。第二，這些訪談能夠在生命循環、自我的發展、以及人們體驗事件的方式等等方面，創造出新的質化資料，而這些資料則可以被建檔歸類並且加入到類似的資料當中（例如，在南緬因州立大學中的生命故事中心，就是此種檔案館）。第三，生命故事訪談能夠讓訪談者深入地去觀察另外一個人的生命，此種經驗通常是一種使彼此生命更為豐富的經驗，能夠創造出親密的個人關係，並且以強化個人完整性的方式來鼓勵自我反思。進行生命史訪談的過程如下：

1. 研究者先進行一些事前的準備，包括閱讀背景資料、強化自己的訪談技巧、與資訊提供者進行接觸、取得進行訪談的許可、並且保證匿名性。
2. 研究者進行一連串的訪談，並且將這些訪談過程加以錄音或是錄影。訪談者先把關於資訊提供者任何先前的歷史懸置在一旁、給予受訪者完全的尊重、而且對於受訪者所說的事情通通表現出感到興趣的樣子。訪談者提出開放式的問題，但是仍然保持彈性，並且絕對不會強迫受訪者一定要說出答案。訪談者所扮演的是指導者的角色，瞭解何時應該提出問題以引出故事。除此之外，訪談者在過程中表現出高度的專注，對受訪者所說的內容完全不加評判，而且完全地支持。通常訪談者會提供照片或是物品來幫助受訪者勾起回憶以及過去的感覺。
3. 研究者依據四個步驟將錄下來的訪談內容謄寫下來：（a）準備每次訪談的摘要；（b）謄寫逐字稿，盡量不作任何編輯（例如，幫句子或是段落加上標點符號等等）以及現場狀況的說明（例如：笑聲、咳嗽聲）；（c）重新檢查整篇逐字稿以澄清意義，並且作更進一步的編輯以及小幅度的重新組合；以及（d）請資訊提供者審閱逐字稿，並針對錯誤之處進行更改或是修正。
4. 研究者寄送一封感謝函給資訊提供者、針對重大的主題準備一份評論、並且／或者把評論內容寄到一個檔案資料中心去留存。

*資料來源：*改編自 A. B. Atkinson & John Hills, 1998. "Exclusion, Employment and Opportunity," *CASE Papers* 04, Centre for Analysis of Social Exclusion, LSE. http://sticerd.lse.ac.uk/case/

之後，結構性問題會佔了極大的比例。至於對照性的問題是在研究的中期開始出現，並且逐漸增加，而到了研究的後期階段時，對照性問題的比重就會超過其他兩種類型的問題。[51]

你提出描述性問題的目的，是要去探索田野環境以及瞭解其中的成員。描述性問題可能是關於時間與空間的問題，例如：「浴室在哪兒？」、「送貨的卡車何時抵達？」、以及「星期一晚上發生了什麼事？」等等。描述性問題也可能與人們和活動

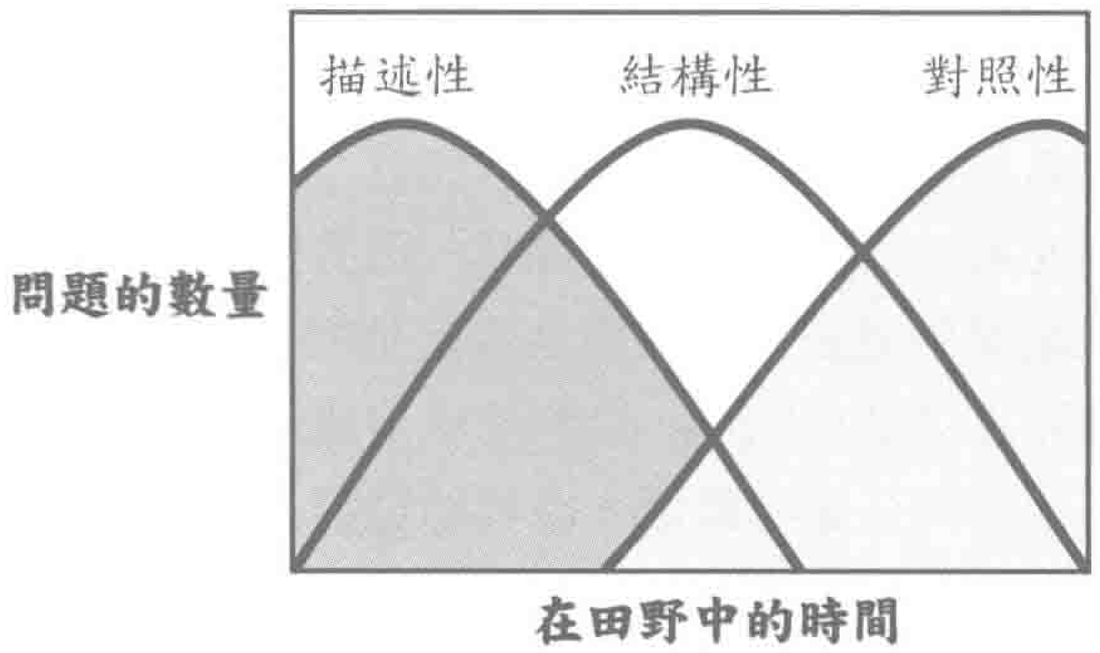


圖 13.6　田野／實地訪談中的問題類型

有關，例如：「現在坐在窗邊的人是誰？」、「你的伯父看起來如何？」、以及「在開幕典禮時發生了什麼事？」等等。描述性問題也可能與物品有關，例如：「你什麼時候會用到線鋸？」、以及「當你從事緊急的漏水維修工程時，你通常會帶著什麼樣的工具？」等等。而要求受訪者提出例子或者是說明經驗的那些問題，也都是描述性問題，例如：「你可不可以提出一個例子來說明所謂的大日子指的是什麼？」、或是「你過去擔任郵局職員的經驗是什麼？」等等。描述性問題也可能會去詢問一些假設的情境，例如：「如果有學生在考試期間打開書本，你會怎麼處理？」而另外一種類型的描述性問題，則是去詢問成員關於該環境中的行話，例如：「你們怎麼稱呼副警長的？」〔答案是「郡騎警」（“county Mountie”）〕。

當你在田野中待了一段時間、並且開始分析資料〔尤其是領域分析（domain analysis），我們將在第十五章裡進一步討論〕之後，你將會開始提出結構性問題。在你把特定的田野事件、情境、以及對話加以組織並且分門別類之後，就應該開始提出結構性問題。例如，在你觀察了一間高速公路休息站的餐廳之後，你發現裡面的員工以非正式的方式將來此光顧的客人加以分類。在初步的分析中，你創造出了一個稱爲「顧客類型」的概念性類別，然後你則利用結構性問題去和成員們談話，以確認有哪些類別。

提出結構性問題的方法之一，就是去詢問成員們，在某個類別裡除了你已經發現的元素之外，是否還包含了其他的元素。例如，你或許會問：「除了常客、老油條、借用洗手間的客人、以及賴著不走的客人之外，是否還有其他類型的顧客？」除此之外，你也可以要求成員幫忙確認，例如：「老油條是你們服務的客人類型之一嗎？」、「你會把……樣的客人稱作是老油條嗎？」、或是「借用洗手間的客人曾經在這兒吃過全套的晚餐嗎？」

對照性問題的基礎，建立在你利用結構性問題所確認過的分析結果之上。對照性問題的焦點，在於區別在同一類別的不同元素之間、或者是在不同的類別之間的相似性以及差異性。而你則是利用對照性問題，來要求成員確認這些相似性或是差異性，例如：「這兒似乎有一些不同類型的顧客來用餐。我曾經聽到你把某些顧客稱作『常客』，而把其他人稱作是『借用洗手間的客人』。常客和借用洗手間的客人之間有什麼

相似之處嗎？」、或者是「在賴著不走的客人以及老油條之間的差別，是不是老油條都不給小費？」、或者是「有兩種客人剛剛進來借用廁所，其中一種是全家人一起進來、另外一種則是單獨一個男性。你會把這兩種客人都稱作是借用洗手間的客人嗎？」

資訊提供者　在田野研究中，資訊提供者是一個田野研究者與其發展出關係的成員，而且這個成員會說出或者是提供關於田野的資訊。[52] 理想的資訊提供者具有四種特質（請參閱資訊箱 13.6 理想的田野／實地研究資訊提供者）。

你可能會訪談數種不同類型的資訊提供者。那些能夠提供有用觀點的對照類型，包括了新來者以及老居民；位於事件中心的人以及位於活動邊緣的人；最近剛剛改變狀態的人（例如，剛剛升官）以及那些沒有變化的人；感到挫折或是貧窮的人，以及快樂或是有安全感的人；掌握權力的領導者、以及在其後追隨的下屬等等。當你去訪談了各種不同類型的資訊提供者時，你應該預期所得到的訊息會是混雜不一的、前後不連貫的。

訪談脈絡　我們都知道，在私人辦公室中的對話，不太可能會發生在一間擁擠的餐廳之中。[53] 因此，訪談通常會選在資訊提供者的家裡進行，如此一來，資訊提供者才能感到舒適自在。然而，這並非總是最好的作法。如果某個資訊提供者在家中會心不在焉、或者是沒有隱私的話，你就必須移動到另外一個環境中去進行訪談（例如：一間安靜的餐廳或者是大學的研究室）。

在訪談中的意義是由完形（Gestalt）所形塑的，所謂的完形，指的是研究者與資訊提供者在一個特定脈絡下的完整互動過程。除此之外，你也應該去留意那些能夠增加意義的非言語的溝通形式（例如：聳肩、手勢等等）。

資訊箱 13.6　理想的田野／實地研究資訊提供者

1. 良好的資訊提供者應該要完全熟悉田野中的文化，並且處於適當的位置，能夠親眼目擊重大的事件。此人應該完全融入於田野中的文化之中，並且不必經過思考就能從事田野環境中的各種例行公事。這個人並非是一個新來者，而且在田野文化中已有多年的深入經驗。
2. 目前正參與田野之中的個人。之前的成員雖然能夠反思田野中的事件並且提供有用的洞見，但是這些人離開直接參與田野的時間越久，他們的回憶就越有可能是他們所重新建構的，而非是真實的情況。
3. 能夠花時間與研究者相處的人。接受訪談可能必須花費數個小時，而且有些成員就是無法接受大量的訪談。
4. 非分析式的個人能夠成為比較好的資訊提供者。非分析式的資訊提供者不但熟知、而且也會運用當地的民俗理論或者是實用的常識。此種情況與分析式的成員形成強烈對比，因為分析式的成員會利用來自於媒體或是教育中的類別來預先分析田野環境。即使是受過社會科學教育的成員，也能夠學習以非分析式的方法來回答問題，但是他們必須把所接受過的教育放在一旁，並且運用成員觀點才行。

資料品質

品質的意義

在田野研究中，高品質資料所指的是什麼？而質化研究者又應該怎麼做才能獲得此種資料？[54] 對於依循著實證主義以及量化取向的研究者而言，高品質的資料意味著具有信度以及效度，也就是對於所有的研究者而言，此種資料都能針對相同的「客觀」事實提供精確的、一致的量數。對比之下，田野研究所依循的是詮釋的取向，並且相信成員們是在一個社會脈絡中以主觀的方式來詮釋經驗，而不是去假定有一個單一的、客觀的事實等著被我們去發現。成員們認爲是真實的事物，來自於社會互動以及詮釋。因此，高品質的田野資料能夠捕捉此種過程，並且讓我們能夠理解成員們的觀點。你所想要獲得的是「豐富」的資料，而這意味著這些資料是具有多樣性的，而且你是在一段長期的時間內以系統性的方式來蒐集資料。我們並不會爲了得到高品質的資料而排除主觀的看法，相反的，高品質的資料應該包含主觀的回應以及經驗。高品質的田野資料是極爲詳盡的描述，而其內容則是你沉浸於成員社會世界的真實經驗中的過程。[55]

田野／實地研究的信度

田野資料的信度所處理的問題，是你對於某位成員或者是某個田野事件的觀察，是否具有內部以及外部的一致性。**內部一致性**（internal consistency）所指的是，以我們對於某個人或是事件所已經知道的一切來判斷，田野資料似乎是正確可靠的，而且排除了人與人之間常見的欺騙形式。換句話說，這些資料能夠契合於一個具有一致性的圖像。例如，某個成員的的行動在一段長時間中、以及在不同的社會脈絡中都是一致而且連貫的。

外部一致性（external consistency）所指的是，田野資料已經被證實過，或者是與其他的、不同的資料來源交叉檢核過。換句話說，這些資料能夠符合於整體的脈絡。例如，其他人可以去證實你對於某個人的觀察結果。而外部一致性所問的問題是：其他的證據是否也能證實你的觀察結果是正確的？

內部一致性　在田野研究中所使用的信度，其決定方式是由研究者去檢驗資料的可信度，以瞭解資料是否形成一個一致而連貫的整體、符合關於某個人或是某事件的其他已知事實、並且避免了常見的欺騙形式。

外部一致性　在田野研究中所使用的信度，其證明方式是由研究者利用多重的資訊來源，以交叉檢核或者是證實質化資料的正確性。

在田野研究中的信度也包括了那些雖然未被說出或做出、但是原本是被期望或者是期待會發生的事物。此種遺漏的部分或者是不存在的資料雖然可能具有很大的重要性，但是卻非常難以察覺。例如，當你在觀察一名收銀員在輪班時間結束所做的動作時，你注意到她並沒有去計算放在抽屜裡面的錢。但是，只有當其他的收銀員在每次在輪班結束時一

定都會去計算抽屜裡的錢的情況下，你或許才會注意到上述的那位收銀員遺漏了這個動作（請參閱在第十五章裡關於反面個案方法的討論內容）。

田野研究中的信度，取決於你的洞見、察覺、懷疑、以及問題。你從各種不同的角度（法律、經濟、政治、個人）來觀察成員以及事件，並且在心裡提出問題，例如：這些錢是從哪裡來的？這些人一整天在做些什麼事呢？

由於你必須依賴成員們所告訴你的資訊，因此，成員們的可信度以及他們的敘述就會成爲信度的一部分。爲了要檢查成員們的可信度，你必須要問：這個人有沒有說謊的理由？他或她是否處於一個能夠得知這些事情的地位上？這個人的價值觀是什麼，而這些價值觀對於他或她所說的話有沒有可能會造成影響？這個人是不是爲了要讓我開心而故意說出這些話？有沒有任何事物可能使他或她不願意把所有的訊息都告訴我？

在你評估可信度時，你應該把主觀性以及脈絡一起列入考量之中。一個人的的觀點以及過去的經驗會影響到他或是她的主觀看法，而主觀看法則會去影響到這個人所做出的陳述或者是行動。與其去評估每一句陳述的真實性，你或許會發現到，陳述本身就具有極大的用處，換句話說，即使是不正確的陳述以及行動，也都可能會揭露出許多訊息。

正如同之前已經提過的，脈絡會形塑了行動以及陳述。在某個情境下所說的話，在其他脈絡中就可能會有所變化。例如，當被問到「你會不會跳舞」這個問題時，如果是在一個有許多優秀舞者的公開場合裡，某位成員可能會回答「不會」，但若是在一個幾乎沒有任何舞者、而且播放著不同音樂的半私人環境裡，則該位成員就可能回答「會」。這並非意味著這位成員在對你說謊，而是答案會受到脈絡所影響。其他四個對於信度的障礙包括了一些可能會誤導你的行爲：錯誤的資訊、逃避、謊言、以及掩護（請參閱資訊箱 13.7 田野／實地資料信度的障礙）。[56]

田野／實地研究中的效度

田野研究中的效度，來自於你對資料的分析是否能夠正確地代表田野中的社會世界。因爲田野研究是幾乎不可能複製的，所以可複製性並不能當作評判效度的標準。田野的各個重要的面向都有可能會發生變化：社會事件以及脈絡會發生變化、成員們會有所變化、不同的研究者之間也會有一些差異等等。總共有四種類型的效度、或者是對研究正確性的測試方式：生態效度、自然歷史、成員確認、以及能夠勝任的圈內人表現。

1. **生態效度**（ecological validity）所指的是，在你所描述的社會世界與成員的世界之間的契合程度有多高。生態效度所要詢問的問題是，你所描述的自然環境是否並未因爲你的出現或是研究程序而受到相對的干擾。如果當你不在場時，事件也一樣會發生的話，該項研究就具有生態效度。

生態效度　某項研究的真實性以及可靠度，證明的方式是表現出研究者對於田野場域的描述能夠符合成員們的描述，而且田野研究者的出現並未造成困擾。

資訊箱 13.7　田野／實地資料信度的障礙

1. *錯誤的資訊（misinformation）*所指的是由於生活的不確定性以及複雜性，因而造成非預期的錯誤。例如，在某家醫院裡的護士們提到了某種像是「醫院的官方政策」之類的話，但是事實上可能根本就沒有此種正式的書面政策。
2. *逃避（evasions）*所指的是刻意不去揭露資訊的行動。常見的逃避方式包括了不回答問題、故意答非所問、轉換主題、或者是刻意以一種模糊或是模稜兩可的方式來回答問題。例如，當談到利用電話應召女郎來引誘顧客參加晚宴的這個主題時，某位業務員顯得不太自在。他說：「沒錯，很多人都會使用這種技倆。」，但是之後在他單獨一人、而且經過我們仔細詢問之後，這名業務員才全盤說出實情，並且承認他自己也用過這種方法。
3. *謊言（lies）*所指的就是刻意去誤導他人或者是給予錯誤看法的假話。例如，某個幫派成員給你一個錯誤的名字和地址，或者是一位教會的牧師刻意浮報信徒的人數，好讓自己的教會看起來比較成功。道格拉斯（Douglas, 1976: 73）曾經提過：「在我所知道的所有其他研究情境中，說謊是相當常見的，不論是在成員之間、或者是對研究者說謊，尤其是在關於成員們覺得相當重要的事情方面。」
4. ***掩護**（fronts）*是共享以及習得的謊言以及欺騙。掩護可能包括了使用實體道具以及同謀者。例子之一，是一間事實上進行非法下注的酒吧。表面上看來，這間酒吧是合法經營的，而且販售一些酒精飲料，但是，只有在經過仔細調查之後，才能發現這間酒吧真正的營業活動。掩護不一定總是帶有惡意的，例如，一個常見的例子是聖誕老人－這是為了讓小孩子們開心而刻意提出的「掩護」。

2. **自然歷史**（natural history）是對於你如何進行研究計畫的詳盡描述。在自然歷史中會完整而且坦白地揭露出你的行動、假定、以及研究程序，以供其他人去進行評估。如果圈外人看到並且接受田野場域以及你的行動的話，那麼該項研究在自然歷史方面就是有效的。
3. 當你把田野研究的結果拿回去給成員們看，並且請他們來判斷研究結果的適當性時，你就是在進行**成員確認**（member validation）的工作。如果許多成員承認、並且瞭解你的描述確實能夠反映他們所熟悉的社會世界時，該項研究就是「通過了成員確認」。成員確認有其限制，因爲在環境中彼此互相衝突的觀點，可能會對你的觀察結果產生意見不一致的情況，而且當結果並未以有利的方式來描述某些團體時，其中的成員也可能會提出反對意見。除此之外，成員們也可能因爲你的描述並不是從他們的觀點出發、或者是不符合他們的目的，因此而不認同你的描述。[57]

掩護　一種在田野中的人們所說的謊言或是欺騙，換句話說，雖然人們實際上從事了某種活動，但是卻刻意說一些誤導的話，以造成與實際發生事物有所差異的印象或是外表。

自然歷史　對於一項計畫進行方式的詳盡描述。

成員確認　一種田野研究者用來展現一項研究的真實性以及可靠性的方法，其作法是要求被研究的人們去閱讀研究者的報告，並且確認其內容是真實的。

4. **能夠勝任的圈內人表現**（competent insider performance）所指的是，某位非成員的人能夠有效地以成員的身分去進行互動、或者是扮演爲一位成員的能力。此種能力包含了說一些圈內人才懂的笑話、以及理解這些笑話的能力。一項有效的研究能夠充分地描述田野中的社會生活、並且提供足夠的細節，好讓一位圈外人也能夠充當爲圈內人。此種效度的限制在於，我們不可能知道在每一種情境之下的社會規則。除此之外，某位圈外人之所以能夠成功地扮演成員，可能僅僅是因爲成員們只是在表現出禮貌，而且不想當眾指出社會錯誤而已。[58]

田野／實地研究中的倫理兩難困境

由於在田野研究的過程中，你個人直接涉入了其他人的社會生活當中，因此也會造成一些倫理上的兩難困境。其中有些會發生在當你獨自一人在田野中、而且幾乎沒有時間去仔細思考倫理問題的時候。你在進入田野之前可能會意識到一般的倫理議題，但是當你在田野中進行觀察以及訪談的過程中，倫理問題通常會意外地出現。我們在此考量了五種在田野研究中的倫理議題：隱密研究、保密性、涉入非法行爲、有權力的人、以及出版報告。[59]

1. *隱密研究*（*covert research*）　最備受爭議的倫題議題，就是隱密與公開的田野研究之間的爭議。[60] 這個爭議牽涉到在完全秘密的或者是隱密的研究中更爲廣泛的欺騙議題，以及當研究者採用了一個僞造的角色、姓名、或是身份時，或者是以某種方式對成員說謊。在研究社群當中，有些人支持隱密的研究或者是欺騙，並且認爲這是在進入田野以及充分瞭解社會生活的某些領域時所必須使用的手段。其他人則是全然地反對此種作法，這些研究者認爲此種作法會傷害到研究者與社會之間的信任關係。[61] 雖然其道德地位仍有疑問，但是有些田野場域或者是活動確實只有以隱密的方式才能去加以研究（請參閱第十五章）。

如果有其他方式的話，盡可能不要選擇進行隱密的研究，而且隱密的研究通常會比公開的研究更加難以進行，這是因爲研究者在維持掩護時會有很大的困難，而且也會一直擔心自己會被成員逮到。羅夫蘭等人（Lofland et al., 2006: 39）提到：「對於倫理敏感、深思熟慮、而且博學多聞的研究者，是判斷隱密性研究是否合理的最佳人選。然而……在你自己熟悉與此種研究有關的問題、爭議、兩難困境，以及熟悉當地勞資關係局的規定與指示之前……我們建議你不要去進行任何的隱密研究。」

2. *保密性*（*confidentiality*）　你可能會得知一些成員們私底下向你透露的私密資訊，而且具有強烈的道德義務去維護資料的保密性。此種義務包括了將資料保密而不透露給田野中的其他成員知道，並且在田野筆記中以假名來稱呼這些成員等等。有些時候你無法在研究報告中直接引述某個人的說法，而處理此種情形的策略之一，就是去找出一份提出相同說法的文件（例如：一份

能夠勝任的圈內人表現　田野研究者用來展現一項研究的真實性以及可靠性的行動，其作法是讓研究者「扮演」成被研究的團體中的一個成員。

舊的備忘錄、一篇新聞報導等等），然後以這份文件去代替成員當作是資訊的來源。

當田野研究者與某位成員在除了他們的研究者與被研究者的關係之外，還發展出親密的個人關係時，就會產生比較嚴重的倫理困境。由於彼此之間有高度的信任關係，因此該位成員就有可能會單獨與田野研究者分享一些比較私人的秘密。何薇兒（Howell, 2004: 346）曾經針對來自於墨西哥瓦哈卡地區的女性進行了一項研究，而正如同她在這項研究中所發現的：

> 一般而言，民族誌學者都會詳盡地描述其研究對象的生活以及狀況，以完整地描繪出文化以及個人事件。然而，資訊提供者可能會試圖在大量的主題中隱瞞一些較為敏感的主題，因為他們不想讓一般大眾知道這些事情……其中包括了對配偶不忠、財富的累積、犯罪活動，以及暴力等等……以陌生人或者是泛泛之交的身分，如果不是不可能、也很難去探問一些在文化方面或者是個人方面較為敏感的主題—像是性侵害之類的問題—……當資訊提供者願意提供這些經驗的相關資訊時，就會引起關於維持倫理準則的問題…當這些資訊是被當作「朋友」之間互相分享的秘密時，情況就會變得更加微妙與棘手，因為這兩位朋友的其中一人同時也是位民族誌學者，其立場是要去發表一些來自於另外一個人的秘密，而且這些秘密可能具有潛在的傷害性……盡可能正確地去呈現出那些會影響到資訊提供者的機會與選擇的真實情況（包括了暴力以及對於暴力的恐懼），是非常重要的，而此種情況也迫使民族誌學者必須去討論這些被仔細保護的秘密，這些秘密在研究者與被研究者的典範當中是不必然要被揭露出來的。然而，在討論這些秘密時，就會強化了研究者去重新思考像是保密性、背叛、以及權力等等議題的重要性……

3. *涉入非法行爲* 當研究者在田野研究中所研究的對象是從事一些非法的、傷風敗俗的、或是不道德行爲的人們時，研究者可能會得知一些非法活動的內容，而且有時候甚至會參與這些非法活動。費特曼（Fetterman, 1989）把此種知識稱爲**罪惡的知識**（guilty knowledge），而且不僅僅是執法人員很有興趣得知此種知識，在田野場域裡的其他成員也會想要知道此種知識。如此一來，研究者將會面臨兩難的困境，也就是一方面要去與成員建立信任以及融洽的關係，但是另一方面卻又不能涉入太多，以免違反了研究者自己的基本個人道德標準。一般而言，研究者通常會與此種行爲偏差的成員之間作出一些明確的協議。

罪惡的知識 關於在田野場域中的人們所從事的非法、不道德的、傷風敗俗行動的資訊，這些資訊並非廣為流傳的資訊，但是研究者卻得知這些資訊。

4. *有權力的人* 許多田野研究者所研究的對象都是那些在社會上處於邊緣地位或者是毫無權力的人（例如，以街頭爲家的街友、貧窮的人、兒童、在科層體制中處於較低階層的工作者等等）。有些人因此而批評研究者刻意忽略了那些有權力的人，然而，在社會中那些富裕或是有權力的人都會有強力

的守門員，並且能夠輕易地阻擋研究者去接近他們。而且與此同時，菁英以及官員們則會去批評研究者們過於偏袒那些比較沒有力量的人。

貝克（Becker, 1970c）利用**可信度的階層**（hierarchy of credibility）來解釋此種情況。可信度的階層告訴我們，那些去研究沒有力量的人、罪犯、或者是低階部屬的人，通常都會被視爲是存有偏見的，而具有官方權威的人則會被假定爲是可信的。許多人會假定位於組織頂端的人有權利去定義事物的發展方式、比低階層的人們擁有更廣闊的觀點與看法、並且擁有權力與地位去做一些重大的事情。因此，「那些偏袒官員的社會學家，就能避免被人指控他們帶有偏見」（Becker, 1970c: 20）。當研究者讓自己沉浸在處於弱勢地位人們的世界裡、發展出對於社會生活另外一面的深入瞭解、並且發表了一個鮮爲人知的觀點時，他們就有可能被指控爲帶有偏見，而這其實單純只是因爲他們正在爲一個鮮少被聽見的社會階層發聲而已。

5. *出版田野報告*（*publishing field reports*）　研究者所獲得並且報告的私密知識，可能會在隱私權以及知的權利之間造成一種兩難的困境。研究者無法總是在不違反隱私或者是傷害名譽的情況下去揭露他們所得知的所有秘密，然而，若是研究者無法公開他們得知的所有資訊，則會使得這些資訊以及細節一直無法爲人所知。當研究者沒有對事件提出完整而且正確的說明時，其他人可能會質疑該份報告遺漏了一些重要的關鍵性細節。

有些研究者建議請成員們去閱讀田野報告，以確認報告的正確性、並且同意研究者在報告中對他們的描繪。然而，若是研究中牽涉到邊緣團體（例如，成癮的人、妓女、吸毒者等等）的話，此種審閱的作法就不太可行，因爲我們必須去尊重成員的隱私。另一方面，審查制度或者是自我審查也可能會是一種危險。因此，折衷的作法之一，是當我們要去揭露真實而且正確的資料時，必須是在這些資料對於提出一個較大的論點、或者是呈現出一幅正確的整體圖像而言是不可或缺時的資料時，才能這麼做。[62]

焦點團體研究

焦點團體（focus group）是一種特殊的質化研究技術，也就是人們會在一個團體討論的情境中以非正式的方式來接受「訪談」。[63] 焦點團體研究在過去的 20 年之間已經快速地成長，而其進行的程序，通常是由一名研究者把 6 到 12 個人聚集在一個房間裡，並且由一位主持人來協助討論一些議題，而進行的時間一般大約是 90 分鐘左右。主持人接受過訓練，在主持討論時必須沒有任何方向性，並且協助所有團體成員進行自由、開放的討論（也就是不能讓某一個人完全主導討論的過程）。團體成員應該

可信度的階層　一種關於可信度的等級的概念，其內容是指當研究者所研究的對象是社會中較為弱勢的成員、而且這些成員的觀點或看法很少被聽見時，研究者就會被指控為帶有「偏見」，而由於有權力者擁有高社會地位，因此他們的觀點或看法就會被接受為「沒有偏見的」。

焦點團體　一群在一個討論的情境中非正式地接受訪談的人們，而且這一群人正在參與一種質化的研究技術。

要具有同質性，但是不應該包括親密的朋友或者是親人。一項典型的研究會使用四到六個個別的焦點團體。焦點團體所討論的主題可能包括了公共態度（例如：種族關係、工作場所的平等）、個人行爲（例如：處理愛滋病的方式）、一項新的產品（例如：早餐的麥片）、或者是一名政治候選人等等（請參閱範例 13.3 討論失去父親以及男子氣概問題的焦點團體）。我們通常會把焦點團體與量化研究結合在一起，而此種結合的作法具有其特定的優點以及缺點（請參閱資訊箱 13.8 焦點團體的優點與限制）。

提供非常清楚的說明與指示、以及謹慎小心地挑選焦點團體的參與者，能夠大幅度地影響焦點團體訪談的結果。正如同威別克、達爾格蘭、以及歐柏格（Wibeck, Dahlgren, and Oberg, 2007: 262）所觀察到的：「由於參與者的詮釋架構以及過去的經驗可能會有所差異，因此在討論開始之前先確保所有的參與者都清楚地知道參與焦點團體的先決條件，就是非常重要的事情。」雖然參與者應該具有適度的同質性，但是此種作法並不總是能夠保證討論具有開放性、或者是成員們一定會願意去坦白地分享自己的信念以及意見。例如，在討論性別敏感的議題時，如果只有單一性別的成員在場的話，那是不足夠的。何蘭德（Hollander, 2004）發現到，許多參與者仍然擔心透露那些被污名化的、創傷的經驗（例如：遭受強暴或是家暴的經驗）。她（p. 626）認爲：「個別的參與者在焦點團體中所說的話，不必然能夠被當作是可靠的經驗指標。參與者可能會根據社會脈絡而去誇大、低調處理、或者是保留一些經驗。」

範例　13.3

討論失去父親以及男子氣概問題的焦點團體

杭特等人（Hunter et al., 2006）曾經與一群年輕的非裔美國人進行過焦點團體研究，其主題是關於在失去父親的情況下成長的經歷。因為只有低於 40%的非裔美國人成長於雙親家庭當中，所以研究者們想要知道青少年男孩以及年輕的男性如何得到他們對於男子氣概的觀念。作者們在一間當地的社會活動中心裡舉行了兩次焦點團體訪談。許多年輕人以及他們的家人都是在這間活動中心裡接受社會服務，而且許多年輕人也都會在活動中心裡打籃球。每一次的焦點團體訪談持續時間大約是 75 到 90 分鐘，而且訪談過程全部都會被錄音下來，之後也會被謄寫為逐字稿。兩次的焦點團體中都分別有 20 名非裔美國男性參與，他們的年齡範圍是從 15 歲到 22 歲之間。作者們是透過諮商師以及與該中心有關的其他管道來召募到這些參與者。大多數的參與者（92%）都還未取得高中畢業文憑，而且當時也都還在就學當中。其中有絕大多數的參與者（91%）是成長於沒有父親的家庭之中的。所有的參與者都是來自於低收入戶或者是低階的勞動階級家庭。焦點團體所討論的主要問題，是參與者們認為「身為男人」的意義是什麼，以及他們想要成為什麼類型的男人。在對逐字稿進行了分析之後，作者們瞭解到，對於年輕男性在成為男人的觀點方面，失去父親是很重要的因素。此種資訊以兩種方式顯現出來：對於父親身分以及男子氣概的一般看法、以及對於父親的自傳式反思，而這些反思會對這些身為年輕男性的參與者們造成重大的影響。失去父親是一個重覆出現的議題，而且連結了一般看法以及自傳。作者們在這些年輕男性所表達的敘述之中發現了數個主題，其中包括了以下的四個主題：（1）有些事情是只有爸爸才能教你的；（2）如果爸爸能夠教你任何事情的話，那麼他應該還在你身邊；（3）媽媽既是媽媽、也是爸爸；以及（4）我將會是家中的男人，而我爸爸不是。

資訊補給站 13.8　焦點團體的優點與限制

優點

- 自然的環境讓人們可以自由地表達意見或是想法。
- 鼓勵那些在被邊緣化的社會團體中的成員去公開地表達自己的意見。
- 人們通常會覺得自己被賦予了一些權力，尤其是在行動導向的研究計畫裡。
- 調查研究者能夠有機會去瞭解人們談論調查主題的方式。
- 有助於量化調查結果的詮釋。
- 參與者們可能會去質詢其他人，並且對彼此解釋他們的答案。

限制

- 可能會產生一種「極化效應」（“polarization effect”）（也就是在團體討論之後，態度變得更為極端）。
- 在一次的焦點團體時間之內，只能討論一個或者是少數的主題。
- 主持人可能會在不知不覺之中限制了團體成員之間以開放、自由的方式來表達意見的過程。
- 焦點團體所產生的想法會少於個別的訪談。
- 焦點團體研究很少會報告關於研究設計／程序的所有細節。
- 成員在只有個人的脈絡、以及焦點團體的脈絡這兩種情況之下所提出的答案可能會有差異，而研究者可能無法去調和此種差異。

此處的脈絡所指的不僅僅是其他的參與者而已，還包括了主持人，以及更大的社會脈絡（例如：大型的社會事件以及潮流）、機構脈絡（例如：焦點團體的地點以及贊助者）、以及地位脈絡（例如：來自於不同社會地位或是職位的人們）。焦點團體應該要依據地位來作出一些區隔。例如，與其把主管和他們的員工混合在一起，你應該要把這兩種不同地位的人分到不同的團體中。同樣的，把教師和他們的學生混合到同樣的團體當中，也是不明智的作法，因爲當地位比較高或是比較低的人在場時，人們通常會以非常不同的方式來回答問題。

結論

在本章中，你學到了田野研究以及田野研究的過程（選擇一個田野場域並且取得進入的權利、在田野中建立關係、觀察以及蒐集資料、以及進行田野訪談）。田野研究者先從資料分析開始，然後在蒐集資料的階段中去形成理論。

有人說田野研究者是直接參與被研究者的生活，並且沉浸在一個自然的環境裡，現在你已經能夠體會這句話的含意了。進行田野研究會對研究者的情緒、個人生活、以及自我意識造成比其他類型的研究更大的衝擊。進行田野研究是一件困難的事情，但是對於一些我們無法以其他方式來研究的社會世界的許多部分而言，田野研究是最佳的研究方式。

要進行良好的田野研究必須結合數種技巧。更明確地說，除了強烈的自我意識

之外，研究者還必須擁有傾聽以及吸收細節的絕佳能力、無比的耐性、敏感性以及對於他人的同理心、卓越的社交技巧、一種「靠你自己」快速進行思考的天賦、看出在人們／事件彼此之間的關聯性的能力、以及在寫作中清楚表達自己的想法與意見的優秀能力。

當我們在研究一小群在當下彼此進行互動的人們時，田野研究是最能夠發揮作用的。對於微觀層次的互動、或者是小型團體的面對面互動而言，田野研究是非常有價值的。然而，如果研究者所關心的是鉅觀層次的過程以及社會結構的話，田野研究的作用就會非常有限。對於發生在遙遠過去的事件、或者是延續數十年之久的過程而言，田野研究幾乎沒有任何的用處。在調查此種類型的事件或者是過程時，下一章將要討論的歷史比較研究會是比較適合的方式。

問題回顧

1. 在芝加哥學派發展過程中的兩個主要階段是什麼？該學派的新聞報導模型以及人類學模型是什麼？
2. 在遵循「方法論實用主義」的田野研究者所做的十項事情當中，請列出其中五項。
3. 對於田野研究者而言，爲什麼在開始田野工作之前先去閱讀文獻是很重要的事？而這件事與去除焦點之間有何關聯性？
4. 對於剛入門的田野研究者而言，一個良好的田野場域應該具有哪些特質？請列出這些特質。
5. 「自我的呈現」會以何種方式影響田野研究者的工作？
6. 陌生的態度所指的是什麼？此種態度爲什麼很重要？
7. 在選擇田野中的角色時，有哪些應該要考量的因素？以及研究者涉入的程度會如何變化？
8. 請指出三種確保高品質的田野研究資料的方法。
9. 請比較田野研究與調查研究訪談之間的差異、以及田野訪談與朋友間對話的差異。
10. 田野筆記有哪些不同的類型或是層次？每一種不同類型的田野筆記所要達到的目的是什麼？

注解

1. See Lofland et al. (2006:2–20).
2. For studies of these sites or topics, see Neuman (2000:345–346). On studies of children or schools, see Corsaro (1994), Corsaro and Molinari (2000), Eder (1995), Eder and Kinney (1995), Kelle (2000), and Merten (1999). On studies of people who are homeless, see Lankenau (1999) and on studies of female strippers, see Wood (2000).
3. For a background in the history of field research, see Adler and Adler (1987:8–35), Burgess (1982a), Douglas (1976:39–54), Holy (1984), and Wax (1971:21–41). On the Chicago School, see Blumer (1984) and Faris (1967).

4. Ethnography is described in Agar (1986), Franke (1983), Hammersley and Atkinson (1983), Sanday (1983), and Spradley (1979a:3–12; 1979b:3–16).

5. See Geertz (1973, 1979) on "thick description." Also see Denzin (1989:159–160) for additional discussion.

6. For more on ethnomethodology, see Cicourel (1964), Denzin (1970), Leiter (1980), Mehan and Wood (1975), and Turner (1974). Also see Emerson (1981:357–359) and Lester and Hadden (1980) on the relationship between field research and ethnomethodology. Garfinkel (1974a) discussed the origins of the term *ethnomethodology*.

7. The misunderstandings of people resulting from the disjuncture of different cultures is a common theme.

8. For a general discussion of field research and naturalism, see Adler and Adler (1994), Georges and Jones (1980), Holy (1984), and Pearsall (1970). For discussions of contrasting types of field research, see Clammer (1984), Gonor (1977), Holstein and Gubrium (1994), Morse (1994), Schwandt (1994), and Strauss and Corbin (1994).

9. See Georges and Jones (1980:21–42) and Lofland et al. (2006:11–15).

10. Johnson (1975:65–66) has discussed defocusing.

11. See Lofland (1976:13–23) and Shaffir et al. (1980:18–20) on feeling marginal.

12. See Adler and Adler (1987:67–78).

13. See Hammersley and Atkinson (1983:42–45) and Lofland et al. (2006:17–32).

14. Jewish researchers have studied Christians (Kleinman, 1980), Whites have studied African Americans (Liebow, 1967), and adult researchers have become intimate with youngsters (Fine, 1987; Fine and Glassner, 1979; Thorne and Luria, 1986). Also see Eichler (1988), Hunt (1989), and Wax (1979) on the role of race, gender, and age in field research.

15. See Douglas and Rasmussen (1977) and Yancey and Rainwater (1970).

16. For more on gatekeepers and access, see Beck (1970:11–29), Bogdan and Taylor (1975:30–32), Corra and Willer (2002), and Wax (1971:367).

17. Adapted from Gray (1980:311). See also Hicks (1984) and Schatzman and Strauss (1973:58–63).

18. For discussions of ascribed status (and, in particular, gender) in field research, see Adler and Adler (1987), Ardener (1984), Ayella (1993), Denzin (1989:116–118), Douglas (1976), Easterday et al. (1982), Edwards (1993), Lofland et al. (2006:22-24), and Van Maanen (1982).

19. Roy (1970) argued for the "Ernie Pyle" role based on his study of union organizing in the southern United States. In this role, named after a World War II war correspondent, the researcher "goes with the troops" as a type of participant as observer. Trice (1970) discussed the advantages of an outsider role. Schwartz and Schwartz (1969) discussed various roles.

20. See Douglas (1976), Emerson (1981:367–368), and Johnson (1975:124–129) on being patient, polite, and considerate.

21. Negotiation in the field is discussed in Gans (1982), Johnson (1975:58–59, 76–77), and Schatzman and Strauss (1973:22–23).

22. On entering and gaining access to field sites with deviant groups, see Becker (1970a:31–38), Hammersley and Atkinson (1983:54–76), Lofland et al. (2006:30–47), and West (1980). Elite access is discussed by Hoffman (1980).

23. See Lofland et al. (2006:22-25).

24. For discussion of "normalizing," see Gans (1982: 57–59), Georges and Jones (1980:43–164), Hammersley and Atkinson (1983:70–76), Harkens and Warren (1993), Johnson (1975), and Wax (1971). Mann (1970) discussed how to teach members about a researcher's role.

25. For more on roles in field settings, see Barnes (1970:241–244), Emerson (1981:364), Hammersley and Atkinson (1983:88–104), Warren and Rasmussen (1977), and Wax (1979). On dress, see Bogdan and Taylor (1975:45) and Douglas (1976).

26. See Lofland (1976) and Lofland et al. (2006) on focusing. Spradley (1979b:100–111) also provides a helpful discussion.

27. See Denzin (1989:71–73, 86–92), Glaser and Strauss (1967), Hammersley and Atkinson (1983: 45–53),

Honigmann (1982), and Weiss (1994:25–29) on sampling in field research.

28.See Gurevitch (1988), Hammersley and Atkinson (1983), and Schatzman and Strauss (1973:53) on "strangeness" in field research.

29.See Gans (1982), Goward (1984b), and Van Maanen (1983b:282–286).

30.See Douglas (1976:216) and Corsino (1987).

31.See Warren and Rasmussen (1977) for a discussion of cross-gender tension.

32.See Wax (1971:13).

33.Also see Adler and Adler (1987:40–42), Bogdan and Taylor (1975:35–37), Douglas (1976), and Gray (1980:321).

34.See Bogdan and Taylor (1975:50–51), Lofland et al. (2006:57–60), Shupe and Bromley (1980), and Wax (1971).

35.See Johnson (1975:105–108).

36.The acceptable incompetent or learner role is discussed in Bogdan and Taylor (1975:46), Douglas (1976), Hammersley and Atkinson (1983:92–94), Lofland et al. (2006:55-57), and Schatman and Strauss (1973:25).

37.See Strauss (1987:10–11).

38.See Georges and Jones (1980:105–133) and Johnson (1975:159). Clarke (1975) noted that it is not necessarily "subjectivism" to recognize this in field research.

39.See Becker and Geer (1970), Spradley (1979a, 1979b), and Schatzman and Strauss (1973) on argot.

40.For more on recording and organizing data, see Bogdan and Taylor (1975:60–73), Hammersley and Atkinson (1983:144–173), and Kirk and Miller (1986:49–59).

41.See Schatzman and Strauss (1973:69) on inference.

42.See Burgess (1982b), Lofland et al. (2006:99-108), and Spradley (1979a, 1979b) on notes for field interviews.

43.See Denzin (1989:87), Lofland et al. (2006: 88), Schatzman and Strauss (1973:34–36), and Stimson (1986) on maps in field research.

44.See Albrecht (1985), Bogdan and Taylor (1975:109), Denzin (1989:210–233), and Jackson (1987) for more on taping in field research.

45.Altheide (1980), Bogdan and Taylor (1975:75–76), Lofland et al. (2006), Maines et al. (1980), and Roadburg (1980) discuss leaving the field.

46.Discussion of field interviewing can be found in Banaka (1971), Bogdan and Taylor (1975:95–124), Briggs (1986), Burgess (1982c), Denzin (1989:103–120), Douglas (1985), Lofland et al. (2006), Spradley (1979a), and Whyte (1982).

47.See Fontana and Frey (1994).

48.On comparisons with conversations, see Briggs (1986:11), Spradley (1979a:56–68), and Weiss (1994:8).

49.See Weiss (1994:78).

50.See Atkinson (1998), Denzin (1989:182–209), Nash and McCurdy (1989), Smith (1994), and Tagg (1985) on life history interviews.

51.The types of questions are adapted from Spradley (1979a, 1979b).

52.Field research informants are discussed by Dean et al. (1969), Kemp and Ellen (1984), Lofland et al. (2006:93-94), Schatzman and Strauss (1973), Spradley (1979a:46–54), and Whyte (1982).

53.Interview contexts are discussed in Hammersley and Atkinson (1983:112–126) and in Schatzman and Strauss (1973:83–87). Briggs (1986) argued that nontraditional populations and females communicate better in unstructured interviews.

54.For additional discussion of data quality, see Becker (1970b), Dean and Whyte (1969), Douglas (1976:7), Kirk and Miller (1986), and McCall (1969).

55.Douglas (1976:115) argued that it is easier to "lie" with "hard numbers" than with detailed observations of natural settings.

56.Adapted from Douglas (1976:56–104).

57.See Bloor (1983) and Douglas (1976:126).

58.For more on validity in field research, see Briggs (1986:24), Bogdan and Taylor (1975), Douglas (1976), Emerson (1981:361–363), and Sanjek (1990).

59.See Lofland et al. (2006), Miles and Huberman (1994:288–297), and Punch (1986).

60.Covert, sensitive study is discussed in Ayella (1993), Edwards (1993), and Mitchell (1993).

61.See Douglas (1976), Erikson (1970), and Johnson (1975).

62.See Barnes (1970), Becker (1969), Fichter and Kolb (1970), Goward (1984a), Lofland et al. (2006), Miles and Huberman (1994:298–307), and Wolcott (1994) on publishing field research results.

63.For a discussion of focus groups, see Bischoping and Dykema (1999), Churchill (1983:179–184), Krueger (1988), Labaw (1980:54–58), and Morgan (1996).

第 14 章

歷史比較研究

歷史比較研究簡史
適合歷史比較研究的研究問題
歷史比較研究的邏輯
進行歷史比較研究計畫的步驟
歷史脈絡中的資料以及證據
比較研究
歷史比較研究中的同義議題
倫理
結論

> 社會學的解釋必然是和歷史有關的。因此，歷史社會學並非是某種特殊的社會學，更確切地說，歷史社會學正是社會學領域的本質。
>
> ——Philip Abrams, *Historical Sociology*, p. 2

> 不以比較的觀點來進行思考，是一件無法想像的事。而且，若是沒有任何比較的話，所有的科學思想以及所有的科學研究就都不會存在了。因此，當我們知道隱性以及顯性的比較都已經滲透到社會科學家的工作之中、並且從一開始就是此種情況的時候，就不應該感覺到驚訝了。
>
> ——Guy Swanson, "Frameworks for Comparative Research," p. 145

雖然基督教的傳教士已經在東亞地區活動了大約 300 年的時間，但是當人們在看著東亞地區時，他們卻只有看見兩個基督徒人口比例較高的國家：菲律賓以及南韓。由於菲律賓群島曾經被西班牙天主教徒統治了長達 200 年的時間，因此可以解釋為什麼菲律賓國內會有許多的天主教徒，但是在南韓國內的基督宗教則是新教。一直到韓戰（1953）之前，韓國幾乎沒有受到什麼西方文化的影響。時至今日，每三個韓國人之中，就有一個人是基督教徒。相較之下，在日本和中國境內，基督教徒的比例只佔了大約 2%而已。肯恩以及帕克（Kane and Park, 2009）想要瞭解為什麼韓國會有這麼多的新教徒。他們認為，宗教皈依的微觀理論無法解釋韓國的情況，所以他們發展出一個地緣政治的文化理論來解釋此種現象。他們的資料來自於大約 150 本關於韓國以及東亞地區的歷史、文化、以及政治的書籍，而且這些書籍都是以韓文或者是英文所寫成的。他們注意到，西方帝國主義強權在十九世紀時於東亞地區強迫各個國家簽訂不平等條約的作法、以及對於這些國家的統治，引發了許多民族主義的、排外的運動。韓國的民族主義者把此種排外情緒的矛頭指向日本，而日本則是懷疑在韓國境內的少數新教徒傳教士煽動了此種排外運動。之後，這些傳教士

加入了韓國的民族主義者，這些民族主義者都迫切地希望韓國能夠脫離日本的掌控。藉由比較基督宗教於十九世紀晚期以及二十世紀前半葉在日本、中國、以及韓國等地的運作情況，肯恩以及帕克證明了韓國的新教如何與追求獨立的民族主義運動融合在一起。在中國以及日本境內，排外情緒阻止了當地人們皈依基督宗教，而在韓國境內的情況則與上述兩國不同，因為韓國的新教徒傳教士抗拒外國（日本）統治，並且成為傳達民族主義情緒的媒介。因此，韓國境內的基督宗教在今日如此興盛，其原因就在於韓國在地緣政治網絡中的位置，以及在歷史中某個特定的時間點上，某些宗教信仰的普及所致。

有時候我們希望能夠檢視鉅觀尺度的議題，或者是長期社會變化的議題。我們無法對整個社會進行實驗，進行調查的話又只能侷限於當今的時間，而且我們也只能在某些國家中針對某些主題才能取得現存的統計量數。而最適合進行田野研究的情境，則是去檢視當下的小規模環境。因此，爲了要對整體文化或者是社會進行比較以瞭解鉅觀的模式、或者是橫跨數十年或數世紀之久的長期趨勢，我們就必須求助於歷史比較研究（也被稱作是*比較歷史研究*）。此種研究法允許我們去處理許多令人興奮的「大型」問題，這些問題與整個社會或是長時間的模式有關。而歷史比較研究的另外一個優點在於，此種研究法能夠去澄清在其他研究技術中所發現的許多關於方法論方面的議題。

社會科學的建立者，像是涂爾幹（Emile Durkheim）、馬克思（Karl Marx）、以及韋伯（Max Weber）等人，想要去解釋整體社會由前工業時代過渡到現代工業形式的轉變過程，而他們所使用的研究方法，就是歷史比較研究法。此種研究法能夠揭露出一些在長時間之內以及在許多不同的社會之間運作的過程，因此可以幫助我們瞭解爲什麼整個社會會以特定的方式來運作。歷史比較研究能夠大幅度地促進概念以及理論的發展，而且通常會造成對於過去思想的重大修正。正如同馬哈尼（Mahoney, 2004: 93）所說的：「在概念的創新方面，歷史比較研究已經爲許多最重要的社會科學概念提供了主要的定義。」歷史比較研究通常會使我們去修正理論以及理解。例如，英格蘭姆以及李弗許次（Ingram and Lifschitz, 2006）曾經研究過蘇格蘭的造船業在過去兩百年之間的發展情況。他們發現到，由家族經營的公司、以及在各公司之間的人脈關係，最有可能會創造出獲利豐富的、蓬勃發展的公司。現代企業的興起以及轉化成爲法人公司的組織形式，過去被認爲能夠將舊式的企業加以「現代化」並且提高獲利。然而，作者們卻發現到，法人公司的形式並未改善獲利或者是促進產業的發展，相反地，在他們所研究的造船業中，此種形式反而產生了反效果，也就是使成長與獲利都減緩或是下降了。

在處理社會理論中的許多中心議題時，歷史比較研究是一項強而有力的工具。基礎研究者通常會比應用研究者更常使用歷史比較研究來檢視許多社會科學中的議題，這是因爲歷史比較研究對於基礎知識的擴展是極爲重要的。

歷史比較研究簡史

在十九世紀時創立社會學的先驅們所進行的研究，正是歷史比較研究，而在那個時候，文化人類學、人文地理學、社會學、歷史學、政治學、以及經濟學等等領域之間並沒有太多的區隔。此種情況使得在這些領域中的學者之間大量地分享了許多概念、理論、以及研究方法。隨著社會科學在二十世紀初期開始專門化以及分化，歷史比較研究也逐漸式微，像是人類學家越來越注重文化以及進行比較，歷史學家則是側重於研究過去的事件等等。歷史比較研究因而移轉到其他學科的小型次領域當中。除了少數的例外情況之外〔像是布洛赫（Marc Bloch）、霍曼斯（George Homans）、莫頓（Robert Merton）、以及博藍尼（Karl Polanyi）等人〕，在第二次世界大戰之前，社會學領域中幾乎看不到什麼關於歷史比較的研究。[1]

在第二次世界大戰之後，由於國際之間的交流日漸頻繁、殖民帝國的瓦解、以及冷戰所導致的新地緣政治等等因素，重新激起了學者們對於比較研究的興趣。到了 1950 年代時，我們才開始看到一些具有重大理論重要性的歷史比較研究出現。[2]

到了 1960 以及 1970 年代時，有數個因素造成了歷史比較研究再度興起。首先，有數位歷史學家從社會科學中引進了量化研究的技術，而此種情形則促進了不同領域之間的交流。對於流動性、鐵路擴張、以及投票行為的統計研究，讓歷史學家見識到了量化資料的力量，而且也讓量化研究者們有機會利用他們的技術去處理一些與過去不同的新問題。其次，使用調查技術的各種研究開始擴散到許多國家之中，而且研究者們也運用了高階的量化技術，以進行跨國的推論。第三，韋伯以及馬克思的一些歷史比較作品首度被翻譯成英文版本。「對於在 1960 年代的歷史學著作方面而言，韋伯作品的英譯版本所造成的影響，很可能大於社會科學中的任何其他單一因素」（Stone, 1987: 13）。第四，在 1960 年代時，歷史比較研究出現了一些理論方面的突破性進展，因而使得研究者們越來越有興趣要把歷史比較研究當作是一種研究技術來看待。[3] 而到了 1970 年代時，則出現了一些如何進行歷史比較研究的新模型。[4]

歷史比較研究在 1980 年代時出現了欣欣向榮的發展趨勢。在 1983 年時，美國社會學學會設立了一個特別的歷史比較研究部門，而在柯恩（Melvin Kohn, 1987）就任該學會會長時所發表的演說中，則提及他已經觀察到跨國的研究正在經歷明顯而且重大的復甦。在主要的學術期刊上，採用某種形式的歷史比較研究的文章數量正在成長當中。例如，在 1970 年代末期時，在主要的社會學期刊當中，歷史比較研究的文章在所有文章中大約佔了 18%。在 1980 年代時，這個比例已經上升到 28%，而到了 1990 年代以及二十一世紀早期時，這個比例則上升到了 40%左右。[5]

適合歷史比較研究的研究問題

在處理像是下列這些大型的問題時，歷史比較研究會是一種強而有力的工具：重大的社會變化是如何發生的？為什麼當今的社會安排會以某種形式在一些社會中

出現，但是在其他社會中則不然？我們可以利用歷史比較研究來處理一些鉅觀層次的問題，就像是在本章一開始的部分所提到的那項研究一樣，也就是關於爲什麼在韓國裡信奉基督宗教的人口比例很高、但是在鄰近的東亞國家中卻沒有此種現象的研究。在第二章裡，你也已經閱讀過關於在 1940 年代的華沙猶太特區的研究，該研究所探討的問題，是爲什麼在面臨大規模屠殺時，某些人會起身抗爭到底、而其他人卻選擇坐以待斃的原因。你也曾經閱讀過一項關於爲什麼一些在拉丁美洲地區的前殖民地仍然相當貧窮、而其他國家則能夠成功轉型的研究。[6]

歷史比較研究相當適合於檢視一些能夠產生特定結果（例如：內戰、某種稅賦制度、某種宗教等等）的社會因素的組合。歷史比較研究也相當適合於對整體社會制度或是系統進行比較，以發現在不同的社會之間的共同點爲何、各個社會的獨特之處爲何，以及去研究長期的社會變遷。歷史比較研究者通常會去建立一個理論、或者是把一個理論應用於特定的個案中。例如，法國擁有高度中央集權的政府，政府高度掌控經濟，而且通常都是由左派政黨贏得大選。然而，美國則擁有地方分權的制度、強大的自由市場、以及力量薄弱的左派政黨。某個理論可能會因此認爲法國對於高收入的公民所課的稅會比美國來得高，但是研究者們卻發現了完全相反的情況（請參閱範例 14.1 爲什麼美國和法國採取不同的課稅方式？）。我們可以利用歷史比較研究來建立一個理論，以解釋在這兩個國家之間的差異以及令人意外的結果。利用歷史比較研究，我們可以去針對文化單位或者是國家進行比較，並且建立跨時間的理論推論以及分析。[7]

歷史比較研究能夠強化概念化的過程以及理論的建立。藉由觀察歷史事件或者是多種不同文化脈絡的方式，我們能夠產生新的概念並且擴展我們的理解。我們不太可能把概念限制於一個單一的歷史時期或者是單一的文化裡、然而卻又同時要把我們的想法植基於特定文化以及歷史脈絡裡的人們的經驗之中。[8]

在閱讀歷史比較研究時所可能遇上的困難之一，是讀者通常需要具備一些關於過去的歷史或者是其他文化的知識，才能完整地瞭解整份研究。一個只知道他或她自己文化的人、或者是一個缺乏歷史知識的人，可能會覺得歷史比較研究、甚至是古典理論家的著作是相當艱澀難懂的。例如，如果某個讀者不瞭解歐洲封建時期的狀況、以及馬克思寫作當時的世界情況，那麼他或是她就會覺得很難去理解馬克思所著述的《共產黨宣言》（*The Communist Manifesto*）。在當時的時空脈絡裡，農奴們生活在極爲嚴重的壓迫之下。實行封建制度的社會，包括了在城市裡依社會等級而訂定的服裝標準，以及一種勞役償債的制度，此種制度迫使農奴們必須把其農作物收成的一大部分繳交給地主們。唯一的基督教會也握有大量的土地，而緊密的家族關係則存在於貴族、地主、以及教會之間。現代的讀者們可能會質疑，既然情況這麼惡劣的話，農奴們爲什麼不逃跑呢？要理解這個問題的答案，讀者們必須要知道，如果農奴逃到當時的歐洲森林裡的話，光是靠著食用樹根、漿果、以及打獵等等方式，農奴們活下來的機會是非常渺茫的。除此之外，也沒有任何人會去協助逃跑的農奴，因爲傳統的社會不但不願意接受陌生人，而且還會對陌生人心生恐懼。爲了理解古典理論家所進行的歷史比較研究，「讀者們必須去體會這些理論家們視爲

範例　14.1

為什麼美國和法國採取不同的課稅方式？

摩根以及普拉薩德（Morgan and Prasad, 2009）檢視了為什麼法國以及美國採用了不同的課稅制度的原因。稅賦的兩個主要類別為消費稅以及所得稅。消費稅（consumption taxes）通常採用遞減稅率（regressive）（也就是收入較低的人要繳比較多的稅），但是能夠促進經濟成長，並且避免政治方面的抗議聲浪，而所得稅（income taxes）則通常採用累進稅率（progressive）（也就是收入較高的人要繳比較多的稅），但是卻會引起政治上的抗議行動。採用消費稅的政府通常會發展成為一個較大的福利國家。對比之下，所得稅雖然是累進制（也就是所得較高的人要繳較多的稅）的，但是卻會阻礙經濟成長、引發政治抗議、並且限制了政府的規模。為了要解釋為什麼某個國家會偏重於某種稅制，作者們比較了法國以及美國的政府在發展出稅賦制度方面的過程。在先進國家中，美國幾乎是最高度依賴國民所得稅的國家，而法國的依賴程度則是最低的。吊詭的是，在先進國家當中，美國擁有最強大的「自由市場」經濟、最少的政府干預、而且幾乎沒有任何左派的政黨。對比之下，法國擁有非常強勢的中央政府、廣泛地掌控經濟活動、而且左派政黨也經常在大選中贏得勝利。人們或許會因此而預期會在法國看到像是所得稅之類的累進稅制，而不是在美國看到此種稅制。相反地，美國的稅收大部分都是依賴某種形式的營業稅。在 1920 年之前，美國逐漸地從間接的消費稅轉移到主要依賴所得稅。作者們檢視了這兩個國家在十九世紀晚期以及二十世紀初期時的政治、經濟、以及政府型態，並藉由此種方法來解釋美法兩國之間在稅制方面的發展如何走向各自不同的道路。作者們證明了，抗拒工業化的抗爭運動、以及政府中央集權化的某種特定形式，可以解釋今日兩個國家之間不同的稅賦制度以及其他特色。作者們發展出了一個可以應用於這兩個個案之外的理論，並且得到了一個結論，也就是除非我們去檢視了在第一次世界大戰之前十年之內所發生的事情，否則的話，我們就無法去解釋今日各個不同社會的重大面向。在此項研究中的主要證據形式，包含了關於這兩個國家的事件以及型態的豐富歷史細節。作者們利用這些證據建立了一個嚴謹的理論論點，而且也提出並且駁斥了各種其他的替代性解釋。他們的參考書目列出了大約一百種以英文和法文寫成的書籍以及文章，而在研究中所引用的作品裡，有大約三分之二都和這兩個國家的歷史主題有關。

理所當然的時代特徵，以及理論家們對於過去的詮釋」（Tuchman, 1994: 310）。

歷史比較研究的邏輯

在歷史研究中，充斥著對於術語感到困惑以及混淆的狀況，而且其中的專門術語也是不一致的。研究者會把他們所使用的研究方法稱爲*歷史的*、*比較的*、*跨國的*、*跨文化的*、或者是*歷史比較方法*。研究者們在進行歷史比較議題的研究時，會採用差異性極大的特定技術。另外一個造成混淆的來源，則來自於是否真的有一個獨樹一格的歷史比較方法，或者是雖然我們使用了相同的研究方法，但卻只是碰巧用來檢視過去或者是在數個社會裡的社會生活而已。馬哈尼（Mahoney, 2004: 81, 註腳 1）把歷史比較研究稱爲是「一種田野研究，其特色在於運用系統性的比較方法、以及對於長期過程的分析，以解釋像是革命、政治體制、以及福利國家等等的大規模結果。」他把歷史比較研究和歷史社會學中的其他研究方法區隔開來，像是理性選擇分析以及詮釋分析、賽局理論、以及民族誌等等。

有一些研究者會使用實證主義的量化取向來研究歷史的或者是比較的議題，而其他人則是依賴質化的、詮釋的、或者是批判的取向。根據拉金以及札瑞特（Ragin and Zaret, 1983）的看法，歷史比較研究中的涂爾幹（或者是實證主義）取向、以及韋伯（或者是詮釋的）取向，所使用的是不同的邏輯。[9]歐顏（Øyen, 1990: 7）總結了此種混淆的情況：

> 用來區別不同種類的歷史比較研究的詞彙，不但是多餘的、而且也不是非常精確的。像是跨國的、跨民族的、跨社會的、跨系統的、跨機構的、跨文化的、以及鉅觀層次的比較等等概念，在一般意義上都可以被當作是比較研究的同義詞，而且也可以表示特定種類的比較。

我們可以依據你在第二章中已經學過的三種面向來對歷史比較研究進行組織與分類：個案、時間、以及資料的種類。歷史比較可以聚焦於單一個案中、或者是在某個國家裡所發生的事件之上；可以去針對一小組的國家（二到六個國家）進行比較，或者是同時檢視許多國家（十二個到一百個國家）。歷史比較研究可以聚焦於現在、過去的某一段時間、或者是橫跨許多年的時間。最後，歷史比較研究中的資料可以主要是量化的或者是質化的。如果我們把這三種面向交叉分類的話，我們可以得到十八種在邏輯上可能的歷史研究類型（請參閱表 14.1），這也就難怪爲什麼關於歷史比較研究的內容會有這麼多混淆不清的狀況了。雖然大多數的社會科學研究的焦點都是在於研究者當前的祖國，但是歷史比較研究的焦點卻可能會具有非常大的歧異性。

表 14.1　在邏輯上可能的歷史比較研究種類

	比較的面向		
時間面向以及資料種類	*單一國家**	*少數國家*	*許多國家*
過去一段時間			
量化	1	2	3
質化	4	5	6
跨時間			
量化	7	8	9
質化	10	11	12
現在			
量化	13	14	15
質化	16	17	18

*所研究的國家是目前和研究者以及研究結果的讀者都不相同的國家。

範例　14.2

在十五個國家中的監禁率

許多利用國家當作分析單位的跨國研究，都是依賴現存的統計資料來進行的。例如，撒頓（Sutton, 2004）就針對十五個國家在 1960 年到 1990 年之間的情況進行了一項量化的、統計的研究。許多研究者長久以來已經觀察到，監禁率並未緊密地跟隨著犯罪率而變化。撒頓測試了一個由羅許以及克許海默（Rusche and Kirchheimer）所提出的論點，該論點認為失業率會造成監禁率上升，這是因為監禁是政府所使用的一種手段，其目的在於控制過剩的失業勞動階級的男性人口，因為這些人有可能會因為失業而不受約束。基本上，當許多工人失業時，監獄就會人滿為患，而當經濟蓬勃發展時，監獄就會變得門可羅雀。撒頓所蒐集的資料，是來自於這十五個國家的政府統計數據年報、由像是世界衛生組織以及國際勞工組織之類的國際組織所發行的刊物、以及之前的一些社會科學研究，這些研究發現了數個國家所擁有的特徵，像是這些國家的工會組織模式、政黨結構等等。就原來的論點而言，撒頓只有發現有限的支持證據而已，但是他詳細記錄了來自於數個其他因素的強烈效果。他認為失業對於監禁所產生的效果很有可能只是一種虛假關係而已（請參閱本書的第二章、第四章、以及第十章中關於虛假關係的相關討論）。他發現，一個國家的政治的特殊特色以及勞動市場的結構，似乎是造成特定的失業模式以及不同的監禁政策的真正原因。簡而言之，相較於低收人的人們以及勞工擁有較多政治力量的時期，當低收人的人們以及勞工在政治上與富裕的人們以及企業主相較之下較為弱勢時，失業率以及監禁就會同時上升。

大部分的歷史比較研究都會落入這十八個細格當中的十個細格（1、4、5、7、8、10、11、13、14、15、以及 16），這其中也包括了整個單一國家的直行。許多歷史比較研究試圖去發現普世的、或者是接近普世的過程。他們想要發現一些能夠在長時間之內（請參閱 Ember and Ember, 2001）橫跨所有人類社會、或者是至少在一種主要類型的社會（例如：高度工業化的、或者是農業的）之中運作的過程或者是關係。然而，大多數的研究只考量了單一社會或者是單一時期的資料，因此其可類推性就會非常地低。我們無法有信心地認爲，依據此種資料所得到的解釋是否只能適用於某個社會、是否能夠在大多數的人類社會中被發現，或者是專屬於某個時間點或是能夠適用於整個人類的歷史時期。

目前在世界上大約有 170 個國家，而有一些研究者則試圖去取得其中大多數國家的資料，像是在蕭弗（Schofer, 2004）針對科學在全世界的發展過程所進行的研究當中，就列出了超過 120 個國家的資料；而在孫（Sung, 2003）針對女性參與政治的情況所進行的研究當中，則是取得了來自於 99 個國家的資料。然而，有許多社會的資料是無法取得的、或者是無法進行比較的，因此大多數的比較研究都被限制在大約 10 到 20 個國家之間，像是馬哈尼（Mahoney, 2003）（我們已經在第二章中討論過該項研究了）研究了十五個南美洲的國家，撒頓（Sutton, 2004）（請參閱範例 14.2 在十五個國家中的監禁率）檢視了十五個國家裡的監禁率，而曼德爾以及夏雷夫（Mandel and Shalev, 2009）則調查了在十七個國家裡的男性與女性收入之間的差距大小（請參閱範例 14.3 在十七個國家中的兩性收入差距）。

範例 14.3

在十七個國家中的兩性收入差距

曼德爾以及夏雷夫（Mandel and Shalev, 2009）研究了在數個國家中，男性與女性收入之間的差距大小。他們提到：

> 在最近幾十年來，女性主義學者已經讓大眾注意到促進女性經濟自主的福利國家政策的重要性……相對於男性而言，福利國家強烈地影響了女性的生活機會，而且也改善了階級不平等的狀況。此項洞見大大地擴展了許多關於性別不平等以及福利國家的研究，尤其是那些採取一種比較觀點的研究。（p. 1873）

曼德爾以及夏雷夫把兩性之間在收入方面的差距，區分為以階級為主與以性別為主的要素。換句話說，兩性之間在收入方面的差距，取決於在一個國家中有多少女性達到階級結構的頂端、以及在該國當中階級不平等的程度。作者們考量了在十七個先進國家中的福利國家政策如何影響了這兩個要素。福利國家有三種不同的形式：自由的、社會民主的、以及保守的。有些政策鼓勵家庭承擔起社會福利的責任，其他政策則是依賴政府提供工作給人民，而還有一些政策則是鼓勵私人市場的解決方式。作者們檢視了在十七個國家中的性別不平等現象，而且這些國家在社會計畫以及階級結構方面都有相當大的差異性。藉由此種檢視的方式，作者們就能夠去區隔出造成性別不平等的數個因素。作者們發現，一個國家的階級結構、以及各種不同社會計畫的混合情況，會以特定的方式結合起來而產生了各種不同程度的兩性收入不平等的現象。福利國家的角色，會以不同的方式去影響了位於不同階級地位的女性。因此，作者們主張我們不能把「家庭政策」（性別）和「社會政策」（階級）當作是彼此互不相干的兩種事物。

其他的研究者則是專注於針對兩個到四個個案進行比較。例如，弗瑞澤爾（Fretzer, 2000）研究了法國、德國、以及美國這三個國家裡的公民對於移民的態度；甘爾（Gangl, 2004）比較了德國以及美國的福利國家政策以及失業情況。維特納以及法倫（Viterna and Fallon, 2008）則是調查了南非、阿根廷、迦納、以及薩爾瓦多等四個國家裡追求性別平等的政治運動，他們選擇這四個個案的原因，是因爲這些國家清楚地顯現出在提升性別平等方面的成功以及失敗，而且在其各自的理論架構的特色方面也存在著明顯的差異。

許多的研究都是把焦點放在現在上面，但是其他的研究則被限制在過去某個時間點上的某個社會，像是艾恩渥納（Einwohner, 2003）針對在 1943 年時的波蘭華沙地區所進行的研究（我們已經在第二章裡面討論過了），以及艾米（Emigh, 2003）針對十五世紀時的托斯卡尼地區所進行的研究（將在第十五章裡進行討論）。有一些研究考量了一段較長的時間，像是拉克曼（Lachman, 2003）研究了從 16 世紀到 18 世紀之間在四個歐洲國家中的菁英份子，馬哈尼（Mahoney, 2003）研究了南美洲在過去兩個世紀裡的變化情況（已在第二章裡討論過），佩堤特以及威斯騰（Pettit and Western, 2004）研究了在 1964 年到 1997 年之間美國黑人與白人男性的監禁率（已經在第二章裡討論過），而羅格斯（Ruggles, 2007）則是調查了從 1850 年到 2000 年之間美國民眾兩代共居比例下降的現象（請參閱範例 14.4 美國在 150 年之間兩代共居比例下降的現象）。

範例　14.4

美國在 150 年之間兩代共居（coresidence）比例下降的現象

在 19 世紀中期，有將近 70%的年長民眾和他們的成年孩子住在一起。到了 2000 年時，年長者與自己的成年孩子住在一起的比例已經低於 15%。而造成此種現象的原因，一直是學者專家們爭辯的主題之一。學者們對於此種現象所提出的主要解釋是，年長者們的經濟狀況已經比過去改善很多，而且在晚年時也能夠負擔得起獨自生活的開銷。羅格斯（Ruggles, 2007）檢視了這個問題，並且發現造成此種情況主要是與年輕的一代比較有關係、而與年長的那一代則是比較沒有關係。兩代共居情形的變化，主要與經濟方面的變化有關，尤其是年輕一代在以家庭為主的生產方面下降、從事較多以工資為主的勞力工作、接受較多的教育、以及擁有較大的地理流動性等等。此外，新的家庭形態（離婚以及分居的情形越來越普遍）以及女性工作機會增加等等，也都是會影響兩代共居比例的因素。羅格斯檢視了許多年來詳盡的人口普查檔案，其主題是關於家庭組成、職業、以及地理位置等等。而為了要取得從 1850 年到 1950 年之間的資料，他還必須從微縮膠片中，把人口調查員手寫的人口普查資料輸入到電腦裡面去。

獨樹一格的、質化的歷史比較社會研究類型，完全不同於實證主義取向以及極端的詮釋取向。歷史比較研究使用了個案研究以及質化資料，因此和嚴格的實證主義原則大相逕庭。個案研究可能是非常重要的，因爲若是沒有個案研究的話，學者們「將可能繼續發展一些理論的論點，而這些論點對於某個特定地區而言是不適合的、過時的、或者甚至是完全不相關的」（Bradshaw and Wallace, 1995: 155）。個案研究能夠詳細地闡述歷史過程、並且清楚地說明具體的歷史細節，然而，那些嚴格遵守社會科學實證主義取向的研究者們，卻批評歷史研究此種只使用少數個案的作法，因爲此種作法無法產生實證科學中關於機率的因果推論類型。[10]

不論是以詮釋取向爲基礎的田野研究或者是歷史研究，都是聚焦於文化之上的。這兩種研究都要求我們透過研究對象的觀點來觀察事物、去重新建構研究對象的生活、以及深入地去檢視特定的個人或者是團體。極端的詮釋取向會假定每一個社會情境都是獨一無二的，因此也完全不可能去進行任何的比較。此種取向會重新創造特殊的主觀經驗，並且會去描述各種特定的細節。正如同史東（Stone, 1987: 31）曾經說過的，傳統的歷史學「所處理的是在一個特定時間、特定地點中的特定問題以及特定的一組行動者。」而獨特的歷史比較取向則是從民族誌以及文化人類學中借用一些優點，而且歷史比較研究的一些變化型態則試圖要去重新創造另外一個時間或是空間的現實，然而，從民族誌當中借用一些優點，並非意味著一定要採取極端的詮釋取向。[11]

藉由使用量化資料來補充質化資料以及分析的方式，獨特的歷史比較研究方法結合了對於特定歷史或是文化脈絡的敏感性以及理論的可推論性。因此，歷史比較研究的邏輯與目標比較接近於田野研究、而比較不像是傳統的實證主義取向（請參閱表 14.2）。

表 14.2　不同研究取向之間的比較摘要：質化與量化之間的區別

主題	田野研究與歷史比較研究	量化
研究者的觀點	當作研究過程不可或缺的一部分而包含進來	從研究過程中移除
資料取向	沉浸在許多細節中以獲得理解	精確地將變項操作化
理論以及資料	利用紮根理論以及資料與概念之間的對話	以實徵資料考驗演繹的理論
呈現研究發現	轉譯意義系統	考驗假設
行動／結構	人們在結構中行動以及建構意義	檢視社會力量如何形塑行為
法則／推論	運用依賴於脈絡的有限推論	發現普世的、不受脈絡影響的法則

獨特的歷史比較研究取向的特質

主題	歷史比較研究的取向
證據	從片段以及不完整的證據當中重新建構
扭曲	防止研究者運用在社會或是歷史脈絡之外的其他因素來進行詮釋
人類角色	包含了在某個脈絡中的人們的意識，並且以人們的動機當作原因因素
原因	將原因視為是由情境所決定的、隱藏於表面之下的、而且是來自於各種元素的結合
鉅觀／微觀	對整體案例進行比較，而且把微觀層次連結到鉅觀層次或者是社會現實的各個層次
跨脈絡	在某個脈絡的具體細節之間來回移動，並且橫跨不同的脈絡以進行更為抽象的比較

與田野／實地研究的相似之處

在歷史比較研究以及田野研究之中，研究者的觀點都是研究中不可避免的一部分。在這兩種類型的研究當中，詮釋都具有舉足輕重的地位。換句話說，研究者—詮釋者在時間、空間、以及世界觀方面的位置都是非常重要的。歷史比較研究並未試圖去產生一組單一的、明確的客觀事實，而是以新的觀念以及證據去面對舊的事物、或者是去考慮多重的世界觀。我們對於歷史的或者是比較的證據所抱持的看法，會受到我們對於過去的知覺以及我們活在當下的這個事實所影響。

其次，田野研究以及歷史比較研究都會檢視許多差異性極大的資料。在這兩種研究裡，身爲研究者的我們必須讓自己沉浸在資料裡，以獲得對於事件以及人們的同理心理解。我們試圖去捕捉主觀的感受，並且去瞭解平常的活動如何表達了重要的社會意義。

我們會在大量的事件、行動、象徵符號、以及文字當中去進行調查、選擇、並

且聚焦在社會生活的特定面向上，而這意味著我們必須組織資料，並且把注意力集中於不斷演變的概念的基礎上。我們會去檢視那些使文化變得戲劇化的儀式以及象徵符號（例如：遊行、衣著、物品的擺放方式等等），並且調查行爲的動機、理由、以及辯護的藉口。例如，布拉吉以及柯里（Burrage and Corry, 1981）就曾經利用在大型公開活動（遊行、慶典、宴會、皇室成員出遊等等）當中官方指定出席的記錄，測量了在十四世紀到十七世紀之間倫敦地區的職業地位的變化情形。[12]

第三，不論是在田野研究或者是歷史比較研究中，我們都會運用*紮根理論*（*grounded theory*）。理論通常會在蒐集資料的過程當中浮現出來。我們是在沒有特定假設的情況下去開始檢視資料的，更確切地說，我們是透過與資料進行對話的方式來發展出並且修改概念與理論，然後再應用理論來重新組織證據。因此，資料的蒐集以及理論的建立之間會不斷地進行互動。湯普森（Thompson, 1978: 39）把此種情形稱作是「概念與證據之間的對話，此種對話一方面是由連續的假設來引導的、而另一方面則是由實徵研究來引導的。」[13]

除此之外，田野研究以及歷史比較研究也都牽涉到轉譯的過程。我們的意義系統通常和研究對象的意義系統不同，但是我們仍然要試著去深入並且理解他們的觀點。當我們深入理解了研究對象的生活、語言、以及觀點之後，我們接下來的工作就是要針對這些事物進行「轉譯」，以讓閱讀研究報告的讀者們也能夠理解這些事物。

第五，田野研究和歷史比較研究都把焦點集中於行動、過程、以及順序上面。這兩種研究都把時間的推移當作是研究中不可或缺的一部分。進行這兩種研究的研究者，對於在能動作用（agency）（也就是積極參與改變的個人以及組織）以及結構（也就是會限制行動以及知覺的固定規律性）之間持續存在的緊張關係都相當敏感。兩種研究的研究者都把社會現實視爲是人們在創造以及轉化生活時的產物、同時也是某種會限制人類選擇性的事物。[14]

第六，在這兩種類型的研究當中，類推性以及抽象理論都是相當有限的。在微觀層次的田野研究以及鉅觀層次的歷史比較研究中，我們所擁有的知識都是不完整的以及暫時性的。複製是不切實際的，而且我們幾乎無法利用演繹命題或者是考驗假設的方式來發現固定的法則。相反地，每位研究者都會發展出一個獨一無二的證據體系，以建立可信的敘述說明並且提供有限的類推性，而此種有限的類推性則會緩慢地促進知識的進展。

歷史比較研究的獨特特質

雖然和田野研究之間有許多相似之處，但是歷史比較研究仍然有一些重要的差異之處。歷史比較研究「在本質上就是一種特殊的領域，在此領域之中的研究者必須整理大量的、長期的資料以推論出因果關係。事實上，有一種普遍的看法是，對於長時間過程的分析，正是因果推論的核心基礎……」（Mahoney, 2004: 88）。我們可以指出歷史比較研究的六個獨特的特質：證據的類型、扭曲的控制、承認人類的能動作用、對於因果關係的詳細說明、連結微觀層次與鉅觀層次、以及脈絡與比較的替代性使用方式。

1. *證據的類型*　歷史比較研究的證據是有限的而且間接的。我們幾乎無法去直接觀察研究對象、或者是密切地涉入研究對象的生活，而只能夠從證據中去重新建構曾經發生過的事情。歷史證據相當依賴從過去留存至今的資料，而且這些資料通常是以文件的形式（例如，信件以及報紙等等）保存下來的。我們會受限於那些尚未被破壞的資料，以及那些會留下痕跡、記錄、或者是其他證據的事物。

我們必須藉由沉浸在某個脈絡的細節當中、以及吸收這些細節的方式，才能夠去詮釋證據。例如，假設我們想要去研究在 100 年前、或者是在一個非常不同的文化當中的家庭運作方式。爲了達到我們的目的，我們必須瞭解家庭的完整社會脈絡（例如：工作的本質、溝通的形式、交通運輸科技等等）。我們也應該去檢視地圖、瞭解在實施當中的法律與政策、並且研究人們的烹調以及飲食的習慣。我們還應該去瞭解兒童們所玩的遊戲、學校教育、醫療照護、預期壽命、以及一般的社會習俗。例如，我們必須要知道「家庭成員來訪」所指的意思是什麼，而此種拜訪的事件會受到一些條件所影響，像是充滿塵土和泥巴的道路、以及無法事先通知拜訪時間的情況等等。對於那些與動物們一起生活在農場上、並且必須時常照顧動物的人們而言，他們的生活節奏以及優先事項可能會與我們所習慣的完全不同。因此，我們必須去瞭解人們日常生活的步調、以及一般生活的經驗才行。

2. *扭曲的控制*　當我們在重新建構過去的事件或者是其他的文化時，很容易會發生扭曲的情況。與我們所研究的人們相較之下，我們通常會比較傾向於去注意到過去所發生的事件、離研究地點相當遙遠的地方所發生的事件、以及在我們研究的時期過後所發生的事件。此種情況能夠產生一種強烈的連貫性的感受，此種感受會比那些生活在過去、或者是在一個與世隔絕的社會環境中的人們所可能感受到的還要更加強烈。「簡而言之，當事件仍然在發生時，歷史解釋勝過於任何的理解。我們所重新建構的過去，比起事件發生的當時還要更具有連貫性」（Lowenthal, 1985: 234）。我們更爲廣泛的知覺可能會創造出一種錯覺，讓我們誤以爲事件發生的原因是因爲這些事件必然會發生、或者是誤以爲所有的事件都可以緊密地契合在一起，而且這可能是、也可能不是一般人們體驗事件的方式。

我們無法輕易地透過研究對象的觀點來觀察事物。當前的知識以及隨著時間流逝所產生的變化，可能會扭曲了我們看待事件、人們、法律、以及甚至是實體物體的方式。例如，那些留存到現在的老舊建築物，比起那些無法留存下來的建築物還要更加持久以及堅固。除此之外，由於出現時的脈絡不同，因此一棟留存至今的建築物，在 2010 年時的外觀一定會與在 1810 年時的外觀不同。當這棟 1810 年落成的建築物才剛剛建好、並且位於其他類似的或者是更古老的建築物之間時，當時的人們看待這棟建築物的方式，一定會我們這些活在二十一世紀的人們看待它的方式不同。我們可能會把這棟建築物視爲是古怪的或是具有歷史意義的，但是當時的人們所經歷過的感覺一定與我們大不相同。

3. *承認人類的能動作用*　在所有的研究當中，我們都承認人類具有去學習、作出決定、依據他們所學習到的知識而採取行動以修改事件過程的能力。當我們在研究長時間的過程時，我們必須去考量到人們具有察覺到事件、以及依據事件來採取

行動的能力，而此種情況將會限制了我們進行推論的能力，使我們無法去主張一些像律則一樣的、在所有社會中都能成立的類推性，[15] 像是某個團體的人們是否能夠察覺到、或者是意識到他們自己過去的歷史之類的。由於擁有了此種知識，他們就能避免再次犯下過去的錯誤，並且有意識地採取行動來改變事件的發展過程。當然，人們不必然會學到教訓、或者是依據學到的教訓來採取行動，而且就算他們真的採取行動了，他們也不必然會成功。然而，我們仍然必須去考量人們學習的能力，而此種情況也會爲歷史比較的解釋引入某種程度的不確定性。

我們應該要去發現涉入事件中的人們是否把各種不同的行動過程視爲是可靠的。人們的世界觀以及知識都是會造成限制的因素，因爲世界觀以及知識會形塑了人們認爲可能的或是不可能的事物。因此，我們必須去詢問人們是否意識到某些事物的存在。例如，如果一支軍隊知道敵軍即將進攻，並且因此決定在半夜時渡河的話，這支軍隊所作出的渡河的決定，其社會意義就會和這支軍隊不知道敵軍即將進攻時的情況有著極大的差異。

4. *對於因果關係的詳細說明*　在歷史比較研究當中，爲了要建立因果聯結，我們通常會採取組合解釋、使用過程分析、以及檢視共同發生的事件。一個**組合解釋**（combinational explanation）就類似於一個化學反應，在其中我們必須在特定的條件之下（溫度、壓力）一起加入數個成分（化學物質、氧氣），以產生某種結果（爆炸）。此種情況迥異於簡單的線性因果解釋。其中的邏輯比較像是「在 G 這個條件下，A、B、以及 C 一起出現在 Z 這個時空當中，並且產生了 D 這個結果」，而不是「A 導致 B、B 引起 C、然後 C 造成 D」。正如同拉金（Ragin, 1987: 13）所總結的：

> 大多數的比較學者，尤其是那些質化導向的學者，所感到興趣的是特定的歷史順序或者是結果，以及在橫跨一組類似的個案當中，造成這些順序或是結果的原因。歷史結果通常需要用到複雜的、組合的解釋，而且我們很難以一種與主流量化社會科學標準一致的方式來證明此種解釋。

例如，在韋伯所提出的解釋中，就把文化因素視爲和經濟、民主、以及社會結構因素具有同等的重要性。他透過理想型而採用了一些因果因素的組合，而此種解釋既不是一種等著接受考驗的、演繹的正式理論，也不是一種歸納式的、針對特定問題所提出的理論。[16]

組合解釋　一種在歷史比較研究中普遍使用的解釋，所強調的是數個因素在某些特定條件下一起發揮作用而產生某種結果的情況，此種情況與簡單線性的原因—結果順序是不同的。

我們對於因果關係的想法通常取決於過程分析。[17] **過程分析**（process analysis）〔又被稱爲*過程追蹤*（*process tracing*）〕會去辨識出特定的中介因果機制、這些機制在過程中的位置／角色以及對於後來事件的衝擊、以及其效果的持續時間等等。簡而言之，我們所要去檢視的，正是一個過程如何隨著時間的經過而展開的方式。你可以回想一下，在

過程分析　對於一個因果過程的檢視，研究者在其中追蹤了整個順序以及中介的機制，包括了事件的持續時間、以及早期的事件對於後來的事件所造成的衝擊等等。

我們發展出一個因果解釋時，詳細說明因果機制是很重要的一件事（已經在第二章裡面討論過了），而且提出對於個案的詳盡分析也是很重要的（已經在第六章裡面討論過了）。過程分析牽涉到概略描述一個隨著時間經過而發展的過程，描述的方式則是去辨識出一些特定的機制，而一個或者是多個原因（自變項）則是透過這些機制而對結果造成影響。

僅僅找出關聯而已，其實是不夠充分的，因此我們必須詳細地說明某個因素如何影響其他因素的方式。在以文件證明了兩個事件之間存在著關聯（例如，離婚率上升與婦女經濟機會的改善）之後，我們應該去尋找在特定個案中連結原因與結果的機制，以找出此種關聯是否爲因果關係。例如，我們觀察到有越來越多的女性外出工作、並且因而獲得了對於財務資源（例如：銀行帳戶）的掌控權，而且與此同時，離婚率也逐漸地在上升當中。我們必須去分析隨著時間經過的過程，才能確定某種現象的確造成了另外一種現象。換句話說，我們必須去具體地找出，到底是在財務資源中的哪項事物，使得財務資源與離婚之間產生了關聯。

歷史比較研究也會去檢視共同發生的事件，也就是某個因素不一定總是會導致另外一個因素發生變化，而是必須取決於其他條件的存在與否。例如，或許只有在具備某些條件的情況之下（例如，法律環境、宗教傳統、經濟與科技的發展程度等等），女性對於資源的掌握權以及離婚率才會同時上升。

5. *連結微觀層次與鉅觀層次*　歷史比較研究的焦點，在於完整個案與跨個案的個別變項上面。我們在處理完整個案時，會把個案當成彷彿具有多重的層次一樣，而這意味著我們除了必須掌握外表的表象之外，還必須去揭露出一般的隱藏結構、未被看見的機制、或者是因果過程。我們會把微觀層次（小規模的、面對面的互動）以及鉅觀層次（大規模的社會結構）整合在一起，而不是把微觀層次和鉅觀層次分開來進行個別的描述。也就是說，我們會同時檢視以及連結這兩個社會現實的層次。[18] 例如，我們會利用閱讀日記或是信件的方式來檢視個人傳記中的細節，以獲得對於這些個人的瞭解，像是他們所吃過的食物、他們所進行的娛樂活動、他們的衣著、他們曾經罹患過的疾病、以及他們與朋友之間的關係等等。然後，我們會把此種微觀層次的觀點和鉅觀層次的過程連結起來，像是移民人數增加、生產過程機械化、無產階級興起、以及勞動力市場緊縮等等。

6. *脈絡與比較的替代性使用方式*　歷史比較研究會在檢視特定的脈絡以及作出一般比較之間來回轉換。我們會去檢視特定的脈絡、留意相似處以及差異處，然後藉由把我們所發現的事物與其他情境以及脈絡進行比較的方式，來作出一般性的推論。然後，利用我們從比較當中所開始發展出的概論，再回頭去檢視特定的脈絡。比較研究者會橫跨文化與地理單位（例如：都市地區、國家、社會等等）來進行比較。[19] 歷史研究者則會去調查過去的脈絡，通常是在單一文化之中（例如：某一個時期、時代、年代、紀元等等），以去尋找順序以及比較。[20] 當然，我們可以把上述兩者結合起來，以研究在一個或者是多個歷史脈絡當中的文化脈絡，然而，每一個歷史時期或者是社會，都會有其獨特的因果過程、意義系統、以及社會關係，而此種情況會在下列兩者之間造成一種特殊的緊張關係：其中之一是某個脈絡中的具體

細節，而另外一個則是我們用來在不同的脈絡之間建立連結性的抽象觀念。

在比較分析當中使用一些跨文化的概念，類似於在歷史研究中使用一些跨歷史的概念。[21] 在比較研究中，我們會把某個脈絡的細節轉譯爲一種共同的、理論性的語言〔請參閱本章稍後部分對於他觀概念（etic concepts）的討論〕。「比較研究者因此可以被視爲是在具體系統類別的『文化束縛』、以及概括系統類別的『內容缺乏性』之間，進行一場持續的掙扎與權衡」（Smelser, 1976: 178）。

年鑑學派　歷史比較研究也從**年鑑學派**（Annales School）中汲取了一些有用的概念，[22] 此學派所指的是一種和一群法國歷史學家〔例如：布洛赫（Marc Bloch）、布勞岱（Fernand Braudel）、費夫爾（Lucien Febvre）以及勒華拉杜里（Emmanuel Le Roy Ladurie）等人〕有所關聯的研究方法，並且是以在 1929 年創立的《年鑑：經濟、社會、文明》（*Annales: Économies Sociétés Civilisations*）這本學術期刊來命名的。這個學派的取向可以總結爲下列四個相互關聯的特質：整體的、綜合的分析；融合**心理狀態**（mentality）；整合具體與抽象；以及注重一種被稱爲**長期**（longue durée）的長時間結構（請參閱資訊箱 14.1 年鑑學派）。年鑑學派在數個方面對歷史比較研究造成了影響，包括了鼓勵把各種不同類型的資料整合爲一個整體、強調對於不同的主觀意識的敏感性、把非常具體的細節與抽象的理論聯結在一起、以及把事件置於一個由長期的底層結構所形成的廣闊脈絡之中等等。

進行歷史比較研究計畫的步驟

在這個部分裡，我們將要把焦點轉移到進行歷史比較研究的過程上面。進行歷史比較研究不會牽涉到一組嚴謹的步驟，而且除了少數的例外情況之外，通常也不會用到複雜的或者是特殊的技術。

將研究的對象加以概念化

爲了要進行歷史比較研究，你應該要先去熟悉情境，並且將你所要研究的事物加以概念化。你可能會從一個鬆散的模型或者是一組初步的概念開始，然後再把這些模型或是概念應用到一個特定的情境當中。這些暫時的概念包含了一些隱性的假定或者是組織的類別，而你則是利用這些假定或是類別來觀察世界、「包裝」（"package"）觀察結果、以及搜尋證據。如果你對於某個歷史時期或者是比較的情境不甚熟悉的話，你應該先進行初步的文獻回顧（也就是去閱讀數種一般性的作品），此種作法可

年鑑學派　一群發展出一種特定研究方法的法國歷史學家，此種研究方法是整體的、融合了對於日常生活具體細節的關注以及抽象的理論建立、並且考量了長期的整體社會結構變化。

心理狀態　一種由年鑑學派所提出的觀念，意思是指在某個特定的歷史時期中，會瀰漫著一種關於普通生活的日常意識與假定的模式。

長期　一種由年鑑學派所提出的觀念，所指的是一段相當長的時間，通常是一個世紀或是更久，而在這段時期之內，社會生活中的基本模式或是結構會維持穩定不變的狀態，並且形塑了人們的日常生活。

資訊箱 14.1 年鑑學派

年鑑學派的取向可以總結為下列四個相互關聯的特質：整體綜合的分析、融合心理狀態、整合具體與抽象、以及長期的結構。

1. *整體綜合的分析*（holistic-synthetic analysis） 年鑑學派有系統地應用一種綜合的、整體的、跨學科的取向。年鑑學派的研究者將地理學、生態學、經濟學、人口統計學和文化因素等等結合起來，以提出關於過去的一幅完整圖像。年鑑學派的研究者把物質生活以及集體信念或是文化的各種不同情況結合起來，以廣泛而全面地重新建構過去的文明。
2. *融合心理狀態*（incorporation of mentalities） 年鑑學派使用了一個法文詞彙「心理狀態」（mentalities），而這個詞彙無法直接地翻譯成英文。心理狀態所指的是一種獨特的世界觀、觀點、或者是一組關於生活的假定，也就是人們用來組織想法的方式，或者是在某個時代中盛行的意識與無意識的認知、信念、以及價值觀所形成的整體型態。因此，我們試圖去發現在某個歷史時期中思想的整體組織方式，而此種組織方式會形塑關於現實基本面向的主觀經驗，像是時間的本質、人類與外在環境的關係、以及真理如何被創造出來的過程等等。
3. *整合具體與抽象*（concrete-abstract integration） 年鑑學派的取向結合了具體的歷史細節以及抽象的理論。而理論所採取的形式是某種模型或者是深層的底層結構。因果原則或者是組織的原則，可以用來說明或解釋各種日常的事件。年鑑學派的歷史學家會去尋找那些形塑了表面事件以及個人行動的底層趨勢。
4. *長期的結構*（long-term structures） 年鑑學派對於辨識出長期的結構或是型態很有興趣。傳統的歷史學家會專注於數年或者是數十年之間的特定個人或是事件，對比之下，年鑑學派的歷史學家則會去檢視在一個世紀或是更長時間內的數個時期當中的長期變化。年鑑學派的歷史學家假定，日常社會生活通常是在基本的組織原則之中運作的，而這些原則在某個地點中會逐年非常緩慢地發生變化。他們用「長期」這個詞彙來指稱在一個地理空間中一段很長的時間或者是歷史時期（例如：西歐的封建制度、或者是從十五世紀到十八世紀的地中海地區）。

以幫助你掌握具體的情境、組織概念、將主要的議題加以細分、以及發展出所要提出的問題。[23] 概念以及證據會產生互動而得以刺激研究向前進展。例如，史寇克波爾（Skocpol, 1979）在剛開始進行她的研究時，她只有對於一個鉅觀社會學理論的疑惑、以及對於特定的革命歷史的瞭解而已。由於她發現在革命的歷史以及現存的理論之間缺乏連結性，因而刺激她去進行這項研究。

不論我們是否意識到並且清楚地加以表明，我們都會把特定的細節分類爲分析的類別。最好的作法，是清楚明白地承認這個過程，並且避免犯下**培根式謬誤**（Baconian fallacy）。此種謬誤是以培根（Francis Bacon, 1561-1626）來命名的，其假定是研究者在進行研究時，錯誤地以爲自

培根式謬誤 一種錯誤的想法，認為研究者在進行研究時，可以不帶有任何預設的問題、想法、假定、理論、或者是推測。

己可以不帶有任何預設的問題、假設、想法、假定、理論、典範、條件、偏見、或者是推測。

找出證據

下一個步驟是透過廣泛地閱讀文獻的方式來找出以及蒐集證據。你將會使用到許多的索引、目錄、以及參考資料，這些文獻資料會列出圖書館中包含了哪些館藏資料。就比較研究而言，這個步驟意味著聚焦於特定的國家或是單位，以及聚焦於每一個國家或是單位之中特定種類的證據。在一個較大型的研究中，你可能必須花上數個月或者甚至是數年的時間在圖書館裡尋找資料來源、拜訪不同的特殊研究圖書館、以及閱讀大量的書籍與文章。比較研究通常也會牽涉到學習一種或者是更多種外國語言。

隨著你掌握了文獻資料並且記下詳盡的筆記，你必須要完成許多具體的任務：製作出一份附有完整引用資訊的參考書目清單（記在卡片上或是電腦裡）、記下既不會太簡陋也不會太廣泛的筆記（也就是至少超過一行、但是不會多於十幾頁的引文筆記）、以及製作出一份關於主題或是可行假設的檔案資料。

根據你在證據中所發現的事物，你可能會調整最初的概念、問題、或者是焦點。當你在閱讀文獻、考量各種不同分析層次（例如，一般的脈絡以及針對特定主題的詳細敘述）的研究報告、以及閱讀了針對某個主題或者是跨主題界線的大量研究時，新的議題以及問題就有可能會逐漸浮現出來。

卡達格諾（Quadagno, 1988）針對老年以及福利計畫進行了一項研究。她一開始時所感到興趣的，是美國對於年長者所提供的福利計畫的歷史。她先從年長者福利計畫的政府記錄開始進行研究。很快地，她發現了南方政治壓力的重要性，因此她花了數個月的時間去瞭解美國南方的歷史。隨著這個議題逐漸展開，她檢視了社會計畫的歷史。然後，隨著她的研究範圍逐漸擴展，她閱讀了許多理論以及經驗的討論，這些討論顯示出與其他社會福利計畫之間的連結，而且也提出了與廣泛的西歐福利計畫之間的比較。在西歐國家中，社會民主黨派是工會的代表，而且與工會密切地合作。這些黨派形塑了在這些國家中大多數的社會計畫。因此，卡達格諾的研究轉向了美國的勞工歷史。而管理勞工的官員所留下的記錄、以及勞工的歷史，又引導她去檢視雇主的行動以及私部門的力量。她說：「我在理論以及檔案資料之間來來回回，每當有新的實徵觀察結果出現，就會引導我去重新推論那些會影響福利政策的因素爲何」（1988: x）。

評估證據的品質

當你在蒐集證據時，你應該要詢問兩個問題：這些證據與逐漸浮現的研究問題以及逐漸演化的概念之間具有多大的關聯性？這些證據的正確性以及有效性有多大？關聯性有多高的問題，是一個相當困難的問題。隨著研究的焦點發生轉移，之前不具關聯性的證據可能會變得具有關聯性。同樣地，有些證據可能會刺激研究的

新途徑，以及刺激你去尋找額外的支持性證據。

當你在閱讀證據的時候，你所尋找的是下列的三項事物：隱性的概念架構、特殊的細節、以及實徵的概論（文獻作者們具有共識的事實性陳述）。你應該去評估證據的替代性詮釋，並且尋找「沉默的人事物」，也就是在一些個案裡，其中的證據沒有去提及某個事件、主題、或者是議題。例如，當你在檢視一群居於領導地位的男性商人時，你或許會發現到，在相關的文件裡都很有默契地刻意不去討論他們的太太以及眾多的僕人。在發展出對於研究對象的完整瞭解時，此種「沉默的人事物」可能會是相當重要的，而且也可能暗示了新的研究領域。

正如同在所有的研究中一樣，你也應該要盡力去避免證據中可能出現的謬誤。費雪（Fischer, 1970）提供了一份關於此種謬誤的廣泛清單。例如，*僞證據*（*pseudoproof*）的謬誤就是一種未能把某個事物放置於完整脈絡中的錯誤。某項證據可能會告訴我們，所得稅增加了 50%，但是在該證據所在的脈絡之外，這件事所造成的衝擊其實是不具任何意義的。研究者必須要問：其他的稅是否下降了？收入是否上升了？所得稅增加是否適用於所有的所得項目？每一個人所受到的影響是否是完全相等的？

時代錯誤（anachronism）是另外一個必須避免的歷史證據方面的謬誤，此種謬誤所指的是某個事件、表現、物品、人物、或者是情境出現的時間似乎早於或者是晚於其真正出現的時間，因而脫離了其真正的時間脈絡。例如，某一部電影的時間設定爲 1939 年，但是在電影中出現的一份文件上面的鉛字字體，卻是在 1972 年之後才被創造出來的，此種情況就是犯下了時代錯誤。雖然大部分的人可能不知道此種鉛字字體是何時被創造出來的，但是此種字體的創造者在看到這部電影的內容時，馬上就發現了此種時代錯誤的情況。[24] 你應該要很精確地瞭解事件的發生順序，並且留意在證據中關於事件發生日期的錯誤之處。

組織證據

當你在蒐集證據以及找出新的資料來源時，你就已經在組織資料了。顯而易見的，瘋狂地寫下筆記、並且讓這些筆記雜亂地堆在一起，是很不明智的作法。你可以利用注意到一些較低層次的推論以及主題的方式，來開始進行初步的分析。例如，在一項關於革命的研究裡，你可能發展出下列這個主題：富裕的農民傾向於支持舊的政體。你可以在筆記裡記下這個主題，並且在之後爲這個主題賦予意義和重要性。你接下來應該要去組織證據，並且利用理論的洞見來激發出組織資料的新方法，並且利用新的問題來質詢所找到的證據。資料與理論之間的*互動*，意味著你必須不斷地去質問證據，以發展出新的概念。換句話說，你必須不斷地檢視證據，並且去探究、質詢這些證據的意義爲何。正如同克萊門斯以及休斯（Clemens and Hughes, 2002: 207）所曾經警告過的：「真正的危險之處，就在於過度依賴一種簡單的『反射』模

時代錯誤 一種錯誤，也就是歷史比較研究者把一個事件、表現、物品、人物、或者是情境發生的時間弄錯了，造成比真正發生的時間更早或是更晚的情況。

型，此種模型把文本或者是敘述當作是證據，認爲此種證據可以用來證明『真正』發生過的事、人們『真正』相信的事物、或者是社會『真正』的運作方式。」

綜合證據

下一個步驟是綜合證據。在這個步驟裡，你必須持續不斷地精簡概念，並且逐漸形成一個一般的解釋模型，但是在你擁有了大部分的證據之後，這個過程的速度會變得越來越快。當你在進行這個步驟時，你應該要去修正舊的主題或者是概念，並且創造出新的主題或是概念。你應該要去尋找那些能夠橫跨不同時間或是單位的模式，並且利用類推的方式來找出相似處以及差異處。你也可以把分散的事件依順序加以排列出來，並且把事件歸類在一起，以創造出更大的圖像。接著，你應該要發展出一個合理的解釋，以把概念和證據整合成爲一個具有一致性的整體。你也應該要去反覆閱讀筆記，並且根據組織資料的策略來把筆記內容加以整理分類。除此之外，你還應該要試著去尋找資料之間的連結，並且以不同的方式來看待證據。

綜合（synthesis）能夠把特定具體的證據，連結到一個關於底層關係或是因果機制的抽象模型。你或許會使用到隱喻，例如，會導致革命的大眾挫折感，就像是一種情緒的雲霄飛車一樣：一開始時，期待會快速地上升（陡然上升），但是，接下來則是突然急轉直下（快速下墜）。這些模型是提高你對於證據敏感性的手段。換句話說，你先發展出一個解釋的模型，然後去尋找新的證據以證實其中的特定連結、評估該模型與證據之間的契合程度有多高、並且依據評估結果來進行調整。而此種情況也意味著你必須在抽象和具體之間不斷地來來回回（請參閱第十五章）。

許多歷史研究者會爲自己所提出的解釋找到關鍵指標以及支持性的證據。所謂的**關鍵指標**（critical indicator），指的是非常清楚明確的證據，而且通常可以當作在推論出一個特定理論關係時的充分證據。你必須爲某個解釋模型中的關鍵部分去尋找出關鍵指標。這些指標能夠清楚明確地證實某個理論性的推論，並且會出現在當許多細節都暗示了一個清楚的詮釋的時候。例如，顯示出在兩個國家之間的敵意的關鍵指標，就是彼此正式向對方宣戰的情況。而在判斷某個社會團體的政治力量逐漸上升時的關鍵指標，就是在社會上有許多公開支持該團體的正式組織，這些組織不但會員人數眾多，而且也都認同這個團體、並且提倡這個團體的主張。對比之下，支持性證據（supporting evidence）所暗示的指標，則是某個模型中比較不是那麼重要的部分，但是此種證據可以建立整體的背景或者是脈絡。此種證據通常比較不充分、或是比較薄弱，而且也缺乏一個清楚而且明確的理論詮釋。

撰寫報告

最後一個步驟，就是把證據、概念、以及綜合都結合起來成爲一份研究報告。在歷史比較研究中，你撰寫報告的方法是極度重要的。把證據、論點、以及結論等部分結合成爲一份報告，一直是相

關鍵指標　在一個特定的文化或是歷史情境中，一種對於某個概念的清楚明確的量數。

當重要的事。比起在量化取向中的情況，證據與解釋的安排以及陳述方式更加有可能會去成就或是摧毀了一項歷史比較研究。你必須將堆積如山的證據淬鍊爲清楚的說明與解釋，並且準備大量的注解、把證據與論點交互編織，以向讀者傳達出一幅有連貫性的、具有說服力的圖像（請參閱第十六章）。

歷史脈絡中的資料以及證據

歷史證據的類型

歷史（*history*）一詞有幾個意思，這個詞可能是指過去的事件（例如，法國從越南撤軍的過程已經是*歷史*了）、過去的記錄（例如，法國在越南參戰的*歷史*）、以及一門研究過去的學科（例如，*歷史*學系）。[25] **史學**（historiography）是進行歷史研究或者是蒐集以及分析歷史證據的方法。歷史社會學是歷史比較研究的一部分。這門學科

> 是一種歷史資料的研究取向、一種史學的型態，試圖以社會學的模型以及理論，解釋及瞭解過去……相對地，歷史資料也有可能會被用來說明及測試社會學概念、原則、以及理論的有效性（Mariampolski and Hughes, 1978: 104-105）。

我們在歷史比較研究中總共可能會使用到四種歷史證據：初級資料、次級資料、連續記錄、以及回憶。[26] 傳統的歷史學家極爲依賴初級資料，而歷史比較研究則通常會使用次級資料、或者是結合不同類型的資料。例如，卡達格諾（Quadagno, 1984）檢視了美國的社會安全法案，以評估有關政治權力的理論。在她所引用作爲證據的 47 筆資料來源當中，有 23 筆資料是初級資料（信件、備忘錄、官方報告、在該時期之內的報紙或雜誌上的文章）、有 3 筆是回憶（回憶錄或者是口述歷史）、而有 22 筆則是次級資料（由歷史學家以及其他研究者所著作的書籍）。

史學　一種研究取向的理論，所指的是當歷史學家在撰寫歷史研究時，會採取一些假定、強調重點、以及理論觀點的作法。

初級資料　關於過去事件或是社會生活的質化或是量化資料，此種資料是在過去的某個時期內被創造出來以及被使用的。

檔案館　一種特殊的圖書館或者是蒐集原始記錄以及歷史文件的地方，讓研究者在進行研究時可以方便取得這些資料。

初級資料　由那些過去的人們所留下的信件、日記、報紙、電影、小說、衣物、以及照片等等，並且能夠留存到今日的資料，就稱作是**初級資料**（primary sources），而我們則能夠在**檔案館**（archive）（一種收藏老舊記錄或是文件的特殊地點）裡、私人收藏品中、在家庭的櫥櫃裡、以及在博物館中找到這些資料（請參閱資訊箱 14.2 使用檔案資料）。而今日的文件以及物品（我們的信件、電視節目、電視廣告、衣物、汽車等等），則將會是未

資訊箱 14.2　使用檔案資料

檔案館是初級歷史材料的主要來源。檔案文件是存放於私人收藏、博物館、圖書館、或者是正式檔案館裡累積的文件材料。

資料所在位置以及如何接近與使用

要找出針對某個主題、組織、或者是個人的資料是否存在，可能會是一個冗長的、令人備感挫折的任務，因為必須經過大量的信件往返、打電話溝通、以及轉介的過程。如果關於某個人或是某個主題的資料確實存在，這些資料也有可能散落在許多不同的地點。在取得使用的權利時，若是私人收藏，可能必須依賴某個家庭成員的善意，或者是必須旅行到遙遠的圖書館去，並且檢視許多裝著信件、沾滿灰塵的盒子。除此之外，研究者也有可能會發現到，接觸這些資料的時間是有限制的（例如，某間檔案館每週只有開放四天，時間是從早上十點到下午五點，但是某位研究者卻必須花上 40 個小時來檢視這些資料）。

儲存以及組織

檔案資料可能沒有好好地整理歸類，或者是以各種不同的方法來安排組織。檔案資料的組織方式，可能會反映出一些與研究者的興趣無關的標準。例如，信件以及文件可能是以時間順序來組織的，但是研究者所真正感到興趣的對象，是那些在三十年的時間內寫給四名專業同僚的信件而已，而不是那些日常的帳單、家庭成員之間通信的信件等等。

技術以及控制

檔案資料可能會以原來的形式、微縮膠片、或者是電子格式（此種情況比較罕見）來存放。研究者可能僅僅被允許去記下筆記、但是卻不能影印複製，或者，研究者也可能只有被允許去看整體收藏品中的某些特定部分而已。當研究者面臨了一些限制時，像是必須在某個特定房間內去閱讀沾滿灰塵的文件、或者是只能在檔案資料在一天當中開放參觀的少數幾個小時內以鉛筆記下筆記等等，研究者很可能會覺得備受挫折。

追蹤與回溯

在進行檔案資料研究時，最困難的任務之一，就是透過資料來回溯共同的事件或是人物。即使所有的資料都存放在同一個地點，同一起事件或是關係可能會以許多不同的形式出現在數個地方。因此，研究者就必須去整理堆積如山的文件資料，以試圖在不同的地方找出關於證據的蛛絲馬跡。

單調乏味的工作、運氣、以及意外的發現

檔案資料研究的進展步調，通常會緩慢到讓人覺得很痛苦的地步。花上許多時間去審視細讀一些字跡難以辨認的文件，可以說是一件冗長而沉悶的工作。除此之外，研究者通常可能會發現在收藏品中有破損、在一連串的文件當中有缺漏、或者是有些文件已經損毀了。然而，仔細地去閱讀以及審視一些之前未曾被查閱過的資料，可能會產生一些令人驚訝的新連結或者是新想法，也就是說，研究者可能會發現一些意外的證據，並且因此而展開的新的調查方向或是線索（請參閱 Elder et al., 1993, 以及 Hill, 1993）。

來的歷史學家眼中的初級資料。經典的初級資料的例子，就像你在一間滿是灰塵的閣樓中找到一整疊已經發黃的信件，這些信件是由一位丈夫在戰爭時寫給太太的信件，而且這兩個人都已經逝世十年之久了。畢靈斯以及布里（Billings and Blee, 2000）曾經針對某個肯塔基社區的貧窮問題寫過一本書，而在他們書裡所使用的資料，包括了來自於一項關於個人以及農業的聯邦人口普查的手寫稿（1985 年到 1910 年）；稅務記錄、行爲、遺囑、以及法院記錄；報紙、肯塔基州會計委員會的記錄；由來訪的傳教士所寫下的信件以及報告；以及在郡和州的官方機構中所留下的正式記錄。作者們把個人、關係、以及家戶都連結起來，以創造出個人與家庭在 60 年期間之內的長期檔案。舒瓦茲（Schwartz, 1997: 493）曾經研究過林肯（Abraham Lincoln）在非裔美國人的社群當中如何成爲代表種族平等的歷史人物的過程，他曾經提過：「我的分析是以各式各樣的材料爲基礎的，其中包括了奴隸的敘述、演講、印刷品、照片、公眾意見調查、傳記、以及教科書。我特別仰賴三份黑人報紙在林肯生日時所發行的特刊：自由派的《芝加哥捍衛者報》（"*Chicago Defender*"）、溫和派的《紐約時代報》（"*New York Age*"）、以及保守派的《亞特蘭大每日世界報》（"*Atlanta Daily World*"）。」

不論是已經出版的、或者是尚未出版的書面文件，都是最重要的初級資料類型。你可能會發現這些資料是以原來的紙本形式保存著、或者是保存在微縮膠片當中。這些文件通常都是過去人們的文字、思想、以及感覺在現今僅存的記錄。在研究一些具有讀寫能力的人們所生存的社會以及歷史時期的時候，書面文件是非常有幫助的。書面的資料，尤其是那些留存至今的資料，通常都是由菁英分子或者是在官方組織中的領導者所寫下來的，而這也意味著此種資料很容易就會忽略了那些沒有讀寫能力的、貧窮的、或者是在官方社會組織之外的人們的意見和看法。例如，過去美國的奴隸若是去讀書或是寫字的話，都算是違法的行爲，因此，關於奴隸經驗的書面資料，通常都是間接的資訊、或者是難以取得。

在人們廣爲使用電訊、電腦、以及影像技術在記錄事件以及想法之前，寫在紙上的文字，是最主要的溝通媒介。事實上，有一些溝通形式是無法留下永恆的實體記錄的，並且已經大量取代信件、書面帳冊、以及報紙等等溝通形式。此種溝通形式的流行或許會使得未來的歷史學家的工作面臨更多的挑戰。在 2010 年 4 月時，美國的國會圖書館已經宣布，將會把在推特（Twitter）這個社群網站從 2006 年 3 月開始的公開推文（tweets）通通都歸類建檔（http://blogs.loc.gov/loc/2010/04/how-tweet-it-is-library-acquires-entire-twitter-archive/）。

次級資料 初級資料具有真實性以及可靠性，但是在時間方面的實際限制，卻有可能把針對許多初級資料所進行的研究，侷限在狹隘的時間框架以及地點之內。爲了獲得一個更爲廣闊的圖像，我們必須使用**次級資料**（secondary sources），也就是由那些花費了多年時間研究初級資料的專業歷史學家所寫出來的著作。

次級資料 一種在歷史研究中所使用的質化以及量化資料，此種資料是由並未直接參與事件或是情境的歷史學家或是其他人所報告或者是寫下的。

連續記錄　由某些組織所定期維護以及保存的檔案或者是現存的統計文件，就被稱作是**連續記錄**（running records）。此種記錄的例子之一，是在某間鄉村教會中的一份檔案，其中包含了從 1910 年至今所有的婚姻和死亡的記錄。例如，當羅伊（Roy, 1983）在研究從 1886 年到 1905 年之間的美國大型企業的董事會時，就使用了連續記錄。他的證據來自於 150 份初級文件、官方報告以及統計數據、以及年度的商業參考書籍等等，其中有些資料至今仍然在印行當中。

連續記錄　以一種相對連貫的方式在一段長時間之內所保留下來的現存檔案、報告、或者是文件，而且被運用於統計研究當中。

回憶　由個人根據記憶而口述或者是寫下的、關於他們過去的生活或是經驗的資料，就稱作**回憶**（recollections）。這些資料的形式可能是回憶錄、自傳、或者是訪談。由於記憶是不完美的，因此回憶通常會出現一些初級資料所沒有的扭曲現象。在蒐集**口述歷史**（oral history）（這是一種回憶的類型）時，你會去和一些年紀較大的人們，針對他們在過去的生活或是事件進行無結構的訪談。對於非菁英團體或者是沒有讀寫能力的人們而言，此種方法是特別有價值的。口述歷史的技術是從 1930 年代開始發展出來的，而且現在已經有專門討論此種技術的協會以及學術期刊。[27] 許多關於記憶的研究都建議，當研究者在使用口述歷史或者是回憶時，應該要抱持著謹慎的態度（請參閱資訊箱 14.3 記憶的七項致命原罪）。正如同薛克特（Schacter, 2001: 9）所說的：

回憶　一種關於事件的年代記錄或陳述，取得口述歷史的方式，通常是去訪談某個人，並請他或她回憶過去曾經經歷過的事件、信念、或者是感受。

> 我們傾向於把記憶當成像是從家庭相簿裡所取出來的照片一樣，也就是說，只要好好保存的話，這些相片在被取出來時的狀況，會和當初被保存時的狀況完全一模一樣。然而，我們現在已經知道，當我們在記錄我們的經驗時，無法和相機記錄這些經驗的方式一樣……我們會從經驗當中提取出關鍵的要素，並且把這些要素儲存在記憶當中。然後，我們會重新創造或者是重新建構我們的經驗，而不是單純地重新取回這些經驗的複本而已。有些時候，在重新建構經驗的過程當中，我們還會去加入一些在事件過後才得到的感受、信念、或者甚至是知識。

曾經有研究者進行了兩項研究，而在研究中所使用的是生活在二十世紀初期時一些美國女性的口述歷史。這兩項研究可以當作範例，以說明研究者如何把各種不同種類的文件和口述歷史結合在一起。布里（Blee, 1991）研究了在三 K 黨裡的女性成員，而盧肯以及佛漢（Luken and Vaughan, 2006）則是訪談了一些女性，而其訪談主題是關於成長、組織家庭、以及搬遷到各種不同房子的經驗（請參閱範例 14.5 三 K 黨裡的女性成員，以及範例 14.6 關於標準美國家庭以及適當的兒童教養方式的口述歷史）。

資訊箱 14.3　記憶的七項致命原罪

薛克特（Schacter, 2001）觀察到，記憶喪失或記憶錯誤會有以下的數種形式：

1. *短暫無常*（Transience）　這是指記憶會隨著時間流逝而緩慢地、持續地衰退的狀態，換句話說，當某個事件發生的時間越久遠，人們能夠回憶起來的細節就會越少。
2. *心不在焉*（Absentmindedness）　一種由於過度專注於某個想法或是事物，因而錯誤地引導了某人的注意力的狀態，如此一來，他或她就會忘記了其他的、通常是簡單的事物（例如：因為過於專注在某個重大的計畫，所以忘記去拿車鑰匙）。
3. *阻礙*（Blocking）　某人試圖去尋找自己所擁有的資訊，但是卻不成功，即使是在當下努力地嘗試，卻仍無法正確地回想起來（通常受訪者會說：「就在我嘴邊，差一點就可以說出來了。」（"It's on the tip of my tongue."）。
4. *錯誤歸屬*（Misattribution）　把幻想誤以為是現實，或者是把某人從朋友那兒聽說的事件、或者是在電影中看到的情節，誤以為是自己的真實經驗。
5. *易受暗示所影響*（Suggestibility）　此種情況所指的是某個人在被問問題時，因為接受到一些暗示，所以開始去扭曲他或是她自己的記憶，並且相信某些事件的確發生過，但是事實上這些事件根本不曾存在過。
6. *偏誤*（*Bias*）　在回憶事件時的扭曲現象，而且通常是投入了一些在事件發生過後才有的想法、感受、或者是信念。
7. *揮之不去*（*Persistence*）　雖然努力地嘗試，但是仍然無法忘記某些事件。

範例　14.5

三 K 黨裡的女性成員

在《三 K 黨裡的女性成員》（*Women of the Klan*）這項研究當中，布里（Kathleen Blee, 1991）注意到，在她進行這項研究之前，沒有任何人研究過在這個美國最大的種族主義的、右派的運動中的大約 50 萬名女性。她認為此種情況可能是因為人們一般都假定女性是不關心政治的、而且是比較被動的。她花了六年的時間去研究這個秘密社團裡超過六十年前的一些不知名的成員，而她的作法也顯示出了在進行歷史比較研究時所需要的獨創性。

布里聚焦於印第安那州，因為在 1920 年代的高峰期時，當地有高達 32%的白人女性都是三 K 黨的成員。除了回顧已出版的關於三 K 黨的研究之外，她也調查了一些文件，其中包括了報紙、宣傳的小冊子、以及未出版的報告。她也造訪了超過一打以上的大學圖書館、政府圖書館、以及歷史圖書館，以進行初級以及次級資料的圖書館研究。而在她書中的歷史照片、素描圖片、以及地圖等等，則為讀者營造出這個主題的氛圍。

因為布里無法取得三 K 黨成員的名單，所以要找到相關的資訊是相當困難的工作。她找出三 K 黨女性成員的方式，包括了去拼湊一些留存至今的名冊片段、找到一些三 K 黨女性成員在報紙上所公告的訃聞、詳細檢查一些公共告示或者是反對三 K 黨的文件以找出女性成員的名字、以及訪談至今仍活在世上的一些三 K 黨女性成員。

為了要找出在三 K 黨活動結束了 60 年之後的倖存者，布里必須要有堅韌的毅力以及獨特的創意。她把一份關於她所進行的研究的通知，寄送到印第安那州的每一個地區性的報社、教會公布欄、廣告夾頁、歷史學會、以及公共圖書館。她獲得了三份書面的回憶，進行了三次未

被記錄的訪談以及十五次有記錄的訪談。她的資訊提供者大多數都已經超過 80 歲了，而且這些資訊提供者在回憶起三 K 黨時，都認為那是她們生命中相當重要的一部分。布里則是透過報紙以及其他的文件證據，來證實這些資訊提供者回憶中的各個不同部分的正確性。

加入三 K 黨一直是一件極具爭議性的事，但是在訪談過程當中，布里並未透露出她對於此事的意見。雖然有些資訊提供者會去試探她的反應，但是她仍然保持中立的態度，而且沒有抨擊三 K 黨。她在研究中提到：「我自己本身曾經住過印第安那州（我從小學開始，一直到大學都住在這兒）、以及白種人的背景，使得資訊提供者們認為——在缺乏反面口頭證據的情況之下——我和她們擁有相同的世界觀」（p. 5）。她發現，三 K 黨的女性成員並不是殘忍、無知、或者是充滿仇恨的。除此之外，當布里去詢問這些女性為何加入三 K 黨的原因時，她也得到了一個出人意表的反應。大多數的資訊提供者對於這個問題都覺得很困惑。對她們而言，這根本是不需要任何解釋的，因為這不過就只是「一種成長的方式」，以及「一種和別人聚在一起並且得到樂趣的方式」而已。

範例　14.6

關於標準美國家庭以及適當的兒童教養方式的口述歷史

盧肯以及佛漢（Luken and Vaughan, 2006）利用口述歷史以及檔案資料研究的方式，來研究三個政府機構（美國的兒童局（Children's Bureau）、擁有自己的房子（Own Your Home）、以及更美好的家園（Better Homes））在 1900 年到 1930 年之間，如何去提倡一項「創造一個養育兒童的適當家庭環境」的活動。這三個機構一起發動了全國性的「標準美國家庭」（Standard American Home）運動，這項運動把適當的兒童教養方式以及創造一個合適的家庭環境這兩件事連結在一起。這項運動通常忽略了人們的實際生活方式，而是去試圖把家長們生活以及教養兒童的方式，連結到商業的住宅計畫以及家庭用品市場。作者們把口述歷史與他們對於標準美國家庭運動的分析銜接在一起。他們把這項運動詮釋為是一種論述，此種論述所表明的是，就教養兒童的問題而言，在適當的住家安排方面有其特定的階級、性別、以及種族標準。例如，家長應該擁有自己的房子、而不是租房子；房子應該是獨棟的，而不是公寓大樓中的一層而已。房子應該位於郊區，或者，如果一定要在城市中的話，也應該要靠近公園。房子應該要通風良好、乾燥、採光佳、溫度維持恆定、而且要易於清理。每一個兒童都應該要有各自的房間，而且房間的要有適當的色彩、傢俱、以及適合兒童年齡與性別的玩具。這項運動也與許多行業的活動互相配合，像是木材公司、存款與借貸機構、不動產協會、建商、兒童發展專家、醫生、房屋裝潢業者、學術界、以及報社和雜誌社等等。

這項研究的兩位作者，把下列這兩件事物連結在一起：一般人們談論以及思考他們的日常活動的方式；以及廣泛的檔案資料研究，其中包括了初級文件、記錄、廣告、以及照片等等。他們曾經在四個場合中，對居住在鳳凰城都會地區的五名女性訪談大約二個小時之久，而這些受訪者都至少超過 60 歲了。盧肯以及佛漢詢問了這些女性一生當中的「居家工作」（"housing work"），並且詢問了在這些女性的生命過程中由她們所塑造的居住環境的情況。兩位作者把訪談過程錄音下來，並且針對每一個女性謄寫出將近 10 到 12 個小時訪談內容的逐字稿。他們提到（p. 302）：

> *「當我們在檢閱謄寫出來的逐字稿時，我們注意到這些女性都會非常詳細地提到她們在少女時期以及身為人母時期的經驗。除了告訴我們她們在兒童時期所住過的地方、以及他們撫養孩子的地方之外，她們也提到了她們做過的事情、她們使用過的物品、她們曾經互動過的人物、以及她們喜歡或者是討厭她們所做的活動以及所處的情況。她們的敘述讓我們大開眼界，讓我們瞭解到在日常居住環境以及養育兒童這兩件事情之間的連結。除此之外，她們也提供了一個機會，讓我們可以透過居住環境來探索兒童的生活，也能夠探索形塑了這些經驗的社會關係。」*

有些人會「重寫」過去的歷史以使其與目前的信念更爲一致、或者是以一種自我強化的方式來回憶過去的經驗（也就是以一種較爲正面的方式來錯誤地回憶過去的經驗）。較爲年長的成人（通常是五十幾歲左右、或者是更爲年長）通常會比那些較爲年輕的成人更可能會忘記了過去事件的特定細節。教育程度較高、心智方面較爲活躍的年長成人，會顯現出較少記憶喪失的問題，但是在另一方面而言，某種程度的個人或是集體記憶的扭曲情況，則是相對較常出現。[28]

以次級資料進行研究

運用以及限制　社會研究者通常會使用次級資料以及由歷史專家所撰寫的書籍與文章來當作過往狀況的證據。[29] 正如同史寇克波爾（Skocpol, 1984: 382）所說的，運用此種資料的方法仍未被系統化，而且「到目前爲止，對於如何有效地運用次級資料來當作證據的方法，比較歷史社會學家仍然尚未想出清楚的、獲得共識的規則以及程序。」次級資料有其限制，而且在運用時也必須要謹慎小心才行。

次級歷史證據的限制包括了不正確的歷史敘述，以及在你感興趣的領域中可能沒有相關的研究。我們不能使用次級資料來考驗假設。*事後*（*post facto*）的解釋無法達到實證主義對於可證僞性（falsifiability）的標準，因爲我們無法去運用統計控制或者是複製來檢驗此種解釋。[30] 然而，在發展出一般性的解釋時，由其他人所進行的歷史研究除了有許多用途之外，也扮演了一個相當重要的角色。例如，此種研究可以用來證明長期趨勢的出現以及演進。[31]

潛在的問題　數量龐大的次級資料所呈現出來的，是由許多細節以及詮釋所形成的複雜迷宮。你必須把大量的描述性研究轉化成爲一幅清晰而且易於理解的圖像，不但能夠與資料的豐富性一致、而且也能夠反映出資料的豐富性。此外，這幅圖像也必須能夠連結許多特定的時期以及事件發生的地點。在使用次級資料時，你可能會遭遇到四個潛在的問題：隱而未顯的理論、經過選擇的證據、敘事寫作、以及歷史學的各種學派。

1. *隱而未顯的理論*　當你在閱讀歷史學家的著作時，可能會遇上一些問題。[32] 歷史學家所呈現出來的並不是不帶任何理論的、客觀的「事實」。他們會以隱而不顯的方式去架構原始資料、將資訊分門別類、而且也會使用概念。歷史學家所使用的概念通常是一種由新聞學、歷史行動者的語言、意識型態、哲學、當前的日常語言、以及社會科學所形成的混合體。因此，這些概念可能會是模糊的、應用的方式可能出現前後不一致的狀況、而且也不是互相排斥或者是全容的。

例如，某位歷史學家把一群在十九世紀時住在一座城鎮裡的人們描述爲上層階級，但是卻從未去對這個詞彙下定義，而且也沒有把這個詞彙連結到某個關於社會階級的理論。此處所出現的方法論問題，是這名歷史學家隱而不顯的理論會對證據造成限制。藉由去仔細閱讀一些毫無組織的概念的方式，你可以試著去找出某些解釋的證

據，而這些解釋可能與那些由歷史學家們在次級資料中所運用的解釋完全相反。

2. *經過選擇的證據*　歷史學家的選擇過程也不是透明的，換句話說，歷史學家會從所有可能的證據中挑選出某些資訊出來。正如同卡爾（Carr, 1961: 138）所提到的：「因此，就歷史意義而言，歷史是一種選擇的過程……從無窮無盡的事實大海當中，歷史學家挑選出那些對於他的目的具有意義的事實」，然而，你卻不知道此種挑選的過程是如何進行的。在不知道挑選過程的情況之下，你也只能去相信歷史學家的判斷，但是此種判斷卻可能會帶有偏見。[33] 例如，某位歷史學家可能閱讀了 10,000 頁的報紙、信件、以及日記，然後在一本大約 100 頁的書籍中，把這些資料濃縮成爲一些摘要以及挑選過的引文。在此種情況之下，你根本無從得知那些被這名歷史學家遺漏的資訊，是否與他或她的研究目的具有相關性。

歷史學家典型的研究作法也會引入個人的偏見。由於歷史學家極度依賴初級資料以及留存至今的人工製品，而且通常也會有一種非理論的傾向，如此一來，就容易使得歷史學家把焦點放在特定人們的行動上面。此種特殊性的、微觀層次的觀點，會把歷史學家的注意力從那些具有整合性的主題或是型態上面移開。對於特定個人行動的強調，也是一種理論傾向的類型。[34]

3. *敘事寫作*　歷史學家會去組織安排各種證據，以撰寫*敘事歷史*（*narrative history*）（請參閱資訊箱 14.4 歷史學的敘事方法），而此種作法會使得未明確定義的概念以及經過選擇的證據這兩個問題變得更爲複雜。在歷史敘事當中，作者會圍繞著一個單一的連貫「故事」，依照年代的順序來安排各種資料與證據。這位作者會依據在事件的時間順序中的位置，把故事的每個部分與其他部分連結在一起。所有的部分會一起形成一個一致的整體。共同發生的情況（conjuncture）以及伴隨發生的事件（contingency）都是敘事形式中的關鍵要素，換句話說，也就是如果 X（或者是 X 加上 Z）發生了，那麼 Y 就會發生，而如果 X（或者是 X 加上 Z）沒有發生的話，就有可能會發生其他的事件。伴隨發生的事件會在較早以及較晚的事件之間創造出一種邏輯上的相互依賴性。由於具有時間方面的邏輯，因此敘事的組織方式與量化解釋之間就會有所差異，因爲在量化解釋中，研究者是利用辨識出統計型態的方式來推論出原因。

敘事的困難之處在於，研究者用來組織證據的工具－也就是在一個以故事爲基礎的事件序列中的時間次序或是位置－並無法單獨指明在理論方面或者是歷史方面的因果關係。換句話說，敘事符合了建立因果關係的三項標準的其中之一，也就是時間順序（請參閱第三章）。當歷史學家在敘事當中包含了一些不具有因果意義的事件時，敘事方法就有可能會模糊了底層的因果模型或是過程。作者有可能會在敘事中包含一些可以使背景或脈絡更爲豐富、或者是使其更爲引人入勝的事件，但是此種事件可能不具有直接的或者是延遲的因果影響。除此之外，極少有敘事歷史學家會去清楚明白地說明各種結合或者是交互作用的效果是如何運作的。當一名歷史學家討論了某個事件的三種條件或是狀況時，你不知道這三種條件是否必須同時運作才能發揮作用、或者是只有單獨一種條件或單獨二種條件也能夠造成完全相同的影響。[35] 敘事的組織方式也可能會產生一些相互矛盾的發現，因此，你必須非常仔細

資訊箱 14.4 歷史學的敘事方法

許多歷史學家都是以傳統的敘事形式來寫作的，而這些著作則可能成為歷史比較研究者能夠加以運用的次級資料。

敘事形式的特徵

1. 所講述的是一個故事，其中有主要情節與次要情節、分水嶺、以及高潮部分。
2. 依循著一種時間順序以及事件的序列。
3. 聚焦於特定的個人，而不是結構或者是抽象的想法。
4. 主要是個殊的以及描述性的，而非分析的或者是一般性的。
5. 所呈現的事件是獨特的、不可預測的、以及偶然的。

敘事形式的優點

1. 此種形式是多采多姿的、有趣的、而且讀起來具有娛樂效果。
2. 提供不同時代的生活的整體感受，使讀者們有一種身歷其境的感受。
3. 此種形式傳達了過去的人們對於現實的主觀感受，並且幫助讀者們能夠在情緒上與過去的人們感同身受。
4. 此種形式以社會現實的許多層面來環繞著個人以及特定的事件。

敘事形式的缺點

1. 隱藏了因果理論以及概念，或者是讓這些理論與概念處於隱晦不明的狀態。
2. 此種形式使用了修辭、日常語言、以及常識邏輯來說服讀者，因此容易受到語意扭曲的邏輯謬誤以及各種修辭手法的影響。
3. 通常只專注於獨特的、戲劇性的、特別的、不尋常的事件，而忽略了一般的、或者是普通的事件。
4. 此種形式很少建立在先前的知識之上，而且也幾乎不會創造出一般性的知識。
5. 此種形式傾向於過度個人主義的，而且也會過度誇大特定人們的角色、以及他們自發地去形塑事件的能力。

地去閱讀這些薄弱的概念、未知的選擇標準、以及含糊不清的因果邏輯。在敘事的底下或許隱藏著這名歷史學家的社會理論，但是這個理論通常是隱而不顯的、而且是被隱藏起來的。

4. *歷史學的各種學派* 歷史學家會受到各種因素的影響，像是歷史學的學派、個人的信念、社會理論、以及當歷史學家在進行研究當時的時事等等。今日的歷史學家在撰寫研究時，他們檢視初級資料的方式，會與 30 年前的歷史學家的檢視方式完全不同。除此之外，在歷史學的領域中，也存在著數種不同的學派（例如，外交的、人口統計學的、生態的、心理學的、馬克思主義的、知識份子的等等）。每一種學派在尋找證據以及提出問題方面都有自己的規則。卡爾（Carr, 1961: 54）曾經警告過：「在你研究歷史之前，先去研究歷史學家……在你研究歷史學家之前，先去研究該名歷史學家所處的歷史以及社會環境。」

以初級資料來進行研究

當你在使用次級資料時，歷史學家會是一個主要的議題。而當你所使用的是初級資料時，主要的議題之一，是在所有在過去被寫下的或者是被使用的事物當中，只有很小的一部分能夠留存到今日。除此之外，對於曾經存在過的事物而言，留存至今的資料是一種非隨機的樣本。羅恩梭爾（Lowenthal, 1985: 191-192）曾經觀察到：「過去的思想以及事物所殘存至今的遺緒，所代表的是先前世代中的當代結構裡的極小部分而已。」當你在試圖閱讀初級資料時，你應該以活在過去時代裡的人們的眼光和假定來閱讀這些資料，而且你也應該把關於該時期之後發生的事件的知識以及現代的價值觀「放入括弧」或者是加以保留。坎特以及施奈德（Cantor and Schneider, 1967: 46）曾經寫道：「如果你在閱讀初級資料時，沒有保持開放的心胸、沒有想要瞭解作者真正想法的企圖、或者是沒有以*當時的人們*看待事件的方式來看待這些事件的話，那麼你就是在浪費你的時間而已。」例如，當你在一份由某個奴隸的主人所寫的資料時，不論你是基於道德的理由而去反對奴隸制度、或者是去批評作者沒有看出奴隸制度的邪惡之處等等，其實都是沒有意義的作法。你應該先保留自己的道德判斷，並且先成爲一名道德相對論者。你必須「像你的研究對象一樣地去思考以及相信……以去發現在他們自己的眼中，他們覺得自己的表現如何」（Shafer, 1980: 165）。

找出初級資料可能會是一件相當耗費時間的工作。你必須搜尋特殊的索引，並且造訪許多檔案館或者是特殊的圖書館。初級資料通常位於一些布滿灰塵的、偏僻的房間之內，裡面放滿了成堆的紙箱，而紙箱內則放著大量逐漸褪色發黃的文件。這些檔案資料可能是不完整的、沒有組織的、而且有不同程度的損壞狀況。一旦你找到了文件或者是其他的初級資料，你必須以外在與內在的評論來評估這些資料（請參閱圖 14.1）。

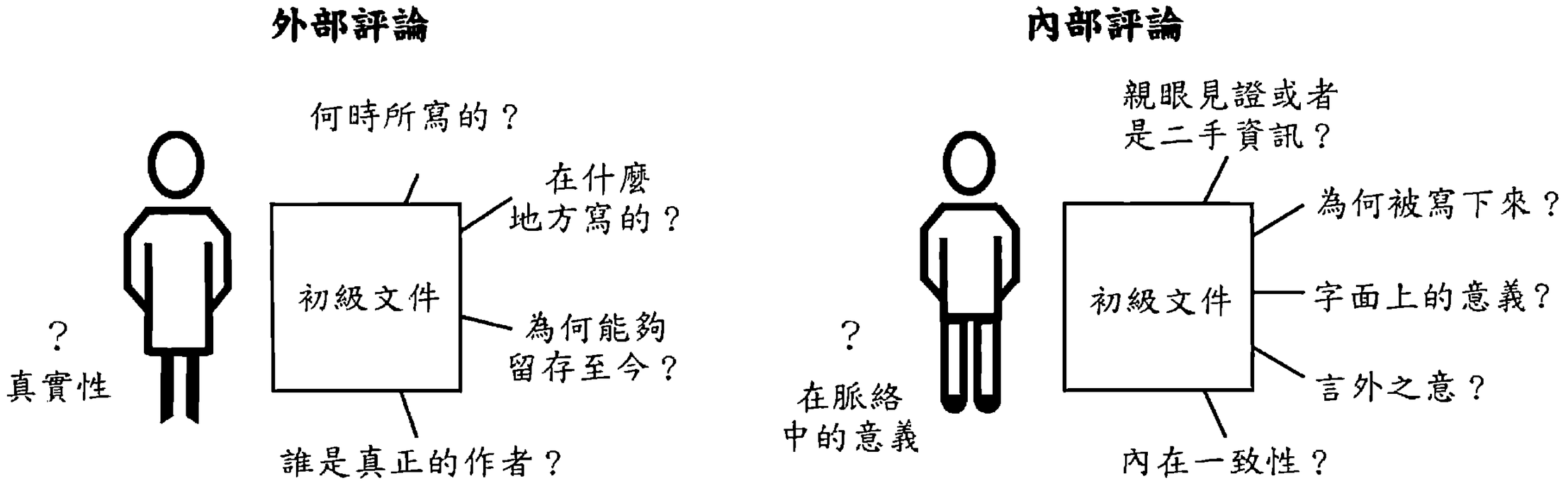


圖 14.1　內部以及外部評論

外部評論（external criticism）所指的是，某份文件本身的真實性必須經過評估，才能確定該份文件不是假的或者是僞造的文件。此種評論包括了去詢問這份文件是否真的是在其所宣稱的時期、所應該在的地方、以及由其所宣稱爲作者人所創造出來的。除此之外，我們也應該要去追究該份文件在一開始時被創造出來的原因，還有該份文件能夠留存至今的方式。

一旦某份文件已經通過了真實性的考驗之後，你就必須使用**內部評論**（internal criticism），也就是去檢驗該份文件的內容，以建立可信度。你必須去評估該份文件的作者是否真的直接目擊了某些事件，或者只是在傳達一些二手資訊而已。換句話說，你必須同時去檢視文件內容在字面上的意義，以及微妙的言外之意或是意圖。你應該要去留意在文件中提到的其他事件、資料、或者是人物，並且問問自己是否能夠去證實這些人事物的真實性。你也應該去檢視隱含的假定或者是價值位置，並且考慮這份文件在被創造出來時的相關條件（例如，在戰爭時期、或者是在一個極權政體當中）。即使是在相同的語言之中，各種字彙的意義也會經常隨著時間的經過而不斷變動，因此你也必須去考量到文件被創造當時的語言用法，以及文件內的敘述所在的脈絡，才能汲取出真正的意義。

任意修改（bowdlerization）是在初級文件中可能會出現的扭曲現象之一，也就是一種爲了維護某種道德標準、或者是呈現出某種特定形象而刻意進行的扭曲。例如，你發現了一張在某棟建築物前面所拍攝的照片。事實上，在該棟建築物周遭散落著大量的啤酒罐，而且牆上的油漆也已經開始在脫落了。然而，那張照片的拍攝地點是該棟建築物某個幾乎沒有垃圾的特定部分，而且在構圖時也刻意避開垃圾，好讓垃圾不會出現在照片當中。除此之外，這張照片也運用了修圖的技術，而使那些褪色的、脫落的油漆看起來就像是新的一樣。另外一個例子，則是當報紙上的社會專欄在報導某位富裕人士所舉行的派對時，把一些實際上並未參加該次活動的一些名人的名字加到來賓名單當中。[36]

除了初級以及次級資料之外，你或許還會使用到托波斯基（Topolski, 1976）所說的**非來自於資料的知識**（nonsource-based knowledge）。此種知識是指你可以取得的關於過去事件的資料，但是其內容並非來自於特定的初級或者是次級資料。此種知識可能是來自於邏輯的推理，例如，A 和 B 是在一個一夫一妻制的社會當中的一對已婚夫婦，而這個社會極爲重視在性方面的忠誠。因此，當 B 有了婚外情時，A 就非常有可能會感到嫉妒與憤怒。非來自於資料的知識也包括了對於之前的重大事件的知識，而此種重大事件會形塑了你所研究的事物所在的脈絡。例如，你研究了在 1920 年代時的法國。你

外部評論　對某份初級歷史資料的真實性所進行的評估，而評估的方法是正確地找到該份文件被創造出來時的地點以及時間（例如：決定該份文件是否是一份偽造的文件）。

內部評論　對初級歷史資料的真實性以及可信度所進行的評估，以決定該份文件對於過去事件的描述的正確性。

任意修改　一種對於過去的刻意扭曲作法，其用意在於保護某種特定的（通常是有利的）形象。

非來自於資料的知識　由研究者根據推理或者是對於歷史情境的深入理解所獲得的一般資訊。

應該要知道，在當時的法國裡，有許多年齡在 18 歲到 40 歲之間的男性在先前幾年發生的第一次世界大戰中失去生命。當前的知識也能夠幫助你瞭解過去的事件。例如，你知道黑死病是一種由鼠類身上的跳蚤所散播的疾病，而且是由惡劣的衛生條件所造成的。然而，生活在過去的人們並不知道造成這種疾病的真正原因。

比較研究

比較研究的類型

比較方法 比較研究的焦點在於不同單位之間的相似性以及差異性，此種研究能夠突顯出在其他類型的研究中所出現的議題。[37] 正如同侯爾特以及透納（Holt and Turner, 1970: 6）曾經說過的：「原則上，在跨文化的比較研究、以及在單一社會所進行的研究之間，並無任何不同之處。更確切地來說，真正的差異性在於某些問題類型的規模大小上面。」比較的觀點能夠暴露出在研究設計中的缺點，並且有助於提升研究的品質。比較研究比較像是一種傾向，而不是一種單獨的研究技術。而在這個部分的內容裡，我們所要討論的是比較研究所具有的優點。

比較研究能夠揭露出社會生活的某些面向，這些面向在運作時能夠橫跨不同的單位（例如：城市、國家、文化等等），而不是那些僅限於單獨一個單位所具有的特質。在所有的研究當中，我們都會想要去進行某種程度的推論。實證主義研究者所感到興趣的，是去發現那些在各種社會裡都能成立的、關於社會行爲的一般律則或者是模式。然而，幾乎沒有任何實證主義的研究能夠具有比較的觀點。拉金（Ragin, 1994a: 107）曾經觀察到：

> 比較研究檢視了橫跨不同個案的相似性以及差異性的型態，並且試著去接受個案的多樣性……量化研究者也會去檢視個案之間的差異性，但是所強調的重點卻是不一樣的，因為量化研究的目標是去解釋某個變項和另一個變項之間的共變，而且此種共變現象通常會橫跨許多個案……一般而言，量化研究者對於個案通常都只有廣泛的熟悉而已。

比較的傾向可以改善測量以及概念化的過程。我們所發展出來、並且用來研究許多社會單位或是環境的那些概念，通常不太可能只有適用於特定的文化或是環境。除非我們能夠把一個概念應用於不同的文化或者是環境當中，否則的話，要察覺到隱藏的偏見、假定、以及價值觀等等，是相當困難的。不同的社會環境能夠提供範圍廣泛的事件或是行爲。在某個文化中的範圍，會比一般人類行爲的範圍來得狹隘，因此，在某個單一文化中的研究，就會聚焦於可能的社會活動的有限範圍之內。例如，假設現在有兩位研究者席萍（Hsi-Ping）以及阿布杜爾（Abdul）想要去檢視兒童斷奶的年齡以及開始出現情緒問題的年齡之間的關係。席萍只有檢視了美國的資料，其中的斷奶年齡範圍是從 5 個月到 15 個月，並且也顯示出，隨著斷奶年

齡增加，出現情緒問題的年齡也會跟著穩定地增加。因此，她所得到的結論是，斷奶時間較晚會造成情緒方面的問題。阿布杜爾所檢視的資料是來自於十個不同文化的資料，而且其斷奶的年齡範圍是從 5 個月到 36 個月。他發現到，到 18 個月爲止，情緒問題的比例會隨著斷奶的年齡而上升，而當兒童在 18 個月時斷奶的話，情緒問題的比例會達到頂峰，但是之後則會下降到一個較低的比例。阿布杜爾因此得到一個較爲正確的結論：當斷奶年齡是在 8 個月到 18 個月之間時，情緒問題的比例與斷奶年齡之間呈現正相關，但若是斷奶時間早於 8 個月、或者是晚於 18 個月，反而會減少產生情緒問題的機會。由此可見，由於在美國的斷奶年齡範圍較爲狹窄，因而使得席萍對於斷奶年齡與情緒問題之間的關係作出了錯誤的結論。

比較研究的優點之一，在於此種研究能夠排除、或者是提供因果關係的替代解釋。例如，魏爾（Weil, 1985）曾經研究過在學校教育年數以及不寬容的態度之間的關係。過去的研究已經發現在美國存在著此種關係，而且大多數的研究者都認爲，教育能夠擴展眼界並且減少不寬容的態度。在檢視了許多國家的資料之後，魏爾（Weil, 1985: 470）所得到的結論是，「與採行自由民主政治已經有一段時間的國家相較之下，在非自由的民主政體中、或是在之前數十年並未採行自由民主體制形式的國家之中，此種關係是薄弱的、不存在的、或者有時候甚至是倒反過來的。」換句話說，此種關係並不是舉世皆然的，而是只有在某種特定類型的政府體制之中才能成立。教育的功能並不是產生一種提高容忍態度的普世效應，反而是使人們社會化並且習慣於他們國家的官方價值觀。若是在官方價值觀中包括了容忍的態度，那麼教育就能提升此種態度；而在其他的國家裡，教育則沒有此種效果。由此可見，過去的研究之所以會得到錯誤的結論，其實是因爲這些研究只去檢視了一些重視容忍態度的國家而已。

與非比較的研究相較之下，比較研究較爲困難、成本較高、而且也必須耗費較多的時間。除此之外，我們在比較研究中也經常會遇到同義（equivalence）的問題（稍後將進行討論）。基於以下的幾個原因，使得比較研究者幾乎無法使用隨機抽樣的技術。首先，如果分析單位是國家的話，全世界也僅有 170 個國家左右，而且並不是所有的國家都擁有充分的資訊可以讓研究者加以取用。對於一個非隨機的次群組而言（貧窮國家、非民主國家等等），研究者根本無法取得充分的資訊。除此之外，當有些國家的人口高達 10 億人、而其他國家的人口僅有 10 萬人時，我們是否能夠把所有的國家都當成是完全相等的單位呢？當我們手中僅有少數的個案時，我們傾向於針對每個個案進行特別處理，並且把每一個個案都當作都是獨一無二的，但是此種作法會限制了推論的可能性。例如，你檢視了五個個案（例如：五個國家），但是其中每一個個案與其他個案之間都有二十種差異。當差異性的數量大於個案的總數時，要去考驗理論或者是決定某種關係是否存在，就會變得非常困難。

比較研究者通常是去應用理論，而不是檢驗理論。雖然比較研究者有能力去使用結合的理論、以及把個案當作完整的整體來加以考量，但是卻很難透過實驗的方式來嚴謹地考驗理論。例如，你對於經濟衰退所產生的效果感到興趣，但是你卻無法讓某一群國家發生經濟衰退的情況、而讓其他國家維持原狀。相反地，你必須先

等到經濟衰退的情況發生了，然後再去檢視該國家或是單位的其他特徵。

比較研究的四種類型　正如同之前已經討論過的，歷史比較研究以及比較研究本身的類型都有極大的差異性。柯恩（Kohn, 1987）已經討論過比較研究的四種類型，其中兩種符合獨特的歷史比較取向：（1）把國家當作研究對象、或者是個案研究的比較研究，以及（2）把國家當作是研究的脈絡。第三種是實證主義的取向，在此種取向中，國家是分析單位。至於最後一種則是多國的研究。[38] 在這個部分裡，我們會針對這四種類型中的每一種來進行討論。

1. *個案研究的比較研究*　**個案研究的比較研究**（case-study comparative research）取向會針對特殊的社會或者是文化單位進行比較，但是並不會作出廣泛的類推。此種類型的比較研究所處理的研究問題範例如下：加拿大與美國之間有何差異？在日本、臺灣、以及韓國裡的老年人的生活經驗如何？美國以及俄羅斯的教育系統之間有何相似性與差異性？在進行個案研究的比較研究時，你必須深入地去檢視數量有限的個案，而這裡所稱的「個案」，是一種由文化所定義的團體，通常是指國家。利用深入檢視少數或者是單一個案的方式，相對而言就比較不必去擔心單位之間的同義問題。在辨識出某些固定的因素、或者是在少數個案之間產生變化的因素時，此種方法是相當有幫助的。[39] 在本章一開始的部分所提到的對於韓國境內基督宗教的研究，以及對於法國以及美國稅制的比較研究，都是屬於此種類型的比較研究的範例。

2. *文化脈絡研究*　針對一些可以當作是不同類型的社會或是單位的代用物的個案所進行的研究，就被稱作是**文化脈絡研究**（cultural-context research）。馬爾克斯（Anthony Marx, 1998）曾經針對在美國、南非、以及巴西等國當中的種族關係進行過一項研究。他檢視了在這三個國家裡的一些共同想法、種族關係、官方政策、以及與種族相關的法律。他並不是任意地或者是隨機地選擇這三個國家來進行研究。更確切地說，他選擇這三個國家的原因，是因爲這三個國家代表了三種不同的種族政體的類型、或者是三種組織種族團體之間關係的不同系統。每一個國家都提供了一個不同的文化脈絡，而在這些文化脈絡當中，他則能夠去檢視共同的主題、型態、以及過程。

3. *跨國研究*　在**跨國研究**（cross-national research）中所使用的分析單位是國家。你在不說明國家名稱的情況下，去測量橫跨許多國家的一些變項，並且把這些國家的獨特特質轉化成爲變項。我們之前曾經討論過的兩項研究，正好符合此種類型的比較研究：其中之一是撒頓（Sutton）對於十五個國家中的監禁率所進行的研究，而另外一項研究，

個案研究的比較研究　一種比較研究，在此種研究中，研究者會去深入地比較一個或者是兩個特殊的文化（或者像是地區之類的文化單位）。

文化脈絡研究　一種比較研究的類型，其焦點在於對少數的社會或是文化進行比較，而這些社會或是文化代表了各種理論類型。如此一來，此種研究就允許對其他屬於同樣類型的社會進行推論。

跨國研究　一種比較研究的類型，此種研究檢視了橫跨許多國家的數個變項的資料（通常是量化資料），並且對這些資料進行分析。

則是由曼德爾以及夏雷夫（Mandel and Shalev）針對十七個國家中的兩性收入差距所進行的研究。雖然目前世界上有許多獨立的國家，但是能夠取得完整資料的國家數目很少會超過 50 個，而且我們通常只會針對其中的少數國家來進行研究。

4. *多國研究*　所謂的**多國研究**（transnational research），就是在研究中使用了多國單位（multinational units）（例如，在全球的某個區域，像是第三世界之類的），而其分析單位是由許多國家所組成的不同的區塊，並且聚焦於這些分析單位之間的關係。此種研究把國家當作是一個國際系統的其中一部分，而不是孤立的實體。此種類型研究的範例之一，是華勒斯坦（Wallerstein, 1974）針對從 15 世紀以來的「世界體系」（"world system"）的長期發展所進行的研究。他的著作衍生出一個新的思想學派，也就是*世界體系理論*（*world system theory*），而此學派則激發了更多的歷史比較研究。

進行比較的單位

文化與國家　爲了方便起見，我們通常會把民族國家當成是分析單位。在思考全世界人們之間的區隔時，民族國家是最常被使用的主要單位。雖然民族國家在當前的時代中是一種屬於主流的分析單位，但是民族國家並不是一種必然的單位、也不是一種永恆的單位。事實上，民族國家在歷史上至今也只存在大約 300 年之久而已。

民族國家是一種由社會以及政治所定義的單位，所指的是一個政府對於某塊有人民居住的領土擁有主權或者是統治權（也就是擁有軍事方面的控制權以及政治方面的權威）。經濟關係（例如：貨幣、貿易等等）、運輸路線、以及通訊系統等等，都在領土界線之內被整合在一起。在領土中生活的人們，通常會共享一種共同的語言與習俗，並且通常會擁有共同的教育系統、法律系統、以及一組的政治符號（例如：國旗、國歌等等）。而政府則宣稱代表了在其控制的領土中所有人民的利益。

我們通常會使用民族國家來當作是文化的代用品，這是因爲文化比較難以去定義爲一個具體的、可以觀察的單位。*文化*所指的是在人們之間根據共享的社會關係、信仰、以及技術而形成的一種普遍認同。而在語言、習俗、傳統、以及規範之間的文化差異，通常是以國家的界線爲分隔點。事實上，共享一種共同的文化，正是在形成獨特的民族國家時的一個主要因素。

民族國家的界線與文化的界線之間可能並不是彼此完全契合的。在某些情境中，某個單一文化可能會跨越數個不同的國家，而在其他的情境中，某個民族國家裡可能包含了超過一個以上的文化。在過去的數個世紀裡，戰爭以及征服已經破壞或者是散播各種獨特而具有生命力的文化，並且也重新安排了不同文化之間的界線。帝國強權任意把土地和人民切割爲殖民地，而戰爭則從不同的民族和文化之中創造出新的國家。例如，歐洲的帝國就曾經對一些殖民地的國家任意強行畫定界線。[40] 同樣地，新的移民或者是少數民族可能無法跟某個國家裡的優勢文化同化。例如，在某個國家中的某個地區裡，其中的人民可能會具有獨特的民族背景、語言、風俗習慣、宗教、各種不同的社會機構

多國研究　一種比較研究的取向，該取向主要是去檢視並且比較各種不同的多國單位。

與制度、以及認同等等（例如，像是加拿大的魁北克省）。由於民族和文化認同是民族主義的基礎，因此，此種在同一個國家之內存在著不同文化的情況，就有可能會造成區域性的衝突。[41]

在進行比較研究時，民族國家並非總是最佳的單位。你應該要問的是：對於我的研究問題而言，什麼才是具有相關性的比較單位？是國家、文化、小型區域、或者是一個次文化？例如，假設你的研究問題如下：收入水準與離婚之間是否相關？（換句話說，收入較高的人們是否比較不容易離婚？）有一群擁有獨特文化、語言、以及宗教的人們住在某個國家中的某個地區。在這一群人當中，收入與離婚是沒有關聯的，然而，在這個國家中的其他地區裡，不但主流的文化與上述的地區不同，而且收入與離婚之間存在著關聯性。如果你把民族國家當作分析單位，那麼你的發現就可能是模稜兩可的，而且你所提出的解釋也會是相當薄弱的。如果不要先假定每一個民族國家中都擁有共同的文化，那麼你或許會發現到，一個小於民族國家的單位，可能會產生不同的結果。

對於許多目的而言，文化可能會比國家更具有相關性，但是我們很難去把文化或者是次文化之間的界線加以操作化。文化難以定義、經常在演變之中、而且也會和其他的文化彼此混合在一起。除了一些國界方面的爭議之外，兩國之間的界線通常比較明確，但是這些界線也可能會隨著時間的流逝而發生變化。因此，關於國家與文化的議題，目前仍沒有一個簡單直接的答案。該使用何種單位較爲合適的議題，仍然是一個重要的議題。

高爾頓問題　關於比較單位的議題，與另外一個以高爾頓爵士（Sir Francis Galton, 1822-1911）來命名的問題有所關聯。高爾頓爵士於 1889 年時，針對泰勒（E. B. Taylor）的一份報告，在皇家人類學院提出了一個議題。高爾頓問題所陳述的是，當我們在針對單位或者是其特徵進行比較時，我們應該要確定這些單位都是獨特的，而且彼此之間是各自獨立的。如果這些單位之間並無不同，而且事實上只是某個單一大型單位的次部分時，我們所發現的將只是一種虛假關係而已。

舉例而言，讓我們現在假設你檢視了 110 個單位，這些單位是各國裡的州與省、其中包括加拿大（10 個省）、法國（40 個省）、比利時（10 個省）、以及美國（50 個州）。你發現在說英語與使用美元當作貨幣、以及在說法語與使用歐元當作貨幣之間有著強烈的關聯性。顯而易見地，此種關聯性之所以存在，是因爲分析的單位（也就是州與省）都是更大單位（也就是國家）的次部分而已。這些分析單位之所以具有某些特徵，是因爲這些單位都是更大單位的一部分，而不是因爲這些特徵之間有任何的關係。社會地理學家也會遇上這個問題，因爲許多社會以及文化特徵都會散布在廣大的地理空間中在比較研究中，**高爾頓問題**（Galton's problem）是一個重要的議題，這是因爲文化幾乎沒有固定的界線。[42]

高爾頓問題　一種在多數個案或是單位之間發現相關或是關聯時所會遇到的潛在兩難困境，這是因為這些個案或單位是來自於一個單一的來源，而且這些單位（例如：國家、文化）並不是真正獨立的個案。

我們很難明確地說出某個文化終止於何處、而另一個文化則是起始於何處；某個文化是否與另一個文化截然不同；或者是某個文化的特徵是否隨著時間流逝而散布到另外一個文化當中。當在兩個不同單位中的兩個變項之間之所以存在著某種關係，事實上是因爲這兩個變項是來自同一個來源、而且這兩個變項並非是真正彼此互相獨立的單位時，就會發生高爾頓問題（請參閱圖 14.2）。

雖然高爾頓問題原本是來自於跨文化所進行的比較，但是也能夠應用到歷史比較上面。當研究者去詢問某些單位在不同的歷史時間中，是否真的完全相同或是不同時，就會發生高爾頓問題。例如，1885 年時的古巴，和在 2005 年時的古巴是否是完全相同的國家？在經歷過西班牙殖民主義結束之後的 120 年、美國影響力的崛起、成爲獨立國家、歷經獨裁統治、以及一場共產主義革命之後，是否已經徹底地改變了這個單位？

跨文化研究中的資料

比較田野／實地研究 許多比較研究者會把田野研究應用於他們本身以外的文化當中。人類學家爲了進行此種類型的研究而特別受過訓練以及作好準備。在某人自己的社會中進行田野研究、以及在另一個文化中進行田野研究，這兩者之間存在著一些小型的差異。在一個不同的文化裡進行田野研究通常會比較困難，而且也會對研究者有較高的要求。

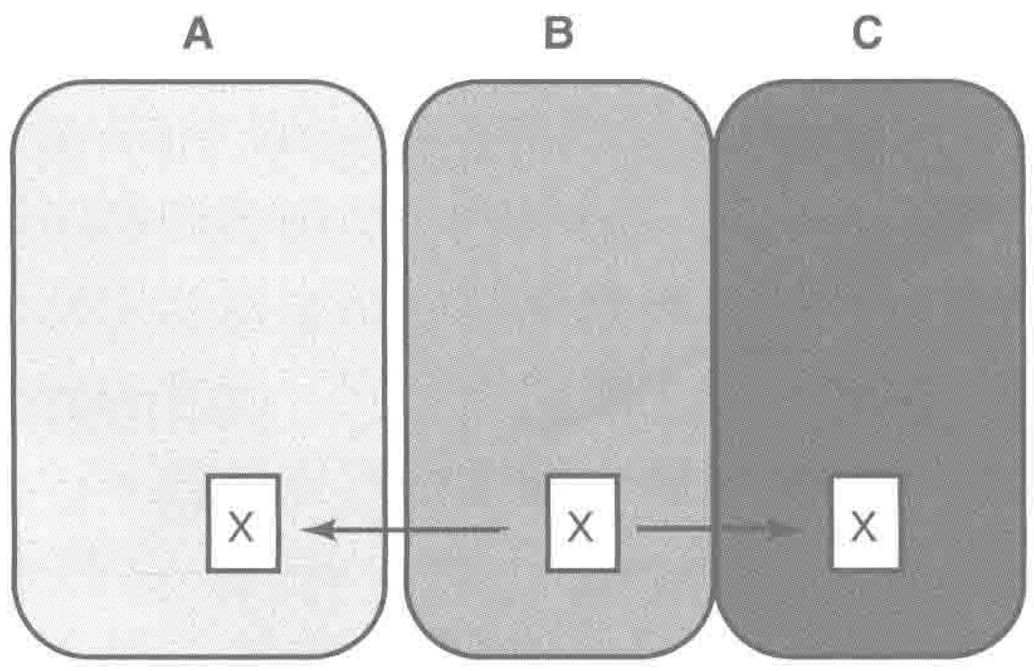


圖 14.2 高爾頓問題 當研究者在不同的環境或者是社會（以 A、B、C 來表示）中觀察到完全相同的社會關係（以 X 來表示），並且錯誤地認為此種社會關係是獨立地出現在這些不同的地方時，就會產生高爾頓問題。這名研究者可能會認為他或她已經在這三個獨立的個案中發現了一項關係，但是造成此種社會關係的真正原因，可能是因為這些個案都是來自於同一個來源，而這個來源已經從一個環境擴散到其他的環境中。此種狀況之所以會造成問題，是因為當研究者在不同的環境或者是分析單位（例如：社會）中發現一項關係（例如，某種婚姻形態）時，該名研究者可能會相信這項關係是獨立地出現在不同的單位之中的，而此種想法則可能會暗示這項關係是一種人類的普遍現象。如此一來，這名研究者或許會因而沒有察覺到，事實上此種關係存在的真正原因，是因為在不同單位中的人們共享了此項關係所致。

現存的質化資料　比較研究者也可以使用次級資料來進行研究。例如，假設有某位研究者想針對巴西、加拿大、以及日本等國的教育系統進行一項比較研究，那麼他或是她可以去閱讀由來自於許多國家（其中包括了巴西、加拿大、以及日本）的研究者所進行過的、描述在這三個國家中的教育系統的研究。

在整個人類的歷史過程中，曾經存在過的人類文化高達 5,000 個以上，而研究者到目前爲止已經研究過其中的 1,000 個文化。關於不同文化的民族誌資料，有一個珍貴的資料來源，叫作**人類關係區域檔案**（Human Relations Area Files, HRAF）以及相關的《民族誌地圖集》（*Ethnographic Atlas*）。[43] 人類關係區域檔案是一種獨特的田野研究報告的集合，而其中的研究報告是由人類學家莫達克（George Murdock）從 1938 年開始蒐集以及組織的。人類關係區域檔案蒐集了針對各種不同的文化所進行的民族誌研究的資訊，而這些文化大多數都是原始文化或者是小型的部落。人類關係區域檔案依據社會特徵以及習俗（例如：哺育嬰兒的方式、自殺、分娩等等）以將大約 300 個文化加以組織安排。針對某個特殊文化所進行的研究會被畫分成不同部分，而該項研究中針對某項特徵的資訊，會與其他研究中的類似資訊被分類在一起。此種作法使得研究者可以輕易地去比較許多文化中的相同特徵。例如，如果你對於繼承權感到興趣的話，你可以從人類關係區域檔案中得知，目前已經被研究過繼承權的文化數目已經高達 159 個，其中有 119 個文化所採行的是父系繼承的形式（也就是由父親傳給兒子）、27 個文化所採行的是母系繼承的形式（也就是由媽媽傳給女兒）、而在另外的 13 個文化中則是採取混合的形式。

我們可以利用人類關係區域檔案來研究在不同文化裡的數個特徵之間的關係。舉例而言，如果你想要瞭解對於女性的性侵害以及強暴行爲是否與父權（也就是由男性掌握權力以及權威）有關的話，你可以去檢視在許多文化當中性侵害行爲的發生情況以及父權的強度之間是否真的具有關聯性。

然而，使用人類關係區域檔案也會有一些限制。首先，原來的研究報告的品質，取決於原本的研究者待在田野中的時間長度、對於當地語言的熟悉程度、先前的經驗、以及研究報告的清楚性。除此之外，研究者所觀察的行爲範圍以及研究的深入程度也可能會有差異。例如，原本的研究者可能會說在 X 文化中，兒童不會受到處罰。事實上，在該文化中的兒童還是會受到處罰，只不過是在私下的場合接受處罰而已，因爲在該文化當中*公開*處罰兒童是一種禁忌。除此之外，在人類關係區域檔案當中，對於特徵的分類方式可能會是相當粗糙的。例如，在某個文化當中，巫術的重要性或許可以在一個從十分重要到不是非常重要的量表中來加以評分，但是在人類關係區域檔案中卻可能只有粗略地加以分類而已。

另外一個限制則是牽涉到人類關係區域檔案中所包括的文化範圍。在某些原始文化與外界有所接觸之前，西方的研究者所曾經接觸過、並且曾經進行過田野研究的原始文化的數量其實是相當有限的。而且這些曾經被研究過的文化，也不能當作是

人類關係區域檔案（HRAF）
一種廣泛分類並且全面性蒐集許多文化（大多數都是還沒有發展出文字的文化）的民族誌研究的集合資料，此種資料允許研究者橫跨許多文化單位來進行比較。

所有曾經存在過的人類文化的代表性樣本。除此之外，高爾頓問題（剛剛已經討論過了）也會是一個問題。

跨文化調查研究 我們已經在第十章當中討論過調查研究了，因此在這個部分裡，我們所要去檢視的，是當我們在不同的文化中使用調查技術時所會出現的議題。[44] 原則上，跨文化調查所會遇到的限制，與在單一文化中進行調查時所會遇上的限制沒有什麼不同。然而，由於在跨文化調查中，這些限制在強度以及嚴重性方面通常會比較大，因此我們必須仔細地去考慮，在某個情境當中，調查是否真的是最佳的方法。

當我們要在一個不同的文化當中進行調查研究時，我們必須要對該文化中的規範、日常行為、以及風俗習慣等等擁有深入的知識。若是沒有此種深入的知識，我們很容易會在程序以及詮釋方面犯下嚴重的錯誤。僅僅瞭解另外一種語言是不足夠的，我們還必須能夠融入多元文化當中，而且除了要熟悉調查方法之外，還必須要徹底地瞭解該文化。在進入另外一個文化、或者是在規畫進行調查研究之前，我們必須要先獲得對於該文化的實質知識。與另一個文化中的當地人密切合作，也是不可或缺的。一般而言，若想要進行一項良好的調查研究，深入地去瞭解某個文化是極度重要的。

例如，溫瑞伯（Weinreb, 2006）比較了在肯亞地區「陌生人」（也就是並非當地的、而是外來的研究者）與「內部人士」（也就是當地的研究者）所進行的調查訪談。他發現到，在內部人士以及陌生人所進行的訪談之間的最大區別特徵，就是訪談者是否認識受訪者以及受訪者的家人、或者是受訪者以及其家人是否認識這名訪談者。他發現到，至少在肯亞的鄉村地區裡，若想獲得居民的高度合作以及正確的資訊，以內部人士來進行訪談是絕對必要的。雖然能夠說當地的語言是很重要的，但是這並不一定總是足夠的。瞭解當地的習俗以及文化、並且與當地社區有一些社會聯結，也已經被證實具有同等的重要性。

當你在選擇要把哪些文化或是國家包含在一項跨文化的研究中時，你應該要同時依據實質的（substantive）（例如：理論的問題或者是研究問題）以及實際的（practical）理由來作出決定。你必須依據你將要進行研究的文化來量身打造調查研究的每一個步驟（問題的措辭方式、蒐集資料的方式、抽樣方式、以及訪談形式等等）。相當重要的議題之一，是來自於其他文化的人們是以何種方式來經歷調查的。在某些文化裡，調查以及訪談本身可能會是一種陌生的、令人害怕的經驗，就好像被警方審問一樣。

文化脈絡也可能會影響到調查的抽樣方法。你必須考量到是否能夠取得正確的抽樣架構、郵件或者是電話服務的品質、以及造訪偏遠地區的交通運輸方式。你也必須去瞭解其他可能會影響抽樣的因素，像是人們搬遷的頻率、人們居住的住宅類型、在住宅裡的人口數、或者是拒絕接受調查的一般比例。你必須依據文化來調整抽樣單位，並且要考量到某個文化對於像是家庭之類的基本單位的定義。除此之外，你也可能必須使用一些特殊樣本、或者是利用一些特殊方法來找出樣本中的人們。例如，埃爾德（Elder, 1973）曾經在印度進行過一項研究，他在研究報告中曾經提到，對於那

些住在佣人區（位於中高收入住宅的後面）的人們而言，他的抽樣數目是不足的。

雖然研究者們進行跨文化以及語言的調查已經有數十年之久了，但是一直到最進幾年，此種調查的出現頻率才變得比較高。除此之外，我們也仍然還在學習跨文化調查設計以及問題撰寫所牽涉到的複雜方法論問題。[45] 當調查研究中包含了兩個或者是多個文化時，那些在某人自己的文化中進行調查時所會遇上的問題或者是關心的議題，都會變得更加擴大（請參閱範例 14.7 在兩個文化中的調查研究）。直接的語言翻譯幾乎都是不足夠的，而且也會牽涉到許多關於同義的問題（本章的稍後部分會討論到同義的問題）。

與跨文化調查有關的五個議題，分別是文化敏感性、溝通方式、情境敏感性、禮貌偏誤、以及進入的權利。

1. *文化敏感性*　進行比較調查的研究者必須在文化方面具有極高的敏感性。在某個文化中沒有任何爭議性的主題，在另外一個文化當中卻有可能會是高度敏感的、甚至可能是禁忌。如同史密斯（Smith, 2004）曾經提到的，在不同文化中，主題敏感性的差異程度是非常大的。飲用酒類在回教文化中就比在猶太教與基督教文化中更爲敏感；在中國，關於共產黨的政治問題是被禁止的；同居的情況在瑞典相當普遍，但是在其他地方仍然是相當敏感的；在泰國，詢問有關國王的問題則是一種禁忌。除此之外，文化環境則暗示了其他的差異性。在日本，你可以詢問關於天皇的問題，而在英國，你則可以詢問關於女王的問題，但是在美國或者德國就不能提出類似的問題，因爲在這兩個國家中並沒有天皇或者是女王這種虛位元首的存在。在一些非洲以及回教的社會裡，你在詢問到某個男人的太太時，應該要以複數形態來詢問。你必須要瞭解地區性的差異。例如，在蘇格蘭的法律裡，一個人可能被發現有罪、尚未被證明有犯罪、或者是清白的。但是在英國的法律裡，一個人若不是有罪、就是清白的。如同柏朗（Braun, 2003）所強調的，受訪者會以他們自己已知的脈絡來詮釋問題，並且運用當地的知識、假定、以及詮釋來填補意義，即使當這個問題是來自於一個不同的文化語言環境時亦然。例如，對於「如果媽媽外出工作的話，是否會對孩子的發展不利？」這個問題的答案，可能會因爲文化的不同而出現極大的差異。一來，受訪者可能會把問題中所提到的孩子想像成是一個嬰兒或者是一個 7 歲的兒童。或許在他們的社會當中，媽媽們可以找到許多工時相當有彈性、而且離家近的兼差工作，而且在每個社區當中也提供了廣泛的、免費的、而且高品質的兒童照護體系。或者，也許在他們的社會裡，媽媽們必須在離家有一段距離的地方擔任全職的工作，而且兒童照護設施不但又少又昂貴，通常也離家很遠。在表面上看起來完全相同的問題，即使是以同樣的語言來表達，在不同的文化中也可能不會具有完全相同的意義。雖然澳洲、英國、加拿大、愛爾蘭、以及美國都使用了*社會安全*（*social security*）這個英文詞彙，但是在每個國家當中，這個詞彙的意義都不相同，而且所指的政府計畫也都各自不同。

2. *溝通方式*　溝通的型態以及方式也可能會因爲文化不同而有所差異。不論是在文化方面以及語言方面，說話的步調以及運用沉默的方式都會有所不同。社會期

範例 14.7

在兩個文化中的調查研究

席爾曼以及杜姆勞（Shearman and Dumlao, 2008）比較了在日本以及美國年輕成人之間的家庭溝通型態。他們總共調查了 173 位美國大學生以及 131 位日本大學生。作者們先雇用一些精通英語與日語的人來把問卷從英文翻譯成日文，然後他們又請人把問卷從日文翻譯回英文，以確認在翻譯之間的落差有多大。雖然過去的研究以及刻板印象都暗示了日本人是「群體導向」（"group-oriented"），但是作者卻發現到，日本的年輕人在群體導向以及集體性的程度上並未高於美國的大學生。作者們也發現在這兩個國家的家庭類型之間有很大的差異。美國的年輕成人報告說他們的家庭比較偏向共識導向（也就是成員之間達到共識或協議），而日本的年輕成人則報告說他們的家庭比較偏向自由放任式的（讓每個人去做他或她想做的事）。在某些方面，這些發現與一般廣泛接受的想法完全相反。作者們也提到了此項研究的一些限制（p. 206）：「我們利用國家當作是文化的代理量數，並且假定某個國家裡土生土長的個人自動地代表了一個相對一致的、普遍的文化。」除此之外，作者們也使用了之前在美國所建構的測量量表，並且將量表翻譯為日文。正如同作者們所提到的，此種作法在比較研究中也可能會是一個問題。「在這項研究中所包含的一種西方人的偏見，或許已經影響了我們的結果」（p. 206）。

望（請參閱第十章）、反應心向（請參閱第七章）、以及回答調查問題的公眾意願，都會隨著文化而差生極大的差異與變化。[46] 除了問題本身之外，反應的量表（例如：非常同意、不同意等等）在不同的語言和文化環境之間也會有很大的變化。

金恩等人（King et al., 2004）描述了一種方法，也就是利用簡短的小圖卡來固定在某個特定文化環境中的反應。例如，在詢問某個人的年紀是老年的或是中年的時候，其答案可能會因爲社會脈絡而有所不同。在某個社會裡，由於死亡率較高的緣故，中年所指的可能是大約 30 歲左右，而老年所指的可能是大約 45 歲左右，但是在另外一個社會裡，人們則可能認爲 50 歲只能算是中年，而 75 歲以上才能算是老年。相同地，與其去詢問某位受訪者他或她所擁有的政治影響力是「完全沒有、一點點、或者是很大」，金恩等人建議提供一組具體的故事或者是小圖卡，並且要求受訪者選擇符合自己答案的那一個故事或是圖卡。例如，在標準的李克特量表中，墨西哥人可能會認爲他們所擁有的政治影響力比中國人還要來得小。雖然墨西哥（人民可以用選票把執政黨趕下台）擁有一個比中國（一黨獨大、專制的國家）還要更爲開放而民主的政治系統，但是仍然可能會發生上述的情況。然而，當研究者提供了五張圖卡來詢問受訪者對於品質低劣的地區飲用水的反應時（這五張圖卡的內容分別是支持反對黨的候選人、發動請願、覺得投票毫無價值、向當地的領導人物請求協助、以及默默地忍受），墨西哥和中國的受訪者在回答問題時的方式，就能夠反映出他們真正的政治影響力。研究者能夠利用一些小圖卡，以統計的方式來將一般的反應加以標準化。巴拉尼克等人（Baranik et al., 2008）檢視了在美國以及七個其他國家中所使用的羅森伯格自尊量表（Rosenberg self-esteem scale），而且發現了系統性的文化差異。這些作者警告說（2008: 1891）：「在進行此種跨文化的研究時，研究者必須要能夠提出證據，以證明不同國家中的受訪者都能以合適且一致的方

式，來回應那些被用來測量構念的工具。」

3. *情境敏感性*　在進行訪談時，研究者必須精通或瞭解跨文化的情境。在選擇以及訓練訪談者的時候，必須依據所要進行研究的文化中的教育、規範、以及禮儀等等因素來作出決定。訪談的情境會引發一些議題，像是隱私的規範、取得信任的方式、關於機密的信念、以及方言之間的差異等等。例如，在某些文化裡，訪談者必須先花上一天的時間來進行非正式的談論，才能獲得像是溫瑞伯（Weinreb, 2006）在肯亞郊區所體驗到的、那種在進行簡短的正式訪談時所需要的融洽關係。

4. *禮貌偏誤*　比較研究者必須要察覺到某種版本的社會期望偏誤，也就是**禮貌偏誤**（courtesy bias）。當強大的文化規範使得受訪者去隱瞞任何不愉快的事情、或者是受訪者刻意去回答一些他或她認爲訪談者想要得到的答案時，就會發生禮貌偏誤。受訪者可能因爲文化規範的緣故，因而過度保守或者是過度誇張地回報了某些特質（例如：收入、成就、教育等等）。除此之外，受訪者提出答案的方式（例如，聲音的音調、情境等等），也可能會改變了答案的意義（請參閱資訊箱 14.5 對於調查問題的跨文化答案）。

5. *進入的權利*　在某些文化中，由於存在著一些限制開放性以及保護隱私的規範，因此，進入的權利就會是一個嚴重的議題。除此之外，你來自於另外一個國家或是文化這件事本身就可能會是一個明顯的障礙。讓當地的受訪者得知你來自於哪個國家，可能會讓他們打開心胸來接納你、但是也有可能會使你非常難以去與他們進行接觸。具體的問題可能包括了去瞭解應該要與什麼機構或是個人進行接觸、進行接觸時的適當程序（例如：一封正式的介紹信函）、如何維持善意（例如：贈送一些小禮物）、以及此種安排對於研究的品質以及可比較性所可能造成的影響。在某些文化當中，若是想要取得抽樣架構、進入鎮上的某些區域、或者是接觸到特定的受訪者時，可能必須要利用賄賂、家族人脈、或者是得到當地政治當局的許可才行。除此之外，在蒐集到資料之後，你可能還必須採取一些特別的預防措拖，以保護資料的機密性以及完整性。

調查問卷以及多種語言　大多數的研究者所採用的都是「詢問相同問題」的方法，換句話說，研究者們利用一個現有的問題，或者是先以某種語言提出一個問題，然後把這個問題翻譯成、或者是改編爲另外一種語言。另外一種替代性的作法，是從一開始的時候就以多種語言來設計出新的問題。在採用現有的問題以及創造出新問題這兩種作法之間，必須有一些權衡取捨（請參閱圖表 14.1），不過，一般而言，發展出新的問題是比較建議的作法，因爲如此一來的話，調查研究的品質會比較好。

改編問題在表面上看起來似乎直接了當而且不會改變意義，例如，像是研究者針對兒童受訪者來改編問題，而把胃（stomach）這個字改成肚子（tummy）一樣，只有做出一些微小的調整而已。然而，改編的作法有許多潛在的陷阱。理想上，翻譯者應該要非常熟悉問卷的設計以及你想要測量的

禮貌偏誤　當某個文化中存在著非常強大的文化規範，要求受訪者必須「保留面子」、或者是對其他人（包括社會研究者在內）隱瞞一些不愉快的資訊時，就會發生此種偏誤。

資訊箱 14.5　**調查問題的跨文化答案**

某句陳述或者是某一個問題的答案的意義，通常必須由某個文化中的習俗、社會情境、以及說出答案時的方式來加以決定。根據說出答案時的方式的不同，即使是完全相同的答案，其意義也有可能會有極大的差異性。

回答答案時的方式	受訪者對於問題的真正答案	
	是	否
	說出的答案	
禮貌的	「否」	「是」
強調的	「是」	「否」

資料來源：調查問題改編自 Fowler, *Survey Research Methods*, Sage Publications. 1992.

概念。在可見的文字之外，你必須去察覺到那些沒有被說出來的、或者是被暗示的事物。例如，像是「你能夠打開門嗎？」這樣一個簡單平常的問題，就可能會有多重的意義。依據語言以及脈絡的不同，這個問題可能是指你有打開門的能力，或者也有可能是一種請求。在許多語言當中，適當的性別（gender）也會是一個議題。德語中有三個性別，而法語、義大利語、以及西班牙語則有二個性別。在某些語言中，文法規則會指明說話者的性別，但是在其他語言當中，文法規則所指明的是某人正在談論的人物或者是事物的性別。例如，你或許不能夠在未指明性別的情況下去詢問關於某個「朋友」的事。在英文當中，「誰是你最好的朋友？」這個問題，其實是對說英語的人提出兩個問題，也就是「誰是你最好的男性朋友？」以及「誰是你最好的女性朋友？」而某一些職業，例如秘書，可能會強烈地與某種性別連結在一起，因此如果詢問他人該項工作是由男性或是女性來擔任的話，就會讓人覺得非常突兀。在某些語言當中，隱含在詞彙當中的除了性別之外，還有地位的意義。例如，在日語當中，當你在詢問關於一個人的兄弟或是姐妹的問題時，如果你不說明你所指的究竟是哥哥、弟弟、姐姐、或是妹妹的話，你根本就無法提出問題。

即使是表面上看起來完全相同的字，也可能會帶有不同的言外之意以及文化意義。例如，英文中的「*liberty*」以及法文中的「*liberté*」雖然可以翻譯爲相同的概念，但是由於兩國具有不同的歷史以及政治觀念脈絡，因此其真正的意義是大不相同的。語言以及文化的差異性可能存在於同一個國家裡。在美國，「*education*」這個英文字主要是指學術學科，而在美國境內說西班牙語的移民所使用的類似字「*educación*」，則包含了學習社會技巧的涵義在內，但是此種涵義通常不會被包含在「*education*」這個字的英文用法當中。

回譯法　先把書面資料翻譯為第二種語言，然後再找不同的翻譯者把該份資料翻譯回原本的語言，並且比較原來的版本以及經過翻譯的版本之間的異同。

在達到同義字方面，最爲普遍的比較研究技術是**回譯法**（back translation）。[47] 回譯法所指的是先把一個片語、問題、或是整份文件翻譯成一種外國語言，然後再找另外一個人把外國語言的版本再翻

圖表 14.1　跨文化調查問題設計

問題來源	優點	缺點
改編或者是採用現有的問題	容易、便宜、而且快速	通常會有語言以及文化適合性的問題
發展出新的問題	高度的文化適合性、幾乎沒有語言的問題	成本較為高昂、必須花費更多時間

*資料來源：*改編自 *Cross-Cultural Survey Methods*, Janet A. Harkness, Fons J. R. van de Vijver, Peter Ph. Mohler, 2003, p. 25. 轉載已獲得 John Wiley & Sons, Inc.同意。

譯回原本的語言。例如，你把一個片語先從英文翻譯成韓文，然後再另外找一個人把這個片語從韓文翻譯回英文。接著，你可以針對第一個和第二個英文版本進行比較，以找出任何不一致之處。不過，即使是使用回譯法，也仍然有其限制。有一些研究者曾經進行過一項比較研究，在研究中針對美國以及日本大學生對於國際議題的知識進行比較。研究者一開始是用英文來撰寫問卷，然後再請一組日本大學教授把該份問卷翻譯成爲日文版本，並且針對日文的問卷進行了一些調整。當這些研究者們後來把日文版本的問卷翻譯回英文時，他們發現了「三十個翻譯方面的錯誤，其中還包括了一些重大的錯誤」（Cogan et al., 1988: 285）。

當關於某個概念的文字並不存在於某種不同的語言當中時（例如，在北印度語當中沒有關於「信任」（trust）的說法、在土耳其語當中沒有關於「忠誠」（loyalty）的說法、而在泰語當中沒有關於「良性爭執」（good quarrel）的說法），回譯法就沒有什麼作用了。因此，如果遇到此種情況時，在翻譯時或許必須要加上複雜的解釋，或者是你根本就無法使用某些概念。

有越來越多的比較研究專家反對使用簡單的翻譯，包括回譯法在內，因爲此種作法無法找出所有的問題。哈克尼斯（Harkness, 2003: 42）注意到，若是把「*Das Leben in vollen Zügen geißen*」這句德文一字一句地照翻的話，在英文中就會變成「享受在擁擠的火車裡的生活」（"Enjoy life in full trains"）。回譯法並未掌握到這句話在英文中的真正意義是指「盡情享受生命」（在英式英文中是"Live life to the full"，而在美式英文中是"Live life to the fullest"）。之所以會發生此種錯誤，是因爲「*Zügen*」這個字有兩個意思，而其中之一正是指「火車」。因此，哈克尼斯（Harkness, 2003: 35）主張：「每當有可能的時候，翻譯應該要被整合到研究設計當中。然而，在實務上，翻譯很少會被視爲是問卷設計的一部分，而且通常會被當成是一種附加物而已。」現代的跨文化傳播與翻譯理論則建議研究者使用一種整合式的方法，也就是由一組翻譯家使用特殊的技術，像是認知訪談以及利用焦點團體等等，來檢查調查的用字遣詞，以討論調查問題的真正意義。你應該要避免犯下**遠征研究**（safari research）的錯誤，也就是一種由來自於某個文化中的一支研究隊伍在自己本身的文化語言脈絡下發展出一個計畫以

遠征研究　由來自於某個文化中的研究者所進行的研究，在此種研究中，他們把自己的觀點、想法、以及議題強加於另外一個文化之上，而且把被研究的文化當作只是一種可以研究的異國物體而已。

及測量方法，然後把這些測量方法強加在另一個文化之上的作法。而另一種替代性的作法，則是與多元文化研究團隊所進行的聯合研究。

回譯法只是數種順序方法的其中之一而已，此種方法是由利用某種語言所撰寫的一份問卷開始，而且此份問卷也是由單一的文化觀點所創造出來的。[48] 這份問卷接著會被翻譯成一種新的語言，以被運用於一個不同的文化環境當中（請參閱資訊箱 14.6 跨文化調查中所使用的翻譯方法，以及圖 14.3）。較新的方法包括了平行翻譯或者是同步調查發展。打從一開始的時候，一支由研究者和翻譯者所組成的隊伍，就以數種語言來撰寫問題，並且以在多重文化環境中使用的前提來設計問卷。此種新的方法很可能可以避免許多過去的問卷翻譯所遇到的問題。

西方文化的偏誤

大多數的社會研究都是由那些在由西方文化主導的少數先進社會中生活、工作、或者是接受過教育的人所進行的。此種情況造成了一種**西方文化的偏誤**（Western cultural bias）以及種族中心主義的危險性。正如同米爾達（Myrdal, 1973:89）所下的結論：「西方的取向必須被視爲是一種有偏誤的取向。」

每一個文化都有其本身的假定、思考模式、對於時間的觀念、以及關於人類生活的基本價值觀。而上述這些事物都會影響到人們的思考方式以及社會關係。如果社會研究能夠完全不受到文化的影響、或者是擁有一種與任何特定文化都不一樣的獨特專業文化，那麼文化的偏誤就不會是一個問題。但是，此種想法是不切實際的。我們很容易就會認爲，我們的假定、概念、發現、以及價值觀－這些事物其實都受到西方文化的影響－能夠普遍地適用於全世界的所有人身上。

比較的取向鼓勵你去提出一些能夠挑戰你自己本身文化傳統的問題。[49] 除此之外，由於採取了比較的觀點，因而使得你必需使自己的目光能夠超越外表的表象。外表的表象可能是更爲深層的信念、價值觀、以及關係所顯示出來的外在徵兆。藉由融合多種文化的方式，你能夠更清楚地察覺到在進行社會研究方面的問題以及西方文化的偏誤，並且能夠因此而提升社會研究的品質。

本末倒置的焦點 在美國之外的研究者、以及一些在美國境內的研究者，批評美國的社會科學是「本末倒置的」（"inverted"）或者是過於自我中心的。在第二次世界大戰過後，美國的理論以及進行社會研究的方法成爲世界的主流。許多美國人自大地假定這是理所當然的事，因爲他們的方法以及理論都是全世界最好的。由於美國在全世界的經濟、軍事、以及政治方面都居於領導地位，而且美國的流行文化以及商業文化在國際間也具有強大的力量，因此也使得上述的問題變得更加錯綜複雜。

西方文化的偏誤 當研究者使用了先進西方社會的看法以及觀點來進行比較研究時所作出的一種判斷。此種判斷通常會忽略了地區性的、非西方文化的議題、價值觀、或者是觀點。

不幸的是，許多研究者都是從他們自己本身的文化價值觀以及信念來看待各種關心的議題。他們只有引用一些在自己的祖國所進行的研究，並且刻

資訊箱 14.6　**跨文化調查中所使用的翻譯方法**

1. *快速翻譯（translation-on-the-fly）*　研究者以某種語言製作一份問卷，並且進行前導測試。之後，研究者雇用會說兩種語言的訪談者，並在訪談過程當中由訪談者直接翻譯每一個問題以及答案。
2. *單次翻譯（single translation）*　研究者以某種語言製作一份問卷，並且進行前導測試。之後由一位技巧嫻熟的翻譯者將問卷由原本的語言翻譯為目標語言。
3. *回譯法（back translation）*　研究者以某種語言製作一份問卷，並且進行前導測試。之後由一位翻譯者將問卷由原本的語言翻譯為目標語言，接著再由另外一位翻譯者把譯為目標語言之後的問卷再翻譯回原本的語言。最後再由研究者對原本的問卷以及翻譯回來的問卷進行比較，並且針對發現的問題進行修正或是調整。
4. *平行翻譯（parallel translation）*　由一個多元文化的團體製作出一份問卷，並且由兩位或是多位翻譯者獨立地翻譯該份問卷。每位翻譯者都在譯稿中併入了來自於多種目標文化語言的資料。之後，這些問卷會根據由一支多元文化／多元語言的團隊在經過審閱以及判斷之後的結果，來進行協調以及調整。接下來，這些不同語言版本的問卷會在多重文化語言脈絡中進行前導測試。
5. *同步發展（simultaneous development）*　一組由來自於不同文化以及語言的翻譯者和研究者所組成的團隊一起分工合作，討論各種概念以及詞彙在數種不同語言中的意義以及可能的替代翻譯方式。問卷中的題目是為了多重的語言／文化脈絡而創造出來的，透過一種對於自觀的／他觀的之間的區別相當敏感的去中心化過程，並且使用了具有文化特殊性以及多重文化的想法與詞彙。一支獨立的多元文化／多元語言團體接著會以多種語言去審閱並且評判這些問卷。最後，這些問卷會針對雙語文化以及單語文化的受訪者進行前導測試。

*資料來源：*請參閱 Smith (2004), Harkness, van de Vijver, and Johnson (2003), Harkness, Pennell, and SchouaGlusberg (2004)。

意忽視了與世界其他地方切身相關的議題，或者是忽略了由世界上其他地方的學者所進行的研究（請參閱 Connell, 1990）。

此種「本末倒置」的焦點造成了兩個問題。首先，全世界實徵社會科學研究的一大部分都是由美國的大量研究機構所產出的。而本末倒置的焦點會扭曲了我們關於全世界的社會關係的理解，因爲此種焦點在顯示出情況或者是事件時，只有透過美國文化的觀點而已。此種作法可能會對大型的全球議題、以及在不同文化環境中的人們所關心的問題產生錯誤的詮釋或者是誤解。其次，許多的社會研究者，尤其是那些採取實證主義取向的研究者，相信他們正在找出人類行爲的普世律則。如果他們只有依據單一文化的觀點來形成研究問題、只有研究單一文化、並且引用了那些只有在單一文化中進行的研究時，其研究結果怎麼能夠放諸四海皆準？這些研究者所建立的是一種單一國家的社會科學，而不是全球性的社會科學。因此，許多美國的研究者無法超越其國家脈絡而作出類推，除非是當他們可以利用比較研究來證明他們的作法是合理的時候，而且他們也必須盡力避免受到種族中心主義的影響。[50]

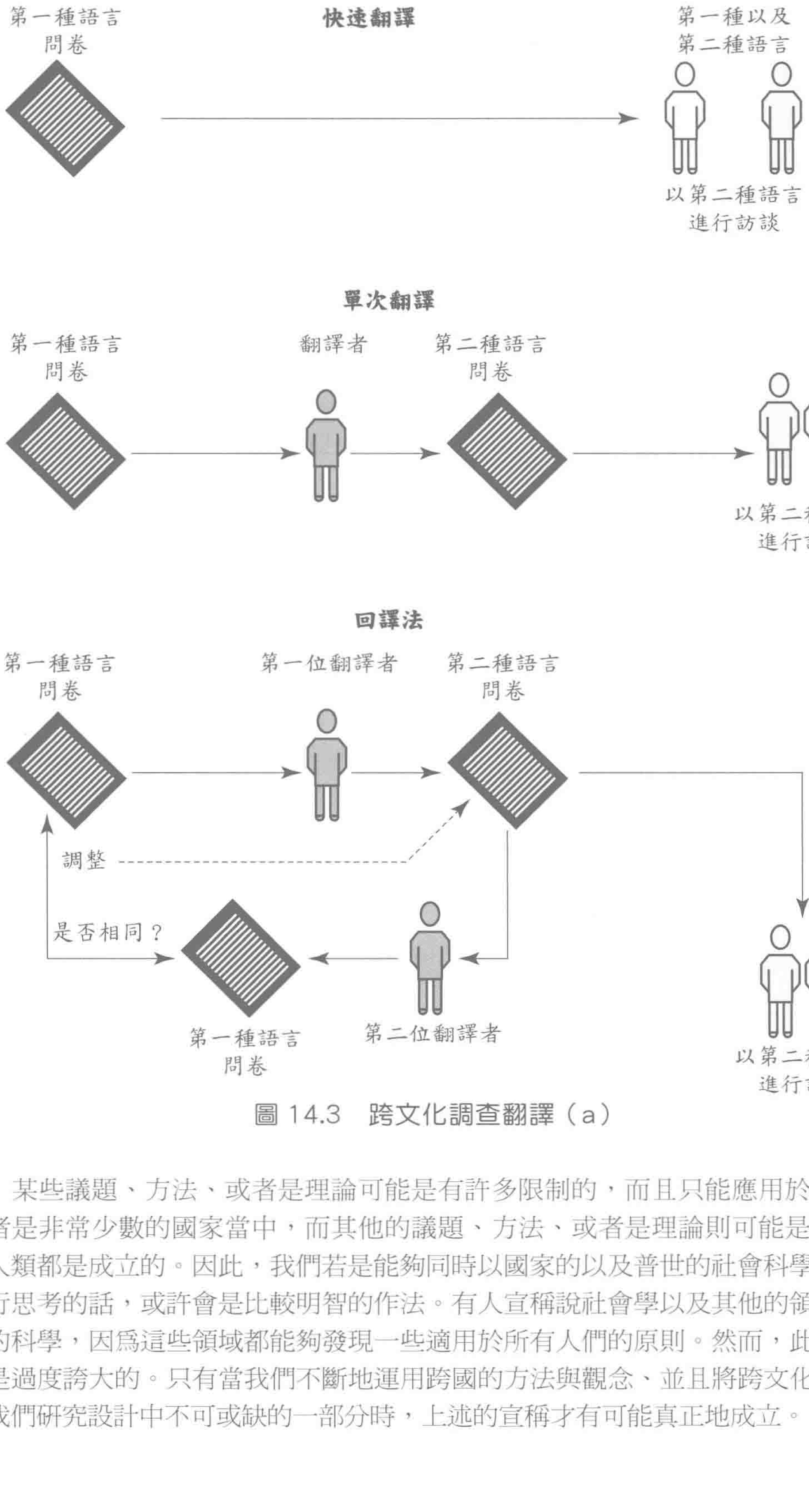


圖 14.3 跨文化調查翻譯（a）

某些議題、方法、或者是理論可能是有許多限制的，而且只能應用於某個國家或者是非常少數的國家當中，而其他的議題、方法、或者是理論則可能是對於所有的人類都是成立的。因此，我們若是能夠同時以國家的以及普世的社會科學的觀點來進行思考的話，或許會是比較明智的作法。有人宣稱說社會學以及其他的領域都是普世的科學，因爲這些領域都能夠發現一些適用於所有人們的原則。然而，此種宣稱通常是過度誇大的。只有當我們不斷地運用跨國的方法與觀念、並且將跨文化比較當作是我們研究設計中不可或缺的一部分時，上述的宣稱才有可能真正地成立。

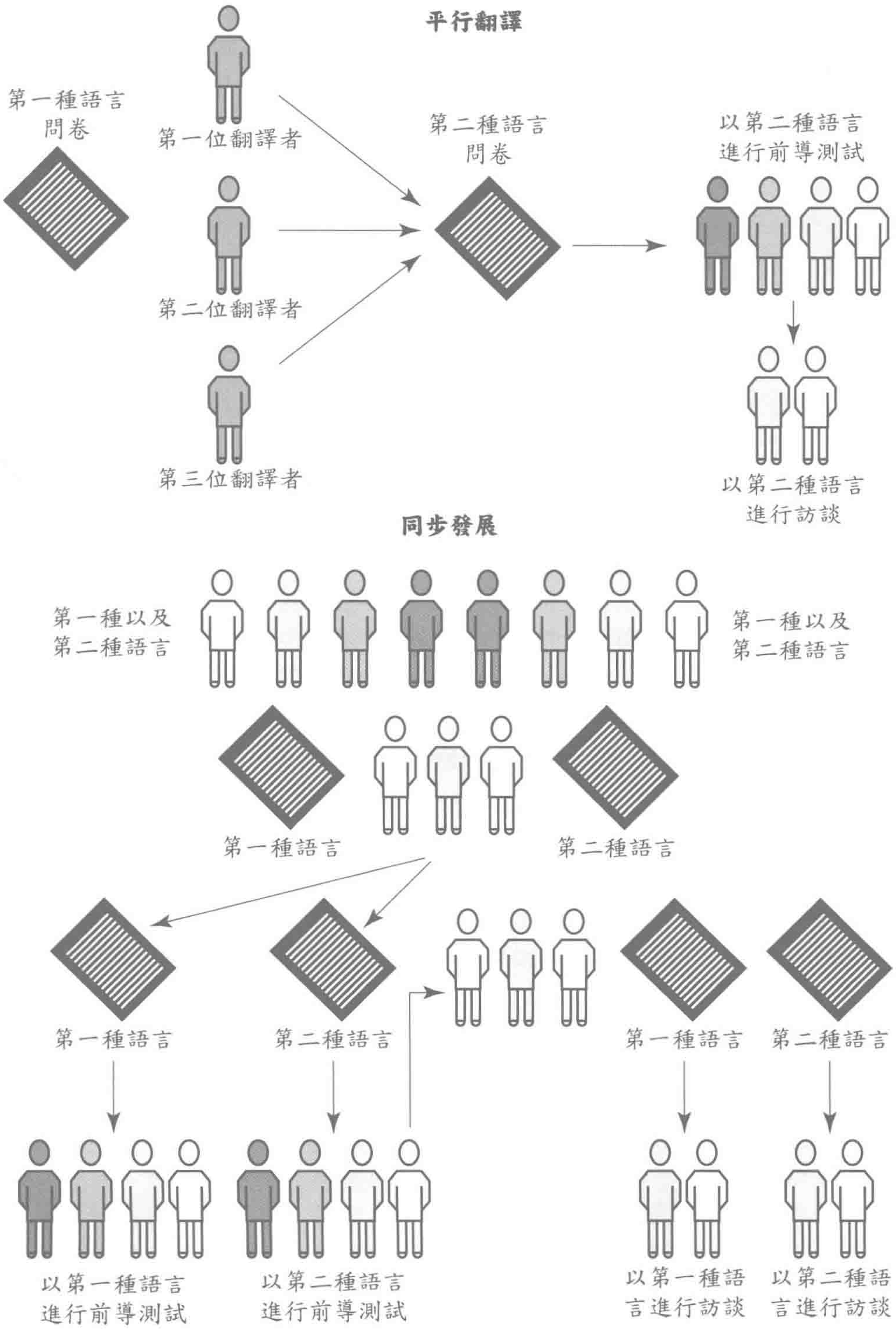


圖 14.3　跨文化調查翻譯（b）

自觀的　在某個特定的文化環境中所發展出來或者是被使用的概念或是方法，但是此種概念或方法可能並不存在於其他地方。

他觀的　在國際科學社群中被廣泛接受與認同、或者是在多重的文化環境中被廣泛共享的概念或是方法。

人類學家會把自觀的以及他觀的概念與取向區隔開來。[52] 雖然這是一種過度簡化的二分法，卻能澄清一個重大的比較研究議題。**自觀的**（emic）取向在分析文化元素時，所依據的是當地人的外顯的或是內隱的類別以及意義體系，換句話說，這是一種「內部者」或者是「當地人」對於習俗或者是信念的詮釋。自觀的概念是在一個特定的文化環境中發展以及使用的，而且可能不會存在於其他地方。因此，自觀的概念是僅僅適合於內部者文化的看法以及理解的。相反地，**他觀的**（etic）取向在分析文化元素時，所依據的是一種文化中立的、比較的類別以及原則。此種取向是以一種去脈絡化的觀點來詮釋事件、習俗、或者是信念，並且聚焦於意義之上。他觀的概念可以橫跨許多文化而被人們廣泛地共享、或者是能夠在國際科學社群中被廣泛地接受與認同。

比較社會研究者承認、並且也會同時應用自觀以及他觀的取向。這些研究者們會試圖從某一個特定文化環境的內部者的觀點來瞭解以及觀察世界，但是同時也會試著超越這個特定的環境，以建立橫跨各種不同文化環境之間的連結，並且在這些不同的文化環境之間進行溝通。全世界各地都存在著種族的差異，但是種族已經逐漸地被視爲是一種自觀的概念。例如，在波多黎各地區存在著許多種族特徵，而且這些特徵與美國其他地區都不相同。葛拉夫利（Gravlee, 2005: 965）曾經針對這些種族特徵進行過詳細的研究，並且認爲：「種族是一種和文化緊密相關的概念，此種概念預設了一組根植於美國經驗的意義……如果把種族的概念當作一種分析的架構，並且將其強加於一個種族並非明顯的自觀構念的社會當中的話，就會造成測量以及詮釋方面的問題。」

還有許多其他自觀取向的例子。也許你想要研究在日本與美國的工會組織。你知道在日文當中指稱工會的詞彙是「rodo kumiai」（譯註：這個詞彙相當於英文當中的 worker union，意即「工會」），然而，正如同摩爾以及河西（Mouer and Kawanishi, 2005: xvii）所主張的：

> 在關於日本勞動工作的著作當中，有許多概念都具有文化的自觀面向（也就是一組專屬於日本環境的特定涵意）。例如，「rodo kumiai」這個詞彙所指的是一種為了勞工、或者是由勞工所組成的組織，而此處的勞工是指在一般的或者是他觀的意義上的勞工。然而，當這個詞彙被用於日本時，對於大多數的日本人而言，這個詞彙所具有的言外之意，暗示了一種特定形式的工會組織、在組成工會比例方面的趨勢、在左翼與右翼團體之間關於意識型態的鬥爭的歷史，而且也會讓人聯想到範圍廣泛的各種公民運動、以及在國家以及地方層次上的特定組織形式。這個詞彙也被用來指稱某些其他自助的或是互助的組織，其中包括了信用合作社以及農業、消費者、或者是保險的合作社。

歷史比較研究中的同義議題

同義的重要性

在所有的研究當中，*同義*（*equivalence*）都是一個重要的議題。[53] 同義的議題牽涉到我們是否能夠橫跨各種不同的脈絡來進行比較。雖然我們活在一個特定的時代以及文化當中，但是我們是否能夠去正確地解讀、瞭解、或者是概念化那些關於一個不同的歷史時代或是文化中的人們的資料？若是沒有同義的話，我們就無法在不同的文化或是歷史時期當中使用完全相同的概念或者是量數了。此種情況很類似於在量化研究中關於測量效度的問題。

同義議題會在一條連續線上變化。在其中的一個極端，我們會發現某些對我們的經驗而言是完全陌生的事物（例如：獵人頭以及吃人的行爲）、或者是某個特定時間以及文化所特有的事物。而在另外一個極端則是一些我們可能會輕易忽略的微妙差異，但是這些差異仍然會對比較造成影響。例如，當埃爾德（Elder, 1973: 127）把「朋友」這個詞彙翻譯成三種不同的歐洲語言時，他就注意到了這個問題：

> 英文中的 friend，在德文中是 Freund，而在西班牙文中則是 amigo。就技術上而言，這三個字翻譯出來的意義應該是完全相同的。然而，德文中的 Freund 所指的是少數深入交往的、私密的朋友；而英文中的 friend 所指的是比較不是那麼密切的、範圍比較廣泛的泛泛之交；至於西班牙文中的 amigo 所指的則是範圍非常廣泛的人們，其中有些人可能是你只有見過一面的人。因此，像是「你有多少個朋友？」這個問題，在這三種語言中所問的是不同的事物。

同義的類型

由於我們可能會誤解或者是錯誤地詮釋在一個不同時代或者是不同文化中的事件，因此，同義的議題對於歷史比較研究而言是相當重要的。假定（自觀的）詮釋是正確的，我們可能會發現到，要橫跨不同的時間或者是地點（他觀的）來進行比較是相當困難的。如果我們完全地掌握了另外一個時間與文化（自觀的），我們仍然可能會發現到，要與那些來自於我們自己本身的時代與文化（自觀的）的人們進行溝通也是一件困難的事。我們可以把同義的議題分類成四種次類型：字彙的同義、脈絡的同義、概念的同義、以及測量的同義（請參閱重點摘要 14.1 同義的類型）。

字彙的同義　**字彙的同義**（lexicon equivalence）所指的是字彙或是片語的正確翻譯，或者是發現某個字所指的意思和另外一個字完全相同。此種情況在兩種不同的語言之間是最爲清楚可見的。例如，許多語言以及文化都有對於密友（intimates）（例如：

字彙的同義　在一個不同的語言中、或者是在把一種語言翻譯為另外一種語言時，可以表達出相同意義的字彙或是片語。

重點摘要 14.1

同義的類別

1. **字彙的同義** 在兩種文化之中是否存在著相同的字彙、溝通的方式、意義、或者是措辭呢？如果沒有的話，我們如何知道我們所發現的差異究竟是來自於溝通或者是翻譯方面的錯誤、或者是真的存在著差異？
2. **脈絡的同義** 在每一個文化環境中的脈絡，是否影響了我們所研究的事物的實質意義？如果真是如此的話，我們如何知道我們所發現的差異究竟是由文化脈絡的差異所造成的、或者是真的存在著差異？
3. **概念的同義** 在兩個文化之中是否存在著相同的概念，或者是在瞭解社會關係、情境、或者是事件時有相同的方式？如果沒有的話，我們如何決定我們所發現的差異究竟是由於概念上的錯誤配對所造成的、或者是真的存在著差異？
4. **測量的同義** 我們可以用相同的方法來測量在兩個文化中的相同現象嗎？如果不可以的話，我們如何知道我們所發現的差異是否是來自於不同的測量方法所造成的、或者是真的存在著差異？

親密的朋友以及家人）以及部屬（subordinates）（例如：年紀較輕而且地位較低的人）的稱呼以及代名詞，而這些稱呼或者是代名詞，與使用於陌生的公共環境或者是用於社會地位較高的人身上的稱呼或是代名詞是完全不同的。雖然關於親密的個人關係與公共關係的觀念確實存在於英語的文化當中，但是在英語當中卻沒有直接對等的語言形式。在像是法語、義大利語、或者是西班牙語等語言當中，當你在說「你今天好嗎？」時變換了代名詞，可能代表了地位或者是社會關係方面的變化。如果是使用英語的話，說話者可能就必須以另外一種可能是非言語的方式，才能表達出這些變化。

在歷史比較研究中，字彙的同義可能會具有極大的重要性，因爲即使是在完全相同的語言當中，字彙的意義也會隨著時間的流逝而改變。時間的距離越遠，同樣一個詞彙產生不同意義及涵意的可能性也就越大。例如，*weed* 這個字在現代所指的是不受歡迎的植物或大麻，但是在莎士比亞的時代裡，這個字所指的卻是布料。

當你在試著瞭解其他的觀點時，對於語言的微妙變化保持敏感性，或許是極度重要的。舉例而言，塞威爾（Sewell, 1980）就發現到，生活在一個半世紀以前的人們在使用某些詞彙方面的差異，幫助他瞭解到這些人們在意識以及社會經驗方面的變化。然而，瓊斯（Jones, 1983: 24）也提到：「利用來自於某些語言理論的基本洞見，以瞭解實質歷史詮釋中的問題，此種作法目前仍然處於……一種極度原始的狀態。」

脈絡的同義 把詞彙或者是概念正確地應用於不同的社會或者是歷史脈絡中，就被稱作是**脈絡的同義**（contextual equivalence）。這是一種嘗試著在特定的脈絡中去達到同義的作法。例如，在一些有不同的主流宗教的文化中，一位宗教的領袖（例如：神父、牧師、猶太教的祭司等等）可能會有不同的角色、訓練、以及權威。在某些脈

脈絡的同義 橫跨不同的文化或者是歷史時期，在社會角色、規範、或者是情境方面的相似性。

絡之中，神父是全職的男性專業人員，而且是富裕的、受到高度尊重的、受過良好教育的社區領導者，並且也擁有政治力量。而在其他的脈絡中，神父可能只是任何一位暫時在集會中起身帶領大家的人，但是在社區中沒有任何的力量或是地位。在此種脈絡中的神父可能比較沒有接受過良好的教育、收入很低、而且可能被視爲是誠摯而虔誠的人物。如果你在沒有注意到脈絡的情況下去詢問關於「神父」的事情，你可能會在詮釋時犯下嚴重的錯誤。

脈絡也可以應用於各個不同的歷史時期。例如，*上大學*這件事在今日的意義，完全不同於在另外一個歷史脈絡中的意義，因爲在後者那個時期當中，只有在人口中最富裕的前百分之一的人才能上大學；大多數大學的學生人數不超過 500 人；所有的大學都是私人機構，只招收男性學生，而且在入學時並不要求具備中學的文憑；除此之外，大學的課程內容主要是古典語言以及道德訓練。在 120 年前，上大學的意義和今日並不相同，因爲現在有絕大多數的高中畢業生都可以就讀於男女合校的公立大學、大學提供了範圍廣泛的科目、而且大學也被視爲是中等學校教育的延伸。因此，歷史脈絡已經改變了*上大學*這件事的意義。

概念的同義　能夠在不同的文化或是歷史時期中使用相同概念的能力，就被稱作是**概念的同義**（conceptual equivalence）。我們生活在特定的文化以及歷史時期當中，我們的概念來自於我們的經驗，而且會受到來自於我們本身的文化以及歷史時期的知識所影響。我們可能會藉由學習其他文化或者是歷史時期的方式，來試圖延伸我們的概念，但是我們對於其他文化或是歷史時期的看法，仍然會受到我們目前的生活情境所影響。此種情況會造成一種持續的緊張對立關係，而且也引起一項特定的議題，也就是我們是否能夠創造出一些概念，這些概念不僅能夠同時真正地反映出在不同文化中的生活經驗，而且也對我們具有意義。

概念的同義是更大的自觀與他觀議題的一個特殊個案，因爲概念在不同的歷史時期或是文化中可能會是不相容的。我們有沒有可能創造出一些概念，能夠真實地、正確地、而且有效地代表在兩個或是多個非常不同的文化或歷史環境中的社會生活？例如，湯普森（Thompson, 1967）就主張，在工業革命之前的時期當中，人們對於時間的主觀經驗以及測量方式是與現代非常不同的。在那個時期當中，遵守時間以及工作日的概念，具有非常不同的意義、或者根本就完全不存在。如果有某位研究者想去比較二十世紀晚期以及十六世紀早期的工作概念，那根本就像是在「拿蘋果跟橘子來比較」。例如，雖然*階級*這個字存在於許多社會當中，但是在不同的社會當中，階級的體系（也就是收入、財富、工作、教育、地位、以及與生產工具的關係等等事物所扮演的角色）、階級的數目、身處某個特定階級所代表的涵意、以及階級類別和界線等等都是非常不同的。而此種情形也使得研究者難以去針對橫跨數個社會中的社會階級來進行研究。[54]

有些時候，完全相同或者是類似的概念會存在於數個文化當中，但是卻有不同的形式或者是不同的強度。例如，在許多亞洲社會當中，在外在的、

概念的同義　在不同的文化或是歷史環境中，某些觀念或是概念的相似性。

公共的自我呈現與定義，以及私人的、個人的自我呈現與定義之間，有著非常明顯的區別。一個人的外在所透露與表現出來的感覺，通常在文化方面是與其真正的、內在的感受截然不同的。有些語言也會以語言學的方式來標示出此種差異。在西方文化中也存在著一種在公共的、非家庭的、或者是非私人情境中的自我的觀念，但是與日本以及亞洲其他國家相較之下，此種觀念是非常薄弱的、而且也比較不具有社會重要性與意義。除此之外，許多西方文化假定內在的自我是「真實的」，而且也應該被揭露出來，然而，此種假定卻不一定總是能夠被其他文化所接納與認同。

在其他時候，可能根本就沒有直接的文化同義。例如，對於日文中的「ie」（譯註：此字在日文中意指「家」），就沒有直接對應的西方概念同義字，而這個字的英文翻譯是家庭系統。然而，正如同亨德瑞（Hendry, 2003: 26）所解釋的：「整個『家庭系統』的概念，是當外界的影響力於一種比較的脈絡中試圖去解釋日本人的行爲時所創造出來的一種概念。」日文中的「ie」包括了可以追溯回到數個世代之前、而且持續不斷的家族血脈傳承，以及由不具繼承權的男性後代（或者是收養的兒子）所建立的家族分支不斷地將家族繁衍下去的概念。這個概念的意義比較接近於歐洲封建時期在貴族階層之間的「家族」血統，而比較不像是現代的家庭或者甚至是大家庭的概念。這個字也可以包含一種宗教的認同以及財產擁有權的面向（也就是代代相傳的某片土地或者是某種事業）。這個字還可以包含對於先人的義務感，其中包括了去維持先人們所曾經作出的任何承諾。「ie」這個字也深植於與其他「ie」所形成階層關係網絡之中，並且也暗示了在一個群體當中的社會位置或是地位。由於種種原因，當美國在第二次世界大戰之後暫時接管日本政府時，在法律上廢除了「ie」的概念。而藉由此種作法，美國試圖強加一種非常不同的、外來的文化概念於日本社會之上，而此種概念就是西方的核心家庭的概念。

概念的同義也能夠應用於那些針對不同歷史時期所進行的研究上面。例如，在一個社會上大多數人都沒有用現金交易的歷史時期當中，收入的概念與現代是非常不同的，因爲在當時的社會裡，大多數人們都是自己栽種食物，並且自己製造傢俱、衣物、或者是以物易物。在人們幾乎不使用任何金錢的情況下，把收入定義爲所賺得的金錢數目，基本上是沒有任何意義的。比較適當的作法是去計算一個人擁有多少頭豬、多少英畝的土地、多少雙鞋子、多少個僕人、以及多少部馬車之類的資產。相同的，在今日，即使是被認爲是貧窮的人們，也可能已經完成了 8 年的學校教育；或許擁有一部黑白電視；可能住在一間狹小而破舊的房子裡；而且可能擁有一輛生鏽的、破舊的、車齡 15 年的汽車。但是在過去的時代裡，貧窮可能意味著必須與動物一起睡在穀倉裡、在街上乞討、每天過著三餐不繼的生活、從未上過學、而且幾乎衣不蔽體。然而，雖然在物質方面有所差異，但是就特定的社會以及貧窮的概念而言，今日的窮人以及過去的窮人可能仍然會是同義的。

測量的同義 量數之間的相似性，此種量數能夠在不同的文化或者是歷史環境中去正確地代表一個構念或者是變項。

測量的同義 **測量的同義**（measurement equivalence）所指的是在不同的環境中使用同一個

概念的相等量數。就算是一名研究者發展出一個適用於不同脈絡的概念，仍然會遇到下面的這個問題：在不同的脈絡當中，是否有需要針對相同的概念使用不同的量數或是測量方式？阿爾莫（Armer, 1973: 52）將這個觀念定義如下：「關於測量的同義，所指的是，不論明顯的內容以及程序是否完全一致，在不同的社會中所使用的工具是否真的能夠測量一個完全相同的概念。」他認爲在不同的脈絡中或許必須使用不同的指標。爲了要測量完全相同的概念，我們可能在某一個文化當中使用態度調查的方式，但是在另外一個文化中則使用田野研究的方式。如此一來，真正重要的問題就變成我們是否能夠對來自於不同指標的結果進行比較。

測量同義的議題，暗示了我們必須去檢視許多部分證據的來源，才能夠去測量或者是確認一個理論構念。當證據是以片段的形式存在著時，我們就必須檢視大量的間接證據，才能夠去確認某些構念。在他對於十九世紀早期的法國作品所進行的研究中，塞威爾（Sewell, 1980: 9）曾經提到過此種確認的過程，他說：

> 我們所追尋的觀念，是部分而且片段的陳述、是在行動的過程中所寫下來的、通常是由一個不知名的人物或者是一群人所寫下來的、而且是只有在最為異質的形式中才能取得的——也就是像宣言、會議的辯論記錄、郵件、諷刺的印刷品、協會的規章、以及宣傳的小冊子等等之類的形式。在此種情境之下，思想的一致性並非在於特定的文本當中……而是在於整體意識型態的論述當中，此種論述是由大量的個別片段以及不完整的陳述、動作、意象以及行動所構成的。

倫理

在歷史比較研究中會遇上許多和在非反應式研究技術中相同的倫理議題。使用初級歷史資料的作法，偶爾會引起特殊的倫理議題。首先，依據初級資料來複製研究，是相當困難的。我們在使用證據以及文件的外部評論時的選擇標準，會對個別研究者的正直與誠實造成負擔。諾維克（Novick, 1988: 220）曾經建議：

> 歷史學家已經以第一手的方式看到大量堆積的證據，而且通常是未被出版的證據。歷史學家以自己多年沉浸於這些資料中的經驗—當然，還有他或是她本身的感知方式以及所抱持的假定—來發展出對於證據的詮釋。歷史學家會採取一些手段（註腳是其中最為明顯的例子），以使他們的作品達到某種接近「可複製性」（“replicability”）的狀態，但是其實此種相似性並不是那麼地完整。

在文件中的錯誤、或者是未能充分地以文件來記錄或是證明初級資料，都有可能會導致他人，尤其是來自於對立的史學學派的學者們，去指控歷史學家刻意造假。[55]

其次，保護隱私的權利，有可能會妨礙了蒐集證據的權利。某個人物的後代子孫可能想要破壞或是隱藏一些私人文件、或者是不名譽行爲的證據。而重要的政治

人物（例如：總統或者是最高層級的行政官員等等）則或許會想要隱藏一些令他們難堪的官方文件。[56]

在比較研究當中，我們必須對那些與跨文化互動有關的文化以及政治議題保持敏感性。我們必須瞭解在某個文化中被認爲是令人不悅的事物。敏感性意味著要對於另外一個國家的傳統、習俗、以及隱私的意義表現出尊重的態度。例如，在某些國家當中，一位男性若在一位已婚女性的丈夫不在場時對該位女性進行訪談的話，就是一種禁忌，而研究者則必須去尊重此種禁忌。

一般而言，我們應該要與進行研究的國家中的政府建立良好的關係。若是在沒有給予某些回報（例如，研究結果）的情況下就把資料帶出該國家的話，是一種不明智的作法。有些時候，你的祖國在軍事或是政治上的利益、或者是你自己個人的價值觀，可能會和你進行研究的國家中的官方政策有所衝突。如此一來，在你進行研究的國家裡的官員有可能會懷疑你是一名間諜，或者，你也有可能遭受到來自於祖國的壓力，要求你去藉機蒐集一些隱密的資訊。

有些時候，某位研究者的在場或者是研究發現，有可能會造成外交方面的問題。例如，假設有一位研究者在某個國家中檢視了墮胎的問題、並且宣布該國的官方政府政策其實是在強迫婦女墮胎，那麼此種情況就會引起嚴重的爭議。相同地，如果研究者對於某些反對政府的團體感到同情，那麼該國的政府或許會將該名研究者關入監獄、或者是驅逐出境。在其他國家進行研究的社會研究者，必須對此種議題以及他們自己的行動所可能產生的後果保持警覺。

許多來自於「富裕」而且高度工業化的國家、並且在「貧窮」而且正在工業化的國家中進行應用經濟發展研究的比較研究者，都會強調使用參與式的行動研究（請參閱第二章）的重要性。採取此種作法既是爲了政治方面的理由、也是爲了倫理方面的理由。發展研究者已經瞭解到，除非他們能夠與當地的人民密切合作、融入當地人民的普遍信念、並且獲得當地人民的配合，否則的話，應用的研究計畫就不可能會成功或者是持續下去。[57] 除了應用研究的常見議題、以及學習去尊重當地的習俗與文化之外，此種情況也意味著研究者必須去教導當地民眾關於研究設計的要素。當研究者在偏遠的、尚未工業化的環境中進行研究時，研究者可能需要以非常基本的、視覺化的方式，來詳細示範說明研究的邏輯，以及某項計畫（例如：使用新的食物儲存方法、消滅寄生蟲、改善飲用水的品質等等）對於當地生活條件所可能會造成的影響。

結論

在本章中，我們向你介紹了進行歷史以及比較研究時的方法論原則，以及此種研究在處理研究議題的一般方式。當你想要提出關於鉅觀層次變化的「大型問題」、或者是想要瞭解在各個不同的時期以及數個社會中運作的社會過程時，使用歷史比較研究是最爲合適的。你可以用數種方法來進行歷史比較研究，但是獨特的質化歷

史比較取向，其實在許多重要的面向上與田野研究的取向相當地類似。

歷史比較研究牽涉到一種對於研究的不同傾向，也就是不僅僅是去應用一組特殊的技術而已。你或許要用到一些特殊的技術，像是對於初級資料的外部評論或者是回譯法等等，然而，歷史比較研究最重要的特色，其實是在於你處理研究問題、探索資料、以及發展出解釋時的所使用的方式。

與其他既非歷史的、也非比較的研究相較之下，進行歷史比較研究的困難程度較高，但是，話說回來，在其他類型的社會研究當中也會遇到這些困難之處，只是困難程度比較低一些而已。例如，關於同義或是等值的議題其實存在於所有的社會研究之中。然而，在歷史比較研究當中，你卻不能夠把這些問題當成是次要的關心事項而等閒視之，因爲此種問題是你在進行歷史比較研究時所必須首先直接面對的問題，而且也會決定了你是否能夠成功地進行研究並且處理研究問題。

在下一章裡，我們會把焦點轉向質化資料的分析過程。許多歷史比較研究以及田野研究的類型，都會產出大量的質化資料。資料的形式、以及蒐集資料的研究取向背後的邏輯，對於資料的分析都會造成影響。

問題回顧

1. 歷史比較研究的一些獨樹一格的特質爲何？
2. 田野研究以及歷史比較研究之間的相似之處爲何？
3. 年鑑學派是什麼？年鑑學派在研究過去時的取向有哪三種特質或是詞彙？
4. 關鍵指標以及支持性證據之間的差異之處爲何？
5. 當研究者採用外部評論時，會提出什麼樣的問題？
6. 使用次級資料的限制是什麼？
7. 何謂高爾頓問題？爲什麼高爾頓問題在比較研究中很重要？
8. 在社會研究中使用比較方法的長處或是優點爲何？
9. 跨國調查研究與在某人自己本身的文化中所進行的調查研究之間有什麼不同之處？
10. 在歷史比較研究中，同義的重要性爲何？同義有哪四種類型？

注解

1. The early works include the following: Marc Bloch, Feudal Society, (1961) transl. L. A. Manyon (University of Chicago Press; original 1939–1940); George Homans (1941), English Villagers of the Thirteenth Century (Harvard University Press); Robert K. Merton (1970), Science, Technology and Society in Seventeenth Century England (Harper & Row; originally published in 1938); and Karl Polanyi (1957), The Great Transformation, Beacon.
2. They included Bellah's Tokugawa Religion (1957), Bendix's Work and Authority in Industry (1956), and Smelser's Social Change in the Industrial Revolution (1959).
3. Three such works are Tilly's The Vendee (1964), Moore's The Social Origins of Dictatorship and Democracy (1966), and Thompson's The Making of the English Working Class (1963). Tilly's study of France in the 1790s combined quantitative logic and new historical data. Moore's study of England, India,

Japan, Germany, the United States, and the former Soviet Union traced how combinations of events and coalitions of social groups caused some nations to develop democractic and others nondemocratic governments. Thompson's study of England prior to 1840 examined the lives, words, and actions of ordinary people. "The inspiration for a good deal of the new social history came from E. P. Thompson's (1963) The Making of the English Working Class. Surely, no work in European history ever so profoundly and so rapidly influenced American historians" (Novick, 1988:440).

4. Some influential works of this period include Anderson (1974a, 1974b), Hector (1975), Paige (1975), Skocpol (1979), Tilly et al. (1975), and Wallerstein (1974).

5. This is a large increase over the previous period (1985–1989) when it was about 28 percent. By contrast, the percentage of historical or comparative articles in the journals between 1976 and 1978 was about 18 percent. Additional information on the history of historical-comparative research can be found in Calhoun (1996), Johnson (1982), Kohn (1987, especially footnote 1), Lipset (1968), Novick (1988), Roy (1984), Skocpol (1984), D. Smith (1991), Warwick and Osherson (1973), and Zaret (1978).

6. See Mahoney (1999) for major works of historical-comparative research.

7. For a discussion of differences between generalizations and analysis across temporal units and cultural units, see Firebaugh (1980) and Smelser (1976).

8. See Calhoun (1996), McDaniel (1978), Przeworski and Teune (1970), and Stinchcombe (1978) for additional discussion.

9. Brown (1978), Johnson (1982), Lloyd (1986), McLennan (1981:66–71), and Nowak (1989) provided discussions of the relationship between positivist and nonpositivist approaches to historical-comparative research and the turn toward a realist philosophy of science. See Murphey (1973) with regard to historical research. For comparative research, see Hymes (1970) and Mehan (1973).

10. See Goldthrope (1991, 1997), Lieberson (1991), and Treiman (1977). Counterarguments are made by Burawoy (1977), Goldstone (1997), Mahoney (1999), and Rueschemeyer and Stephens (1997).

11. For more on borrowing from anthropology, see Biersack (1989), Desan (1989), Johnson (1982), Sewell (1980), Stone (1987), and Walters (1980).

12. See also Desan (1989), Griswold (1983), and Ryan (1989) for discussions of ritual and cultural symbolism.

13. Also see Carr (1961:35, 69), McDaniel (1978), Novick (1988:604), and Ragin (1987:164–166) on the dialogue metaphor.

14. For additional discussion, see Sewell (1987).

15. See Roth and Schluchter (1979:205).

16. See Kalberg (1994).

17. See George and Bennett (2005), Gerring (2007), Mahoney (2004) and Munck (2004) on process analysis or process tracing.

18. For an additional discussion of the penetration of surface events, see Bloch (1953:13), Lloyd (1986), McLennan (1981:42–44), and Sewell (1987).

19. See Naroll (1968) on the difficulties in creating distinctions. Also see Whiting (1968).

20. See the discussion on periodization in Chapter 15.

21. Bendix (1963), Przeworski and Teune (1970), and Smelser (1976) discuss transhistorical concepts.

22. For more on the Annales School, see Braudel (1980), Darnton (1978), Hunt (1989), Lloyd (1986), and McLennan (1981).

23. Orientation reading is discussed in Shafer (1980: 46–48).

24. See Wilkinson (2005).

25. Shafer (1980:2) discussed this in depth.

26. See Lowenthal (1985:187).

27. For additional information on oral history, see Dunaway and Baum (1984), Sitton et al. (1983), and P. Thompson (1978). Also see Prucha (1987:78–80) for a guide to major collections of oral histories in the United States.

28. See Schacter (1995) and Schacter (2001:20–21).

29.Bendix (1978:16) distinguished between the judgments of historians and the selections of sociologists.

30.Merton (1957:93–94) discussed the limitation of post facto interpretations.

31.For a discussion of law versus tendency in historical social theory, see Applebaum (1978b) and McLennan (1981:75). Murphey (1973:86) provided a useful discussion of the issues.

32.As used here, read, means to bring a theoretical framework and analytic purpose to the text. We read "past" specific details and the historian's interpretations (i.e., passed by, but not without notice) to discover patterns in underlying structures. See Sumner (1979) for discussion. This relates to the objectivity question in historiography in Novick (1988) and Winkler (1989).

33.Bonnell (1980:161), Finley (1977:132), and Goldthorpe (1977:189–190) discussed how historians use concepts. Selection in this context is discussed by Abrams (1982:194) and Ben-Yehuda (1983).

34.See Barzun and Graff (1970), Braudel (1980), Cantor and Schneider (1967), Novick (1988), or Shafer (1980). Most focus on the assembly of historical details that are documented in artifacts, including those of collective biography. This focuses attention on specific historical actors and their actions and motives. It takes on an individualistic-voluntaristic slant, and studies become ideographic accounts. See also Block (1977), Laslett (1980), and MacIver (1968).

35.Abbott (1992), Abell (2004), Gallie (1963), Gotham and Staples (1996), Griffin (1993), McLennan (1981:76–87), Runciman (1980), and Stone (1987:74–96) discuss the narrative.

36.For more on primary sources, see Barzun and Graff (1970:63–128), Cantor and Schneider (1967:22–91), Dibble (1963), Mariampolski and Hughes (1978), Milligan (1979), Platt (1981), Shafer (1980:127–170), and Topolski (1976). Bloch's (1953:79–137) discussion of historical criticism is still valuable today.

37.On the strengths and limitations of comparative research, see Anderson (1973), Holt and Turner (1970), Kohn (1987), Ragin (1987), Smelser (1976), Vallier (1971a, 1971b), Walton (1973), and Whiting (1968).

38.Bollen et al. (1993), Chase-Dunn (1989:309–333), and Ragin (1994a) provide lists of similar types. Also see Ragin (1989) for a critique of Kohn's typology.

39.See Ragin (1987:49–50).

40.See Wolf (1982) on civilizations around the world between 1400 and 1900. He noted many separate cultures and civilizations that existed prior to European colonization and the rise of nation-states.

41.For examples, see Hector (1975) and See (1986).

42.See Elder (1973) and Whiting (1968) on Galton's problem.

43.For more on the Human Relations Area Files and the Ethnographic Atlas, see Murdock (1967, 1971) and Whiting (1968).

44.For more on comparative survey research, see Burton and White (1987), Elder (1973), Frey (1970), Verba (1971), Warwick and Lininger (1975), and Williamson et al. (1982:315–319). For an additional discussion of access issues, see Armer (1973:59) and Form (1973).

45.See Harkness, van de Vijver, and Johnson (2003) and Smith (2003, 2004).

46.Couper and de Leeuw (2003) discuss nonresponse, and Johnson and van de Vijver (2003) describe social desirability issues.

47.For more on back translation, see Anderson (1973), Grimshaw (1973), and Hymes (1970).

48.See Harkness, van de Vijver, and Johnson (2003).

49.See Frey (1970), Grimshaw (1973), and McDaniel (1978).

50.See also Bradshaw and Wallace (1996).

51.See Hiller (1979).

52.See Ember (1977), Harris (1976), and Headland and associates (1990).

53.For additional discussions of equivalence, see Anderson (1973), Armer (1973), Frey (1970), Holt and Turner (1970), Przeworski and Teune (1970, 1973), and Warwick and Osherson (1973).

54.See Hazelrigg (1973).

55.See Novick (1988:612–622) for an extensive discussion of the David Abraham case.

56.For a discussion of archived data, see Odette and Mautner (2004) and Richardson and Godfrey (2003).

57.See Mikkelsen (1995).

第 15 章

質化資料分析

資料分析方法的比較
編碼以及概念形成
質化資料的分析策略
其他的技術
結論

在社會學中許多最棒的研究，都是在沒有統計考驗的情況下，利用質化方法來完成的。此種情況在許多研究領域之內都是成立的，其範圍包括了針對組織以及社區的研究，到對於面對面互動的微觀研究以及對於世界體系的鉅觀研究等等。而此類的研究也不應該被認為是較為薄弱的、或者是在處理這些主題時一開始的「探索性」取向而已。

——Randall Collins, "Statistics versus Words," p. 340

在田野研究、歷史比較研究、以及一些其他的研究領域裡，我們通常會去蒐集大量的質化資料，以描述關於在社會生活中的人們、行動、以及事件的細節。這些文本形式的質化資料是來自於文件、觀察筆記、開放式訪談的逐字稿、實體的人工製品、錄音帶或是錄影帶、以及影像或是相片等等。光是蒐集這些資料，其實是不足夠的，因爲我們還必須要去分析這些資料。在研究的質化取向裡，分析的過程在蒐集資料的時候就已經開始了，只不過此種分析通常是暫時性的、而且是不完整的。

要去分析資料，意味著我們必須以系統性的方式去組織、整合、以及檢視資料，而當我們在進行分析時，我們會去尋找那些存在於特定具體的細節當中的模式以及關係。爲了要進行分析，我們必須把特定的資料與概念銜接起來、發展出概論、並且辨識出廣泛的趨勢或者是主題。而在分析之後所得到的結果，則是能夠讓我們增進理解、擴展理論、並且深化知識。

在量化研究中所使用的資料，幾乎全部都是數字的形式。與矇矓不清的質化資料的龐大數量、多樣性、以及變化性相較之下，數字所承載的資訊是精確的、一致的、標準化的、而且簡潔的。在應用數學當中有一個大型的、高度發展的領域是專門在分析數字的（你在第十二章中已經讀過關於統計的相關內容了）。除此之外，我們用來分析量化社會科學資料的統計方法，與所有量化科學以及許多應用領域（例如：商業、教育、醫學、農業、以及工程學等等）中所使用的統計方法是完全相同的。由於電腦科技在過去四十年之間已經有了極大的進步，統計學家與電腦科學家已經發展出許多精密複雜的軟體、以及可以廣泛取得的程式，以協助人們進行量化資料的分析。

然而，數量龐大的統計知識以及相關的電腦軟體，卻幾乎完全無法應用於質化資料的分析當中。質化資料允許我們使用系統化的方式來進行分析，而且在邏輯上有一定的嚴謹性，但是，質化資料的分析方法與統計分析卻有許多不同之處。

一直到大約二十年之前，進行質化研究的研究者仍然很少會去解釋他們分析資料的方法。事實上，對於質化研究最常見的批評之一，就是質化資料的分析是不夠清楚的、而且也無法公開接受檢驗，但是，目前質化資料的分析已經變得越來越清楚以及具有系統性。[1]我們現在已經有質化資料分析的電腦軟體了，其中有些是以數學或者是其他的邏輯關係爲基礎而設計出來的。然而，我們在進行質化資料分析時，仍然還是會用到許多不同的方法。

本章包含了四個部分。我們首先會針對質化與量化資料分析的相似之處以及相異之處進行比較。接著，我們會討論如何使用編碼以及建立概念／理論的方式來協助分析質化資料。第三，我們會回顧質化研究者們曾經使用過的一些主要的分析策略，並且顯示出這些策略如何把資料與理論連結在一起。我們也會去檢視當我們缺乏直接、可觀察的證據時，此種情況在解釋中所扮演的角色。最後，我們則是會去回顧一些可以用來檢視質化資料當中的模式的特定技術。

資料分析方法的比較

質化形式以及量化形式的資料分析之間有著相似之處以及相異之處，而在這個部分裡，我們要來看看四個相似之處以及相異之處。

相似處

首先，在這兩種形式的資料分析中，我們都是以社會生活的實徵細節來進行推論。*進行推論*（*infer*）意味著我們必須以證據爲基礎而作出判斷、運用推理、並且得到結論。在兩種形式的資料分析中，我們必須謹慎小心地檢視實徵資訊，並且依據推理以及將資料中的複雜性加以簡化的方法來得到結論。此種過程需要運用一些抽象思考的能力，也就是與具體證據中的特定細節保持一定距離的能力，但是應該要與細節保持多遠的距離，則會因情況不同而有所變化。不論如何，我們都必須忠於原始的資料。

在這兩種形式的資料分析中，關於社會生活所作出的陳述，都是以具有「適當性」（adequacy）的調查爲基礎的。正如同摩爾斯（Morse, 1994: 230）所觀察到的：「在量化研究中，*適當性*指的是所蒐集到的資料數量，而在質化研究中，適當性所指的是主題的數目。當研究者已經蒐集到充分的資料而達到飽和時，就達到了適當性。」

第二個相似之處，則是兩種形式的資料分析都包含了公開的方法或是過程。由於我們蒐集了大量的資料，因此我們必須讓他人能夠理解我們的行動。我們會去描述蒐集以及研究資料與文件的方法，而且我們會把這些方法公諸於世，以接受科學社群中的其他成員的檢驗，而此種作法的標準化程度以及公開程度可能會有所變化。如同金

恩等人（King et al., 1994: 118）曾經提到過的：「雖然在質化研究中不一定會清楚明白地說明其研究設計，但是研究設計至少會內含在研究的每一個部分裡。」

第三，不論是質化分析或者是量化分析，在所有的資料分析當中，進行比較都是極爲重要的。我們會把我們所蒐集到的證據進行內部的比較、或者是與其他的相關證據進行比較。我們會去探索資料，並且辨識出在資料中的多重過程、原因、性質、或者是機制，並且尋找一些特定的模式，也就是相似處以及相異處，以及那些相同或者是相異的面向：

> （質化）研究者檢視了橫跨不同個案之間的相似性以及差異性的模式，並且盡力去接受多樣性的存在……雖然量化研究者也會去檢視不同個案之間的差異性，但是他們所強調的是不同的重點，因為他們的目的是去解釋某個變項與另一個變項之間的共變關係，而且通常是存在於許多個案之間的共變關係……一般而言，量化研究者對於個案只有廣泛而籠統的熟悉度而已。

第四，在這兩種資料分析的形式當中，我們都會極力去避免犯下錯誤、作出不正確的結論以及具有誤導性質的推論。我們會保持警戒之心，並且謹慎地留意可能的謬誤或者是假象。我們會仔細地去爬梳各種不同的解釋、討論、以及描述，並且去評估各種不同的描述方法以及解釋方法的優點。我們會在各種可能的描述或是解釋之中，去尋找最真實的、最有效的或者是最有價值的描述或是解釋。

相異處

量化研究者可以在一些特殊的、標準化的資料分析技術中進行選擇。不論是在自然科學或者是在社會科學中，假設考驗以及統計方法都是相當類似的。更確切地說，量化分析已經發展得很完備，而且也建立在大量的應用數學基礎之上。對比之下，質化資料分析的標準化程度較低。甚至可以說，有多少種質化研究方法，就有多少種資料分析的方法。而除了方法眾多之外，使得質化資料分析更爲複雜的是，質化研究通常是歸納式的。換句話說，當我們在開始進行一項研究計畫時，我們並不知道資料分析的具體細節爲何。夏茲曼以及史特勞斯（Schatzman and Strauss, 1973: 108）曾經說過：「量化分析者通常能夠預測他們本身的分析過程，但是質化分析者卻無法享受此種操作方面的優點，因此，他們就無法依據一開始內含在研究設計中的操作過程來修正以及安排其原始資料。」

第二個相異之處是，在量化研究中，我們會一直到蒐集資料的過程結束之後，才去開始進行資料分析的工作。只有到了那個時候，我們才會去操弄各項數字以找出模式或者是關係。但是，在質化研究中，當我們在蒐集資料時，我們就會開始去尋找模式或者是關係。我們會利用早期資料分析所得到的結果，來引導後續的資料蒐集過程。因此，與其說分析是研究過程中一個明確的最終階段，不如說是一個延伸到所有階段之中的研究面向。

另外一個相異之處，則是與社會理論之間的關係。量化研究牽涉到操弄一些代

表實徵事實的數字，以去考驗一些由變項構念所組成的假設。對比之下，質化研究經常藉由融合實徵證據以及抽象概念的方式，以創造出新的概念與理論。質化研究並不是要去考驗理論，相反地，我們可能會以說明證據或者是使證據更加豐富生動的方式，來顯示出某個理論、概化、或者是詮釋是可信的。

第四個相異之處，則是抽象的程度，或者是與社會生活的細節之間的距離（請參閱重點摘要 15.1 量化與質化資料分析的比較）。在所有的資料分析中，我們都會把特定的原始資料放入更爲廣泛的類別當中。接著，我們會去檢視以及操弄這些類別，以辨識出特定的模式。在量化分析中，我們是以統計、假設、以及變項等等方式來進行這個過程的。我們假定我們可以利用數字來捕捉以及進行測量，然後再以統計來操弄這些數字，以揭露出社會生活的關鍵特質。

對比之下，質化分析中的資料相對而言較不精確、較爲模糊、以脈絡爲基礎，而且可能有超過一種以上的意義。不過，此種情況並不必然是一項缺陷。

> 文字不僅僅是在智識方面更為基本，我們或許可以說，在社會科學的社會結構當中，文字必然會更勝於數學。這是因為作為一種表達的方式，文字具有更大的開放性、更多將各種不同論述以及經驗連結在一起的能力、以及接觸到更多知識份子的能力。（Collins, 1984: 353）

解釋以及質化資料

我們不必在嚴格的「個殊的／律則的」二分法當中被迫去作出二者擇一的抉擇，換句話說，我們不必在描述具體細節以及證明普世法則這二種作法當中去進行非此即彼的選擇。當我們在分析質化資料時，我們會發展出接近於具體資料以及脈絡的解釋或者是概論，而我們通常會使用比較不是那麼抽象的理論。我們或許會建立新

重點摘要 15.1

量化與質化資料分析的比較

相似處	相異處
兩者都是從實徵資料推論到抽象觀念	*量化*　使用一些共同的、標準化的技術。 *質化*　使用許多各自歧異的、非標準化的技術。
兩者都是使用公開的過程，並且會詳細描述分析過程	*量化*　在所有的資料都已經蒐集完畢之後才開始進行分析。 *質化*　當研究者仍在蒐集資料時，就開始進行資料分析。
兩者都會進行比較	*量化*　考驗先前就已經存在的理論以及假設。 *質化*　進行概念化並且建立一個新的理論。
兩者都會避免犯下錯誤或是作出不正確的結論	*量化*　使用精確而且簡潔的抽象資料。 *質化*　使用不精確的、分散的、相對具體的資料。

的理論，以創造出社會生活的寫實圖像並且深化關於社會生活的理論，而不是去考驗因果假設。我們所提出的解釋通常會有豐富的細節、對於脈絡很敏感、而且能夠顯現出社會生活的各種複雜過程或是次序。這些解釋可能是具有因果關係的、也可能不具有因果關係。我們的目標是將特定的具體細節組織成爲一幅連貫的圖像、模型、或是一組彼此緊密相關的概念。

質化的解釋若不是高度不可靠、就是高度可信的。我們提出支持性的證據，以把一些理論解釋排除在考量之外，並且增加其他解釋的可信度。質化分析可能會藉由顯示出某個解釋與大量證據相互矛盾的方式，以排除該解釋。資料可能會支持超過一個以上的解釋，但是並非所有的解釋都會與資料一致。除了排除較不可信的解釋之外，我們通常也會想要證明事件的順序或者是某個過程的各種步驟。此種時間方面的順序，正是在變項之間找到關聯性的基礎，而且也支持了因果關係的論點。

少數的質化研究者幾乎完全只採取描述性的解釋，而且避免進行理論性的分析。一般而言，我們應該總是清楚明白地說明理論以及概念。在沒有分析性的詮釋或者是理論的情況之下，質化研究的讀者們或許會使用一些他們本身在日常生活中的、認爲理所當然的概念，而此種概念幾乎無法促進一般知識的進展。除此之外，讀者們的常識框架也可能會包含了一些未經檢驗的假定、偏見、種族中心主義、以及一些來自於主流文化價值觀的定義模糊的概念。[2]

編碼以及概念形成

質化研究通常牽涉到使用一般觀念、主題、或者是概念來當作工具，以作出概化。其中有許多觀念、主題、或是概念等等都是非變項的概念，或者是簡單的名義尺度變項。

概念化

當我們在進行量化研究時，我們會將變項加以概念化，並且將概念加以詳細定義，以當作測量變項的一個步驟。對比之下，當我們在進行質化研究時，我們會形成新的概念，或者是去詳加定義一個紮根於資料的概念。概念形成是資料分析中--個不可或缺的部分，而且在蒐集資料的期間就開始了。因此，概念化是組織以及瞭解資料的一個方法。

進行質化研究的研究者在進行資料分析時，通常是把資料依據主題、概念、或者是類似的特質而加以分門別類。而在進行分類的過程當中，他們或許也會發展出新的概念、形成概念性定義、並且檢視概念之間的關係。到了最後，這些研究者將會依照順序而把概念彼此連結起來，以形成對立的群組（例如：X 與 Y 相反），或者是形成相似類別的群組，並且把這些群組交織於理論陳述當中。

當你在閱讀資料並且提出一些關於資料（例如，田野筆記、歷史文件、次級資料等等）的重要問題時，你或許已經在開始形成概念了。這些問題可以是來自於像

是社會學之類的學術領域裡的抽象詞彙，例如：這是不是一個階級衝突的案例？在那個情境當中，是否出現了角色衝突的狀況？這算是一種社會運動嗎？問題也可以是邏輯方面的，例如：事件的發生順序是什麼？事件在這兒發生的情形和在其他地方發生的情況之間有何異同之處？這些是相同的或是不同的個案、是一般的或者是特定的個案？[3]

概念以及證據之間是彼此相互依賴的，尤其是在個案研究的分析之中。個案並不是既定的、事先存在的實徵單位，也不是與資料無關的理論類別。更確切地來說，個案是由資料與理論所共同定義的。當你在組織資料以及應用觀念時，你就是在創造或者是詳細說明一個個案。當你把資料和理論連結在一起時，就會創造出一個個案，而此種過程又稱爲*個案化*（*casing*）（我們已經在第七章裡討論過個案化的觀念了）。決定把什麼事物當作個案的過程，能夠幫助你解決在下列兩種事物之間的緊張對立關係：其中之一是你實際觀察到的事物，而另外一個則是你對於觀察到的事物所抱持的想法。「個案化被視爲是一個方法論的步驟，而且可能會發生在研究過程的任何一個階段當中，但是特別容易出現在研究計畫的開始以及結束的階段裡」（Ragin, 1992b: 218）。

質化資料編碼

當你在對質化資料進行編碼時，你必須把變項的量數安排爲一種機器能夠閱讀的格式，以便於進行統計分析。資料編碼在質化研究中的意義，與在量化研究中的意義是不同的。在質化研究中，你會把原始資料組織成爲概念的類別，並且創造出主題或者是概念。質化資料編碼並不是一種資料管理的文書工作，而是資料分析中不可或缺的一個部分。你的研究問題會提供指引，但是分析的過程通常會引導你去發現新的問題。資料分析的過程可以讓你從原始資料盤根錯節的細節當中解放出來，並且鼓勵你從一個更高的層次去思考這些細節，以形成理論以及概化。

> 編碼是一種標籤或是貼紙，能夠把意義單位分派給在研究過程中編輯而成的描述性或者是推論性資訊。編碼通常會被指派到各種不同大小的「區塊」（“chunks”）上——文字、片語、句子或者是整個段落，而且可能與某個特定的背景有所關聯、也可能沒有關聯。（Miles and Huberman, 1994: 56）

史特勞斯（Strauss, 1987）定義了三種類型的質化資料編碼過程，並且建議研究者應該在三種場合中去回顧資料，而且每一次都使用不同的編碼。他（p. 55）警告說：「對於經驗不足的研究者來說，編碼是最難以理解、而且最難以精通的工作。」[4]

開放式編碼　對於質化資料所進行的初次編碼，此種編碼過程會去檢視資料，以把資料濃縮成為初步的分析類別或者是編碼。

1. *開放式編碼*　針對最近才剛剛蒐集到的資料，你首先要進行的是**開放式編碼**（open coding）。你先找出一些主題，並且指派一些初步的編碼，以作爲你把大量資料濃縮成各種類別的第一步。當你在緩慢地閱讀田野筆記、歷史資料、或者是其他的

資料時，你同時也在尋找一些關鍵性的詞彙、重要的人物、關鍵的事件、或者是主題。接著，你將會在筆記卡的邊緣或者是電腦記錄中寫下初步的概念或是標籤，並且用不同的顏色或是一些其他的特殊方法來突顯出這些概念或是標籤。對於創造出新的主題、以及在後續的分析中改變最初的編碼，你應該要抱持著開放的態度。如果你能夠以一種具有彈性的方式來應用理論架構的話，理論架構就能對你產生幫助。當你在使用開放式編碼時，你會讓深藏在資料中的主題浮現出來。此時這些主題只具有最低的抽象程度，而且是來自於你一開始的研究問題、文獻中的概念、或者是因爲沉浸在資料中而產生的一些新想法。正如同夏茲曼以及史特勞斯（Schatzman and Strauss, 1973: 121）所警告的，你應該要在具體的資料中看出抽象的概念，並且應該要在抽象概念以及具體細節之間來來回回。

在抽象概念以及具體細節之間來來回回的例子之一，是勒梅斯特斯（LeMasters, 1975）針對一家勞動階級的小酒館所進行過的一項田野研究。當時他發現到，在許多的對話當中，都會出現婚姻這個主題。如果他對田野筆記進行開放式編碼的話，他很有可能會把田野筆記中的一大部分都以「婚姻」這個主題來進行編碼。以下這個例子是一個假想的田野筆記，而且能夠以婚姻這個主題來進行開放式編碼：

> 當我在星期四到了酒吧時，脖子上還打著領帶，因為那一天我開會開到很晚。山姆（Sam）馬上就注意到我的領帶並且說：「該死，博士！我只戴過那玩意兒一次——在我結婚的時候——但是看看我現在變成什麼樣子！看在上帝的份上，下一次就將會是葬儀社的人幫我打上領帶了。」我點了一罐啤酒，然後問他：「你當初為什麼會想要結婚呢？」他面帶微笑、對我眨了眨眼，然後回答說：「要不然你還想要怎麼樣呢？你總不能一輩子一直在跟女人同居吧——當我還是單身的時候，確實和女人們同居過許多次。」他暫停了一下，又點了一罐啤酒，並且點了一支煙，然後接著說：「男人遲早都會想要擁有一個屬於自己的家庭和幾個孩子，而想要得到這些，你就得要結婚啊。這是沒辦法逃脫的——女人們會把你套牢的。」我說：「海倫（Helen）（這是山姆的太太）看起來是一個很不錯的人。」他回答道：「噢，該死，她是個還不錯的女孩，但是卻是個討厭的女人，而且這一點讓我很受不了。她常常讓我抓狂。每當我去參加一場派對，而且當我正要開始狂歡時，我太太就會說『我們回家吧！』」（改編自 LeMasters, 1975: 36-37）

歷史比較研究者也會使用開放式編碼。例如，我曾經研究過勞工騎士團（the Knights of Labor），這是一個十九世紀時在美國所發起的運動，其目的在於經濟以及政治改革。我閱讀了一份次級資料，其內容是關於在一個特定城鎮裡，該運動的地區分支所進行的活動。當我在閱讀以及作筆記時，我注意到禁酒黨（Prohibition Party）在當地的選舉當中具有舉足輕重的地位，而且當地分支的成員會激烈辯論關於禁酒（temperance）的問題。我主要的興趣在於勞工騎士團運動的內部結構、意識型態、以及發展狀況，而禁酒則是一個新的、意料之外的類別。我把相關的筆記內容編碼爲「禁酒」，並且把它當作是一個可能的主題（請參閱資訊箱 15.1 主題以及質化資料編碼）。

資訊箱 15.1　主題以及質化資料編碼

「一個好的主題編碼，能夠捕捉到某個現象中的質化豐富性。此種編碼可以被運用在分析、詮釋、以及研究的呈現當中」（Boyatzis, 1998: 31）。在把資料編碼成為主題時，研究者首先必須學習如何在資料中「看見」或者是辨認出主題。要能夠看見主題，必須仰賴四種能力：（1）辨認出在資料中的模式，（2）以系統以及概念的方式來進行思考，（3）擁有隱性的知識或者是深入的背景知識（例如，知道希臘神話故事有助於理解莎士比亞的劇戲），以及（4）擁有相關的資訊（例如，若是想要針對關於搖滾音樂會的主題進行編碼的話，研究者必須要非常瞭解搖滾歌手以及搖滾樂才能辦得到）（請參閱 Boyatzis, 1998: 7-8）。

當你在進行編碼時，必須盡量避免以下的三種錯誤（請參閱 Schwandt, 1997: 17）：（1）只停留在描述的層次（而未進展到分析的層次），（2）把編碼視為純粹機械性的過程，以及（3）使用僵硬固定而沒有彈性的編碼。編碼具有五個部分：（1）一個由一個字到三個字所組成的標籤或者是名字，（2）一個具有主要特徵的定義，（3）一個如何在資料中辨識出編碼的描述「旗幟」，（4）任何排除在外的情況或者是符合資格的條件，以及（5）一個範例。

編碼的五個部分的實例說明

標籤　性別角色爭議。

定義　人與人之間的口頭爭論是一個例子。另外，當兩性在進行互動或者是個別獨處時，關於男性或女性的適當的或者是可以接受的行為的爭議或者是衝突。

旗幟　例子之一，是說出諷刺的言論或者是笑話，或者是對於男性或女性所應該表現出的行為產生爭論（從非常溫和的爭論到火爆的論戰）。

符合的資格　只有在同樣性別的人們之間所產生的爭論才列入考慮。任何類型的行為（不論是言語的或者是非言語的）都可能是爭論的目標。在明顯的同性戀者或者是變性者之間的互動則不包括在內。

範例　在教室外面，同為 16 歲的莎拉和潔西卡正在談論她們昨晚的約會情況。莎拉說：「我們出去吃披薩——當然是他付錢囉！」潔西卡則說：「當然？你的意思是你本來就覺得應該是由男生來付錢嗎？」莎拉回答說：「噢，算了，別提了。」

葛蘭海姆以及盧德曼（Graneheim and Lundman, 2003）曾經針對二十名罹患第一型糖尿病的成人進行過訪談，並且對於訪談資料進行過質化內容分析。他們在研究報告中描述了開放式編碼的過程。在訪談當中，訪談者會詢問關於罹患第一型糖尿病在生活中的各種不同面向。然後，研究者會去閱讀訪談內容的逐字稿好幾次，以得到對於整體內容的概念。接著，研究者把資料中關於參與者血糖過高症（hyperblycemia）的經驗的部分擷取出來，並且把相關的段落都放到文本當中。如此一來，他們就建立了分析單位。研究者們接著把文本畫分爲意義單位（也就是一群具有相同中心意義的文字或者是敘述），然後再把這些意義單位加以濃縮。最後，他們把這些濃縮過的意義單位進行摘要，並且把每一個摘要都加上了編碼的標籤。

雖然我們可以利用一份概念的清單來開始進行編碼，但是當我們在閱讀資料筆

記時，通常會產生大多數的編碼主題。不論我們是否利用一份關於主題的清單來開始進行編碼，我們都會在完成開放式編碼*之後*，把所有的主題列成一份清單。這份清單可以達到三個目的：

主軸式編碼　質化資料編碼過程中的第二個階段，在這個階段裡，研究者會去組織編碼、把不同的編碼連結起來、並且發現一些關鍵的分析類別。

1. 幫助我們一眼就能夠看出正在浮現的主題。
2. 刺激我們在未來的開放式編碼當中去發現新的主題。
3. 我們可以利用這份清單來建立一項研究中所有主題的範圍，我們可以在進一步的分析當中去重新組織、整理、結合、揚棄、或者是延伸這些主題。

在編碼的完整度、以及要對多少細節進行編碼等方面，每個人都會有不同的作法。有些研究者會把每一行、或者是每隔幾個字就進行編碼，其他人則是會針對段落或是一整頁的內容來進行編碼。有些資料不會被進行編碼，此種資料就是無用的資料。編碼的細節程度則是取決於研究問題、資料的「豐富性」、以及研究目的（請參閱資訊箱 15.2 質化資料的編碼過程）。

開放式編碼可以延伸到你在蒐集資料時所寫下的分析筆記或者是備忘錄。除此之外，你也應該寫下關於編碼的備忘錄（請參閱稍後部分關於撰寫分析備忘錄的討論內容）。

主軸式編碼　這是對於資料的「第二次過濾」。在開放式編碼的過程當中，你的焦點在於實際的資料、以及爲所有的主題指派編碼標籤。也就是說，你比較不會去關心在主題之間作出連結、或者是去詳細闡述主題所代表的概念。對比之下，當你在進行**主軸式編碼**（axial coding）時，是從一組已經整理過的最初編碼或者是初步概念來開始的。在這第二個階段裡，你比較著重於最初已經編碼過的主題，而不是在資料本身上面。在這個階段裡，或許會浮現出其他的編碼或者是新的想法，而且你也應該要把這些編碼或是想法記下來，但是，你最主要的任務是要去回顧並且檢視最初的編碼。你應該要開始去組織概念或者是主題、並且要去辨識出在分析中的關鍵概念的主軸。

邁爾斯以及哈伯曼（Miles and Huberman, 1994: 62）曾經警告過：

> 基本上，編碼到底是在研究早期或是後期階段被創造出來或是被修正的，其重要性比不上編碼是否具有一些概念上以及結構上的順序。編碼應該以一種一致的、對研究具有重要性的方式而與彼此產生關聯，而且編碼應該是整體結構的一部分。

當你在進行主軸式編碼時，你會去詢問原因以及結果、條件以及互動、策略以及過程。你會去尋找那些群集在一起的類別或者是概念。你應該提出像是下列的這些問題：我能夠把現有的概念分割爲次面向或者是次類別嗎？我能夠把數個密切相關的概念結合成爲一個更爲廣泛的構念嗎？我能夠把不同的類別組織成爲一個順序（也就是先是 A，然後是 B，之後是 C），或者是依據其實體的位置（也就是發生的地點）來加

資訊箱 15.2 質化資料的編碼過程

雖然史特勞斯（Strauss, 1987）以及其他學者努力地嘗試著把質化資料編碼的過程加以系統化以及簡化，並且使得開放式編碼、主軸式編碼、以及選擇性編碼的過程看起來像是一個固定的三步驟順序，但是不論質化資料的形式是觀察的田野筆記、錄影帶或者是錄音帶、開放式的訪談、或者是詳細的歷史文件等等，要對質化資料進行編碼都是一個挑戰。有一些研究者會依賴一種文件編碼的軟體程式（請參閱本章稍後的相關討論），雖然此種軟體會強迫他們去創造出編碼，但是此種軟體充其量只是在較大的編碼過程當中的一種工具而已。

魏斯頓等人（Weston et al., 2001）曾經很詳細地描述了他們的編碼過程。魏斯頓是一個六人研究團隊的其中一員，並且發現到團隊合作有助於使編碼過程變得更為清楚明白。與紮根理論有關的一個理想狀況是，研究者在開始進行編碼時應該要抱持著完全開放的心胸，而且沒有任何先入為主的期望。但是，此種理想狀況真的只是一種理想而已。在實際情況裡，一個人所受過的學術訓練、對於概念以及理論假定的知覺程度、以及來自於研究報告讀者群的期望，都會對資料編碼造成影響。在魏斯頓的研究裡，資料編碼過程的第一個步驟，是先由團隊當中的一位研究者先去閱讀開放式訪談的逐字稿資料，並且根據閱讀後的結果創造出一個具有四種編碼的編碼系統。這個編碼系統對每一個編碼的概念以及規則都有提供定義，並且也有把原始資料轉化成為編碼的實際範例。研究團隊中的其他成員則是利用這個編碼系統來針對原始資料中一些挑選出來的部分進行編碼。依據他們對於這個初步編碼系統的經驗，他們會對這個系統進行修正，並且在原來的編碼裡加入一些次類型。這個過程會重覆好幾次，而團隊中的每一個成員都會個別地去對原始資料進行編碼、與其他成員開會討論編碼方式、並且修正編碼系統。在進行了數個月的編碼以及會議之後，最初的四個編碼變成了三個主要概念，其中有兩個概念各自包含了兩個類型，而每個類型之中又包含了四到七個更為精細的編碼。結果總共產生了 34 個編碼的類別。在接下來的兩年之中，研究團隊把這個編碼系統應用於數百頁的原始資料上。而團隊成員則是持續地進行反思編碼、開會討論編碼、以及修正系統的過程。到了最後，他們的編碼系統有四個階層——三個主要概念，在主要概念之下有七個類型，在七個類型中有三個類型包含了兩個次類型，而在每個次類型當中又包含了數個更為精細的編碼。加總起來的話，他們總共創造了 58 個編碼。

在接下來的兩年之內，當他們持續地檢視資料並且把研究發現呈現給科學社群時，這個團隊仍然繼續在修正以及調整編碼系統。他們所依循的是一種漸次接近（successive approximation）的策略（請參閱本章稍後部分的討論）。一些新的編碼出現了，而且編碼系統的結構也會有一些改變，但是在研究進行了四年，經過了數百個小時的開會討論以及對原始資料不斷地重覆過濾之後，整個編碼系統逐漸趨於穩定。如同你可以看到的，編碼系統可能不只是一個對原始資料進行編碼的方法而已。編碼系統還提供了一個分析的系統，該系統能夠提供一個結構式的詮釋。順便一提，魏斯頓等人的研究主題是提升大學的教學品質，而該團隊的原始資料則是來自於在一個學期當中對於六位教授進行詳細的開放式訪談的內容。

以安排，或者是把不同類別之間的關係安排成爲一個主要的、我感到興趣的主題嗎？

例如，當他在研究勞動階級的生活時，勒梅斯特斯（LeMasters）可以把婚姻這個一般性的議題再細分成數個次部分（例如：訂婚、婚禮等等）。他可以把所有與婚姻有關的筆記都標註起來，然後再把婚姻連結到其他的主題，像是性別、家事的勞力分工、以及對於孩子的看法等等。當這個主題在不同的地方重覆出現時，他也可以去進行比較，以找出新的主題（例如，男人與女人對於婚姻的態度不同）。

在前面提過的關於勞工騎士團的歷史研究範例當中，我尋找了與禁酒有關的主題。我尋找了關於酒吧的討論內容、喝酒或是酒醉、以及該項運動與其他支持或是反對禁酒的政黨之間的關係。圍繞著禁酒的一些主題，包括了把喝酒當作是一種娛樂活動的形式、把喝酒當成是民族文化的一部分、各種宗教對於喝酒的不同看法、以及與喝酒有關的性別差異等等。

葛蘭海姆以及盧德曼（Graneheim and Lundman, 2003）針對他們研究中關於糖尿病的訪談資料進行了主軸式編碼。他們依據相似處以及相異處來對不同的編碼進行比較，並且把這些編碼整理成爲六個次類別以及三個類別。這兩位研究者討論了暫時性的類別，並且修正了這些類別。在經過了一個反思以及討論的過程之後，他們終於在如何整理編碼方面達到共識。最後，兩位研究者在他們用來形成一個更廣泛主題的類別當中，辨識出了隱藏的意義，也就是隱性內容。

主軸編碼不只是能夠刺激我們去思考關於在概念或是主題之間的連結，也能夠提出新的問題。主軸式編碼也能夠建議我們是否應該要放棄某些主題，或者是更加深入地去檢視其他的主題，此外，主軸式編碼也能夠強化證據與概念之間的連結。當你在加強編碼時，你或許會在許多地方找到關於核心主題的證據，並且在質化資料中建立一張關於核心主題的綿密支持性網絡。此種情況很類似於我們之前提過的、在信度以及測量變項方面使用多重指標的概念。此種在主題以及資料之間的連結，會因爲有許多實徵證據的例子而被大幅度地強化。[5]

當我在研究勞工騎士團時，我把該運動未能與其他政治團體形成聯盟的失敗情況當作是一個主要的主題。我審視了筆記，尋找在勞工騎士團與其他政黨之間的妥協與衝突，這些政黨包括了禁酒團體以及禁酒黨等等。在主軸式編碼中，與禁酒有關聯的大量概念與主題，幫助我看出禁酒這個議題如何促進或是阻礙了各團體之間的結盟情況。

選擇性編碼　到了你準備好要進行資料的最後過濾階段時，你已經辨識出了主要的主題了。**選擇性編碼**（selective coding）牽涉到快速瀏覽所有的資料以及先前的編碼，選擇性地尋找那些能夠舉例說明主題的個案，並且在大多數或者是所有的資料蒐集過程都已經完成之後去進行比較。研究者應該要等到概念都已經發展完善、以及已經辨識出數個核心推論或者是想法之後，才開始進行選擇性編碼。

選擇性編碼　質化資料編碼的最後一個階段，在此階段中，研究者會去檢視先前的編碼，以辨識並且選擇出一些特定的資料，這些資料將能夠支持已經發展出來的概念性編碼類別。

例如，當勒梅斯特斯在研究一家小酒館內的勞

動階級生活時，他決定要把性別關係當成是一個主要的主題。因此，當他在進行選擇性編碼時，他可能會快速地瀏覽他自己的田野筆記，尋找男人和女人在談論到約會、訂婚、婚禮、離婚、婚外情、或者是丈夫／太太關係時的差異之處。接著，他就可以去比較一下在婚姻這個主題的每個部分裡，男性和女性的態度之間的異同之處。

葛蘭海姆以及盧德曼（Graneheim and Lundman, 2003）也針對他們研究中關於糖尿病的訪談資料來進行選擇性編碼。他們提供讀者關於他們研究中的一些關於編碼、次類別、類別等等的範例，其中還有一個主題的範例是取自於關於血糖過高症的文本敘事。而且這兩位研究者對於每一種編碼以及類別都提供了非常清楚明白的範例。

在選擇性編碼的過程當中，最終會由主要的主題或者是概念來引導研究的過程。你會去重新組織在先前的編碼當中所辨識出來的特定主題，並且針對超過一個以上的主要主題進行詳細的說明。例如，在他針對一家勞工階級的小酒館所進行的研究中，勒梅斯特斯可以去檢視關於婚姻的意見，以同時瞭解性別關係的主題以及生命循環中的不同階段的這個主題。同樣地，在勞工騎士團的研究中，我利用禁酒來瞭解該運動無法與其他政治團體形成結盟的這個主題，而且也同時去瞭解另外一主題，也就是在這個運動中造成分裂的原因，而這些原因則是來自於成員之間的種族或者是宗教方面的差異。

撰寫分析備忘錄

在進行質化研究時，你必須不斷地撰寫筆記。你會在筆記中記錄資料、以及寫下對於方法或者是研究策略的評論等等。你必須使自己成爲一個不斷撰寫筆記的人，把筆記依次歸類建檔，並且創造出許多記錄筆記不同主題的檔案，例如，一個關於方法論議題（例如：資料來源或者是倫理議題）的檔案、一個存放地圖或圖表的檔案、一個關於最後的研究報告或是章節的可能整體綱要的檔案等等。

分析備忘錄是一種特殊類型的筆記（我們已經在第十三章裡討論過了）。[6] 這是一種你寫給自己看的備忘錄或是討論，內容是關於你自己對於編碼過程的想法或觀念。每一個已經編碼的主題或者是概念，都會形成一個獨立備忘錄的基礎。因此，分析備忘錄就是一種關於概念或是主題的討論內容。而最初的分析備忘錄的內容，則是你對於理論的初步想法或者是看法。

分析備忘錄可以把具體的資料或是原始證據連結到抽象的、理論的思考（請參閱圖 15.1），其中包含了你對於資料以及編碼的反思以及思考過程。當你利用各種類型的編碼來過濾資料時，你應該在分析備忘錄中加入一些新的內容，並且善加利用分析備忘錄。分析備忘錄可以形成你在研究報告中分析資料的基礎。事實上，從品質良好的分析備忘錄中所重新撰寫過的部分，往往可以成爲最終研究報告中的部分內容。

在撰寫分析備忘錄時所使用到的工具相當簡單：筆和紙、一些筆記本、一疊的檔案夾、以及一些筆記的影本。有一些研究者會使用電腦，但這並不是必要的。撰寫分析備忘錄的方法有許多種，因此你應該發展出你自己的風格以及方法。關於一些來自於其他人經驗的具體建議，請參閱資訊箱 15.3 撰寫分析備忘錄的建議。有些

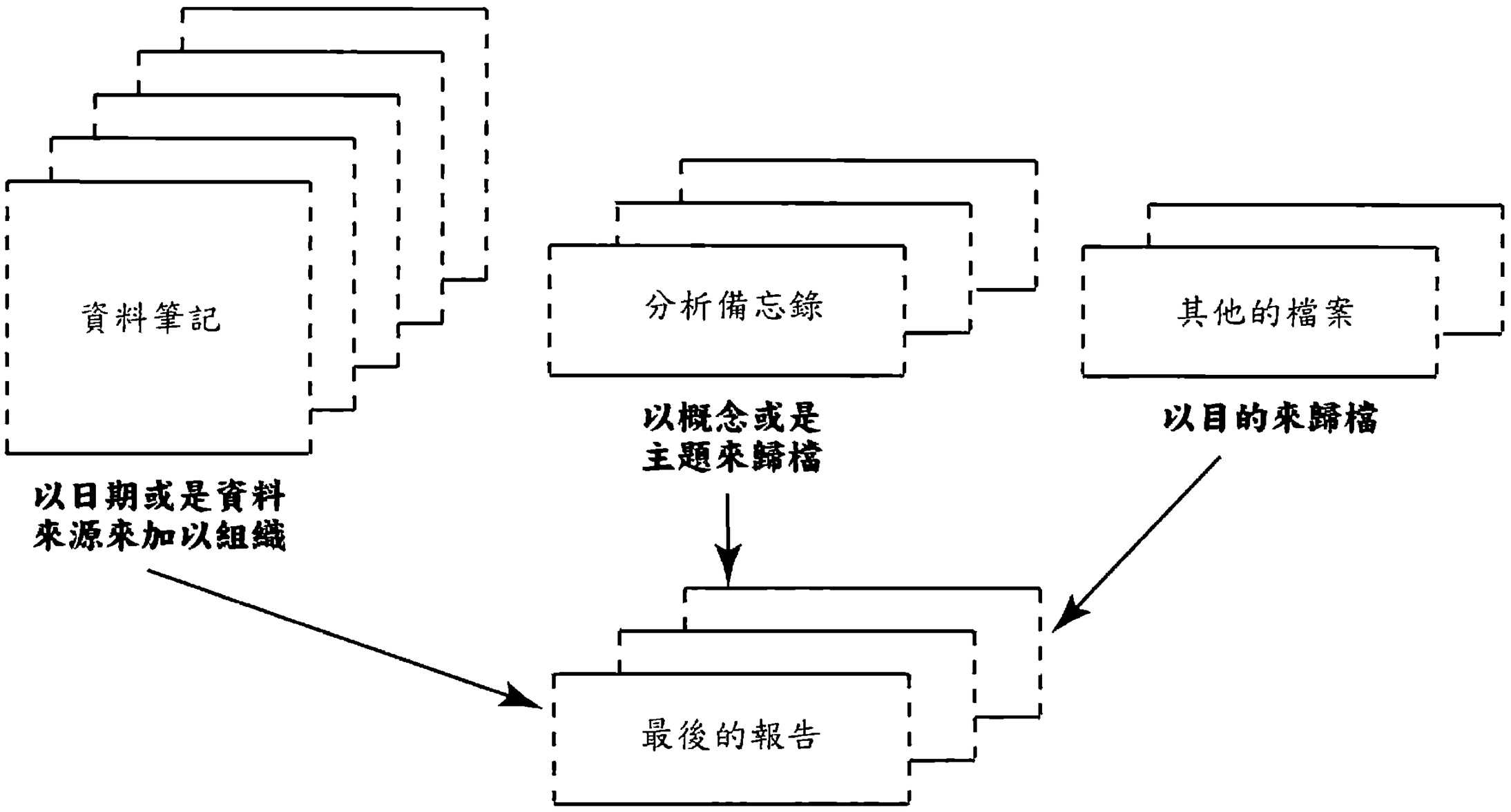


圖 15.1　分析備忘錄以及其他的檔案

研究者會把筆記影印好幾份，然後把這些影本剪下，並且把不同的影本放入分析備忘錄的檔案當中。如果實體的檔案很大，而且分析備忘錄在檔案中有特別標示出來的話（例如，用不同顏色的紙來標示，或者是放置在檔案的最前面部分），那麼此種作法確實是有用的。其他的研究者則會在分析備忘錄的檔案中，列出某個主題在資料筆記中出現的位置。如此一來，就可以輕易地在分析備忘錄以及資料之間來回進行交叉比對。因爲資料筆記包含了特別標註的主題，因此可以輕易地找到資料中的特定部分。另外一個策略則是維持一份持續更新的清單，內容是主要主題在資料中出現的位置，並且在清單裡加入一些筆記中關鍵部分的影本，以便於查詢參考之用。[7]

當你在回顧以及修正分析備忘錄的時候，你應該要和同儕們討論一些相關的想法，並且要記得回去查閱相關文獻以及把焦點放在一些新的議題上面。分析備忘錄也可能會有助於產生一些潛在的假設，而你則可以視情況需要而在備忘錄當中加入或者是排除這些假設。這些筆記也可以幫助你發展出新的主題，或者是修正編碼系統。

表面事實

我們所蒐集的特定實徵證據，與理論的觀念以及隱藏在可觀察的事實之下的結構有關。在圖 15.2 中的模型所展現的關係，顯示了資料只是發生在可見的、表面的層次上的一切事物的樣本而已。我們利用資料來產生以及評估理論與概化，並且同時假定，在外在事實的表面之下，隱藏著更爲深層的社會結構或者是關係。

我們能夠輕易看見的表象事實，只能夠部分地反映出那些在表面之下發生的、未被看見的事物。借用在地質學中一個術語，我們可以把發生在表面層次的事件稱

資訊箱 15.3　撰寫分析備忘錄的建議

1. 在你開始蒐集資料之後不久，就應該開始撰寫分析備忘錄，而且你應該持續撰寫分析備忘錄，一直到最後的研究報告完成之前都不應該停止。
2. 在每一筆備忘錄中都加上日期，如此一來，你才能夠看到思考過程的進展以及發展。當你在重新閱讀冗長而且複雜的備忘錄時，加上日期的作法將會對你有所幫助，因為隨著研究不斷地往前進展，你將會定期地修正備忘錄並且增加內容。
3. 當有靈感時，你應該立刻中斷編碼或者是記錄資料的工作，並且趕緊寫下分析備忘錄，不要耽擱或者是讓創意的火花或是新的洞見就這樣消逝無蹤——立刻就寫下來！
4. 定期地閱讀備忘錄，並且對於那些談論類似編碼的內容進行比較，以看看這些編碼是否能夠被合併、或者是在不同編碼之間的差異處是否能夠被解釋得更加清楚。
5. 針對每一個概念或是主題的備忘錄都保留一份獨立分開的檔案。所有關於同樣主題或者是概念的備忘錄內容都應該被保存在同一個檔案、資料夾、或者是筆記本當中。在這些檔案、資料夾、或者是筆記本的外面標上概念或是主題的標籤，如此一來，你就能夠輕易地找到某個概念或者是主題。隨著分析的過程向前進展，能夠實際地去整理或者重新組織備忘錄是很重要的，因此，你應該要能夠以某種方式來整理備忘錄才行。
6. 你應該把分析備忘錄和資料筆記分開保存，因為這兩者的目的是不同的。資料是證據，而分析備忘錄則具有一種與概念有關的、建立理論的意圖。分析備忘錄的目的並不是去報告資料，而是要去評論資料如何連結在一起、或者是某一群資料如何當作一個一般主題或者是概念的實例。
7. 參照在分析備忘錄中的其他概念。當你在撰寫備忘錄時，你應該去思考某個概念與其他概念之間的相似處、相異處、或者是因果關係。你應該在分析備忘錄中記下上述這些事物，以利於之後的整合、綜合、以及分析。
8. 如果同時產生兩種想法的話，你應該把每一種寫在個別不同的備忘錄當中。你應該試著把每個個別的主題或者是概念放在獨立的備忘錄以及檔案當中。
9. 如果在一份備忘錄中已經沒有什麼新的資料可以再加入了，而且你已經達到一個飽和點，無法再獲得關於某個主題或者是概念更進一步的資料時，你應該在備忘錄中明確地指出來。
10. 為分析備忘錄製作一份編碼或是標籤的清單。如此一來，你只要看一下清單，就能夠看到備忘錄中的所有主題。當你定期地整理以及重新組織備忘錄時，記得也要重新整理這份備忘錄標籤的清單，以使清單內容能夠符合整理之後的結果。

*資料來源：*改編自 Miles and Huberman (1994: 72-76)，Lofland and Lofland (1995: 193-194)，以及 Strauss (1987: 127-129)。同時也請參閱 Lester and Hadden (1980)。

表面事實　質化資料分析的一個面向，也就是去辨識出代表更深層結構關係的某個事件或者是特徵。

作是**表面事實**（outcroppings）。[8] 在地質學當中，這個術語所指的是岩床的一部分，而且是暴露於地表之外，讓人們可以看到的部分（譯註：這個術語在地質學中的是指「岩層的露頭」或是「礦脈」），也

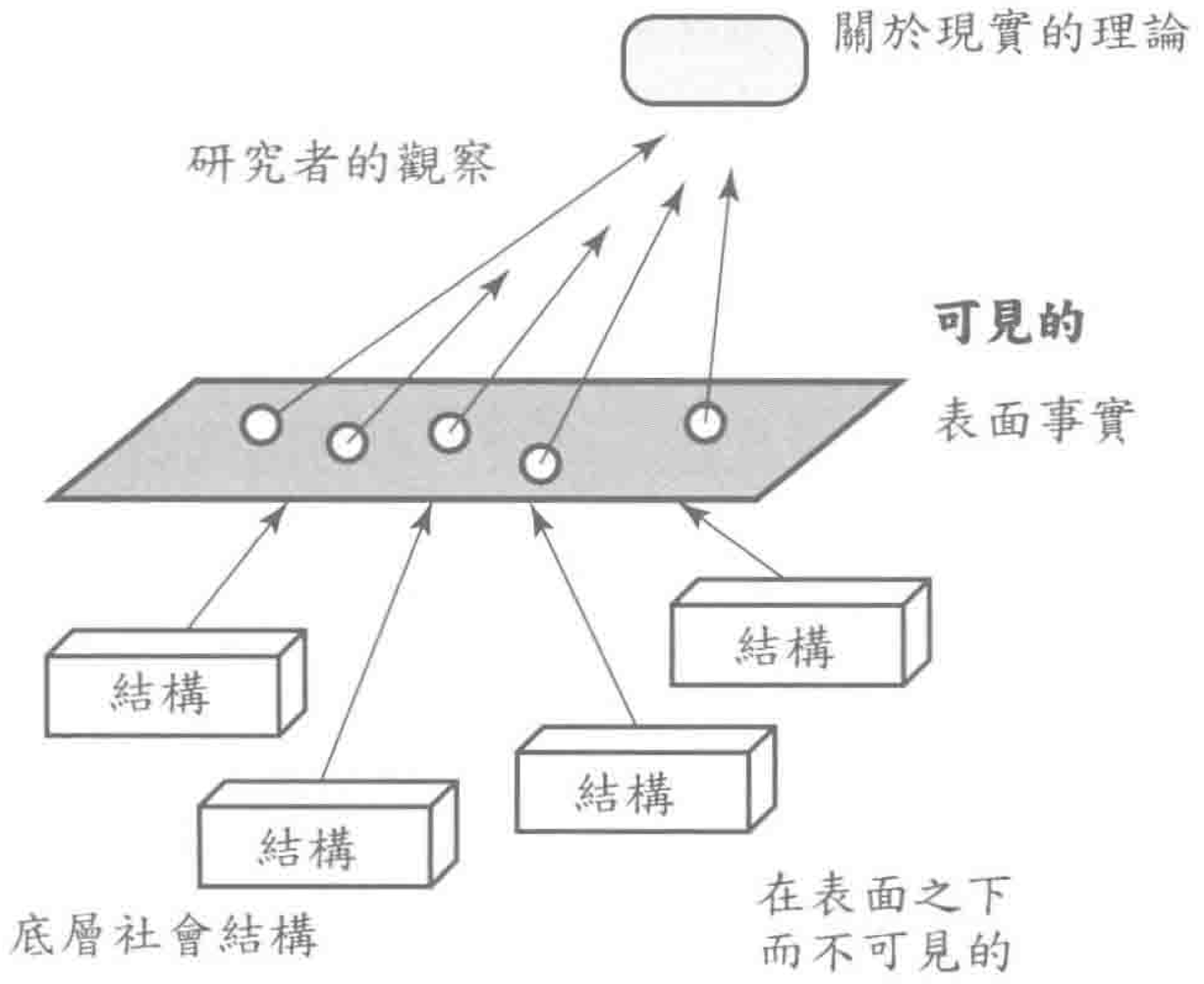


圖 15.2　理論、表面事實以及底層結構

就是指在陸地中央的、堅硬的部分在地表上所顯露出來的表象。地質學家會利用研究岩層露頭的方式，來得到關於隱藏在地表之下的事物的線索。

我們通常無法去直接觀察社會世界的各種特徵。例如，我們無法觀察在兩個人之間互相深愛的關係，因此，我們只能去觀察此種關係外在所顯現出來的行爲，像是親吻、表達感情的具體行爲、以及關愛的舉動等等。相同地，我們也無法去直接觀察像是社會階級之類的社會結構。然而，我們卻可以看到外在的徵象，像是人們行動方式的差異、對於職業生涯的展望、以及他們的物質財產等等。不過，有時候我們也會被外在的觀察結果所誤導。因此，我們必須同時去分析表面層次的現實、以及可能隱藏在表面之下而未被看現的底層結構以及力量。

質化資料的分析策略

大部分的質化研究者都會在某種程度上使用編碼、撰寫備忘錄、以及尋找表面事實的技術。這個部分將會介紹你可以用來分析質化資料的七種不同的策略：（1）理想型，（2）漸次接近，（3）舉例法，（4）領域分析，（5）分析比較，（6）敘事分析，以及（7）反面個案方法。

正如同在本章前面部分就已經提過的，與量化研究相較之下，質化資料的分析策略是比較多樣化的、比較不標準化的、而且也是比較不明確的。如同馬哈尼（Mahoney, 1999: 1192-1193）曾經提過的：「由於缺乏方法論方面的清楚性，因而使得許多讀者難以去完全理解或是欣賞〔質化資料〕研究者所提出的論點。」除此之外，有些研究者只使用一種策略，而其他研究者則可能會結合數種不同的策略。

一般而言，*資料分析*意味著去尋找在資料中的特定模式，像是重複出現的行爲、物體、詞語、或者是觀念等等。一旦你發現了一個模式，你必須依據社會理論或者

是發生背景來進行詮釋，而此種作法讓你能夠從關於一個歷史事件或是社會背景的特定具體描述，移動到一個更爲廣泛的詮釋。

在質化研究中，資料可能會有許多不同的形式。例如，田野研究資料就包括了研究者所體驗到的原始感官資料、在田野筆記中所記錄的資料、以及在最後的研究報告中會出現的那些經過選擇的或者是處理過的資料（請參閱圖 15.3）。資料分析牽涉到去檢視、整理、分類、評估、比較、綜合、思考編碼過的資料，以及重新探討原始資料和記錄下來的資料。

理想型

在質化分析中最常見的策略之一，就是韋伯（Max Weber）所提出的*理想型*（*ideal type*），這是一種社會關係或是過程的模型或是心智抽象概念。理想型是純粹的標準，而資料或者是「現實」則可以拿來與理想型進行比較。理想型是一種用來進行比較的人爲策略，因爲沒有任何現實狀況可以完全符合於理想型。例如，我發展出一個關於理想的民主制度或是一個理想的大學啤酒派對的心智模型。這些帶有各種特徵的抽象概念並不是在描述任何一個特定具體的民主制度或者是啤酒派對。然而，當我們把這些抽象概念應用到許多特定的個案上，以瞭解每一個個案符合理想狀況的程度時，這些抽象概念就會變得相當有用。

韋伯的理想型方法，也能夠補充彌爾的一致法（method of agreement）（請參閱分析比較的討論內容）的不足之處。一致法的焦點集中在橫跨不同個案之間的共同點，並且尋找在不同個案之間造成共同結果的共通原因。就其本身而言，一致法暗示了一種對於實際個案的比較。而此種個案之間的比較，也可以拿來與理想化的模式進行比較。換句話說，你可以發展出關於一個社會過程或者是關係的理想型，然後再拿特定的個案來與這個理想型進行比較。

在質化研究裡，我們能夠以兩種方式來運用理想型：對照脈絡以及類比。

1. *對照脈絡*（*contrast contexts*）　採取強烈詮釋取向的研究者，可能會以一種對於脈絡以及成員的文化意義相當敏感的方式來使用理想型以詮釋資料。此種研究者使用理想型的目的並不是在於發展出假設、或者是創造出一個可以概化的理論，而是要展現出每一個個案的具體細節，並且強調獨特的脈絡所造成的衝擊。[9] 當他們在脈絡之間進行對比時，他們或許會選擇一些具有戲劇性強烈對比的、或者是具有明顯特徵的個案。例如，在《工業裡的勞動與權威》（*Work and Authority in Industry*）一書中，班迪克斯（Bendix, 1956）比較了在兩個非常不同的脈絡之中的管理關係，而這兩個脈絡分別是沙皇時代的俄羅斯以及工業化之後的英國。在比較脈絡時，有些研究者並不會使用理想型來具體說明不同個案之間的共通理論、或者是去發現規律性。相反地，他們所強調的是個案的特殊性以及獨特性。對比之下，其他研究者則是利用理想型來顯示出獨特的特徵如何形塑了一般過程的運作方式。正如同史寇克波爾以及索默斯（Skocpol and Somers, 1980: 178）所解釋的：

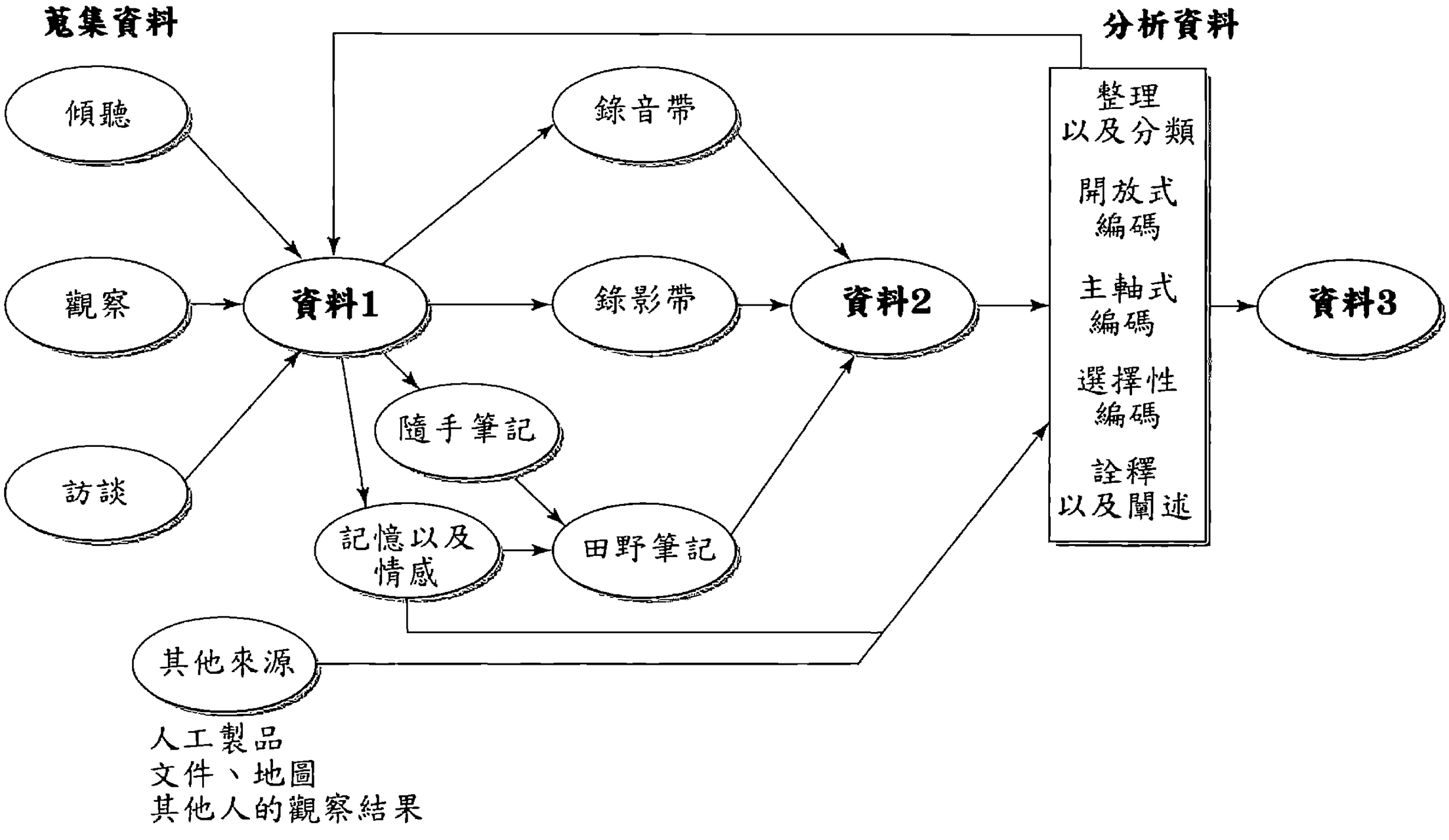


圖 15.3　田野／實地研究中的資料　（資料 1 = 原始感官資料，也就是研究者的經驗；資料 2 = 記錄下來的資料，也就是經驗的實體記錄；資料 3 = 在最後報告中選擇過的、處理過的資料）*資料來源*：改編自 Ellen (1984a: 214)

> 「最重要的是，對照是在各個獨立的個案之間所提取出來的。通常在發展出此種對比時，都會去參照廣泛的主題、引導的問題、或者是理想型。主題以及問題或許可以充當架構，以指出個案之間的差異之處。而理想型或許可以被當是一種敏感化的工具——也就是每一個個案在建立其特殊的特徵時所參照的基準。」

你或許也能利用理想型來顯示出，特定的情境、文化意義、以及特定個人的觀點，對於理解一個社會環境或是過程所具有的重要性。如此一來，理想型就變成一種陪襯物，而你則可以利用這個陪襯物來突顯出獨特的脈絡特徵。

2. *類比*　我們也可以把理想型當作是類比，以用來組織質化資料。所謂類比（analogies），就是指兩個物體、過程、或者事件是彼此相似的。我們利用類比來傳達觀念以及幫助進行邏輯方面的比較。類比能夠利用參照某個已知的事物或者是研究者熟悉的經驗的方式，來傳達關於資料中的模式的資訊。類比也能夠去描述被深埋在許多細節裡的關係。換句話說，類比是一種簡便的方法，可以幫助我們在具體事件所形成的迷宮中看見特定的模式。如此一來，在橫跨不同的個案或是背景以進行社會過程的比較時，就會變得比較容易。[10] 例如，你或許會說，在 X 這個人說話之後，整個房間頓時都安靜了下來，「就像一陣冷風吹過一般」。這句話的意思並不是說房間裡的溫度真的下降了、或者是房間裡的人真的感覺到有一陣風吹過，但是

這句話卻能夠簡潔有力地表達出在情緒方面的急遽變化。同樣地，你也可以報告說，在 Y 社會裡的性別關係，是女性被「視爲財產而且被當成像是奴隸一般來對待」。這句話的意思並不是說在 Y 社會裡兩性之間的法律與社會關係，真的就和奴隸主人與奴隸之間的關係完全一模一樣。這句話所暗示的是，如果把一個關於奴隸與主人之間關係的理想型應用於 Y 社會之中的話，將能夠顯現出該理想型與 Y 社會中的性別關係之間有許多重大的相似之處。理想型類比的運作方式就像是一種啓發式的方法（也就是一種幫助人們學習或是看見的手段）。當你試圖藉由指出一個深層結構或者是基礎機制的方式而去理解或者是解釋資料時，類比將會是特別有幫助的。[11] 理想型並未提供一個關於解釋的決定性測試。更確切地說，理想型引導我們將大量細節進行概念性的重新建構，以轉變成爲一個系統性的形式。

漸次接近

漸次接近（successive approximation）是一種牽涉到不斷重覆的過程。你在步驟之間不斷地來回循環，並且逐漸移往最終的分析。隨著時間經過，或者是在數次的重覆之後，你會從模糊的想法以及資料中的具體細節，逐漸移向一個具有概化性的廣泛分析。此種情況與之前討論過的編碼過程相當類似。你先從研究問題以及一個關於假定與概念的架構開始。接著，你深入去研究資料，提出關於證據的問題，以看看概念與證據之間的契合程度，並且揭露出資料的特徵 。你也可以藉由對證據作出摘要、以及調整概念以使其更加符合證據的方式，來創造出新的概念。接下來，你可以去蒐集額外的證據，以處理那些出現在第一階段中的、尚未解決的議題，然後再重覆這個過程。在每一個階段裡，證據與理論都會彼此互相產生影響。這個過程之所以被稱作爲漸次接近，是因爲修正過的概念以及模型會越來越接近完整的證據，而且也會被不斷重覆地修正以漸漸地變得更加精確。

每一次對於證據的過濾都是暫時性的或者是不完整的。雖然概念是抽象的，但是卻植基於具體的證據，並且反映出脈絡。當分析逐漸移向會受到條件與附帶狀況所影響的概化時，你可以仔細修正這些概化以及連結，以更清楚地反映出證據。[12]

漸次接近 一種質化資料分析的方法，此種方法會在實徵資料以及抽象概念、理論、或是模型之間重覆地來來回回，每次都會調整理論以及修正資料的蒐集方式。

舉例法 一種質化資料分析的方法，此種方法把理論概念當作是空盒子，並且以具體的實徵範例以及描述來填滿這些空盒子。

空盒子 在某個解釋當中的概念類別，被當作是舉例法的一部分來使用。

舉例法

另外一種分析的方法，則是以實徵證據來舉例說明理論概念。**舉例法**（illustrative method）是把理論應用於一個具體的歷史情境或是社會環境，並且依據理論來組織資料。先前已經存在的理論可以提供概念方面的**空盒子**（empty boxes），而你則可以利用實徵證據來填滿這些空盒子。[13] 在盒子裡的證據可能會證實、修正、或者是拒絕某個理論，而該理論

的形式可以是一個一般模型、一個類比、或者是一系列的步驟（請參閱資訊箱 15.4 舉例法的三種變化形式）。[14]

使用舉例法的單一個案研究，並非是對某個解釋的有力考驗或者是驗證，因爲來自於一個個案的資料可能成爲數個不同解釋的空盒子中的證據。除此之外，只利用一個個案來找出填滿某個空盒子的證據，並無法建立一個可以概化的解釋，因爲此種解釋必須要有來自於許多個案的證據才行。

領域以及基模分析

認知人類學（cognitive anthropology）所研究的是人類文化與思想之間的關係，而這門學科對於質化資料分析有著極大的貢獻。認知人類學把文化當作是心智的創造物，或者是對於實體的、物質的世界的認知組織系統。認知人類學家會去研究人們如何理解以及組織實質的物體、事件、以及經驗。他們認爲我們是依據認知類別來理解現實的，而且我們會依據文化類別去整理事件、物質生活、以及觀念等等。認知人類學試著去發現並且詳細記錄行爲的規則、或者是我們所使用的思考邏輯系統。爲了要達到這個目的，認知人類學所描述的，是人們在各種不同情境下認爲在文化上所預期的或是適合的事物，而不是人們實際上所做的事。認知人類學是較廣泛的資料分析以及理論化類型的一部分，此種類型的分析重點在於辨識出人們或者是機構如何去把世界進行分類（通常是以一種隱而未顯的方式）。而在分類完成之後，這些分類就會「自己有了生命」，並且組織了人類的經驗（請參閱 Bowker and Leigh-Star, 1999）。

早期的認知人類學家會要求人們把色彩以及植物分派到不同的類別中，或者是把親戚們整理成爲一個親屬系統。此種作法幫助認知人類學家去發現隱藏在人類社會行爲背後的組織原則。認知人類學從在 1950 年代到 1960 年代的研究（當時稱作「民族科學」（"ethnoscience"），演變爲 1970 年代對於「民俗模式」（"folk models"）或者是「領域」（"domain"）的研究，之後則又演變成爲「基模分析」。*基模*（*schemes*）

資訊箱 15.4　舉例法的三種變化形式

1. *個案澄清（case clarification）*　一個理論模型被用來舉例說明或者是澄清一個特定的個案或是單一的情境，並且藉由應用理論的方式來使得該個案更加容易理解。
2. *平行示範（parallel demonstration）*　將多重個案（也就是單位或是時期）並置在一起，以顯示出相同的理論在多重個案中都能夠成立。培吉（Paige, 1975）曾經在一項關於農民階級衝突的研究中使用了平行示範的方法。他首先發展出一個關於造成階級衝突的條件的詳細模型，然後提出來自於秘魯、安哥拉、以及越南等國的證據來舉例證明這個模型。
3. *模式配對（pattern matching）*　這個方法會把來自於單一個案的觀察結果，與來自於理論或是其他研究的模式或是概念進行配對。此種方法允許部分的理論證偽（partial theory falsification），換句話說，此種方法會把一些觀念、變項、或者是模式從考量中加以排除，因而使可能的解釋數量變少。

領域分析　一種質化資料分析的方法，能夠描述並且揭露一個文化領域的結構。

文化領域　一個文化環境或者是場域，人們在其中會定期地進行互動，並且發展出一套可以被分析的共享理解或者是「迷你文化」（"miniculture"）。

民俗領域　一種文化領域，其基礎是來自於在一個田野場域中被研究的人們所使用的行話以及類別。

意思是指抽象的實體以及關於世界的無意識模型，而我們則是利用基模來組織我們的經驗。在基模分析裡，我們並不把一個文化的各個部分視爲物質的或是象徵的，相反地，我們把文化當作是由許多部分所建構組合起來的。這些部分並不是靜態的、或者是被整合爲一個單一的整體，相反地，我們是利用基模來組織這些部分。基模是認知單位，像是原型、命題、以及認知類別等等。我們可以去分析文化的各個部分，以看看基模是否被這些部分所共享，我們也可以去檢視這些基模是否分散在不同的人們身上，並且去尋找基模與行爲之間的關係。

人類學家史培德利（Spradley, 1979a, 1979b）發展出了**領域分析**（domain analysis），而我們將在這個部分去檢視此種質化資料分析的系統。對於史培德利而言，在一個文化環境中的基本單位就是**文化領域**（cultural domain），也就是一種組織的觀念或是概念。資料分析系統的焦點就在於分析各個領域。之後，我們可以結合不同的領域以成爲分類學以及更廣泛的主題，而這些分類學以及主題則能夠提供我們對於某個文化場景或是社會環境的詮釋。

文化領域有三個部分：表面詞彙、內含詞彙、以及語意關係。*表面詞彙*（*cover term*）所指的就是領域的名稱。*內含詞彙*（*included terms*）是指在該領域內的次類別或是部分。至於*語意關係*（*semantic relationship*）則是告訴我們內含詞彙在邏輯上如何與領域連結在一起。例如，在一個司法情境中的證人這個領域裡，表面詞彙就是*證人*（*witness*）。而內含詞彙則包括了*辯護證人*（*defense witness*）以及*專業證人*（*professional witness*）。而語意關係則是「某一種的」。因此，不論是專業證人或者是辯護證人，都是某一種證人的類別。其他的語意關係則列於圖表 15.1 當中。

史培德利分析了在民族誌田野研究中的成員之間的行話，並且利用此種方式發展出了領域分析，不過，我們仍然可以把此種方法延伸到其他的質化資料中。例如，札麗澤（Zelizer, 1985）曾經研究過不斷變動的兒童社會價值，而她所用的方法是去檢視在十九世紀晚期時，那些與對兒童死亡的態度與行爲有關的文件。她也可以利用領域分析的方法，其中「對於兒童死亡的態度」就是一個領域，而她在文件中所發現的關於各種不同態度的陳述，就是內含詞彙。而各種態度則可以用「是……的一種類型」的語意關係來加以組織。史培德利發現了三種類型的領域：民俗領域、混合領域、以及分析領域。

1. **民俗領域**（folk domain）包括了來自於某個社會環境中的成員所使用的行話。要使用此種領域分析方法的話，你必須非常注意語言以及其用法。此種領域使用了次文化的行話之間的詞語關係、或者是以歷史行動者的語言，來確認文化意義。

圖表 15.1　在文化領域中的關係形式

語意關係	運用的範例
是……的一種類型	公車是交通工具*的一種類型*（交通工具的類型）
是……的一部分／是在……的某地方	輪胎是車子*的一部分*（車子的各個部分）
是一種……的方法	作弊*是一種*在學校獲得高分*的方法*（學生獲得高分的方法）
被用於……	火車*被用於*運輸貨物（運輸貨物的方式）
是……的原因	高失業率*是*公眾不安*的原因之一*（造成公眾不安的原因）
是……的一個階段	進攻是戰爭*的一個階段*（戰爭的階段）
是……的結果／是……的原因	燃煤工廠是造成酸雨*的原因*（造成酸雨的原因）
是一個……的地方	市鎮廣場*是一個*民眾聚集*的地方*（民眾聚集的地方）
是……的特徵	高聳的髮型以及染髮*是*龐克*的特徵*（龐克的特徵）

2. **混合領域**（mixed domain）包含了民俗詞彙，但是你則在其中加入了你自己的概念。例如，各種不同類型的賽跑者會以賽跑者專門的術語來稱呼（例如：長距離賽跑者、徑賽跑者等等），但是你觀察到還有其他類型的人們，而且在現有的行話中沒有可以稱呼這些人的詞彙，因此你就自己分派新的標籤給這些人（例如，偶然造訪的人、新來的人、業餘者等等）。
3. **分析領域**（analytic domain）則包含了來自於研究者以及社會理論的詞彙。當在某個環境中的意義是參與者心照不宣的、未言明的、或者是未被辨認出來的時候，此種方法是最爲有用的。你從觀察結果以及人工製品中推論出意義類別並且辨識出各種模式，然後再分配適當的詞彙給這些類別或是模式。

你也可以從資料筆記中建構出領域，而其過程如下：閱讀你自己的筆記，並且尋找共同的語意關係（例如：是某種類型的地方、是某種類型的人、是某種類型的感覺），以發現社會關係的組織方式。接著，確認一份表面詞彙的清單，舉例而言，在一個司法情境中的證人，就可能是一個表面詞彙。一旦你擁有了一份表面詞彙的清單之後，接下來你應該把來自於筆記中的資訊加以組織爲內含詞彙。然後爲每一種領域關係準備一張工作表。工作表中應該要包含表面詞彙、內含詞彙的清單、以及語意關係（請參閱範例 15.1 領域分析工作表的範例）。

接著，你必須從你的筆記中找出領域關係的範例。這個分析工作會不斷進行直到所有相關的領域都已經被確認爲止。然後，你藉由比較差異處以及

混合領域　一種文化領域，此領域結合了被研究的成員所使用的行話與類別，以及由研究者為了分析而發展出來的類別。

分析領域　一種文化領域，其基礎是研究者為了要瞭解某個社會環境而發展出來的類別或詞彙。

範例　15.1

領域分析工作表範例

1. 語意關係：完全包含
2. 形式：*X*（是某一種）*Y*
3. 範例：橡樹（是某一種）樹

內含詞彙	語意關係	表面詞彙
自助洗衣店，飯店大廳 拖車車廂，果園	是某種 →	棲身之處
廉價旅館，橋下 棚車，小巷弄 公共廁所，壁爐 結構問題：你會稱小巷弄是一個棲身之處嗎？		

內含詞彙	語意關係	表面詞彙
模範囚犯，管理員 廚房裡的助手，拖地的人	是某種 →	監獄牢友
頭號模範囚犯，關在拘留所的人 不怕子彈的人，掃地工人 清理草皮的人，犯人的理髮師 結構問題：你會把模範囚犯稱作是某種監獄牢友嗎？		

相似處的方式，來組織這些領域。最後，把所有領域都再重新組織成爲類型學或是分類學，並且重新檢視所有領域，以創造出新的、更廣泛的領域，這些新領域則把其他領域當作內含詞彙而包含在內（請參閱資訊箱 15.5 領域分析步驟摘要）。這個過程是從筆記中的具體細節開始，逐步建立起一組整體的邏輯關係。[15]

分析比較

英國的哲學家以及理論家彌爾（John Stuart Mill, 1806-1873）發展出一個至今仍被廣泛使用的比較邏輯方法。他所發展出來的一致法以及差異法形成了**分析比較**（analytic comparison）的基礎。[16]我們可以使用理想型、漸次接近、舉例法、以及領域分析來檢視單一個案或者是多重個案的質化資料，然而，在使用分析比較時，則一定要有多個個案。分析比較採用一種準實驗的方法，此種方法結合了演繹以及歸納的理論化過程。基本上，你先在一組個案中辨識出許多因素，整理出這些因素在邏

分析比較　一種質化資料分析技術，此種技術利用一致法以及差異法來找出在一組個案當中影響結果的原因因素。

資訊箱 15.5　領域分析步驟摘要

領域分析有六個正式的步驟，而這些步驟在其他類型的質化資料分析中也可以找到：

1. 閱讀並且重新閱讀充滿細節的質化資料筆記。
2. 在心中把所有細節重新整理為十多個組織性的觀念。
3. 利用主觀意義或者是組織性的觀念，從筆記當中發展出新的觀念。
4. 尋找各種觀念之間的關係，並且依據邏輯相似性把這些觀念集合在一起。
5. 對各種不同觀念群組進行比較以及對照，以組織成更大的群組。
6. 重新組織，並且把群組連結起來，以成為一個更廣泛的整合性主題。

輯上的可能組合，然後跨個案進行比較。在某些方面，分析比較與統計推理所共有的特徵，比起與量化資料分析所共有的特徵還要來得多。分析比較甚至還被用於理性決策模式當中，在此種模式中，某些因素的特定組合，或許會使某些選擇對於人們而言似乎是理性的，而其他的組合則沒有此種作用。

分析比較有時候會被稱作是*名目比較*（*nominal comparison*），因爲在質化資料中的因素通常是名目測量尺度，但是也可以是順序尺度。[17] 你先把一組個案（通常是三到十個）的資料組織成爲彼此互斥以及全容的因素。當你使用電腦程式（質化比較分析（Qualitative Comparative Analysis, QCA），本章稍後部分會進行討論）來正式地進行分析比較時，你會建構出一份邏輯學家以及數學家所稱的*真值表*（*truth table*）。真值表中包含了各種因素在邏輯上可能出現的所有組合，以及在個案之間的結果。這項資訊通常會被組織成爲一張圖表（請參閱範例 15.2 一致法以及差異法的範例：史寇克波爾的革命理論），而此種圖表看起來就類似於古特曼量表（在第七章裡已經討論過了）。分析比較幫助你確認因素的結合方式，而這些因素通常是以名目尺度來測量的，而且也與一小組個案之間的結果有所關聯。

拉金（Ragin, 1994b）將個案導向的分析比較，與傳統的變項導向的統計分析進行了對比。他提到，個案導向的比較「將個案視爲有意義的、但是複雜的配置方式，此種配置方式是由事件與結構所形成的，而且也把個案當作是刻意挑選出來的、單一而完整的實體」（p. 300）。比較分析牽涉到來自於少數個案的質化資料，並且採取一種密集的（也就是少數個案裡的大量深入細節）、而非廣泛的（也就是大量個案中的少數細節）資料分析策略。除此之外，在分析比較中的解釋，通常傾向於詮釋的或者是結構的，而不是律則的。分析比較強調的是在個案或者是脈絡中各種條件的特定配置方式所產生的效果。分析比較也允許不同的原因因素產生同樣的結果，並且也會考量具有質化差異的複雜結果。[18]

一致法　**一致法**（method of agreement）將焦點集中於不同個案之間的共同之處。你先確定不同的個案之間擁有共同的結果，然後試著去找出一個共同的原因，即使個案的其他特徵可能會有所不同。在

一致法　一種質化資料分析的方法，此種方法會針對具有一個共同顯著重要結果的許多個案，去比較這些個案之間的相似特徵。

範例 15.2

一致法以及差異法的範例：史寇克波爾的革命理論

	原因因素		結果
個案	*政府解體*	*農民起義*	*是否發生革命？*
法國	是	是	是
1917 年的俄羅斯	是	是	是
中國	是	是	是
英國	是	否	否
1905 年的俄羅斯	否	是	否
德國	否	否	否
普魯士	否	否	否
日本	否	否	否

*資料來源：*改編自 Mahoney 1999, Table 1。

進行此種方法時，所使用的是一種淘汰的過程。如果在擁有共同結果的個案當中，某個特徵並非這些個案所共有的，那麼你就應該把該特徵從可能的原因當中淘汰。例如，你檢視了四個個案。這四個個案當中有兩項共同的特徵，但是在許多方面卻都彼此互異。你試著去尋找一個或者是多個共同的原因，以解釋在所有個案中的共同結果。同時，你也必須淘汰其他的可能性，並且確認少數的主要原因因素，如此一來，你才能夠宣稱，雖然這些個案之間有許多差異性，但是仍然存在著非常重要的相似之處。

差異法 你可以單獨使用**差異法**（method of difference），也可以結合一致法一起使用。差異法通常比較具有說服力，而且相當於「雙重應用」一致法。首先，找出在許多方面相似、但是在少數重要面向有差異的個案。接下來，明確指出一組個案在某個結果以及因果特徵方面的相似特徵，並且找出另外一組個案，這些個案不論在結果或是因果特徵方面都各自不同。差異法會強化來自於正面個案（例如，那些擁有共同因果特徵以及結果的個案）的資訊，其方式是把此種資訊和反面個案（例如，沒有上述結果以及因果特徵的個案）進行對比。因此，你所尋找的個案，是那些擁有正面個案的許多因果特徵、但是缺乏少數關鍵特徵，而且也具有不同結果的個案（關於分析比較的範例，請參閱範例 15.3 關於街友安置組織成敗相關研究的分析比較）。

差異法 一種質化資料分析的方法，此種方法會去比較個案之間的各種特徵，而在這些個案當中，有一些個案共同具有一個重要的結果，而其他個案則沒有。換句話說，此種方法的焦點在於個案之間的差異之處。

敘事分析

敘事以及分析一系列事件的相關概念，都具有多重意義，而且也都被使用於人類學、考古學、歷史學、語言學、文學批評、政治學、心理學、以及社會學當中。[19] 我們在第十四章中討論歷史比較研究時，曾經把敘事當作是一種撰寫歷史的形式。除

範例　15.3

關於街友安置組織成敗相關研究的分析比較

克雷斯以及史諾（Cress and Snow, 1996）利用比較分析的技術，分析了他們所蒐集的大量田野研究資料（總共有約一千五百頁的田野筆記），而其內容則是關於在美國八個城市裡十五個幫助街友的社會運動。他們發現了這些社會運動可能具有四種廣泛類型的資源，也就是道德、物質、資訊、以及人力資源。他們衡量某個社會運動組織所具有的資源的方法，是去看看該組織是否擁有十四種具體的資源，而在上述的四種廣泛類型的資源當中，每種類型中至少要有兩個。例如，具體的道德資源之一，就是由一個外部組織所發表的公開支持的聲明；物質支持則包括了像是報紙或是電話之類的服務；資訊支持則包括了一些對於主持會議很有經驗的人；而人力支持則包括了那些經常自願犧牲時間以及遵守命令的個人。

這兩位研究者依據這些社會運動組織是否能夠持續下去（viable）來進行分類（其中有七個可以，而有八個則不行），所謂的持續下去，意思是指該組織能夠存活超過一年以上，而且在這期間每個月至少要聚會兩次。他們發現，有九種具體的資源是必要的，否則的話，組織就會失敗。而影響組織成敗的因素，還包括了其他五種具體資源的組合方式。這十五個組織會依循著三條「路徑」的其中之一來發展，而這些路徑則是以九種必要資源以及五種「其他」資源的組合為基礎的。

此之外，*敘事*（*narrative*）也可以用來指稱一種質化資料的類型、一種研究以及蒐集資料的形式、一種討論以及呈現資料的方法、一組質化資料分析的技術、以及一種理論性的解釋。正如同葛里芬（Griffin, 1992a: 419）所觀察到的：「敘事既是一種修辭的形式，也是一種一般的、邏輯的解釋形式，也就是將某個事件的理論化描述以及對於該事件的解釋融合在一起。」

作爲一種檢視世界的方式，敘事具有數種特色：在各個部分之間的一種連結關係、一種形成「情節」的事件因果順序、一種刻意選擇某些事物而突顯出重要與較不重要部分爲何的方法、以及一種時間與空間的特定混合。我們利用敘事來達成數種目的，舉例而言，敘事可以處理我們身爲個人的「我們是誰」的議題，或者，也可以是把我們與更大的團體、社群、或者是國家連結起來的公共敘事。有些敘事所描述的則是那些在我們身上產生作用的社會力量。最後，*後設敘事*（*metanarratives*）則是一種具有主要概念的整體架構，此種敘事能夠在好幾個世代的時間內影響了整體人們的思想（例如，像是進步、工業化、或者是全球化等等觀念）（請參閱 Somers and Gibson, 1994）。雖然敘事具有許多種不同的用途，各種敘事之間仍然共享一些核心要素（請參閱資訊箱 15.6 敘事的六項特色）。[20]

接下來，我們將要簡短地討論一下數種不同的敘事類型，然後再接著討論敘事分析，因爲這也是一種質化資料分析的類型。

作爲原始資料，敘事所指的是社會生活中的文本以及活動。敘事所指的是人們如何去組織他們的日常活動以及主觀理解。敘事出現在口頭或者是書面的文本當中，以表達人們所理解的事物以及生命經驗的品質。敘事也是一種形式，人們依據此種形式而建構認同，並且依據此種形式而在他們周遭微觀以及鉅觀層次上所發生的事情之間找到自己的定位。[21]

資訊箱 15.6　敘事的六項特色

1. 敘事所訴說的是一個故事或傳說（也就是從一個特定觀點來呈現出逐漸展開的事件）
2. 敘事具有一種運動感或者是過程感（也就是一種在之前或是之後的情況）。
3. 敘事包括了在一個複雜的、詳細的脈絡當中的各種交互關係或者是連結。
4. 敘事牽涉到一些採取行動並且作出選擇的個人或是群體。
5. 敘事具有連貫性，也就是說，整體內容是一致地結合在一起的。
6. 對於一連串的事件而言，敘事具有一種時間上的順序。

敘事文本（*narrative text*）所指的是一種像故事一般的格式的資料，而人們則是利用敘事文本來組織並且表達在社會生活中的意義以及理解。「在許多儲存經驗的其他場所當中，像是學校、診所、支持團體、以及自助組織等等，都能提供敘事架構，以把個人的經驗代代傳承下去」（Gubrium and Holstein, 1998: 164）。我們可以在許多地方發現敘事，像是小說裡的故事、詩歌、神話、史詩故事、戲劇表演、電影、報紙或者是媒體報導、佈道、口述歷史、訪談、以及述說某個人物生命中的事件等等。敘事除了是一種表達的形式之外，同時也是一種實踐。

敘事實踐（*narrative practice*）是一種像故事一般的形式，而人們則是透過此種形式來主觀地經歷其日常生活以及行動，並且爲其日常生活以及行動賦予意義。敘事能夠組織那些在時間流逝中所發生的資訊、事件、以及經驗，並且能夠提供一種來自於特定觀點的故事情節。而此種特定的觀點，是屬於一個表達出意圖的、有動機的行動者。由於敘事情節是內嵌於許多特定細節所形成的集合當中，因此，要利用敘事情節來作出普世的推論是相當困難的。

在史邁爾德（Smilde, 2003）對於委內瑞拉首都卡拉卡斯（Caracas）所進行的一項研究中，他強調了一個他在當地的聖靈降臨教派（Pentecostal）的信仰中所發現的敘事。一個當地的男性團體利用來自於聖靈降臨教派敘事裡的故事，來重新詮釋他們的生活經驗，而且該敘事也形塑了他們的日常生活。這些男性改編並且利用該項敘事來重新組織他們對於持續進行的生活事件的理解，而且該項敘事也賦予這些事件一種新的連貫性。因此，由聖靈降臨教派所提供的敘事，就把宗教皈依以及一種新的自我理解融合在一起了。這些當地的男性利用此種敘事來重新詮釋他們過去的行動，並且引導他們目前的活動。如此一來，聖靈降臨教派的敘事就不只是在訴說一個故事而已，還幫助這些男性建構他們的認同以及找到生命的意義。

敘事探究（*narrative inquiry*）是一種調查以及蒐集資料的方法，此種方法保留了一種來自於社會生活的、像是敘事一樣的特質（Chase, 2005）。當我們把敘事當作一種探究方式來加以運用時，我們會試圖在不打斷、不破壞、或者是不減少其敘事特質的情況下，去捕捉人們的日常生命經驗。此種探究具有自我反思的性質，也就是說，你把自己放置在一個事件的流動過程中，並且有意識地成爲「情節」中的一部分。探究的本身－在一個田野環境中進行參與觀察，或者是去檢視歷史比較文件－是以敘事的形式出現的，換句話說，是以一個故事的形式出現的，而且其中具有

一種運動感以及一連串連貫的事件，而這些事件則是與在一個特定脈絡中的積極社會行動者有關。

敘事呈現（*narrative presentation*）是由詮釋的社會科學取向中所發展出來的。此種呈現方式通常被稱作是*說故事*（*story-telling*）（Berger and Quinney, 2004），而且融合了描述、同理心的理解、以及詮釋。此種呈現方式消除了存在於研究者以及研究對象之間的空間，因此使得研究者變成研究中的描述、討論、以及詮釋裡不可分割的一個部分。研究者以及被研究者一起共同參與了創造/蒐集資料的過程，並且一起對資料進行反思。這樣的過程將研究者的生命與被研究者的生命相互交織在一起。作爲一個個別的社會行動者，研究者變成研究過程以及資料呈現過程中不可或缺的一部分。正因爲如此，研究者的個人傳記以及生活情境，通常都會被包含在故事以及資料的呈現、討論、以及詮釋當中。除了讓被研究的人們「發聲」之外，研究者本身的聲音、在場、以及主體性也都會出現。說故事者／研究者並不是一個脫離現實的聲音、或者是一個疏離的觀察者，相反地，他或她是一個說故事者，而且他或她的情感、個人經驗、以及生命事件已經變成了故事中的一部分。

敘事分析（narrative analysis）是一種分析資料以及提供解釋的方法，而且具有數種不同的形式。敘事分析也可以被稱爲*分析敘事*（*analytic narrative*）、*敘事解釋*（*narrative explanation*）、*敘事結構分析*（*narrative structural analysis*）、或者是*次序分析*（*sequence analysis*）。[22] 除了辨識出敘事的核心要素（之前已經列出來了）之外，你或者也會使用敘事分析的技術來詳細規畫一個敘事，並且爲該敘事賦予一個正式的文法／結構。你不但能夠辨識社會生活的敘事特質，而且也能夠以一種保留以及揭露出該特質的方式來分析資料。敘事是一種用來組織資料的概要或是模型，但是同時也可以充當一種解釋的類型。

有些研究者在質化資料中只應用了少數的分析概念，而其他的研究者則採用了能夠概述敘事結構的複雜邏輯系統，而且通常會利用電腦軟體的輔助。當你在檢視以及分析質化資料以尋找其中的敘事形式與要素時——不論是個人的生命故事、一個特定的歷史事件、某個組織在過去幾年當中的演變過程、或者是一個鉅觀層次的歷史過程等等——你的焦點是在於事件（而非變項、個人、或者是個案）與事件之間的連結關係。你會發現到時間特徵（例如，順序、步調、持續時間、頻率等等）是非常重要的組織概念，而且你也會很快地開始把事件次序本身當作是一個探究的對象。

弗蘭佐西（Franzosi, 1998）主張，一旦我們在資料中辨識出了敘事的存在，我們應該在不破壞其產生意義的能力或者是結構的情況下，試著去提取並且保存該敘事。我們也應該去尋找阿貝爾（Abell, 2004: 293）所稱的「行動連結」（"action linkages"）——也就是說，一個社會行動者如何從事一些行動以把某個條件或情境轉變成爲另外一個條件或情境，或者，簡單地說，就是如何使某些事情發生。當我們在描繪出某個敘事次序的結構時，此種過程的運作方式同時是一種資料分析

敘事分析　同時是一種述說故事的歷史寫作類型、以及一種質化資料分析的類型，所呈現的是一連串以時間順序連結在一起的事件，而在這些事件當中，個別的或者是集體的社會行動者都扮演了重要的角色。

模式、也是一種解釋的類型。此種過程是對於下列這個問題的解答：爲什麼會發生事件？有些研究者相信敘事解釋並不具有因果關係的性質，但是其他研究者則相信敘事分析是一種因果關係的解釋，只不過此種解釋或許與傳統的實證科學取向的因果關係分別屬於不同的類型而已。[23]

敘事分析的工具 我們接著要來檢視三種分析的工具：路徑相依、時期劃分、以及歷史偶然性。

1. *路徑相依* 一個獨特的開端能夠引發一連串事件並且造成一個決定性路徑的方式，就被稱作是**路徑相依**（path dependency）。此種路徑在一連串連續的事件中是明顯可見的，並且會限制了進行中的事件能夠發展的方向。當我們利用路徑相依來解釋某個結果時，此種結果對於在過程中非常早期的階段所發生的事件是非常敏感的。路徑相依的解釋所強調的，是某個時期的選擇會如何限制了未來抉擇的可能性、形塑了之後的選擇、甚至是加速了事件朝向未來危機的發展速度，而且在這些未來的危機當中，選擇的機會或許是相當有限的。[24]

在建立一個路徑相依的解釋時，你應該先從結果開始。接著，你展現出該結果會如何依循著一個由之前的一連串事件所形成的路徑。當你在追溯並且論證每一個事件對另外一個事件的影響時，你會依循著整個過程往前回溯到最初的事件或者是條件。你所確認的初始條件或情境，就是一個「路途上的歷史分叉點」（Haydu, 1998: 352）。

利用路徑相依所建立的解釋，會假定產生初始事件（一種社會關係）或者機制的過程，可能會與那些使事件繼續發展的過程有所不同。因此，可能會有一個解釋是針對「起始事件」的，而另外一個解釋則是針對後續事件的路徑的。研究者通常會將起始事件解釋爲某種偶然過程的結果（也就是說，在一個特定的時空當中，各種因素以一種特定而且獨一無二的方式組合在一起，而且無法再重覆）。除此之外，在某個歷史時期中的因果過程，可能無法在另外一個歷史時期當中運作。「沒有任何良好的理由，可以讓我們假定在某個時期所得到的發現，能夠支持另外一個時期裡的因果宣稱」（Haydu, 1998: 345）。

路徑相依有兩種形式：自我強化的、以及反應順序。[25] 如果你使用了*自我強化的*（*self-reinforcing*）路徑相依解釋，你所要檢視的，是事件在開始轉動之後，會如何自己持續產生作用、或是把後來的事件推往一個抗拒外在因素的方向。一個初始的「觸發事件」（“trigger event”）會侷限、或是施加限制在過程的發展方向上面。一旦一個過程開始運作了，「慣性」（“inertia”）就會發揮作用，以使該過程沿著完全相同的路徑繼續發展下去。

路徑相依 一種被使用於敘事分析中的分析概念，其目的在於把一個過程或者是一連串的事件解釋為具有一個能夠引起一個結構化順序的開端，如此一來，一連串的事件就會依循著一個可以辨識的軌跡而逐漸發展。

一個有關於慣性的經典例子，就是在鍵盤上的 QWERTY 的字母配置模式。其實這種配置模式是沒有效率的，因爲此種模式會使手指比其他配置方式花更多時間才能觸及按鍵，而且也比較難以學習。在超過一個世紀之前，工程師爲了配合早期粗糙的、緩慢的、機械式的打字機而發明了 QWERTY

的配置模式。他們設計出此種鍵盤配置模式的目的，是爲了減緩人類打字員的速度，以防止原始的打字機發生卡住的故障狀況。之後，機械式的打字機不斷地改善，而且後來被電子打字機以及電子鍵盤所取代。老舊的鍵盤配置模式是不必要的、而且是過時的，但是卻一直被持續使用到今日。此種使用過時的、無效率系統的慣性是相當強大的。雖然也有人曾經試圖改變現有的機器與人們，以使用一種更爲理性的、快速的鍵盤，但是強大的慣性卻戰勝了此種推動改變的努力。社會機構與制度也有類似的情況，也就是說，一旦某種社會關係或是機構以某種特定的形式（例如，一種具有許多地區辦公室的地方分權形式）被創造出來之後，即使這些關係或是機構在現今的情況下已經不再有任何效率可言，要改變它們也會是一件很困難的任務。

至於*反應順序*（*reactive sequence*）的路徑相依，所強調的則是一個不同的過程。此種形式的路徑相依把焦點集中在每一個事件對一個相連在一起的先前事件作出反應的方式。因此，此種形式的路徑相依不會去追溯一個過程的源頭，而是去研究在過程中的每一個步驟，以瞭解每一個步驟如何影響了緊接在後的下一個步驟。此種路徑相依所關心的，是事件的發展順序是否會轉變或者是逆轉了事件從最初所發展的方向。路徑不一定是單一方向或者是線性的，而是可以「彎曲」或者甚至是倒轉整個發展過程，以抵消之前的發展方向。

我們可以把反應順序路徑相依想像成爲一個像是鐘擺一樣的一連串事件，也就是會來來回回地擺動。一個單一事件可能會引發一個反應，而這個反應可能會改變或者是倒轉發生在該事件之前的事件發展方向。例如，作爲在美國民權運動冗長過程中的一部分，金恩博士（Martin Luther King Jr.）遭到暗殺的事件反而使得民權法律的執行更爲積極，而且也使得福利計畫迅速擴張。在此之前，各種事件一直朝著增加社會平等、減少歧視、以及擴大法律權利的方向發展前進，然而，民權的快速進展以及福利擴張，卻打亂了現存的地位以及權力關係。此種情況產生了緊張對立的氣氛，並且使得一些充滿怨恨的白人開始進行反擊。白人的反擊行動試圖要限制或者是逆轉民權法案的執行，並且減少社會福利計畫。因此，對於一系列事件的反應，就逆轉了這些事件原本的路徑方向。

2. *時期劃分*　在歷史比較研究中，我們知道歷史事實就像是不連續的階段一般地在流動著。爲了辨識這種狀況，研究者或許可以使用**時期劃分**（periodization）的方法，把社會現實中的時間流動，劃分爲一些片段或者是時期。例如，我們或許可以把 100 年的歷史劃分爲數個不同的時期。我們可以把連續的時間切割成爲數個不連續的時期，而我們則是透過時期劃分的技術，以理論性的方式來定義這些時期。理論可以幫助我們確認在每個時期裡、或者是在不同的時期之間，什麼事物是重要的、而什麼事物又是共通的。如同卡爾（Carr, 1961: 76）所說的：「把歷史劃分爲各個時期並不是一個事實，但卻是一個必要的假設。」在各時期之間的切割點都是人爲的。雖然這些切割點在歷史中並不是自然發生，但卻也不是我們可以任意指定的。

時期劃分　把在社會現實中的時間流動劃分為片段或是時期的作法。田野研究者或許會在一個持續的過程（例如：平常的一天、一年的循環等等）當中發現一些部分或是時期。

你必須一直到已經檢視完所有的證據之後，才

能夠去決定時期的數量、每個時期的時間長短、以及各個時期之間的切割點。你在一開始時，對於應該要劃分爲多少個時期、以及各個時期之間有何特色等等，可能只有廣泛籠統的概念，而在你檢視了證據之後，你應該調整時期的數量與時間長短、以及切割時期的時間點位置等等。之後，你也許會以新增加的資料來重新檢視證據、以及調整時期的劃分方式等等，以此進行概化推論。在你進行了上述的循環數次之後，你就可以在連續理論化以及檢視證據的基礎上，把 100 年的時間劃分成數個不同的時期。

3. *歷史偶然性*　**歷史偶然性**（historical contingency）所指的是特定因素或是特定條件的某種獨一無二的組合方式，而且此種組合方式可能無法再重覆。此種組合的方式是獨特的，而且無法從先前條件的流動中來加以預期的。正如同馬哈尼（Mahoney, 2000a: 513）所解釋的：「偶然性所指的是，不論是以決定論的方式、或者是以機率論的方式，理論都無法去預測或者是解釋某種特定結果的出現。因此，一個偶然的事件就是一個沒有人預料到會發生的事件。」一個偶然的情境可能是在意料之外的，但是一旦發生了，就可能會對後來的事件造成深遠的影響。因爲有可能會發生許多獨特的事件組合方式，因此我們必須運用理論去確認對於一項解釋具有重要意義的偶然事件。

關鍵時刻（*critical juncture*）通常是歷史偶然性的一個部分（請參閱範例 15.4 路徑相依、關鍵時刻、以及歷史偶然性）。我們利用這個概念來解釋，數個可行的選擇可能會以何種方式共同存在於一個特定的時間點上。在其中某個選擇被挑選了之後，許多獨特的事件就會匯集在一起，而此種情形通常會造成強烈的持續影響。我們可以把歷史偶然性和路徑相依結合在一起。

羅伊（Roy, 1997）結合了歷史偶然性以及路徑相依，來解釋美國大型企業的興起。他認爲，在十九世紀中期時存在於投資者以及政府官員之間的權力關係，並未使得大型私人企業興起並取得重要地位。相反地，在一個特定時空中的一組獨特的因素，卻促使大型企業的出現（也就是歷史偶然性）。一旦大型現代企業的制度出現了，此種情況就會使得某些團體的地位提升，並且促成了新的權力安排方式。接著這些團體以及權力安排方式則會繼續運作，以維持組織的企業形式。金融家、富有的投資者、以及高級官員裡的菁英的權力逐漸增加，並且從商業組織的私人企業形式中獲得利益。這些菁英積極地透過新的法令、政府規定、金融關係、以及其他的方式來支持此種形式。此種企業形式則維持了這些菁英們逐漸增加的權力以及特權。因此，一些特定事件在某個時間點上「偶然」匯集在一起，並且在各種可能的形式當中選擇了一種特定的商業組織形式。而此種情況並非是不可避免的。然而，一旦此種特定的商業形式被確立了，就會引發許多新的動力，使得此種商業形式存續下來，並且改變了周圍的條件。此種特定的商業形式使得其他可能的形式變得越來越不可行，因爲此種商業形式會強化某種社會政治安排方式、並且重新調整經濟權力，因此削弱了其他可能形式的力量。如此一來，此種組織的企業形式就創造出了一條路徑，而在其

歷史偶然性　一種在敘事分析中的分析概念，此種概念所處理的是在某個特定時空中集合在一起的因素的特定組合方式，並藉此來解釋一個過程、事件、或者是情境。

範例　15.4

路徑相依、關鍵時刻、以及歷史偶然性

研究者在敘事分析中結合了路徑相依以及關鍵時刻的概念，以發現一個特定的情況短期組合方式如何能夠使後續的事件沿著一條新的軌跡發展，而且他們也試著要去確認這些「關鍵時刻」或者是歷史的轉捩點。奇瑟以及林頓（Kiser and Linton, 2002）就曾經在他們對從 1515 年到 1789 年的法國所進行的研究中使用過這個觀念，並且提到：「特定的歷史轉捩點改變了變項之間的關係」（p. 905）。他們把焦點集中在法國人民對於稅制的叛亂上面。在 1515 年到 1789 年之間，約有 20%的年度裡發生過因為不滿稅制而引發的叛亂，而政府徵收這些稅賦的原因，主要是為了支付正在進行當中的戰爭（在上述的期間之中，約有 65%的時間都發生過戰爭）。投石黨之亂（Fronde）是一連串大規模的叛亂活動（從 1648 年到 1653 年），而法國國王的軍隊則成功地鎮壓了這些叛亂活動。在投石黨之亂發生前，法國政府不斷增加稅賦以及主動對他國發動戰爭的作法，經常引起地區性的叛亂，但是在投石黨之亂被鎮壓了之後，地區性的叛亂就變得非常罕見了。此種情況在理論方面的暗示，是研究者或許會發現一組穩定的、並且能夠在一段時期內運作的因果關係，但是在其他時期內卻幾乎找不到此種關係的證據。除此之外，研究者或許會發現一個就像是關鍵時刻或是引爆點的特定事件或是短暫的時期，而在這個關鍵時刻或是引爆點過後，重要的關係會發生戲劇性的轉變，然後開始以不同的方式來運作。也就是說，沿著某種路徑上的持續模式，在一個關鍵的時刻被打斷了，之後就被重新導向一條新的軌跡。

後所發生的事件則會依循著這條路徑而發展下去。

路徑相依可能會是自我強化的，也就是以慣性而沿著某一個方向而繼續下去。或者，特定的事件可能會引發一個反應，而改變了發展方向。在隨著時間而流動的事件順序當中，也可能會出現周期性的關鍵時刻。一開始被引發的過程或者是條件可能會抗拒改變。或者，偶然發生的條件也可能強大到足以引起方向的重大變化，並且開啓一條事件的新路徑。

反面個案方法

我們通常會把注意力集中在資料中較爲明顯可見的部分，然而，我們也可以去研究在資料中*不*明顯的部分，或者是*沒有*發生的事情。首先，去研究一些不存在的事物，表面上看起來似乎是違反直覺的作法，但是一個能夠察覺到所有線索的機警的觀察者，除了會注意到存在的事物，也會去注意到那些遺漏掉或者是消失的事物。在「銀色馬」（“Silver Blaze”）這個故事裡，偵探福爾摩斯（Sherlock Holmes）注意到，當一匹昂貴的賽馬在被竊走之時，看門的狗竟然沒有吠叫，此種情況暗示了看門狗可能認識竊賊，因此，他就依據這條線索而解決了一樁案件。當原本應該發生的事情卻沒有發生時，此種情況本身就是一項重要的資訊。

反面證據有許多種形式（請參閱資訊箱 15.7 反面證據的類型），其中包括了沉默、缺席、以及遺漏。例如，一位田野研究者注意到，某個特定年齡、種族、或者是性別的人，並未出現在某個社會情境當中。而此種缺席的狀況可能會揭露出許多關於該情境本質的資訊。同樣地，你注意到有一些錢掉在地上，但是卻沒有人去把

資訊箱 15.7　反面證據的類型

1. *沒有發生的事件*　依據過去的經驗，有些事件被預期應該會發生，但是卻沒有發生。例如，對於美國歷史上的進步時期（Progressive Era）所進行的研究，發現到大型的企業並未反對溫和的勞工改革法案。在大型企業已經對勞工顯現出敵意多年之後，人們普遍預期大型企業應該會反對此種改革。相反地，大型企業實際上反而鼓勵此種改革，因為此種改革可以使逐漸增加的勞工騷動的情況平息下來。

   同樣地，當擁有權力的團體並未直接參與事件時，就可能會發生無法作出決策的情況，這是因為這些團體所握有的權力地位會影響到哪些議題能夠浮出檯面。例如，某個城市的空氣污染非常嚴重，但是當地的民眾卻沒有針對這個問題發起任何的公眾行動，這是因為「所有人」都默默地認可了這些造成污染的產業在工作、稅收、以及當地經濟方面所具有的影響力。這些造成污染的產業完全不必擔心當地民眾會要求立法來懲罰污染，因為根本沒有人會提出此種反污染法令的草案。
2. *民眾沒有察覺的事件*　有些活動或者是事件，根本沒有被某個環境中的人們或是研究者所注意到。例如，在某個時期裡，雇主們普遍認為受過高等教育的女性只適合擔任文書工作，而且，因為人們沒有特別去注意到此種現象，所以此種現象並未成為一項議題。一直到整體社會對於性別偏見以及性別平等的意識逐漸提高了之後，才有少數人認為上述的作法會限制了女性的工作機會。另外一個例子，則是美國西部鄉村歌曲的作曲者否認他們寫歌時是在套用一個既定的公式。雖然他們本身並沒有察覺到，但是透過對於歌詞所進行的內容分析，就能夠明顯地看出一個公式。某個環境裡的成員或是參與者未能察覺到某項議題，並非意味著研究者就應該忽視這項議題、或者是不去發掘這項議題所產生的影響力。
3. *民眾想要隱瞞的事件*　人們或許會刻意地扭曲某些事件，以保護他們自己或是其他人。例如，菁英份子通常會拒絕討論不道德的行為，而且或許會銷毀一些文件、或者是長期隱匿一些文件而不讓大眾取得這些文件。同樣地，在很長的一段時間內，關於亂倫的案件都沒有被報導出來，部分原因是因為亂倫的行為觸及了一個非常嚴重的禁忌，所以完全不能張揚出去。
4. *被忽視的平常事件*　每一天，我們都會預期遇到一些日復一日的事件，而這些事件也使我們產生了一種「視為理所當然」（taken-for-granted）的態度。例如，由於電視節目經常出現在人們的對話之中，因此這些節目就很少被人特別注意到。由於大多數人都擁有電視，並且經常收看電視，因此，只有那些很少看電視、或者是身為謹慎仔細的分析家的人，或許才會注意到這個主題。或者，某個研究者觀察了一個歷史時期，在該時期當中，吸煙是非常普遍的。只有當這名研究者本身不吸煙、或者是生活在一個吸煙已經成為公共健康議題的時期裡，他或是她或許才能察覺到此種現象的存在。
5. *研究者先入為主觀念的影響*　研究者必須要小心，不要讓自己先前的理論架構或者是先入為主的觀念，使得他們看不見在某個社會情境裡的相反事件。關於到何處去尋找證據、以及何種資料才具有相關性的先入為主的強烈觀念，會妨礙了研究者，使研究者無法注意到其他相關的或者是反面的證據。例如，一名研究者預期在毒癮者以及其孩子之間會出現暴力衝突，因此，當此種衝突出現時，該研究者就會馬上注意到，但是卻沒有看到雙方其實

也在試圖維持彼此之間的親情關係。

6. *無意識地未加以報導*　在研究者心中，有些事件似乎不具有任何重要性，而且也不值得報告，然而，如果進行詳細的觀察並且記錄下來的話，只要對筆記重新仔細閱讀以尋找反面個案，或許就能揭露出一些被忽略的事件。例如，一開始時，某位研究者並不認為某間公司所舉辦的野餐活動具有任何的重要性。然而，在重新閱讀資料筆記以及仔細思考之後，他或是她可能才會瞭解到，在建立一種群體感方面，野餐活動其實扮演了一個相當重要的象徵性角色。
7. *有意識地未加以報導*　研究者可能會故意遺漏情境或者是事件的某些面向，以保護在該情境中的某些個人或者是關係。例如，某位研究者發現了一個知名人士有婚外情，但是為了想要保全這位知名人士的良好名聲以及形象，因此就不會把這件事報導出來。另外一個更嚴重的問題，則是違反研究倫理。當研究者無法提出那些未能支持他或她的論點或是對於資料詮釋的證據時，就會發生此種情況。研究者應該要同時提出能夠支持、以及無法證實某種詮釋的證據，也就是應該同時陳列出正反兩面的證據。如此一來，讀者們就可以自己衡量兩種類型的證據，並且判斷是否要支持研究者所作出的詮釋。

*資料來源：*Lewis and Lewis (1980)。

這些錢撿起來，而此種現象可能就是一個重要的線索。或許，在歷史比較研究中，你注意到在某些地區或是時期內，沒有關於某種犯罪行爲（例如：仇恨罪、虐待兒童等等）的報告。你或許會發現到，某種報告或是事件沒有出現的情況，與這些報告或事件出現的情況具有同等的重要性。

反面個案方法（negative case method）是一種以系統性方式來檢視原本應該存在、但是卻沒發生的事物的方法。[26] 此種方法結合了分析比較中的差異法，以及異常個案分析。異常個案分析（deviant case analysis）把注意力集中於大量個案（包括量化資料群組）中的少數個案上面，因爲這些個案並不符合一般廣泛的模式。我們利用這些異常的個案來瞭解過程、或者是創造出一些新的觀念。

反面個案方法所運用的是對某個特定個案的詳盡知識，而這個個案並不符合於依據理論所預期應該出現的狀況。除此之外，該理論則是已經擁有來自於許多其他個案的支持性證據。你利用這個單一的反面個案來重新檢視這個理論，並且注意到其中的錯誤或是問題等等。接著，你則是可以應用來自於反面個案的洞見，來修正這個理論。

例如，艾米（Emigh, 2003）觀察到，十五世紀時的托斯卡尼正處於高度發展的義大利北方文藝復興文化的巔峰，而且具備了所有關於快速「起飛」到工業資本主義的主要理論所預測的一些先決條件，像是有效率的農業、發展良好的商業製造、沒有封建時期的貴族、大型的城市經濟、以及一個穩定的政治組織等等。然而，托斯卡尼並沒有如預期般地發展出工業資本主義。艾米想知道爲什麼托斯卡尼會成爲反面個案的原因，並且因此而獲得了對於此種個案的深入知識。她後來發現了一些先前

反面個案方法　一種質化資料分析的方法，該方法的焦點在於某個不符合理論預期的個案上面，並且利用來自於該個案的細節以修正理論。

沒有人知道的因素（與當地鄉村投資有關），而那些主要理論並未把這些因素列入考量。在質化分析中所使用的各種類型的分析策略，以圖表的方式摘要總結於圖 15.4 當中。

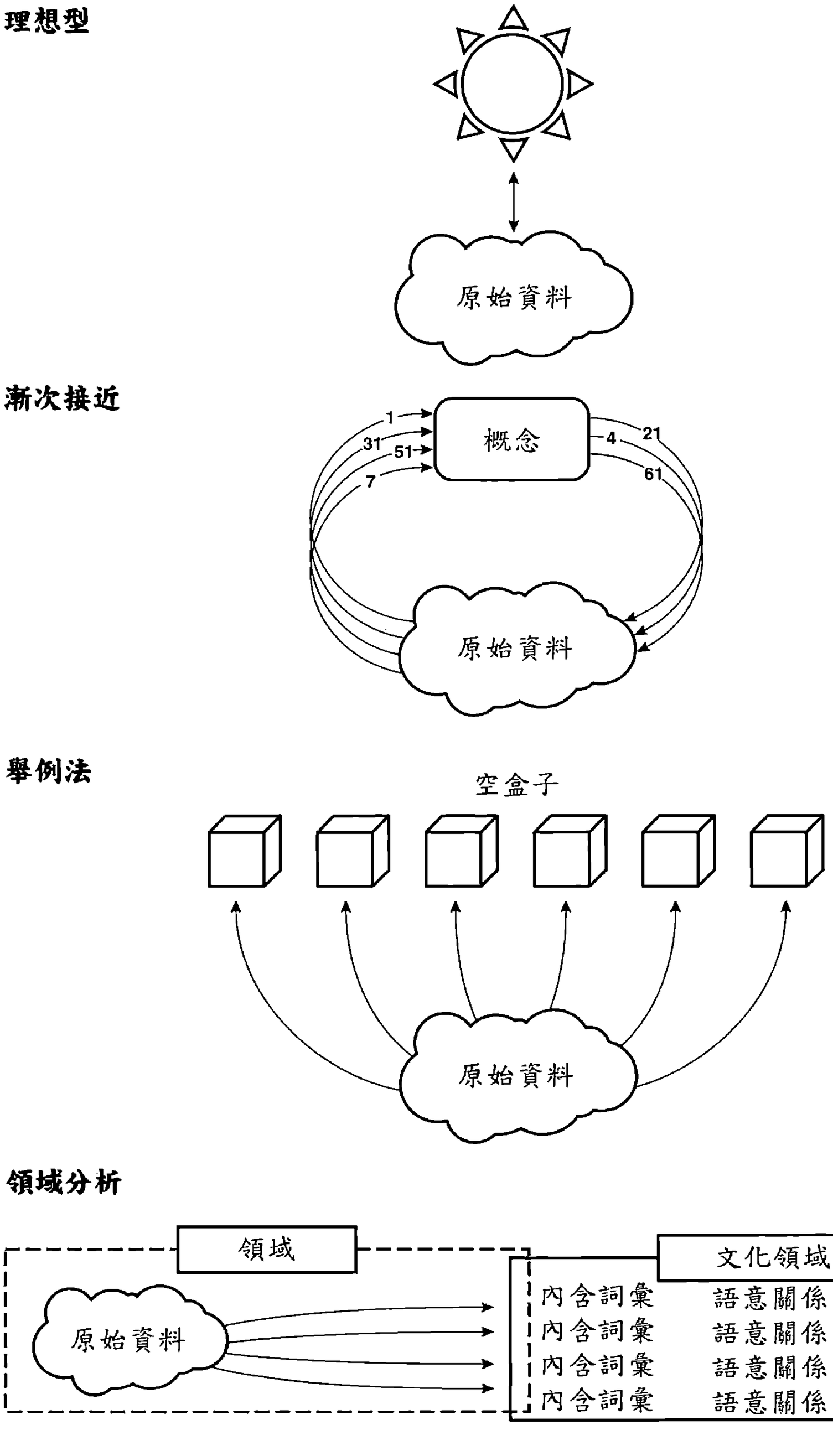


圖 15.4　質化研究中使用的分析策略摘要（a）

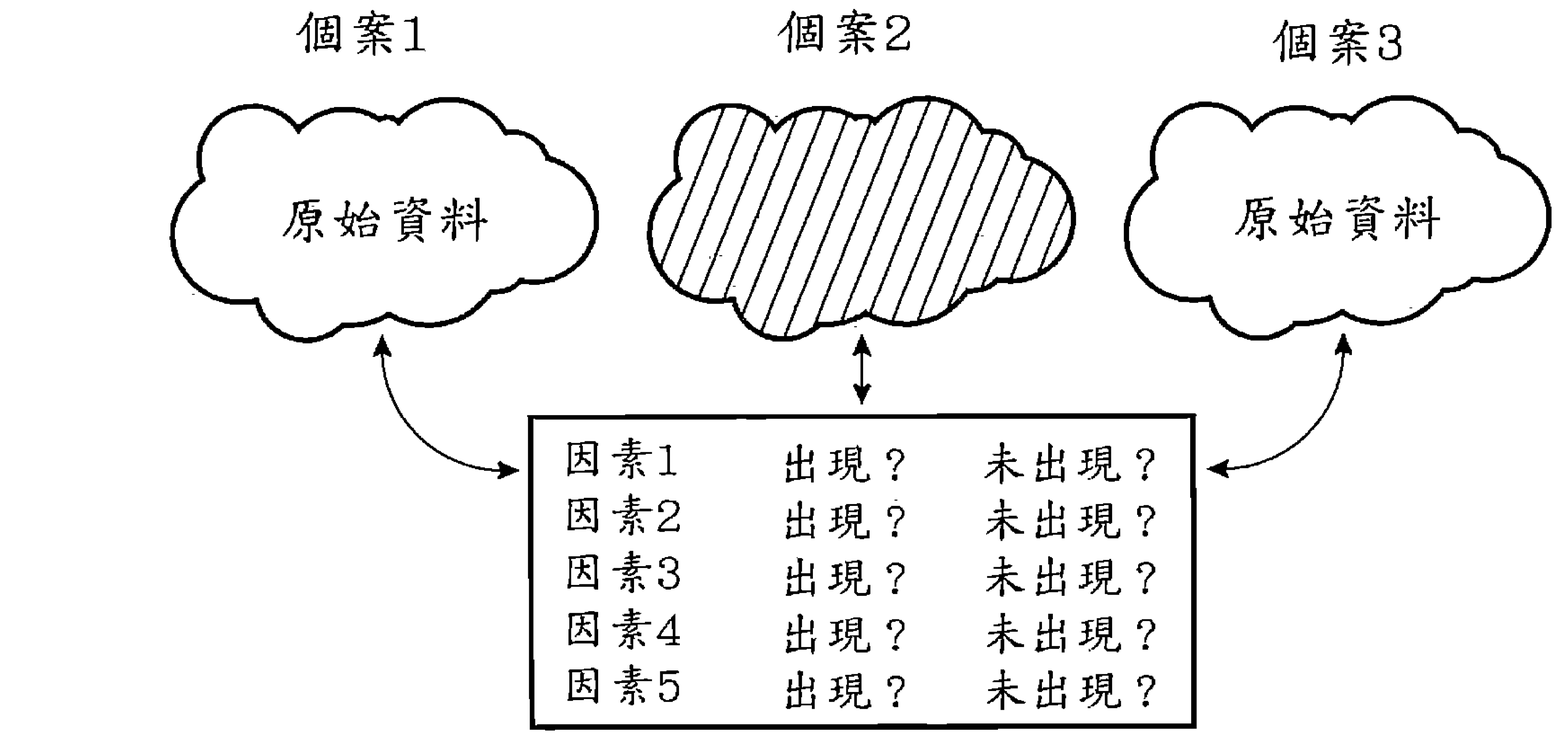


圖 15.4　質化研究中使用的分析策略摘要（b）

其他的技術

質化研究牽涉到使用許多不同的分析技術，以下我們將簡短地介紹一些其他的技術，以證明其多樣性。

網絡分析

在第三章中的網絡理論中、以及在第八章的滾雪球抽樣中，我們就已經討論過

社會網絡的概念了。在質化研究當中，我們通常會去「繪製」（"map"）在一組人群、組織、事件、或是地方之間的關聯性的「地圖」。利用社會關係圖以及類似的繪製地圖技術，我們就能夠發現、分析、以及展現數組不同的關係。例如，在一家公司裡，哈利（Harry）下達命令給蘇（Sue）；蘇和山姆（Sam）彼此互相商量以及互相幫忙。山姆從珊卓（Sandra）那兒得到一些資料。而珊卓則與瑪莉（Mary）有所往來。因此，我們發現網絡圖能夠幫助我們看清楚以及理解複雜社會關係的結構。[27]

時間分配分析

在研究當中，時間是一項重要的資源。我們檢視人們或是組織花費或是投資時間的方式，以揭露出隱性的行爲規則或者是優先次序。我們會詳細記錄人們投入於各種活動的持續時間以及時間量。我們在第十章中曾經討論過時間預算調查，而與此種調查類型相似的是，質化研究也會去檢視投入於各種活動的持續時間以及時間量。針對人們、團體、或者是組織如何去分配他們所控制的珍貴資源（像是時間、空間、金錢、名聲等等）所進行的分析，能夠揭露出關於他們真正的（相對於官方正式宣稱的）優先次序的許多資訊。通常，人們不會察覺到、或者是明確地承認某個他們花時間從事的活動所具有的重要性。例如，你注意到在與經理見面之前，某些人會被要求先等候一段時間，而其他人則不必等候。你或許可以去分析等候的時間長短、什麼人必須等候、當他們在等候時所做的事、以及他們覺得等候是不是公平的等等。或者，當人們說公司裡的某個慶祝活動根本不重要、但是每個人卻都參加了這個活動並且在活動中待了兩個小時的時候，你也可以把此種情況詳細記錄下來。在忙碌的一個星期當中，人們集體分配了兩個小時的時間在這個慶祝活動當中，這件事本身可能暗示了該活動在公司文化中具有一種隱性的、或者是未言明的重要性。[28]

流程圖以及時間次序

除了投入於各種活動的時間長短之外，我們也會去分析事件或者是決定的順序。歷史研究者早已經把焦點集中於詳細記錄事件的發生次序上面，但是比較研究者以及田野研究者也會去檢視事件的流程或者是次序。除了事件的發生時間之外，我們也可以使用決策樹狀圖（a decision tree）或者是流程圖來概略描繪出決策的順序，以瞭解某個事件或是決策與其他事件或是決策之間的關聯性。例如，我們可以概略描繪出一個像是製作蛋糕一樣的簡單活動的流程圖（請參閱圖 15.5）。此種以圖形來描繪出步驟、決策、或者事件，並藉由圖形來調查這些事物之間關係的概念，被研究者廣泛地應用於許多情境當中。例如，布朗以及坎特（Brown and Canter, 1985）就曾經針對購屋行爲發展出一個詳盡的流程圖。這兩位研究者把購屋行爲畫分爲一條時間線上的五十個步驟以及許多行動者（例如，相關的買方、財務官員、調查員、買方的律師、廣告公司／房地產經紀人、賣方、賣方律師等等）。[29]

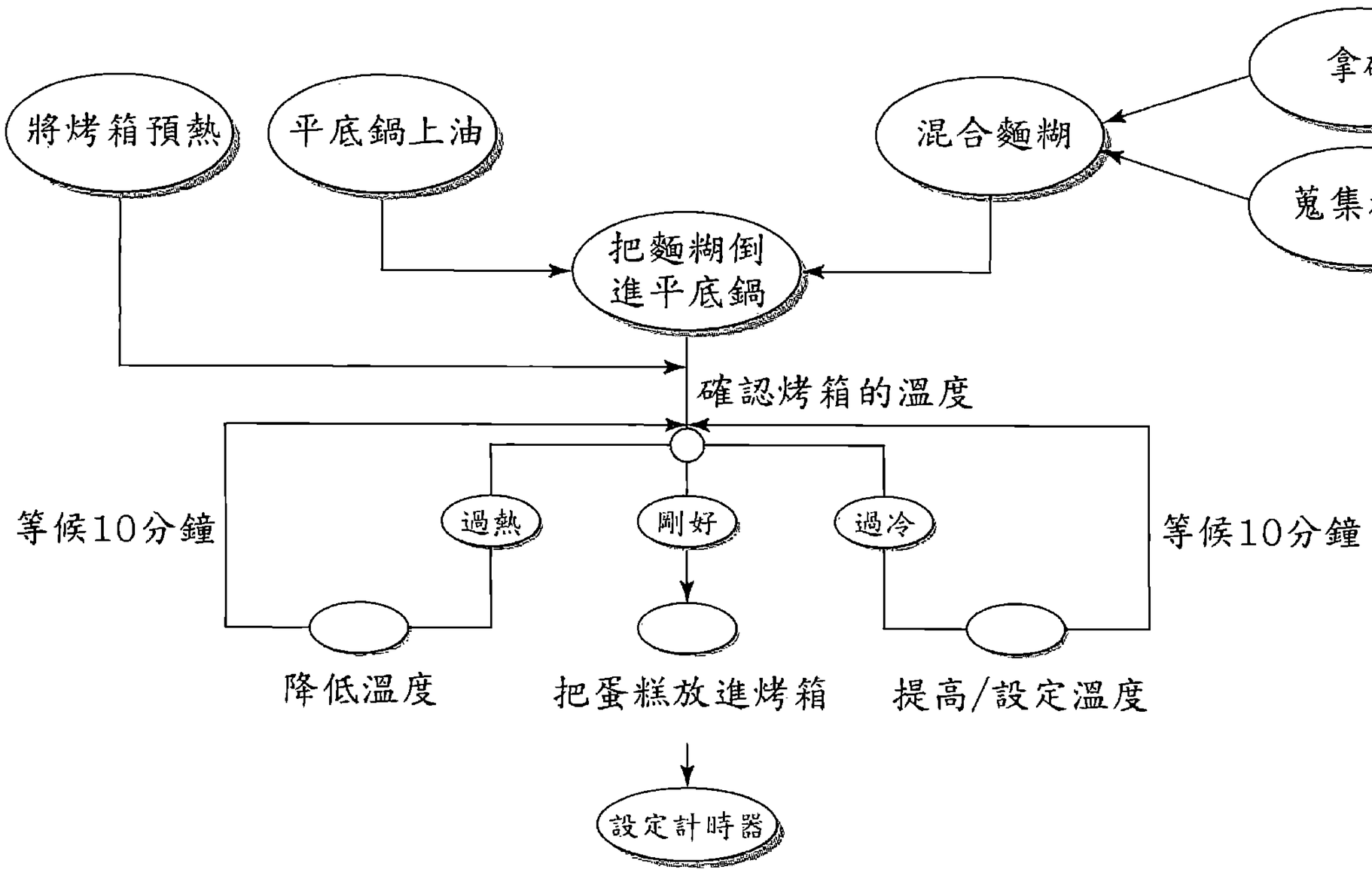


圖 15.5　製作蛋糕的部分流程圖

多重分類程序

多重分類是一種類似於領域分析的技術，而且通常使用於田野研究或者是口述歷史當中。此種技術的目的，是要去發現人們如何將自己的經驗分門別類，或者是如何把物品分類爲相似的以及不同的。認知人類學家以及心理學家通常會使用一 種多重分類程序，而你則可以利用多重分類程序來蒐集、確認、或者是分析資料。以下是此種技術的使用方式：你先給予研究對象一張關於詞彙、照片、地點、或是人物名字等等的清單，並且要求他們把這清單的內容加以組織成爲不同的類別或是堆疊。他們所使用的是他們自己所創造出來的類別。在完成分類過程之後，你詢問他們所使用的分類標準爲何。接下來，你再把同樣的清單給予研究對象，並且要求他們以他們可能想到的其他方法再進行一次分類。此種方式有點類似於賽斯通量表法（我們已經在第七章中討論過了），因爲人們同樣也要去進行分類的過程，但是在此處，堆疊的數目以及物品的類型都是不同的。進行分類的目的，並不是在製作出一個統一的量表，更確切地來說，其目的是要去發現人們如何理解這個世界。坎特等人（Canter et al., 1985: 90）提供了一個例子，在這個例子裡，有一個賭徒對八個賭博場所進行分類，而且這個分類過程總共重覆了五次。在其中一次的分類過程裡，他所用的標準是依據「賭場的等級」（從高到低）。其他幾次的分類標準則是依據「華麗的裝飾與設施」、「籌碼的大小」、「能夠讓我賺到錢」、以及「個人的喜好」。藉由檢視各種不同的分類方式，你就可以看到人們以何種方式來組織他們眼中的社會現實。[30]

圖表

質化研究者通常會以視覺的方式來呈現資料分析的結果，像是圖表以及曲線圖等等。圖表以及曲線圖能夠幫助組織概念，並且以系統性的方式來協助調查資料。除此之外，圖表以及曲線圖也能夠把結果傳達給讀者。我們可以使用空間或是時間地圖（請參閱第十三章）、類型學（請參閱第三章）、以及社會關係圖。因此，在一項針對美國職棒的小聯盟所進行的研究中，范恩（Fine, 1987）運用了社會關係圖來顯示出在球員之間的關係。除了分類學、地圖、以及清單之外，我們也會使用流程圖、組織圖、因果關係圖、以及各式各樣的清單以及表格，來強化分析結果、並且舉例說明研究發現（請參閱圖 15.6）。

地圖

量化研究者以及質化研究者都會把資料放在地圖上，以幫助他們看出資料之間的空間關係，並且補充或是強化來自於其他資料分析的結果。例如，貝倫以及理查森（Ballen and Richardson, 2002）就曾經利用法國以及美國的地圖來檢視自殺率地

範例 1

人物	*上大學前曾經工作過*	*在大學時打工*	*現在懷孕*	*有自己的車子*
約翰	是	是	N/A	否
瑪莉	是	DK	否	是
馬汀	否	是	N/A	是
優子	是	否	是	是

DK = 不知道，N/A = 不適用

範例 2

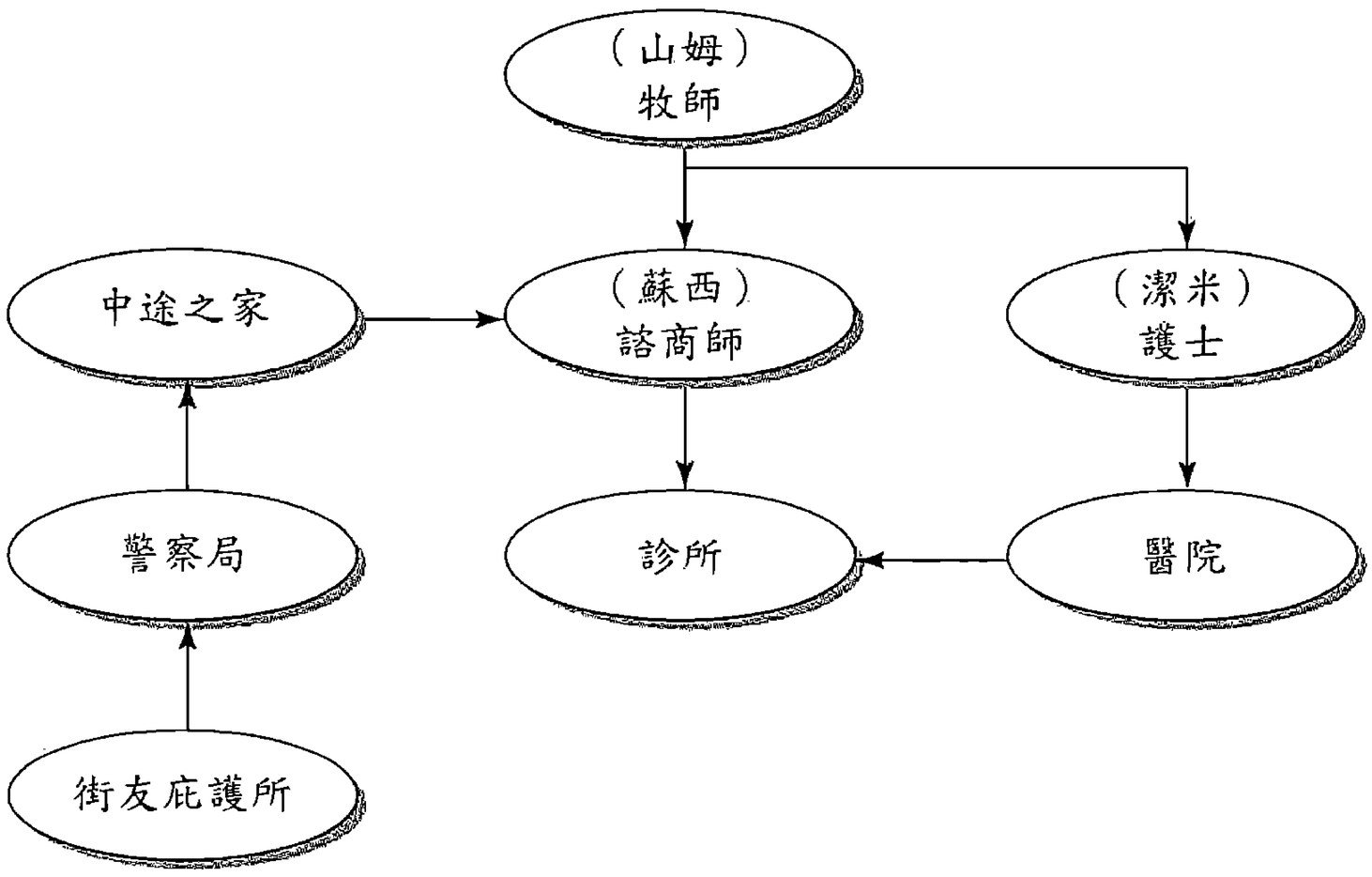


圖 15.6　在質化分析中使用圖表的範例

理分布模式的資料，並且用來支持涂爾幹的社會整合以及模仿的理論。奇瑟以及林頓（Kiser and Linton, 2002）則是呈現了一幅法國的地圖，並且在圖上標示出在他們研究中的叛亂發生的位置（這項研究已經在範例 15.4 中討論過了）。維拉瑞爾（Villarreal, 2002, 2004）也曾經在他針對暴力以及社會政治變遷所進行的研究中使用了一幅墨西哥的地圖。麥可貝以及同事（McVeigh and colleagues, 2003）曾經針對地區性仇恨罪執法情況的差異進行過研究，在他們的研究中，他們也提出了一幅橫跨全美國的州郡地圖。梅爾斯以及卡尼格利亞（Myers and Caniglia, 2004）則在他們對於報紙是否報導了抗議事件的研究中，使用了一幅美國各地區的地圖。在葛理斯渥德以及萊特（Griswold and Wright, 2004）針對獨特地區文化的存續時間長短所進行的研究當中，他們也在一幅美國地圖上標示出特定的地區。在分析以及呈現資料以支持某個解釋時，地圖會是非常有幫助的。然而，作爲一種資訊的視覺呈現方式，地圖也可能會產生誤導的效果，因此，我們在使用地圖時，仍然要多加小心（請參閱 Monmonier, 1996）。

質化資料的軟體

自從 1960 年代中期開始，研究者就已經開始使用電腦科技來產生表格、圖形、統計考驗、以及曲線圖等等來分析數字資料。對比之下，質化研究一直到 1980 年代中期之後才開始採用電腦科技。[31] 如果你把筆記輸入一個文書處理軟體裡，你就能夠快速地尋找某個字或是片語。在資料編碼當中採取此種搜尋方式、或者是把編碼與分析筆記連結在一起，只是一小步而已。文書處理軟體還能夠幫助你修改並且移動編碼以及田野筆記的各個部分。

已經有一些軟體是特別針對質化資料分析的用途而設計出來的，而且新的電腦程式也一直持續地在發展或者是修改當中。其中大多數的軟體都會附有極爲詳細的、針對程式進行說明的使用者手冊，因此，此處的回顧並不會去深入探討關於特定軟體的細節。此處只是概略地描述質化資料分析的主要方法而已。

文本檢索（text retrieval）　有些程式可以針對文本文件進行搜尋，此種功能類似於文書處理軟體中的搜尋功能。特殊的文本檢索程式的執行速度更快，而且擁有找到近似字詞、輕微的拼字錯誤、發音相似的字詞、以及同義詞的能力。例如，如果你尋找「小船」（boat）這個關鍵字，此種程式或許也會告訴你下列這些字詞之中的任何一個是否也有出現在文件中：「船」（ship）、「戰艦」（battleship）、「護航艦」（frigate）、「划艇」（rowboat）、「縱帆船」（schooner）、「船艦」（vessel）、「遊艇」（yacht）、「輪船」（steamer）、「遠洋輪船」（ocean liner）、「拖船」（tug）、「獨木舟」（canoe）、「小艇」（skiff）、「單桅快速帆船」（cutter）、「航空母艦」（aircraft carrier）、「救生艇」（dinghy）、「平底駁船」（scow）、「單層甲板大帆船」（galley）、「平底船」（ark）、「巡洋艦」（cruiser）、「驅逐艦」（destroyer）、「旗艦」（flagship）以及「潛水艇」（submarine）。除此之外，有些程式還會利用邏輯字彙（以及、或、沒有）（and, or, not）來找出複合的字彙或是片語，此種搜尋方式正是所謂的*布林搜尋*（*Boolean searches*）（以布

林（George Boole, 1815-1864）的名字所命名的搜尋方法）。例如，你或許會在很長的文件中去搜尋「大學生」、「喝酒」、「抽煙」這幾個關鍵詞出現的段落，但是你所設定的條件是，這些關鍵詞彼此之間的相隔距離必須在四個句子之內，但是在文件的同一個區塊之內不能出現「兄弟會」一詞。此處的布林搜尋使用了「以及」（and）來尋找出「大學生」這個關鍵詞與其他兩種行爲之間任何一種的交會點，而其他兩種行爲之間則使用「或」（or）來連結在一起。另外，你還必須使用「沒有」（not）這個邏輯搜尋字彙來把「兄弟會」這個詞出現的情況排除在外。

大多數的程式都會顯示出一個關鍵字或是關鍵詞，以及在這些字詞前後的文字。這些程式或許還會允許你另外寫下備忘錄、或者是在文件中添加簡短的註記。有些程式會計算所找到的關鍵字，並且提供關鍵字所在的位置。大多數程式都會單獨依據你感到興趣的字詞而爲文件創造出一個非常具體的索引。

以文本為基礎的管理程式（text-based managers） 以文本爲基礎的管理程式與文件檢索程式相當類似，兩者之間的主要差異之處就在於它們組織以及整理搜尋結果的能力。許多程式會創造出文本資料的次集合，以幫助你針對一個關鍵概念進行比較以及整理筆記，或者是增加一些事實資訊。例如，在關於訪談的詳盡筆記上面，你可以加上日期以及訪談的長度、受訪者的性別、以及進行訪談的地點等等。然後，你可以利用複合關鍵字以及增加的資訊，來整理以及組織每次的訪談或者是訪談筆記的部分內容。

除此之外，有些程式還具有能夠把詞彙連結到其他資訊的*超文件*（*hypertext*）能力，因此，當你用滑鼠點了一個詞彙之後，就會開啓一個新的視窗，而視窗內所顯示的就是相關的資訊。你可以先找出關鍵字或者是主題，然後再把這些關鍵字或主題連結到文件當中。例如，在一個田野研究當中，你想要檢視蘇珊這個人以及頭髮這個主題（其中包括剪髮、髮型、染髮、以及帽子或是頭髮上面的覆蓋物等等）。你可以利用超文件的功能，把所有出現蘇珊名字的地方，都連結到所有關於頭髮的討論內容。如此一來，只要用滑鼠點了蘇珊的名字，文件中的某個區塊就會快速到跳到筆記中的另一個區塊，讓你可以一目瞭然地看到所有蘇珊與頭髮這個主題一起出現的部分。

有一些以文本爲基礎的管理軟體能夠從文件中的資訊創造出交叉列表或者是跨類別的散佈圖。例如，有一些學生們針對某一門課持續地寫下日誌。他們利用四種類別（無聊的、刺激的、有挑戰性的、或者是有創意的）的其中一種來寫下他們每一天的感覺。這些學生也描述了每一天當中的重大活動（例如，團體活動、討論、觀賞影片、上課、或者是上台示範等等）。因此，你可以依據學生的感覺以及活動來交叉分類。藉由增加其他資訊（例如，男性或是女性、主修科目）的方式，你就可以知道，具有不同人格特質的學生是否會對各種活動產生不同的感覺，並且也可以瞭解學生們的感受是否會隨著課堂中所呈現的不同主題而有所變化。

編碼以及檢索程式（code-and-retrieve programs） 我們通常會爲質化資料（也就是文字的田野筆記、訪談記錄、以及錄影帶或是錄音帶的逐字稿等等）賦予一些編碼或者是抽象的詞彙。編碼以及檢索程式讓我們可以把編碼附加在文件中的某一

行、句子、段落、以及區塊上。此種程式允許我們對相同的資料使用多種不同的編碼。除了加上編碼之外，大多數的程式也能夠幫助我們組織編碼。例如，某個這類的程式可以幫助我們製造出編碼的大綱、或者是繪製出在編碼之間以及在編碼所指涉的資料之間的關聯性的「樹狀圖」（例如：主幹、分枝、更小的分枝等等）。此種程式可以依據所使用的編碼以及在編碼之間的關係，來重新編排文件資料。

以編碼為基礎的理論建構程式（code-based theory builders） 進行質化研究的研究者，通常對於理論的評估以及建立有興趣。爲了能夠評估以及建立理論，以編碼爲基礎的理論建構程式會要求使用者首先要爲資料賦予編碼。此種程式提供了數種方法，讓使用者能夠去操弄或者是得到編碼之間的對比或是比較。之後，編碼之間的關係就會成爲測試或者是建立理論的基礎。在編碼之間所創造出來的關係類型，可能會因程式不同而有所變化。某個程式可能會允許使用「若……則……」（if-then）的邏輯關係。例如，柯爾沙洛以及海斯（Corsaro and Heise, 1990）就曾經描述過他們把關於年幼兒童的田野研究資料編碼成爲個別事件的過程。接著，他們檢視了在事件之間邏輯順序以及關係，以尋找隱性規則的原則或者是「文法」。他們所尋找的是一些引導了在事件之間的先後順序、結合、或者是分離的規則。而 ETHNO 這個電腦軟體所要求的則是事件之間的邏輯關聯性（例如：時間順序、必要的先決條件、共同發生等等），然後顯示出在事件之間的模式。

與其他的質化程式相反的是，以編碼爲基礎的理論建構程式擁有強大的能力，能夠操弄編碼以揭露出特定的模式、或者是顯示出在資料當中一些並非明顯可見的關係。如此一來，研究者在對資料的類別進行比較或者是分類時，就會變得比較容易了。

質化比較分析（Qualitative comparative analysis, QCA）是一種使用了布林邏輯或是代數的分析策略、也是一種軟體的類型。[32] 拉金（Charles Ragin）在 1987 年時開創了 QCA。之後，一個由邏輯的、像數學一般的關係所形成的完整系統，就成爲了電腦軟體以及數位電子學的基礎，其中包括了集合論（set theory）、二進位關係（binary relations）、邏輯閘門（logical gates）、維恩圖（Venn diagram）（譯註：這是一種用來表示集合與集合之間關係的圖形）、以及真值表。邏輯的原則讓你可以把概念組織成爲各種集合。例如，當你在搜尋一個電腦資料庫時，你通常會使用到一個關鍵字以及像是*或*、*以及*、*沒有*等等的布林運算子。

QCA 的長處在於其分析多重的、共同發生的因果關係的能力，此種情況暗示了某種條件的組合方式產生了某個結果，而各種條件的其他組合方式也可能會產生相同的結果。依據脈絡（條件或情況）的不同，一個特定的情境可能會對結果產生不同的影響。QCA 會一併考量條件與結果，並且確認不同的因果路徑可能會產生相同的結果。作爲一種分析的方法以及軟體，QCA 的運作方式是一種反覆的過程，並且也要求研究者必須積極參與這個過程。正如同里奧克斯（Rihoux, 2003: 354）所描述的：

質化比較分析（QCA） 以布林邏輯為基礎的質化資料分析方法以及電腦軟體，其用途在於檢視解釋性因素的組合方式以及各種不同的結果量數，以幫助研究者確認複雜的、有附帶條件的因果關係。

> 簡而言之，研究者首先必須先製作一份原始資料表格，在表格當中，每一個個案都必須表現出一種條件的特定組合方式（其值為 0 或 1）以及結果（其值為 0 或 1）。然後，軟體就會產生一個真值表，表中會把資料顯示為一份各種配置的清單。所謂的配置（configuration），指的是某些條件（每一種條件的值都是 0 或 1）以及某種結果（其值為 0 或 1）的特定組合方式。某種特定的配置方式可能會對應到數個觀察到的個案，而不同的個案也可能會展現出相同的配置方式。接著，此種分析的關鍵步驟，就是布林最小化（Boolean minimization），也就是說，把長串的布林表式（也就是藉由真值表所表達的長串描述）縮短成為最短的可能表式（最小的方程式），而且此最短的表式能夠揭露出資料中的因果規律性。之後就必須由研究者來詮釋這個最小的方程式所代表的意義。

QCA 有助於分析數個個案的特徵，並且應用差異法以及一致法。除此之外，QCA 也能夠進行邏輯運算，以確認在一組個案之間的共同特徵以及獨特特徵。雖然其中所牽涉到的代數並不困難，但是在沒有使用此種程式的情況下，卻可能會相當耗費時間，而且也容易出現人為錯誤（請參閱範例 15.5 QCA 的範例）。

概念網絡建構程式（conceptual network builders） 此種類型的程式，能夠藉由呈現出圖表或者是網絡圖的方式，幫助研究者建立以及考驗理論。此種圖表或者是網絡圖所顯示出來的資訊，比單純的圖表資料還要多。這些圖表或網絡圖能夠幫助研究者組織概念或者是思考資料。此種程式使用了節點（nodes），也就是研究者在資料中所發現的關鍵概念。然後，程式會顯示出在節點之間的連結或者是關係。大多數此類程式在呈現圖形時，都會使用方塊或是圓圈，而且其間以帶有箭號的直線連結起來。因此，程式所產生出的圖形看起來與流程圖相當類似，並且用網絡圖來顯示出概念之間的關聯性。例如，資料可能是一個家庭關係的樹狀圖，其中呈現出了數代家庭成員之間的關係。而家庭成員之間的關係（X 是 Y 的兄弟姐妹、Z 與 Y 有婚姻關係、而 G 則是 X 的孩子等等），則可以被用來討論以及分析該網絡的特質。

事件結構分析

事件結構分析（ESA） 一種經常以電腦軟體來進行的質化資料分析方法，此種方法強迫研究者去詳細說明一連串的許多事件之間的連結。此種方法也會去澄清事件之間的因果關係，而其方式是去詢問某個事件在邏輯上是否必須發生於另一個事件之後、或者只是剛好發生於另一個事件之後而已。

許多質化研究者會在敘事分析中以按照時間順序的方式來組織資料。**事件結構分析**（Event-structure analysis, ESA）就被用來組織事件的順序，而且是以一種能夠幫助研究者瞭解因果關係的方式來進行組織與整理。研究者最初是使用這個方法以及 ETHNO（一種與事件結構分析一起使用的電腦程式）來分析田野研究的資料，但是此種方法也能被使用於歷史資料當中。ESA 會先把資料組織成為事件，然後再把事件放入一個時間的順序當中。[33]

範例　15.5

QCA 的範例

羅西格諾以及哈德森（Roscigno and Hodson, 2004）利用 QCA 分析了來自於工作場所的質化資料。他們所感到興趣的是工人們的抗拒行動，其中包括了工會活動和罷工等形式的集體行動，以及破壞、偷竊、與怠工等形式的個人行動。他們所提出的問題是：發生不滿以及抗拒的情況，是否為工作場所組織的功能之一，或者是由工作場所的人際虐待情況所造成的？研究者所蒐集的資料（82 個工作場所的民族誌），代表了他們所能夠取得的、關於組織的民族誌證據的整體人口。正如同作者們所提到的（p. 18）：「我們有足夠的理由去預期，組織結構以及社會關係會彼此互相影響，而且也因而會對工人的不滿以及抗拒策略造成難以預料的效果。」

QCA 強迫羅西格諾以及哈德森必須去詳細說明並且聚焦於那些在理論上被認為具有重要性的變項上。QCA 具有理論上的嚴謹性、個案導向的邏輯、以及對於具有潛在複雜性以及制約性的配置方式的詳細說明。QCA 有助於確認那些代表了在資料中各種屬性的獨特組合方式的類型學。利用把質化資料加以編碼的方式，作者們確認了六種作為解釋因素的工作場所條件：科層體制、良好的組織、衝突、虐待、工會代表在場、以及罷工的歷史。接著，作者們依據這六個因素的所有可能組合方式製作了一份「真值表」，其中每一種因素若不是編碼為 1 =出現，就是編碼為 0 =未出現。真值表當中的配置方式，表示了在涵蓋資料中所有正面的、反面的、以及矛盾的個案時，所必需的最少因素數量。研究者們也發現了六種可能的結果或者是工人們的抗拒形式：罷工、社會破壞、怠工、裝聾作啞、曠職、以及偷竊。依據在 QCA 中所揭露的各種不同的組合方式，作者們最後確認了三種類型的工作場所，而且其中每一種都有各自的工人抗拒類型：充滿爭端的工作場所、有向心力的工作場所、以及散漫無序的工作場所。在充滿爭端的工作場所當中，各種形式的抗拒程度是最高的；有向心力的工作場所的抗拒程度則較低；至於散漫無序的工作場所，主要則是個人形式的抗拒行動。請參閱下列關於本項研究的真值表的摘錄。（請注意，在所有 36 種可能的組合方式當中，此處只有顯示出最前面的 3 種）。

解釋性的量數

B = 科層體制（Bureaucracy）。工作場所是以科層體制的方式來組織的，而且明文規定各種日常程序的操作性控制方法。
G = 良好的組織（Good Organization）。在生產過程中具有一致性以及整合性。
C = 衝突（Conflict）。在工人以及管理者之間的持續性衝突是很常見的。
A = 虐待（Abuse）。管理者對於個別員工在口頭上、情緒上、或是肢體上的虐待或是傷害。
U = 工會代表在場（Union Presence）。在特定的工作場所裡，有工會代表的存在。
H = 罷工的歷史（History of Strikes）。工作場所在過去曾經發生過罷工事件。

抗拒的量數

S = 罷工（Strike）。在進行觀察期間曾經發生過罷工事件。
SS = 社會破壞（Social Sabotage）。透過嘲笑以及揶揄的方式來破壞管理者的形象。
AV = 怠工（Work Avoidance）。當有工作或是任務時，刻意推託或是逃避。
PD = 裝聾作啞（Play Dumb）。工人們假裝無法瞭解特定的工作任務或者是組織程序。
AB = 曠職（Absenteeism）。曠職是對於工作場所問題的一種反應方式。
T = 偷竊（Theft）。工人們趁著在工作時順手牽羊。

解釋性的量數（所有可能的組合方式）						個案數目	在每一種組合方式中的抗拒量數					
B	G	C	A	U	H		S	SS	AV	PD	AB	T
0	0	1	1	0	0	2	0	2	2	1	2	0
1	0	0	1	1	0	4	1	2	3	1	3	1
1	1	1	1	1	0	1	0	0	0	0	0	0

ESA 有助於進行敘事分析，因爲此種方法能夠幫助研究者描繪出一組在已發生事件之間的連結。你把在其他事件之前*必須*發生的事件、與原本*可能*發生的事件區隔開來，而電腦程式則會要求你回答一些關於事件之間邏輯關係的問題。例如，在某個情境中有 A、B、C、X 以及 Y 這幾個事件。而電腦程式會問你：在 X 事件導致 Y 事件發生的過程當中，A 事件是否一定要發生在 X 事件之前？（換句話說，對於 X 與 Y 之間的因果關係而言，A 是否爲一個必要的先決條件）？或者，X 是否能在沒有 A 的情況下去對 Y 產生作用？如果答案是肯定的話，那麼在 X 再次對 Y 產生影響之前，A 就必須再次發生才行。此種過程強迫你去解釋，在兩個事件之間的因果關係是否爲獨特的、一次性的關係，或者是一種會再次發生的關係，而此種關係可以無限次地重覆或者是重覆有限的次數。

事件結構分析具有一些限制。此種方法並未提供理論或者是因果的邏輯，因此你必須自己提出理論或是邏輯。ESA 只能提出地圖或是圖表（利用電腦程式的輔助），讓你能夠更輕易地看出事件之間的關係。當你在決定一些在邏輯上可能成立的關係時，ESA 能夠澄清一連串的事件，並且突顯出那些或許可能會有不同發展的事件。然而，在那些能夠塑造事件順序行動的社會結構當中，ESA 並未佔有任何地位。

葛里芬（Griffin, 1993）對於一起執行私刑（lynching）案件的分析，能夠當作說明 ESA 的例子。在 1930 年 4 月，美國密西西比州的波利瓦郡（Bolivar County）地區曾經發生過一起執行私刑的案件，而受害者名叫哈利斯（David Harris）。葛里芬依據許多的口述歷史、一本書、以及一些報紙的報導，重新建構了這起案件周遭的事件順序。在回答了許多關於一系列事件之間的可能連結的是非題、以及分析了這些連結之後，葛里芬最後得到的結論是，造成這起案件的關鍵性因素，正是原本可以阻止這個過程的當地執法機關並未採取任何行動。關於這起案件的 ESA 圖形的精簡摘要，請參閱圖 15.7。

結論

本章所討論的是我們應該要如何去分析質化資料。在許多方面，質化資料比數字形式的資料還要更加難以處理，這是因爲數字具有數學的特質，因此可以讓我們使用統計程序來加以處理。而質化分析則要求我們必須更加努力地去閱讀以及重新閱讀資料筆記、仔細反思我們所讀到的內容、並且根據邏輯以及判斷來進行比較。

大多數質化資料分析的形式，都會牽涉到進行編碼以及撰寫分析備忘錄，而這兩項都是勞力密集以及時間密集的工作。在進行這兩項工作的過程中，研究者都必須仔細地閱讀資料，並且認真地去思考資料。除此之外，本章也提出了一些我們用來進行質化資料分析的方法。在本章中所提出的技術，只是各種不同的質化資料分析技術中的一部分而已。此外，本章同時也討論了思考反面證據以及在資料中未出現的事件所具有的重要性。

本章同時也是在本書中與研究設計、資料蒐集、以及資料分析相關內容的最後

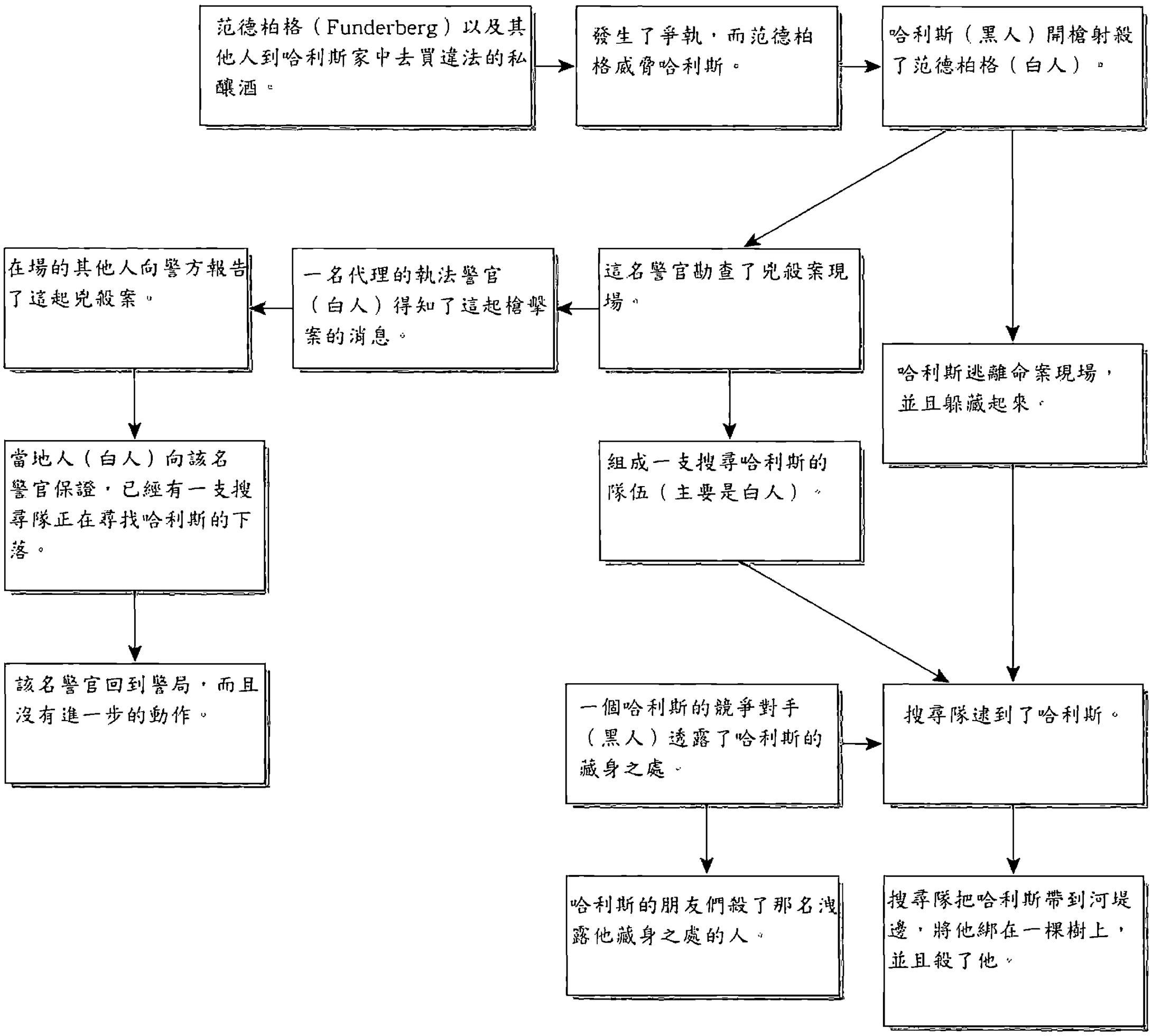


圖 15.7　事件結構分析的範例：哈利斯被處以私刑的案件

*資料來源：*改編自 Griffin (1993)。

一章。社會研究另外還包括了爲研究計畫準備一份報告的步驟或是過程，而我們將在下一章、也是本書的最後一章裡討論相關的議題。

問題回顧

1. 請指出在量化以及質化資料分析之間的四個差異之處。
2. 對於質化以及量化研究而言，概念化的過程有什麼不同之處？
3. 在量化以及質化研究中，資料編碼的過程有何不同之處？質化研究者所使用的三種編碼的類型爲何？
4. 在質化資料分析中，撰寫分析備忘錄的目的是什麼？
5. 請描述何爲*漸次接近*。

6. 在舉例法當中的*空盒子*是什麼？我們應該如何使用空盒子？
7. 一致法與差異法之間的不同之處是什麼？研究者能夠同時使用這兩種方法嗎？請解釋爲何可以或是爲何不可以。
8. 一個領域的各個部分爲何？又該如何使用於領域分析當中？
9. 敘事的主要特色爲何？
10. 對於一個完整的分析而言，爲什麼尋找反面證據或者是未出現在資料中的事物是很重要的？

注解

1. See Miles and Huberman (1994) and Ragin (1987). These should not be confused with statistical techniques for "qualitative" data (see Haberman, 1978). These are sophisticated statistical techniques (e.g., logit and log linear) for quantitative variables in which the data are at nominal or ordinal levels. They are better labeled as techniques for categorical data.
2. Sprague and Zimmerman (1989) discuss the importance of an explicit theory.
3. See Hammersley and Atkinson (1983:174–206) for a discussion of questions.
4. See Boyatzis (1998), Lofland and Lofland (1995:192–193), Miles and Huberman (1994:57–71), Sanjek (1990:388–392), and Wolcott (1994) for additional discussions of coding.
5. See also Horan (1987) and Strauss (1987:25) for multiple indicator measurement models with qualitative data.
6. For more on memoing, see Lester and Hadden (1980), Lofland and Lofland (1995:193–197), Miles and Huberman (1994:72–77), and Strauss (1987:107–129).
7. Also see Barzun and Graff (1970:255–274), Bogdan and Taylor (1975), Lofland and Lofland (1984:131–140), Shafer (1980:171–200), Spradley (1979a, 1979b), and Schatzman and Strauss (1973:104–120) on notes and codes.
8. See Fetterman, 1989:68.
9. See Skocpol (1984) and Skocpol and Somers (1980).
10. For a discussion of analogies and models, see Barry (1975), Glucksmann (1974), Harré (1972), Hesse (1970), and Kaplan (1964).
11. For a discussion of the importance of analogies in social theory, see Lloyd (1986:127–132) and Stinchcombe (1978).
12. For more on successive approximation and a debate over it, see Applebaum (1978a), McQuaire (1978, 1979), P. Thompson (1978), Wardell (1979), and Young (1980).
13. For a discussion of empty boxes, see Bonnell (1980) and Smelser (1976).
14. For a discussion of the illustrative method, see Bonnell (1980) and Skocpol (1984). Bogdan and Taylor (1975:79) describe a similar method.
15. See Coffrey et al. (2002) for an example of domain analysis.
16. For a discussion of methods of difference and agreement, see Ragin (1987:36–42), Skocpol (1984), Skocpol and Somers (1980), and Stinchcombe (1978:25–29).
17. See Mahoney (1999) on a nominal comparison.
18. See Griffin (1993) and Mahoney (1999).
19. On various uses see Abbott (1995) and Franzosi (1998)
20. The six core elements are derived from the following: Abell (2001, 2004), Abbott (1995, 2001), Büthe (2002), Franzosi (1998), Griffin (1992, 1993), Gubrium and Holstein (1998), Haydu (1998), Mahoney (2000), Pedriana (2005), Sewell (1992, 1996), and Stryker (1996).
21. On narrative as a condition of social life, see Abbott (2001) and Somers (1994).

22.Abell (2004:288) remarked, "Although the term narrative and cognate concepts . . . are widely used . . . no settled definition is yet established." Some of the terms used include: analytic narrative (Pedriana, 2005), causal narrative (Sewell, 1996), comparative narrative (Abell, 2001), event structural analysis (Griffin, 1993), historical narrative (Mahoney, 2000b), narrative explanation (Abell, 2004), sequence analysis (Abbott, 1995), and structural analysis of narrative (Franzosi, 1998).

23.On debates about causality in narrative analysis and narrative as explanation, see Abbott (2001:290), Abell (2004), Büthe (2002), Griffin (1993), and Mahoney (2000b). For debate about the narrative, see Haydu (1998), Mahoney (1999), Sewell (1996), and Stryker (1996). Researchers such as Goldthrope (1991, 1997) and Lieberson (1991) question the narrative approach whereas Goldstone (1997) and Rueschemeyer and Stephens (1997) defend its utility.

24.See Haydu (1998:353).

25.Mahoney (2000a) gives a detailed description of the path dependency method and provides many examples of its use. Altman (2000) provides a discussion from the economics literature. Also see Blute (1997) and Pedriana (2005).

26.See Becker and Geer (1982) and Emigh (1997) on the negative case method. Blee and Billings (1986) discuss analyzing "silences" in ethnographic or historical text.

27.See Sanjek (1978) and Werner and Schoepfle (1987a).

28.See Gross (1984) and Miles and Huberman (1994:85, 119–126).

29.See Lofland and Lofland (1995:199–200) and Werner and Schoepfle (1987a:130–146).

30.See Canter and associates (1985) and Werner and Schoepfle (1987a:180–181).

31.See Dohan and Sanchez-Jankowski (1998) and Weitzman and Miles (1995) for a comprehensive review of software programs for qualitative data analysis. Also see Fielding and Lee (1991) and Richards and Richards (1994).

32.See http://www.u.arizona.edu/~cragin/fsQCA/software.shtml.

33.For a more in-depth discussion of event-structure analysis, see Abbott (1992), Griffin (1993), Griffin and Ragin (1994), Heise (1991), and Issac et al. (1994).

第16章

撰寫研究報告以及社會研究的政治

研究報告
社會研究的政治
客觀性以及價值中立
結論

但那是我們份內的工作：把觀念以一種理性的順序加以安排，好讓另一個人可以理解這些觀念。我們必須在兩個層次上來處理這個問題。我們必須依據理論或者是敘事來安排這些觀念、去描述那些造成我們想要解釋的效果的原因以及條件、並且以一種合乎邏輯而且具有實徵正確性的方式來完成這個任務……最後，我們也想讓我們所寫出來的文章能夠清楚地闡明我們所建構的理性順序。我們不想因為我們的文章不夠完美而妨礙了讀者們的理解。這兩項工作是一體的兩面，而且是不能任意區隔開來的。

——Howard Becker, *Writing for Social Scientists*, p. 133

在先前的章節裡，我們已經討論過關於設計研究、蒐集資料、以及分析資料的方法，但是，一直要到我們把研究結果與其他人分享時，一項研究才能算是大功告成了。向他人傳達研究結果以及詳細描述你進行研究的方式，是在研究過程中極爲重要的最後幾個步驟，而其形式通常是以一份書面報告的方式來呈現。第一章中曾經提過科學社群中的研究成果共享主義規範（norm of communalism），此種規範強調的是，我們必須把我們進行研究的方式以及研究的完整發現通通公諸於世。第五章則討論了如何在學術文獻當中找到之前的研究。而本章的焦點則是在於如何撰寫研究報告上面。

進行研究以及報告研究結果，有可能會引起一些爭議。如同我們已經在第五章中提過的，進行研究有可能會引起一些具有爭議性的倫理議題，這些議題主要是牽涉到保護研究參與者、在進行研究時如何維持正直與誠實、以及處理來自於研究贊助者的壓力等等。社會研究也會牽涉到一些政治議題，而且此種議題有可能會引起更大的爭議。社會研究的政治或許會影響到進行研究以及散布研究發現的可能性，而且也會影響到其他人可能會濫用研究發現的方式。

在本章之中總共有兩個重要的主題：撰寫研究報告、以及社會研究的政治。在撰寫研究報告時，你必須去精通相對較爲直接的、不具爭議性的規則以及技巧。然而，在社會研究中的相關政治議題就不是那麼直截了當的，其中包括了像是進行研究的自由、以及在沒有來自於強大的社會團體的干擾下準備一份研究報告的議題等

等。雖然撰寫研究報告有一定的規則以及倫理規範，但是在研究的政治方面就沒有任何的規範或者是規則可循了。

社會研究或許是不完美的，但是社會研究的終極目標就是要去發現知識、擴展理解、以及尋找真相。我們想要探索所有的主題，並且也希望在沒有任何阻礙的情況下，去與科學社群中的成員以及其他人完整地分享我們的研究方法和研究發現。當一些具有強大權力的團體或者是機構試圖去妨礙研究、阻止新知識的自由流通、限制了對於真相的追求、或者是濫用以及選擇性地忽視了研究發現時，就會產生政治方面的爭議。這些團體或是機構之所以這麼做，通常都是爲了推動他們本身非科學方面的目標以及目的。

本章的第一個部分所要討論的是撰寫量化以及質化研究報告的方法。在撰寫研究報告時，你必須以非常嚴肅認真的態度來看待撰寫的過程，而且你也必須解釋你進行研究的方法以及所得到的研究發現。而本章的第二個部分，則是著重於社會研究的政治。我們將會討論到具有強大權力的團體或者是政府會如何去試圖限制研究者所能研究的主題、他們會以何種方式來進行研究、以及他們散布研究結果的方式。而在本章的結尾，則會討論到客觀性以及價值中立的概念。

研究報告

撰寫研究報告的理由

在你完成了一項研究或者是一個大型計畫中的重要階段之後，就應該要透過研究報告把研究發現傳達給其他人知道。你可以藉由一些方式來學習許多關於撰寫研究報告的知識，像是大量閱讀已經出版的文章、以及選修一門關於科學以及技術寫作的課程等等。

研究報告是一種書面文件（或者是根據一份書面文件所做的口頭報告），此種文件能夠把一個研究計畫中的研究方法以及研究發現傳達給其他人知道。研究報告不僅僅是一份關於研究發現的摘要而已，而是整個研究過程的詳盡記錄。

不要一直等到研究結束了才去思考如何撰寫研究報告，你必須在正式撰寫報告之前就先進行思考，並且在進行研究的過程當中保留仔細而完整的記錄。除了研究發現之外，研究報告中還應該包含開始進行研究的動機、對於研究計畫中各個步驟的描述、資料的呈現、以及關於資料與研究問題或是主題之間相關性的討論。

研究報告能夠告訴別人你身爲研究者所做的以及所發現的事物，而這是一種散播知識的方式。如同我們已經在第一章裡提過的，在加強科學社群的凝聚力方面，研究報告扮演了一個相當重要的角色。關於撰寫研究報告的其他理由，則包括了完成一份課堂作業或者是工作任務、完成某個贊助研究的組織所提出的要求、說服一個專業團體去處理某個問題的特定層面、或者是把研究發現公布給一般民眾瞭解等等。然而，與大眾溝通幾乎都不會是傳達科學結果的主要方式，一般而言，通常都

是在把結果傳達給其他研究者之後，研究者才會把研究結果傳達給一般民眾。

撰寫過程

你的讀者群　專業的作者都會告訴你，你必須先知道你的寫作對象是誰，這是因爲只有當你是針對某些特定的讀者群而寫作時，才能達到最佳的溝通效果。根據你的主要讀者群是教授、學生、專業的社會科學家、實務工作者、或者是一般大眾等等，你應該要以不同的方式來撰寫研究報告。當然，更不用說你的報告必須要清楚、正確、以及有條有理。

教授們可能會因爲各種不同的理由而指派學生撰寫報告，而且也可能會對撰寫的方式提出一些要求。一般而言，教授們想要看到的，是一份報告能夠反映出一種清楚而且符合邏輯的思考方式。學生們所撰寫出來的報告應該要展現出自己能夠正確地掌握實質概念以及方法論概念，而達到這個目標的良好方法之一，就是*在適當的時候*明確地使用專業詞彙，但是不要過度地或者是不正確地使用這些詞彙。

當你的寫作讀者群是學生時，你必須爲專業詞彙下定義，並且爲報告中的每一個部分加上標籤。相關的討論應該要以一種符合邏輯的、循序漸進的方式來進行，並且附上許多具體的範例。你應該使用明確的語言來解釋你進行研究計畫中各種不同步驟的方式以及原因。可行的策略之一，是先從研究問題開始，然後把研究報告組織成爲對於研究問題的一個解答。

學者們不需要關於專業詞彙的定義、也不需要你去解釋爲什麼你要使用標準的程序（例如：隨機抽樣）。他們真正感到興趣的部分，是研究者如何去促進理論或者是在文獻裡先前研究發現的發展。他們想要看到的，是對於研究設計的濃縮的、詳盡的描述。他們也會密切關注你蒐集資料、測量變項、以及分析資料的方式。學者們希望你在資料分析的部分是簡潔、緊湊、但是又兼容並蓄的，而在關於結果的討論部分則是嚴謹而且一絲不苟的。

實務工作者則比較偏好你針對進行研究的方法提出一個簡短的摘要，並且以少數簡單的圖表以及圖形來呈現出研究結果。他們對於研究設計、測量、資料蒐集、或者是研究發現等等的細節比較不感興趣。他們喜歡看到的，是研究結果所暗示的替代行動路徑的概要，以及每種路徑所產生的實際結果。重要的是，研究者必須去提醒實務工作者不要從單一研究的結果去進行過度的概化推論。雖然幾乎沒有什麼實務工作者會去要求你列出關於研究設計以及研究結果的細節，但你還是應該這些相關的資訊以附錄的形式放在研究報告之後。

當你的讀者群是一般大眾時，你應該要使用簡單的語言、提供具體的範例、並且把焦點放在研究發現對於當前社會問題的實際應用上面。在爲一般大眾撰寫報告時，你不必把研究設計或是研究結果包含在內，而且也要小心別作出一些沒有證據支持的宣稱。把研究成果告知社會大眾，也是一種重要的服務，因爲此種服務能夠幫助一些並非專家的人們對於公共議題作出更好的判斷。

風格與筆調　當我們在撰寫研究報告時，只能使用少數特定的風格以及一種獨特的

筆調，這是因爲研究報告的目的是要以一種清楚的、直接的、以及誠實的方式來傳達研究方法以及研究發現。

風格（*styles*）所指的是撰寫者所選擇使用的文字類型，以及他或她所使用的句子或是段落的長度與形式。至於*筆調*（*tone*）所指的，則是撰寫者對於主題的態度或是關係。例如，一種非正式的、像對話一般的風格（例如，口語化的文字、慣用語、陳腔濫調、以及不完整的句子等等），搭配著一種個人的筆調（例如：這些都是我的感受），或許適合用於寫信給一個親密的朋友，但是卻不適合用於研究報告中。研究報告的風格是正式的、而且簡潔的（以少數的文字來傳達大量的資訊），而筆調則要能夠表達出作者與主題之間的適當距離，而且必須既專業又嚴肅。田野研究者有時候會使用一種非正式的風格以及一種個人的筆調，但是這只是少數的例外情況而已。研究者也應該避免使用像是說教一般的、或者是過度華麗裝飾的語言，因爲研究報告最主要的目標是要提供資訊，而不是去宣揚某種立場、說教、或者是帶來娛樂效果。

研究報告應該要客觀、正確、以及清楚。你應該要一再確認細節（例如：在引用資訊中的參考資料頁碼），並且完全地揭露你進行研究的方式。如果讀者在文章中察覺到有任何的草率之處，他們或許會對研究本身產生懷疑。研究計畫的細節可能會相當複雜，而此種複雜性則意味著永遠都會有造成混淆的風險，因此，以清楚的方式來寫作是十分重要的。要使文章內容清晰易懂的最佳方法，就是要擁有清楚明白的思考方式，而這也意味著你必須要仔細地再三思考研究問題以及研究設計、明確地定義各種詞彙、以簡短的敘述句子來寫作、並且只寫出有證據支持的結論。

組織想法 雖然有些人對於寫作仍然抱持著一些錯誤的觀念，但是寫作並不會神奇地自動發生、也不會是當某個人把筆放到紙上時，就會自動從筆尖流露出他或她的想法。相反地，寫作是一種辛苦的工作，不但需要不斷的勤奮努力，而且也牽涉到遵循一連串的步驟，最終才能產出最後成果。撰寫研究報告與其他類型的寫作之間並沒有太過極端的差異。雖然有些步驟是不同的、而且複雜程度也可能會比較高，但是一個良好的寫作者在撰寫一封冗長而複雜的信件、一首詩、一組說明、或者是一篇短篇故事時所做的事，大多數也都能應用於撰寫研究報告上面。

首先，撰寫者需要某個可以書寫的題材。研究報告中的「題材」包括了主題、研究問題、研究設計以及測量方法、資料蒐集技術、結果、以及意義。由於要撰寫的部分很多，因此，良好的組織方式是相當重要的。在組織撰寫內容時，最基本的工具就是大綱。大綱可以幫助撰寫者確保所有的觀念都被包含在內，而且觀念之間的關係也都能夠被清楚地呈現出來。大綱是由主題（單字或是片語）或是句子所組成的，而我們之中大多數人都相當熟悉大綱的基本形式（請參閱圖 16.1）。

雖然大綱對於撰寫者有所幫助，但是如果運用不當的話，也可能會變成一種阻礙。一份大綱只是一種幫助作者組織想法的工具而已，它有助於（1）把各種觀念依照次序排列出來（例如：第一點、第二點、第三點分別是什麼內容），（2）把相關的觀念歸類在一起（例如：把一些彼此相似的觀念歸類在一起，並與其他觀念作出區隔），以及（3）把較爲廣泛的、或是較高層次的觀念，與較爲特定而具體的觀念區

圖 16.1　大綱的形式

I. 第一主要標題	最重要的之一
A. 標題 I 的次標題	第二重要的
1. A 的次標題	第三重要的
a. 1 的次標題	第四重要的
b. 1 的次標題	同上
(1) b 的次標題	第五重要的
(2) b 的次標題	同上
(a) (2)的次標題	第六重要的
(b) (2)的次標題	同上
i. (b)的次標題	第七重要的
ii. (b)的次標題	同上
2. A 的次標題	第三重要的
B. 標題 I 的次標題	第二重要的
II. 第二主要標題	最重要的之一

隔開來，或者是把具體觀念與非常具體的細節區隔開來。

有些學生認爲，他們在撰寫研究報告之前必須先擬出一份完整的大綱，而且一旦提出大綱之後，就不能偏離大綱的內容。其實，幾乎沒有任何一位優秀的作者是從一份完整的、詳盡的大綱來開始進行寫作的。最初的大綱通常都是比較粗略的，因爲一直到研究者把所有的內容都寫出來之前，他們根本不可能把所有的觀念都依照順序排列出來、把觀念歸類在一起、或者是把廣泛的觀念與具體的觀念區隔開來。對於大部分的研究者而言，新的觀念會在寫作的過程中逐漸發展出來、或者是變得越來越清楚。

一份初始的大綱與最後的定稿之間的差異，不只是在完整程度方面的區別而已。撰寫的過程不只會揭露出、或者是爲研究者澄清一些觀念，而且還會激發出新的想法、各種觀念之間的新連結、不同的順序、或者是在廣泛觀念與具體觀念之間的新關係。除此之外，撰寫的過程也可能會促使研究者去重新分析或是重新檢閱文獻或是研究發現。然而，此種情況並非意味著研究者必須因此而全部從頭來過。更確切地說，此種情況所代表的意義是，研究者應該要對新的洞見保持開放的態度，並且對於研究計畫的所有面向都必須坦誠以對。

回到圖書館　在開始進行一項研究計畫之前，你對相關文獻應該都已經很熟悉了，話雖如此，在完成資料蒐集以及資料分析的工作之後，你仍然很有可能必須回頭去再次檢視文獻內容。之所以會發生此種情況，其實是有數個原因的。首先，從一項研究計畫的開始到結束，會經過一段時間，而在這段時間裡，有一些新的研究可能已經出版了。第二，在完成了一個研究計畫之後，你將會比較清楚地知道，對於研究而言，哪些事物是重要的、而哪些事物又是不重要的，而且當你在重新閱讀文獻時，你或許會在腦海中形成新的研究問題。最後，當你在撰寫研究報告時，你可能會發現到你的筆記不夠完整、或者是在一份參考資料的引述資訊中遺漏了一些

剽竊 在沒有適當地說明原始出處的情況下，以完全照抄某人的文字或是想法的方式來偷竊他人想法的行為。

改寫 以某人自己的文字來重新陳述某位作者的想法，並且給予原來的作者適當的肯定。

細節。比起你在研究一開始時到圖書館尋找資料的過程，當你在蒐集完資料之後重新回到圖書館時，你比較不需要去廣泛地閱讀文獻，而是比較有選擇性的、而且也比較能夠聚焦於你所需要的資訊上面。

在撰寫研究報告時，我們大多數人都會丟棄掉一些在完成研究計畫之前所蒐集到的筆記或者是參考資料。此種情況並非意味著最初到圖書館找資料以及文獻回顧的工作是一種浪費時間與體力的作法。我們可以預期到，隨著研究的焦點變得越來越明確，有一些（例如，大約 25%）我們在完成研究之前所寫下的筆記可能會變得毫無關聯性。我們不應該在文獻或是參考資料中放入那些沒有關聯性的筆記，因爲這些筆記可能會妨礙了想法的流暢性，並且也會影響到內容的清晰性。

回到圖書館裡去確認並且擴展參考資料的作法，有助於使你的觀念更爲聚焦，而且也能夠幫助你避免剽竊的問題。**剽竊**（plagiarism）是一種很嚴重的欺騙行爲，意思是指在沒有適當引述原始出處的情況下，完全照抄別人所寫下的文字內容。許多大學會把被逮到做出此種行爲的學生處以退學的處份。如果某位專業人士在學術期刊內所發表的文章被發現有剽竊的情況時，那麼整個學術科學同儕社群在對待此人時的方式，就好像是這個人已經犯下了非常嚴重的過錯一般。[1] 因此，你應該要仔細地寫下筆記，並且確認某些字句或是概念的正確出處，以避免無意間犯下了剽竊的錯誤。不論是直接引用的文字、或者是你改寫過的觀念，都應該要註明出處。而在直接引述的情況中，你應該要在引用資訊中列出該句引文所在位置的頁數。

毫無疑問地，使用他人的書面文字而沒有給予該有的功勞，是一種錯誤的行爲，但是改寫的作法就比較不是那麼明確了。**改寫**（paraphrasing）並非完全引用他人的文字，而是以你自己的文字來重新陳述他人的概念，同時加以濃縮。雖然我們經常會去改寫他人的文字，但是如果我們想要做好改寫的工作，就必須真正地去瞭解我們所改寫的內容，而這也意味著我們不僅僅是以同義字來取代他人所寫下的文字而已。真正的改寫應該是借用他人的想法、去蕪存菁、並且適度地給予原作者應得的肯定。[2]

撰寫的過程

撰寫是一種過程，而要學習如何撰寫的唯一方法，就是親自去練習撰寫。[3] 撰寫需要付出時間以及努力，而且也會隨著練習而漸漸進步。雖然撰寫沒有唯一正確的方法，但是卻有一些方法可以幫助我們撰寫出品質良好的文章。一般而言，撰寫的過程總共可以分成三個步驟：

1. *事前準備* 在準備開始寫作之前，你應該先整理關於文獻的筆記、列出觀念的清單、草擬大綱、完成參考書目的引用資訊、並且組織關於資料分析的評論。
2. *正式撰寫* 先把你的想法寫在紙上當作是初稿，而且是一份從頭到尾的完整報告，而不是一些粗略的筆記或者是一份大綱而已。在撰寫時，你可以採用自由寫

作的方式、寫下參考書目以及註釋、準備好要呈現的資料、並且完成導論和結論的部分。

3. *重寫*　你可以利用增強連貫性、校對內容以找出細節上的錯誤、檢查引用資訊、以及重新審視語態和用法等等方式，來評估並且改善研究報告的品質。

事前準備　在撰寫過程中的一個早期步驟，在此步驟當中，寫作者會組織筆記、列出觀念的清單、擬出想法的大綱、並且確認參考書目的引用資訊是完整無誤的。

自由寫作　在撰寫過程中的一個早期步驟，在此步驟當中，寫作者試著盡快地把自己的想法寫在紙上，而先不用去管文法或者是拼字的問題。

許多人發現，一開始要下筆寫作是相當困難的。許多新手通常都會直接跳到第二個步驟，然後就不再往前進行了，此種情況使得他們寫出的文章大多會有品質不佳的問題。**事前準備**（prewriting）意味著你應該從一個裝滿了筆記的檔案夾、大綱、以及清單開始。你必須去思考報告的形式以及讀者群。而花費在進行思考上面的時間是十分重要的，只要你持續不斷地進行思考，那麼在你正式開始進行繁重的撰寫工作之前的一段時間裡，通常會突然出現靈光乍現的一些時刻。

當有些人坐下來開始寫作時，他們會受到一種奇怪的病症所困擾：一種暫時無法寫出任何事物的狀態，而此種情況通常被稱作是寫作者的障礙（writer's block）。當發生此種狀況時，你可能會覺得腦中一片空白、手指僵硬、而且會慢慢地感到驚慌失措。不論是初出茅蘆或者是經驗豐富的作家，都可能會偶爾經歷到此種情況。如果你也面臨到此種情況的話，你應該先冷靜下來，並且試著努力去克服它（請參閱資訊箱 16.1 克服寫作者障礙的建議）。

爲數眾多的作者在開始寫作時，都是利用一種被稱爲**自由寫作**（freewriting）的方式，這是一種把進入你腦海中的所有事物盡快寫下來的過程。自由寫作能夠把在腦海中迅速流動的想法與寫作連結在一起。當你在進行自由寫作時，你不需要停下來重新閱讀你寫下的東西、不必思考最佳的用字遣詞方式、也不用去擔心文法、拼字、以及標點符號是否正確。你只需要盡快地把想法寫在紙上，好讓你的創意或想法源源不斷地湧現出來就可以了，因爲你可以稍後再去整理所寫下來的東西。

寫作和思考是相互交織在一起的，因此不可能去明確地把這兩者畫分開來。而此種情況也意味著，如果你打算先坐下來盯著牆壁、電腦螢幕、天空、或者任何事物，直到所有的思緒都釐清之後才要開始寫作的話，你幾乎不可能寫出任何東西。換句話說，思考的過程可能會在寫作過程中被激發出來。

重寫　或許在一百萬名作者當中，會出現某個具有創造力的天才，第一次寫出來的初稿就能夠展現出驚人的正確性以及清楚性。然而，對於我們這些平凡人而言，寫作就代表了必須要重寫——而且是一再地重寫。例如，據說海明威（Ernest Hemingway）在撰寫《戰地春夢》（*Farewell to Arms*）這本小說時，在結局的部分就重寫了高達 39 次之多。[4]而對於專業的研究者而言，重寫一份研究報告十幾次，根本可以說是司空見慣的事。所以，你不必因爲必須要重寫研究報告而感到沮喪。話說回來，如果真的有任何沮喪感的話，其實重寫反而是可以減少壓力的，這是因爲

資訊箱 16.1　克服寫作者障礙的建議

1. *及早開始*　不要拖拖拉拉或者是等到最後一刻才開始動手。及早開始不但能夠讓你有時間可以再回到寫作的工作上面，而且也能減少緊張的感受，因為你知道你之後還有時間可以回頭把品質不佳的初稿重新寫過一遍。薛弗爾（Shafer, 1980: 205）曾經責備道：「寫作是一件困難的工作，而作者們用來拖延寫作的藉口也多到令人瞠目結舌的地步。」你應該為初稿訂下一個最後期限，時間應該在訂在最後完稿前的一個星期，而且一定要確實遵守！
2. *休息一下再重新開始*　有些寫作者發現，如果他們去散個步、吃些點心、看看報紙、並且大約半個小時後再回頭來寫作的話，原先的障礙就會消失無蹤了。小小的消遣或是娛樂活動，只要是小型而且短期的話，有時候是對你有所幫助的。
3. *從中間開始*　你不一定要從最前面的部分開始。你也可以從中間部分開始，然後儘管開始去寫，即使此種作法似乎沒有直接的關聯性也無所謂。一旦撰寫／思考的過程開始進行了之後，或許你會比較容易進入主題。
4. *進行一些個人的神奇儀式*　有些人在開始寫作之前，會有一些特殊的習慣或者是儀式（例如：洗盤子、清理書桌、把鉛筆削尖等等）。這些事物可以充當一種心靈的啟動裝置，以幫助你開始進行寫作。你應該進行任何能夠幫助你開始寫作的活動。
5. *把寫作內容分成數個小部分*　不要覺得你必須一坐下來就得一口氣完成整個寫作的任務。你應該從一些你覺得較為容易的部分開始，之後再把各個部分拼湊起來。
6. *不要期待完美的作品*　先寫下一份草稿，而這意味著你可以丟棄、修改、以及改變你所寫下的東西。修改一份粗略的草稿，總會比第一次就寫出完美的作品來得容易多了。

重寫意味著你可以先進行寫作並且寫出一份約略的草稿，然後你隨時可以回頭來修正這份初稿，而不必在一開始時就承受了必須寫出完美作品的壓力。一般而言，你應該要規劃至少重寫草稿三到四次左右。

重寫能夠幫助你以更爲清晰、順暢、精確的方式以及更爲精簡的用字，來表達你自己的想法。當你在重寫時，你應該把焦點放在清楚的溝通上面，而不是去使用華麗的或者是複雜的語言。如同雷吉特等人（Leggett et al., 1965: 330）曾經說過的：「千萬不要羞於使用簡單的語言來表達一個簡單的概念。記住，使用複雜的語言這件事本身並不代表智慧。」

重寫意味著緩慢地閱讀你已經寫下的東西，而且，如果有需要的話，你也可以把這些內容大聲念出來，以聽聽看是否順暢。和別人分享你所寫下來的內容，是一個不錯的點子。專業的寫作者通常會請其他人去閱讀並且批判他們的作品。而新手的寫作者很快就會瞭解到，友善的、具有建設性的評論是非常珍貴的。一開始，當你要和別人分享你的寫作內容時，你可能會覺得相當困難，因爲這代表了你要把自己寫下來的想法呈現給別人看到，並且鼓勵他人對這些想法提出評論，然而，這些評論的目的其實是爲了幫助你使內容變得更加清楚，而且這些評論者也是在幫你的忙。

重寫包括了兩個過程：修改以及編輯。**修改**（revising）是指插入一些新的想法、增加支持性的證據、刪除或是改變某些觀念、移動一些句子以使意義更加清楚、或者是加強各種概念之間的轉換以及連結等等。**編輯**（editing）則是指去整理寫作中較爲技術性的層面，像是拼字、文法、用法、動詞一致性與時態、句子長度、以及段落的組織等等。當你在重寫時，你應該仔細檢查草稿並且大刀闊斧地進行修改，以提升文章的品質。如果你在重寫時離完成草稿已經有一段時間的話，要進行重寫會比較容易一些。在草稿當中原本看起來似乎令人滿意的一些字句，在經過了一、兩個星期之後，看起來可能會在意義上變得模糊不清或者是沒有太多的關聯性（請參閱資訊箱 16.2 重寫的建議）。

修改　一種修正的過程，此種過程是重寫的一部分。在此過程當中，寫作者會添加一些想法或是證據，並且刪除、重新安排、或是改變一些觀念，以改善文章的清楚程度並更為明確地傳達意義。

編輯　撰寫過程中的一個步驟，而且也是重寫的一部分。在此過程中，作者會整理內容並使用更為簡潔的語言、檢查文法正確性（例如，動詞的一致性、用法等等）、調整句子長度、並且重新組織段落，以提升溝通的效果以及強化風格。

即使你可能不擅長打字，如果你會使用文書處理軟體的話，最好還是能夠用打字的方式把內容輸入電腦或者是列印出來，最少在最後定稿的前一版要這麼做，這是因爲在一份乾淨的、打字出來的稿子上，會比較容易去找出錯誤以及文章組織方面的問題。除此之外，在列印出來的稿子上，你也可以自由地去剪剪貼貼、刪去某些文字、或者是移動字句。

在撰寫報告以及其他文件時，擁有優秀的打字技巧以及使用文書處理軟體的能力，是極度重要的。認真的專業人士都會發現到，他們投資在磨練打字技巧以及學習使用文書軟體上面的時間，之後都會帶來巨大的回報。文書處理軟體可以使得編輯的工作變得簡單許多，你甚至還可以利用文書處理軟體來檢查拼字、以軟體中附加的同義字典來找到同義字、還有檢查文法正確性。雖然你不能完全依賴電腦程式來完成所有的工作，但是電腦程式卻能使得撰寫的工作變得較爲輕鬆一些。文書處理軟體能夠幫助你更快速而且更容易地完成寫作的工作，因此，一旦熟悉了使用此種軟體的方法之後，幾乎沒有人會回頭去使用手寫的方式或者是使用打字機來撰寫文章了。

最後一個建議是，在你完成了一份報告的草稿之後，你應該要重寫導論以及標題的部分，以使這些部分能夠正確地反映出你所表達出來的內容。[5] 標題應該要簡短而且具有描述性，如此一來才能把主題以及重要的變項傳達給讀者知道。標題可以描述研究的類型（例如：「一個關於……的實驗」），但是不應該包含不必要的字詞或是片語（例如：「一項對於……的調查」）。

量化研究報告

良好寫作的原則可以應用在所有的報告上，但是一份報告中的各個部分，可能會因爲該項研究是量化或是質化研究而有所不同。在你開始撰寫任何報告之前，你都應該先去閱讀相同類型研究的報告，以當作模仿的範本。

資訊箱 16.2　重寫的建議

1. *技術性的細節*　在每次重寫時，你都應該要記得去檢查文法、拼字、標點符號、動詞一致性、動詞時態、以及動詞／主詞的區隔等等。記住，每當你增加了一些新的內容時，就有可能會在不知不覺當中產生新的錯誤。錯誤不但會分散了讀者的注意力，而且也會減弱讀者對你表達的觀念所具有的信心。
2. *用法*　當你在重寫時，記得要重新檢視詞彙，尤其是關鍵詞彙，以看看你是否使用了正確無誤的詞彙來表達你原本想要傳達的意義。不要使用多餘的專門術語或者是冗長的單字，也就是說，你應該使用最能夠表達意義的平凡語言。除此之外，你也應該找一本同義字典來使用，這是一種像是字典一樣的重要參考書，裡面會列出具備相似意義的單字，並且能夠幫助你找到表達意義的正確詞彙。你必須要使用精確的語言，才能夠精確地表達你的想法與意義。如果你要使用的字是「平均」（“mean”），就不要說「一般」（“average”）。而當你所想表達的是「人們」（“people”）或者是「警官」（“police officer”）時，就不要說「人類」（“mankind”）或是「警員」（“policeman”）。此外，也不要把「原則」（“principle”）這個字錯誤地拼為「校長」（“principal”）。
3. *語態*　研究報告的作者經常犯下的錯誤之一，就是使用被動語態來取代主動語態。雖然被動語態或許看起來比較客觀而且有說服力，但是被動語態卻會模糊了行動者或者是行動的主體。例如，以下這句以被動語態來表達的句子：「在校成績以及更為明確的生涯規劃之間的關係，已經被資料所證實了」，最好能夠以主動語態來表達：「我們的資料已經證實，在校成績以及更為明確的生涯規劃之間確實存在著某種關係」。而另外一句以被動語態來表達的句子：「受訪者對於墮胎的態度，是被訪談者所記錄下來的」，如果用主動語態來表達的話，讀起來會感覺比較流暢一些：「訪談者記錄了受訪者對於墮胎的態度」。除此之外，你也應該避免使用不必要的修飾語言，像是「似乎是……」（“seems to”）或者是「看起來……」（“appears to”）。
4. *連貫性*　順序、步驟、以及段落之間的銜接等等都應該具有緊密的邏輯關係。試著一次只讀一整份報告中的一個段落看看。這個段落是否包含了一個一致的概念？是否有主題句？在報告裡的不同段落之間是否能夠順暢地銜接在一起？
5. *重覆*　你應該把重覆的觀念、多餘的贅字、以及不必要的詞語通通都刪掉。最好的作法是以具有說服力的方式把觀念敘述一次就好，而不要以不清不楚的方式一再地重覆。當你在重寫時，應該要刪去空洞的字眼（沒有增加任何資訊的空泛字詞）以及迂迴的說法（用許多字來表達原本只要用一個精確的字就能說清楚的意義）。直截了當的敘述方式，會比使用一堆無意義的贅字好多了。以下是一句冗長的句子：「總結上面所述，我們從資料中所得到的結論是，*X* 對於 *Y* 的發生具有正面的、極大的、強烈的效果，雖然 *Y* 的發生機率非常低」，而這個句子最好能夠改為：「總之，我們的結論是，*X* 對 *Y* 具有大型的正面效果，但是 Y 發生的頻率不高」。正如同塞爾文以及威爾森（Selvin and Wilson, 1984）曾經警告過的，冗長以及過多的文字或是修飾語，都會使得讀者難以理解文章的內容。
6. *結構*　研究報告應該要有透明的組織架構。如果有必要的話，你應該要移動一些部分，以更加符合整體的組織架構，而且也應該要使用標題以及次標題。你應該要讓讀者能夠輕

易地理解報告中的邏輯結構。

7. *抽象化*　一份良好的研究報告會混雜了抽象概念以及具體範例。如果在一長串的抽象概念當中沒有加入任何具體細節的話，將會變得難以閱讀。相同的，在一大堆特定具體細節當中，沒有適時加入一些抽象的概化推論的話，也會使讀者覺得暈頭轉向。
8. *隱喻*　許多寫作者會使用隱喻的手法來表達一些概念。 像是*最新的（the cutting edge）*、*底線（the bottom line）*、*直指核心（penetrating to the heart）*等等詞語，都是借用了來自於其他脈絡的意象來表達一些概念。雖然隱喻可以是一種有效的溝通方法，但是卻只能偶爾使用，並且要謹慎小心。少數適當的、別具新意的隱喻能夠既迅速又有效地傳達概念，然而，浮濫地使用隱喻，尤其是那些已經被過度使用的隱喻（例如：底線），就會變成只是一種草率的、沒有想像力的表達方式而已。

我們先從量化研究報告開始進行討論。此種報告中的各個部分，會與研究計畫中各個步驟的順序大約一致。[6]

摘要或執行摘要　量化研究報告通常是從簡短的摘要開始，但是摘要的篇幅則是長短不一，可以少到僅有五十個字（光是本段的字數就超過 100 字），或者是長達一整頁。大部分的學術期刊文章，都會把摘要放在文章的第一頁。而在摘要裡則會包含了關於主題、研究問題、基本發現、以及任何特殊的研究設計或是資料蒐集技術的資訊。

針對實務工作者所撰寫的應用研究報告，其摘要的篇幅會比較長，而此種摘要則被稱爲**執行摘要**（executive summary），其中包含了比一般普通的文章摘要更多的細節，而且還包括了研究的意義以及在報告中所提出的主要建議。雖然執行摘要的篇幅比一般的摘要更長，但是仍然很少超過四到五頁的篇幅。

摘要以及執行摘要具有以下數個功能：對於比較不感興趣的讀者，摘要可以告訴他們關於報告內容的資訊；對於尋找特定資訊的讀者，摘要能夠幫助他們決定在完整的報告裡面是否包含了重要的資訊。讀者們可以利用摘要來篩選資訊，並且決定他們是否要去閱讀完整的報告。而對於打算閱讀完整報告內容的認真讀者而言，摘要可以讓他們在心裡對於整份報告的內容有大致的瞭解，如此一來，他們在閱讀完整報告時就會比較輕鬆而且快速。

呈現問題　在研究報告的第一個部分裡，研究者會先對研究問題下定義。而這部分的內容可以放在一個或者是多個部分當中，並且會有像是「導論」、「問題定義」、「文獻回顧」、「假設」、或者是「背景假定」之類的標題。雖然次標題可能會有所不同，但是報告中的第一個部分都應該包含對於研究問題的陳述，以及研究某些事物的理由。這部分也應該要解釋研究問題的重要性以及背景。在解釋研究的重要性時，此部分應該要顯示出，同一個問題的不同解決方式，會如何導致不同的應用或者是理論方面的結論。導論的部分通常會包括脈絡的文獻回顧，並且會將研究問題連結到理論。

執行摘要　一種放置在報告開頭部分、關於研究計畫發現的概要，所針對的讀者群是應用的、非專家的人士，而且其篇幅通常會比一般的研究摘要長一些。

導論的部分也會定義關鍵概念，並且提出概念性的假設。

描述研究方法 研究報告的下一個部分，則是要描述你設計研究以及蒐集資料的方法。這個部分可能會有數種不同的名稱（例如，「研究方法」、「研究設計」、以及「資料」等等）。在評估一項研究計畫的方法論時，這是最重要的一個部分。除此之外，這個部分也爲讀者回答了以下的數個問題：

1. 所進行的是何種類型的研究（例如：實驗研究、調查研究）？
2. 蒐集資料的確切方法爲何（例如：研究設計、調查的類型、蒐集資料的時間與地點、所使用的實驗設計等等）？
3. 變項是如何被測量的？量數是否具有信度與效度？
4. 樣本是什麼？研究中有多少個參與者或者是受訪者？這些人是如何被挑選出來的？
5. 研究者如何處理倫理議題以及關於研究設計的特定問題？

結果與表格 在描述了如何蒐集資料、抽樣的方法、以及測量過程之後，你接著就要呈現資料了。雖然接下來的這個部分會呈現出資料，但是並不會去針對資料進行討論、分析、或者是詮釋。有些研究者會把「結果」的部分以及接下來稱爲「討論」或是「發現」的部分合併在一起。

你必須決定要以何種方式來呈現資料。[7] 在分析資料時，你必須看著數十個單變項、雙變項、以及多變項的表格以及統計數據，以獲得對於資料的感覺。但是，這並非代表了你應該把每一個統計數據或是表格通通都放進最後的研究報告裡。相反地，你應該在能夠使讀者獲得完整資訊的情況下，去挑選出最少的圖形或是表格。除此之外，你也應該運用資料分析技術來總結資料以及考驗假設（例如：次數分配表、平均數以及標準差的表格、相關係數以及其他的統計量數等等）。

你應該爲讀者們提供一幅關於資料的完整圖像，但是注意不要提出過度詳盡的資料或者是不相關的資料，以免使得讀者被資料淹沒而迷失方向。讀者們可以作出他們自己的詮釋。至於詳盡的摘要統計資訊，則應該放在附錄當中。

討論 在討論的部分裡，你應該要針對資料的意義，給予讀者一個簡潔的、清楚的詮釋。討論的內容並不是一種選擇性的強調重點或者是偏袒某種立場的詮釋，更確切地說，應該是針對結果那個部分的內容進行坦誠公正的討論。討論的部分應該要與結果的部分區隔開來，如此一來，讀者才能夠去檢視資料並且提出不同的詮釋。葛羅索夫以及沙帝（Grosof and Sardy, 1985: 386）曾經警告道：「你所呈現的資料的安排方式，應該要能夠反映出在下列兩種事物之間的嚴格區分：其中之一是資料（你的觀察記錄）以及資料的摘要與分析；而另一方則是你的詮釋、結論、以及評論。」

新手研究者通常會發現要組織一個討論的章節是件困難的事。可行的方法之一，是根據假設來組織討論的內容，以討論資料與每一個假設之間的關聯性。除此之外，你也應該討論意外的發現、關於結果的可能替代解釋、以及缺點或者是限制等等。

結論 你應該在結論的部分中重新陳述研究問題以及總結研究發現。結論這個部分的目的，是要爲整份報告作出總結，因此這個部分有時候會被稱作是「總結」。而會

被放置在結論部分後面的，則只有剩下參考書目以及附錄而已。參考書目的部分只有包含了你在報告中的正文或是筆記中所引用的參考資料的來源或出處。如果有加上附錄的話，附錄通常包含了關於資料蒐集方法（例如：問卷中的措詞方式）或者是結果（例如：描述性的統計資料）的額外資訊。在量化研究報告中的註腳或者是章節末尾的註釋，其用途則是在於詳細闡述在正文當中的相關資訊。你只能偶爾使用註腳或是註釋來提供能夠澄清文本的次級資訊，但是註腳或是註釋不應該妨礙了閱讀的流暢性。

質化研究報告

與量化研究相較之下，大多數人會發現撰寫一份質化研究的報告是更加困難的，這是因爲質化研究報告比較沒有什麼規則和結構可以依循。然而，這兩種研究報告的目的其實都是相同的，也就是這兩種研究報告都是要去傳達研究過程以及在過程中所蒐集到的資料。量化研究報告會以一種合乎邏輯以及簡潔的風格來呈現出假設以及證據。對比之下，質化研究報告的篇幅通常比較長，而且也常見到篇幅等同於書本的報告（請參閱資訊箱 16.3 爲什麼質化研究報告的篇幅比較長）。

田野／實地研究　田野研究報告幾乎沒有什麼既定的格式以及標準的章節，而且理論的概化推論以及資料也不會被分別放在不同的章節裡。[8] 研究者所作出的概化推論會與證據緊密交織在一起，而證據的形式則是極爲詳盡的描述以及頻繁的引述。雖然在撰寫田野研究報告時沒有一個特定的方法（請參閱資訊箱 16.4 撰寫民族誌的四種類型以及修辭形式），但是大多數的研究者都還是會遵循某種一般的模式。

田野研究報告一般都會試圖在資料的呈現以及分析之間找到平衡點，以避免過度把這兩者區隔開來而造成**隔離的錯誤**（error of segregation）。當我們把資料以及分析隔離得太遠而使得讀者無法看出這兩者之間的關聯性時，就會發生此種錯誤。[9]

與量化研究相較之下，田野研究報告的筆調比較不那麼客觀或正式，而且也比較個人化。田野研究報告通常都是以第一人稱（也就是使用「我」這個代名詞）來描述的，這是因爲研究者直接參與了研究情境、與研究對象進行互動、而且自己本身就是測量「工具」。研究者的決定或是猶豫、感覺、反應、以及個人經驗，都是田野研究過程中不可或缺的特質。

田野研究報告通常會面臨比量化研究報告更多的批評。因此，評估讀者群對於證據的要求以及建立可信度，就變成極爲重要的事。關鍵在於提供讀者足夠的證據，好讓讀者們相信研究者所敘述的事件，並且接受研究者的詮釋。在田野研究中，讀者們會預期出現某種程度的選擇性觀察的現象，因此，至關重要的議題，就在於其他的觀察者如果也在相同的田野場域裡、並且檢視了相同的資料的話，是否也會作出相同的結論。[10]

在呈現田野研究的證據時，作者通常會遭遇到如何減少資料量的兩難困境。大多數的資料形式都

隔離的錯誤　在撰寫質化研究報告時所可能犯下的一種錯誤，意思是指研究者把實徵細節以及抽象理論隔離得太遠，因而使得讀者無法看出兩者之間的關聯性。

資訊箱 16.3　為什麼質化研究報告的篇幅比較長

1. 與量化研究報告相較之下，質化研究報告裡的資料比較難以濃縮，因為質化資料的形式包括了文字、照片、或者是句子，而且也可能包含了許多引述以及範例。
2. 除了呈現出事實證據以及分析的詮釋之外，質化研究者還試圖在讀者心中創造出一種同理心以及理解的主觀感受。對於具體背景環境的詳盡描述，能夠幫助讀者更加瞭解背景環境、或者是獲得對於背景環境的感受。研究者試圖讓讀者能夠進入某個社會情境的主觀世界觀以及意義系統當中。
3. 與量化研究者相較之下，質化研究者在蒐集資料、發展出分析的類別、以及組織證據時，所使用的是比較非標準化的技術，換句話說，在質化研究中所使用的技術，或許只能適用於個別的研究者或者是獨特的情境而已。因此，研究者就必須解釋他們所做的事情以及原因，因為他們所採取的方法可能在之前根本沒有前例可循。
4. 質化研究的共同目標，是去探索新的環境或者是建構新的理論。新概念的發展、以及對於概念之間的關係所進行的檢視，都會增加報告的篇幅長度。理論會由證據之中自然地流露出來，而且詳盡的描述也可以論證研究者發展出詮釋的方法以及過程。
5. 質化研究者或許會採用更加多樣化的寫作風格，而此種情況也可能會增加篇幅的長度。除此之外，質化研究者也擁有更多的自由，能夠去運用一些文學的手法來敘述一個故事。

是數量極爲龐大的田野筆記，但是作者卻不能把所有的觀察結果或者是錄下來的對話直接與讀者分享。例如，在《穿白袍的男孩》（*Boys in White*）這項針對醫學院的學生所進行的研究中，貝克等人（Becker et al., 1961）蒐集了大約 5,000 頁不留空行的田野筆記。田野研究者通常會在研究報告中放入大約 5%的田野筆記，以當作引述。至於剩餘 95%的筆記並沒有被丟棄掉，只不過是沒有足夠的空間可以容納而已。因此，作者會挑選出一些引述的內容，並且以間接的方式來把資料的其他部分呈現給讀者。田野研究報告沒有可以依循的固定組織方式，不過，文獻回顧的部分通常會出現在接近開頭的部分。目前有許多種可以接受的組織形式，而羅夫蘭（Lofland, 1976）則建議採用以下的格式：

1. 導論
   a. 情境中最爲廣泛的面向
   b. 一般情境的主要輪廓
   c. 資料的蒐集方式
   d. 關於田野場域的細節
   e. 報告的組織方式
2. 情境
   a. 分析類別
   b. 不同情境之間的對照
   c. 情節隨著時間經過的發展

資訊箱 16.4　撰寫民族誌的四種類型以及修辭形式

阿德勒以及阿德勒（Adler and Adler, 2008）確認了撰寫田野研究民族誌裡所可能會使用到的四種類型以及修辭形式：古典的、主流的、後現代的、以及公眾民族誌。

*古典的（classical）*風格是年代最久遠的，而且也常出現於田野研究的專業學術期刊當中。此種風格所強調的是可信賴性以及可接近性，目的在於避免因為使用過於抽象的詞彙或者是冗長而且複雜的句子而使讀者產生無法負荷的感覺。採用此種風格的作者們通常會使用主動語態而非被動語態，並且盡量以簡易的詞彙來寫作，好讓受過教育的一般讀者都能夠理解報告的內容。此種報告的開頭部分，通常是一個主題或者是理論議題，以及簡短的文獻回顧。研究方法的部分則像是一個研究者旅程的個人故事，其內容是關於研究的環境、遇到的人物、以及建立的關係等等。作者會描述他們在田野中所遇到的事物，以及他們在一個特定的時間與地點當中蒐集資料的方式。讀者們通常會感受到一種主觀的觀點，而且也覺得研究者「真的處於田野當中」。有時候作者會提出關於資料分析的討論，但是許多古典風格的作品並未提供詳盡的分析。而資料的部分則經常會依循著一種漸進的發展方式：先呈現出一個特定的概念、接著詳細闡述這個概念、然後再提供關於資料的描述。至於資料則通常是以敘事的形式出現。結論的部分會以數種方式來促進知識的進展，像是增加、超越、以及／或者修正現存的理論，並且經常會牽涉到轉變成為一個較為正式的風格。作者通常會依據從頭開始建立理論的方式來組織報告。

*主流的（mainstream）*風格則是出現在主流的學術期刊當中。由於在這些期刊裡的同儕審查者可能並不熟悉質化或是詮釋取向的社會學，因此，他們可能會勉強把實證主義的取向套用在民族誌作者身上。與古典風格相同的是，此種風格的作者也是利用蒐集深入的、第一手的、以及自然主義的資料的方式，來描繪他們所研究的世界。然而，主流的風格在提出討論內容時，其方式與古典風格有所不同。與古典風格相較之下，主流的風格通常與讀者的距離較為遙遠，而且也比較會採取一種專家權威的筆調。一般而言，導論的部分會比在古典風格中更為緊密、嚴謹、正式。主流的風格並不強調可接近性，相反地，此種風格所強調的是要符合於標準的社會科學修辭方式，而且通常會去模仿實證主義以及量化研究報告。導論的部分會比在古典的風格當中還要更長。而在這個部分裡，作者會為各種詞彙下定義，並且提供一種不同類型的文獻回顧，此種類型的文獻回顧比較冗長、比較詳盡、而且通常會再區分成許多較小的部分。廣泛的文獻回顧意味著知識是以一種統一的、線性的方式向前進展，並且也是建立在先前的學術貢獻上，而這些貢獻則是符合於一種較偏向實證主義的傾向。在主流的風格裡，研究方法的部分也會比在古典的風格裡還要更長。這個部分可能會有詳盡的闡述，以證明作者使用質化方法是合理而正當的，並且也解釋了質化方法的認識論基礎。研究者幾乎不會去討論他們自己與主題、參與者、以及環境之間的個人連結性，因為主流的讀者群可能會把此種陳述詮釋為研究者帶有偏見的證據。作者們在呈現研究過程時，會使讀者覺得這些過程彷彿都是事先規劃好的，而不是歸納的或者是偶然進行的過程。採取此種風格的作者，會使用被動語態以及一種客觀與中立的筆調。除此之外，通常還會包括一些關於特定技術或者是所使用的電腦程式的討論內容，而不會像是在古典的風格中只對研究方法進行一種模糊的、印象式的討論而已。主流的風格中的資料或是結果的部分通常會有次標題，而且

也會包含某種形式的圖形或者是表格。此種修辭的形式會讓研究者從資料中消失，並且以一種超然的方式來呈現出資料。

*後現代（postmodern）*的風格是從 1990 年代之後才開始被採納的，而且通常只出現在少數具有後現代傾向的學術期刊之中。與古典風格相較之下，此種風格拒絕嘗試去達到客觀性、信度與效度的原則、以及研究者權威的概念。相反地，此種風格所依據的信念是，在進行以及撰寫田野研究時，根本沒有固定的或者是單一的標準。就算後現代風格真的有任何原則可言，也是一些像是實質的實徵貢獻、美學的優點、反身性、對於讀者的衝擊、以及個人生命經驗的可信度之類的原則而已。此種寫作風格傾向於一種人道主義的、或者是藝術的形式，而且其內容通常是以一種口語的方式寫成的說故事敘事，其中包含了故事情節、道德寓意、以及強調的重點。研究者/作者的主觀意見在其中很常見，而且也帶有高度的自我揭露以及自我覺察、或者是反身性。後現代風格可能會有關於作者經驗的第一人稱敘述，而且在其中點綴著對於這些個人經驗的半超然討論。資料的主要來源、或者甚至是唯一來源，可能就是作者的個人經驗而已。後現代風格通常會以一種非線性的、無法預測的方式來描述事物，其中的筆調和方向會經常轉變，而且整份報告也沒有明確概述的結構或者是組織。

*公眾民族誌（public ethnography）*的風格是最新的形式。此種風格以自覺的方式試著把社會科學的發現傳達給受過教育的一般讀者群。其目標是要教育社會大眾關於社會科學的知識。此種風格通常相當仰賴民族誌或者是田野研究，因為此種形式的研究是最容易被社會大眾所理解的。此種風格通常會出現在以書本形式出版、而且在書店裡被當成非小說類來販售的研究當中，或者是以社論的形式出現在全國性報紙的特約作家專欄裡，或者是以文章的形式出現在文化水準較高、以知識份子為主要讀者的雜誌裡。在以書本形式出現的版本中，作者會試圖利用地圖、照片、以及豐富的描述來吸引讀者。關於研究方法的討論通常是既簡短、而且也不正式的。有時候可能也會有文獻回顧的部分。此種風格的作者通常會把方法論的內容或者是文獻的引述資訊放在註腳或是附錄當中。而撰寫的風格是像小說一般的，其中有很長的引文以及非常詳盡的描述。在此種風格當中，理論若不是完全沒出現，就是相關的討論篇幅是很有限的。

3. 策略
4. 總結以及啓示。

在不同的研究者之間，用來組織證據與分析證據的策略也會有很大的差異。[11] 例如，研究者可以依據*自然歷史*（*natural history*）（請參閱第十三章）的方式來組織報告，也就是事件會隨著研究者的發現而逐漸展開，或者，報告也可以依據編年史（chronology）的方式來加以組織，也就是依循著場域中某個面向的發展周期、或者是場域中人們的生涯發展順序來組織證據。另外一種可能的作法，則是以**變焦鏡頭**（zoom lens）的方式來組織報告，換句話說，就是先從廣泛的主題開始，然後越

變焦鏡頭　一種組織田野研究報告的方法，在使用此種方法時，作者先廣泛地由一個主題開始，然後漸漸地、更為具體地聚焦在特定的主題上。

來越聚焦於一個特定的主題上面。報告中的陳述可以從一個關於所有文化的普遍性陳述、轉移到關於一個特定文化的一般性陳述、再轉移到關於該文化中某個面向的具體陳述、最後則是關於特定事件的具體陳述。[12]

研究者通常會圍繞著關鍵概念以及主題來組織田野研究報告。他們或許會在使用抽象的分析主題、或者是取自於研究對象的主題這兩者之間去作出選擇。後者可以給予讀者關於研究情境的生動描述，並且展現出研究對象關於語言、概念、類別的知識，以及他們所抱持的信念。[13]

雖然田野研究者也會在報告當中討論他們所使用的方法，但是此種討論內容在報告中出現的位置以及形式都各自不同。常見的技術之一，是把對於情境的描述、進入田野場域的方式、研究者的角色、以及參與者與研究者之間的關係等等，交織在關於證據以及分析的討論當中。如果作者採用了一種馬能（Van Maanen, 1988: 73）稱爲「自白式」（"confessional"）的撰寫風格的話，上述的方法所產生的效果會更爲強烈。至於編年史的、變焦鏡頭的、或者是以主題爲基礎的組織方式，都允許作者把關於資料蒐集方法的內容放置在報告的開頭或者是結尾的部分。而在書本形式的報告裡，作者通常會把方法論的議題放在單獨的附錄當中。

許多田野研究的報告會包含錄音帶的逐字稿、地圖、照片、或者是說明分析類別的圖表。爲了要對資料進行補充說明，我們通常會把上述這些內容放在靠近資料討論內容的地方。照片能夠爲文本中所描述的場景提供視覺的影像，並且以研究對象的觀點來呈現出場景的意義。舉例而言，曾經有些田野研究的文章是以全部都是照片、戲劇的腳本、以及紀錄片的方式出現的。[14]

由於研究者會以直接而且個人的方式去涉入某個社會情境的私密細節，因此，此種情況也會突顯出倫理方面的議題。我們會以一種保護研究對象隱私的方式來撰寫報告，而且也會想辦法避免研究對象因爲報告的出版而受到傷害。[15] 因此，在田野研究的報告當中，我們通常會去改變成員的姓名與確切的地點。由於在田野研究中，研究者通常會親身參與許多過程，因此，許多研究者也會在報告中加入簡短的個人自傳。例如，在懷特（William Foote Whyte, 1955）所寫的《街角社會》（*Street Corner Society*）這本書的附錄當中，作者就詳細地描述了他的爸爸以及祖父的職業、他自己的嗜好以及興趣、他曾經做過的工作、他後來怎麼會進了研究所念書、以及結婚對他的研究所產生的衝擊等等。

歷史比較研究　在撰寫歷史比較研究的報告時，並沒有任何單一的方法。最常見的情況是，研究者會「述說一個故事」或者是描述在一般分析類別裡的細節。在撰寫歷史比較研究的報告時，通常會超越了單純的描述，而且會包括有限的推論以及抽象的概念。

很少有歷史比較研究報告會詳細地描述研究中所使用的方法。不論是在報告中或者是附錄中，以明確的部分來描述研究方法，都是相當罕見的情況。偶爾會有像是書本形式的報告會附上關於參考書目的文章，其中會描述所使用的主要資料來源。然而，在大多數的情況中，歷史比較研究報告裡會列出數量龐大的、詳盡的註

腳或是註釋，這些註腳或是註釋會指出資料來源以及其他證據。例如，在一份約 20 頁的量化或者是田野研究報告中，一般會有 5 到 10 個註釋，而在相同頁數的歷史比較研究報告中，則可能會有高達 50 到 60 個註釋。

歷史比較研究的報告中可以包含照片、地圖、圖表、以及整體的統計數據表格，而這些事物會出現在討論相關證據的部分裡。這些圖表、表格等等事物能夠補充說明討論的內容，而且也能夠讓讀者對於報告中所描述的地方以及人物有一些感覺。除此之外，這些圖像也可能會與頻繁的引述一起出現。與量化研究不同的是，幾乎沒有什麼歷史比較研究報告會包含對於特定假設的考驗。相反地，作者們會試圖去建立一張由意義或是描述性細節所編織而成的網，而且也會利用一種能夠傳達詮釋以及推論的方式來組織證據。

組織歷史研究報告的兩個基本模式，分別是依據主題的順序以及時間的順序的模式。大多數的作者都會把這兩種類型混合使用。例如，他們可以把某個主題裡的資訊依據時間先後順序來組織，或者是把某個歷史時期內的事件依據主題來加以組織。他們偶爾也會使用其他的組織形式，像是依據地點、個別人物、或者是重大事件等等來組織報告的內容。如果該份報告是真正的比較研究報告的話，作者還可以有其他額外的選擇，像是在主題之間進行比較等等（請參閱資訊箱 16.5 在歷史比較研究報告中應該考量的面向）。[16]

有一些歷史比較研究者會去模仿量化研究報告，並且在撰寫研究時使用量化研究技術。他們此種做法比較像是在延伸量化研究，而不是採取一種獨特的歷史比較研究方法。因此，他們所寫出來的報告就會遵循著量化研究報告的模式。

你已經在第十五章中讀到關於敘事分析的討論了。使用此種策略的研究者在撰寫研究報告時，通常會採取一種敘事的風格。他們會依時間順序來組織資料，並且試著「訴說」關於特定個人以及事件的「故事」。

研究計畫書

什麼是研究計畫書？ 研究計畫書是一種文件，其內容提出了進行一項研究的計畫以供審查者進行評估。研究計畫書可以是一份提交給指導教授的方案，以作爲獲得教育學位的一部分（例如：碩士論文或者是博士論文），或者也可以是針對一個研究經費贊助機構所提出的研究計畫。研究計畫書的目的在於說服審查者相信，研究者有能力去成功地進行所提出的研究計畫。如果研究計畫書的內容寫得很好、組織完整，並且也展現出仔細縝密的規劃過程，那麼審查者就會比較有信心地認爲這項規劃好的研究會被成功地完成。

雖然研究計畫書與研究報告感覺相當類似，但是研究計畫書卻是在開始進行研究之前就必須先寫好的。研究計畫書描述了研究問題以及該問題的重要性、提供文獻回顧、並且也會針對將在研究中使用的技術以及方法進行詳盡的敘述。此外，在研究計畫書當中也會說明這些技術與方法爲何適合用於該項研究當中。

資訊箱 16.5　在歷史比較研究報告中應該考量的面向

1. *順序*　歷史比較研究者對於事件的順序是相當敏銳的，而且也會把事件放入一個順序當中，以描述一個過程。例如，一個研究某些法案的立法過程，或是社會規範演變過程的研究者，可能會把整體過程打散成為一組循序漸進的不同階段。
2. *比較*　歷史比較研究的核心，正是針對相似處以及相異處進行比較。例如，假如有一個研究者想要比較在兩個不同的歷史時期或國家中的家庭狀況，那麼他或她可以先由列在不同情境中的家庭所共享以及沒有共享的特徵開始。
3. *條件*　研究者經常發現，某個事件、行動、情境會取決於或是受到其他事物所影響。描繪出某個事件如何對於其他事件產生作用、以及事件之間的關聯性，是相當重要的。例如，一位檢視地區性報紙興起過程的研究者可能會發現，地區性報紙的興起過程是受到識字率的提升所影響的。
4. *起源以及結果*　歷史比較研究者會去追溯一個事件、行動、組織、或社會關係在時間上的起源，或密切注意這些事物在之後的時期裡所產生的結果。例如，一名想要解釋奴隸制度如何結束的研究者，可能會發現此種結果是來自於之前 50 年當中的許多運動、演講、法律、以及行動所導致的。
5. *對於矛盾意義的敏感性*　意義會隨著時間而改變，而且在不同的文化之間也會有許多變化。歷史比較研究者會問問自己，某個字或是社會類別在過去以及當前的社會中，是否具有相同的意義。或者，在某個文化中的某個詞彙，在另一個文化中是否有可以直接翻譯對應的詞彙。舉例而言，在某個特定的歷史時期中，上大學是非常昂貴的，而且在 18 歲到 22 歲的人口當中，只有不到 1%的人可以取得大學學位。因此，在那個時期裡的大學學位，其意義就會與在二十世紀晚期有非常大的差異，這是因為在二十世紀晚期時，進入大學念書已經變得相對容易許多了。
6. *有限的概化推論*　在歷史比較研究中，過度概化一直是一個潛在的問題。幾乎沒有什麼研究者會在歷史比較的解釋當中，去尋找嚴密的、固定的法則，他們會限制陳述或是避免提出嚴格的決定性說法。例如，研究者不會作出像是下列這樣的總括性陳述：在歐洲白人殖民地區的原住民文化遭到破壞，是先進科技文化所造成的必然結果。相反地，研究者可能會列出各種具體的因素，並且將這些因素結合起來，以解釋在某些特定的社會歷史情境中，原住民文化遭到破壞的情況。
7. *關聯性*　關聯性的概念被使用於所有形式的社會研究當中。如同在其他領域中，歷史比較研究者也會確認那些在特定時空中一起出現的因素。例如，假定有某位研究者，正在檢視某個城市在十九世紀時的犯罪率。那麼他或她或許會想知道，多年以來在該城市中不斷增加的移民，是否與犯罪率升高有所關聯，或是那些被逮捕的罪犯是否比較可能是新來的移民者。
8. *部分以及整體*　把事件放入脈絡當中是非常重要的。歷史比較研究的作者通常會概略地描述某個過程、組織、或是事件的各部分之間的連結，以及這些過程、組織、或事件所處的較大脈絡。例如，一位研究在十八世紀時某種特定政治儀式的研究者，會描述此種儀式如何契合於十八世紀的政治系統之中。

9. *類比*　雖然類比是相當有用的，但是過度使用或是不當使用類比，則是相當危險的。例如，某位研究者檢視了在 X 國家中關於離婚的感覺，並將這些感覺描述為「就像是在 Y 國家中對於死亡的感覺一樣」。如果真的要作出此種類比，那麼該位研究者就必須描述在 Y 國家中「對於死亡的感覺」才行。
10. *綜合*　歷史比較研究者經常把許多具體事件以及細節綜合起來，以成為一個完整的整體。此種綜合的作法，來自於把許多較小的概化推論以及詮釋交織在一起，以成為連貫的主要主題。例如，某位研究法國大革命的研究者，可以把關於社會結構的變化、國際壓力、農業的混亂、不斷變化的大眾信念、以及政府財政問題的許多具體推論，綜合起來成為一個簡潔的、一致的解釋。採用敘事形式的研究者，會在導論或是結論的部分裡，加入關於論點的摘要，也就是內嵌於描述中的一個中心主旨抑或是主題。因此，理論的概化推論會與證據交織在一起，而且似乎是以歸納的方式，從詳盡的細節中所推導出來的。

在量化研究計畫書當中，大部分的內容都和研究報告相同：標題、摘要、關於研究問題的陳述、文獻回顧、討論研究方法與研究設計的部分、以及參考書目。而與研究報告相較之下，量化研究計畫書中所缺乏的是關於結果、討論、以及結論的部分。量化研究計畫書中也會包含一個關於蒐集以及分析資料（例如：統計數據的類型）的計畫。除此之外，量化研究計畫書通常也會包括一個關於進行各種步驟的時間表，以及對於每一個步驟所需時間的估計值。

質化研究計畫書通常會比較難寫，因爲質化研究過程本身就是比較沒有結構的、也是比較無法事先進行規劃的。你應該先準備一個關於主題／問題的陳述、文獻回顧、以及參考書目。你可以用兩種方式來證明你有能力去完成一項質化研究計劃。首先，你可以準備一份寫得很好的研究計畫書，在計畫書當中包含了關於文獻、研究問題的重要性、以及資料來源的廣泛討論等等，而此種作法能夠讓審查者看到你對於質化研究的熟悉程度，以及使用此種方法來研究某個問題的適當性。第二，你也可以去描述一項質化的前導研究。此種作法可以展現出動機、對於研究技術的熟悉度、以及完成一份無結構研究的報告的能力。

申請經費的研究計畫書　對於研究的補助經費，可以提供完成一項有價值的研究所需要的資源。那些想要利用贊助的資金來獲取個人利益或是名望、逃避其他活動、或者是建立一個「帝國」的研究者，通常都比較無法成功地申請到研究經費。撰寫研究計畫書以及獲得贊助經費的策略，現在已經變成一種被稱爲**申請補助費技巧**（grantsmanship）的產業了。

目前有很多贊助研究計畫的經費來源，像是大學、私人基金會、以及政府機構等等，都有一些提供經費給研究者的相關方案。研究者可以利用這些經費來購買設備、支付薪資或者是雇用助手、支應研究開銷、支付蒐集資料時的旅行費用、或者是用來支付出版研究結果的費用等

申請補助費技巧　運用策略以及技巧以找到適合的經費來源，並且準備優秀的研究計劃書以成功申請研究經費。

等。獲取研究經費補助的競爭程度會因爲贊助經費來源的不同而有很大的差異。在向單一機構提出申請的所有研究計畫當中，有些機構所贊助的研究計畫比例可能高達四分之三，但是也有些機構的贊助比例只有二十分之一而已。

雖然目前有許多贊助社會研究的資金來源，但是卻可能沒有任何一個來源願意贊助單一的特定計畫。你必須先去調查這些贊助機構，並且先去思考下列的這些問題：這些機構所贊助的是什麼類型的研究：應用研究或是基礎研究、特定的主題、或者是特定的研究技術？什麼類型的研究計畫書（例如：長度、詳細的程度等等）是必要的？研究計畫中的哪些面向（例如：設備、人事、旅行）是會以及不會獲得補助的？目前有許多關於贊助機構的資訊來源，而在大學中負責研究經費補助的圖書館員或者是行政人員，都是很好的資訊來源。例如，在每年出版一次的《基金會名錄》（*The Foundation Directory*）中，會列出私人基金會的相關資訊。而在《社會科學家的聯邦基金指南》（*The Guide to Federal Funding for Social Scientists*）中，則會列出美國政府中的各種補助經費的來源。在美國，只要是這份指南的訂閱者，都可以在關於經費贊助機構的許多通訊資料、以及全國電腦資料庫當中，去尋找各種提供研究經費的機構。有些機構會定期地發行**研究計畫書徵求訊息**（requests for proposals, RFPs），以徵求針對某個特定議題進行研究的計劃書。研究者必須熟悉關於經費來源的各種資訊，因爲研究者必須把研究計畫書寄給適合的贊助機構，才能夠成功地獲得研究經費。[17]

你應該在研究計畫書當中展現出過去的成功記錄，尤其是如果你準備要負責進行該項研究計畫的話。負責研究計畫的人就是**主要研究者**（principal investigator, PI），有時候也會被稱爲研究計畫主持人（project director）。研究計畫書裡通常會包含主要研究者的簡歷或者是學術背景、來自於其他研究者的支持信件、以及過去研究的記錄等等。審查者一般會認爲，把資金投注在一個由具有實質研究經驗的人所主導的研究計畫上，會比投注在一個毫無經驗的新手所進行的研究上來得安全。在你以主要研究者的身份來尋求研究經費的贊助之前，你可以先進行一些小型的研究計畫來建立研究的記錄、或者是去協助一些經驗豐富的研究者進行研究。

評估研究計畫書的審查者，會去判斷某個研究計畫是否符合於該贊助機構的目標。大多數的贊助機構都會有一些指導方針，其中陳述了這些機構所贊助的研究計劃類型。例如，對於那些贊助基礎研究的機構而言，它們的目標通常是爲了促進知識的發展。至於贊助應用研究的機構，其目標則是要提升它們所提供的服務的品質。在關於提交研究計畫書的說明當中，會詳細說明頁數長度、複本數量、以及申請期限等等。你應該要確實地遵守所有的規定。如果一名研究者甚至都無法遵守關於研究計畫書頁數長度的規定時，審查者爲什麼要同意提供數千美元的資金讓這名研究者去進行一項複雜的研究計畫呢？

研究計畫書看起來應該要簡潔而且專業。在申請規定當中通常會要求申請者必須列出對於時間運

研究計畫書徵求訊息　由一個贊助組織所發布的訊息，其內容是徵求一些需要經費贊助的書面研究計畫書。

主要研究者（PI）　在由某個組織所贊助的研究計畫當中，主要負責進行研究的人。

用、服務、以及人員編制的詳盡計畫。上述這些資訊應該要清楚地陳述出來，並且要切合研究計畫的實際情形。過高或是過低的估計、不必要的擴充、或者是遺漏掉一些基本的部分，都會降低審查者對於一份研究計畫書的評價。爲研究計畫書擬定預算，是一件相當複雜的任務，而且通常需要得到專業人士的協助。舉例而言，像是薪資比率、福利比率等等資訊，都是不容易取得的。最好的作法，是去請教在大學裡負責研究經費事宜的行政人員、或者是經驗豐富的研究計畫書作者。除此之外，你通常還必須得到某些機構的贊同或者是許可（例如，像是機構審查委員會（IRB）的同意，請參閱第五章）。研究計畫書也應該要包含發表研究結果的具體計畫（例如：出版研究結果、或是向專業團體提出研究結果等等），以及一個可以評估研究計畫是否達成預定目標的計畫（請參閱圖表 16.1）。

研究計畫書是一種在研究者以及經費贊助機構之間的合約。經費贊助機構會要求一份最後的報告，在報告中必須包括經費使用方式的細節、研究發現、以及該研究是否達成預期目標的評估內容。如果研究者未能適當地使用經費、無法依照研究計畫書的內容完成研究、或者是無法提出一份最後的報告，都可能會導致嚴重的後果，像是該名研究者可能被禁止再向同一個機構申請經費、或者甚至可能必須面臨法律訴訟等等。若是嚴重地濫用經費，可能會導致整個機構（例如：大學、企業、醫院等等）被罰款，而且還會被禁止再接受任何經費的贊助。

在向贊助機構提出研究計畫書之後，依據贊助機構的不同，整個審查的過程可能會花上數個星期到幾乎一年之久。在大多數的情況裡，審查者會把許多的計畫書進行排名，並且只把經費提供給排名最高的前幾份計畫書。研究計畫書通常也會經過同儕審查的程序，在這個過程當中，審查者會從計畫書中的簡歷資料而得知計畫書的作者是誰，但是作者卻不知道審查者是誰。有時候也會由一些非專家或者是非研究者的人士來審查研究計畫書。在關於如何準備一份研究計畫書的說明裡，也會明確指出讀者群是某個領域裡的專業人士或者是受過教育的一般大眾，或者是兩者都有。一般而言，那些申請贊助的經費數目較大的研究計畫書，通常審查過程也會更爲嚴格。

如果你的研究計畫成功獲得贊助的話，你可以稍微慶祝一下，但是只能短暫地慶祝而已。如果被拒絕了（這種情況是比較可能發生的），也不要就此絕望。大多數的研究計畫書在第一次或者是第二次被提出時，都會被拒絕。許多經費贊助機構都會針對審查者在評估計畫書時所依據的準則提供書面文件。如果你打算申請經費的機構有提供此種文件的話，記得一定要向該機構索取。有些時候，如果你很客氣地在電話中與某個經費贊助機構裡的人交談的話，或許可以得知被拒絕的理由。你應該根據審查者的意見與評論來提升計畫書的品質，並且重新提出申請。大多數經費贊助機構都願意接受研究者把計畫書修改之後再重新提出申請。經過審查的計畫書，在後續的競爭當中，通常都會比較具有優勢。

圖表 16.1　與成功的研究計畫書有關的因素

1. 成功的計畫書處理一個重要的研究問題。成功的計畫書建立在先前的知識上，並且代表了在基礎研究知識方面的實質進展。此種計畫書詳細記錄了一個重大的社會問題，並且能夠為應用研究提出一些可能的解決之道。
2. 成功的計畫書符合所有的規定、文筆流暢、而且具有清楚陳述的目標。
3. 成功的計畫書完整地描述了所有的研究程序，其中包括了高標準的研究方法論，而且也應用一些適合於研究問題的研究技術。
4. 成功的計畫書包含了發表研究結果以及評估是否達到預期目標的具體計畫。
5. 成功的計畫書明確指出該項研究計畫具有完善的研究設計，並且展現出嚴謹的規劃過程。除此之外，成功的計畫書也會提出合乎實際的預算以及時間表。
6. 成功的計畫書會註明研究者具有能夠成功完成研究所需要的經驗或者是背景。

社會研究的政治

一種對於社會研究的天真的、單純的看法認爲，進行研究以及撰寫研究報告是一種純粹的過程，此種過程會在社會政治的真空狀態中運作，並且與外在社會的壓力與關切的議題完全隔絕開來。另外一種更爲合乎實際的看法則是，當我們在進行社會研究時，我們無可避免地會面臨許多倫理以及政治方面的議題。合乎倫理道德的研究者，會去保護研究參與者、以符合倫理原則的方式誠實地進行研究、避免來自於贊助者的干涉、並且以公開而清楚的方式來發表研究結果。社會研究的政治與接受贊助的研究（我們已經在第五章中討論過此種類型的研究）中的許多議題是重疊的。除此之外，許多社會研究者也必須去面對一些在經濟上或是政治力具有強大力量的團體，這些團體會試著去限制研究者所能研究的事物、研究者進行研究的方式、或者是研究者發表研究結果的方式。

對於研究者探討的事物所施加的限制

對於研究的直接限制　政府或者是在社會上具有權力的團體可能會試圖去限制自由的科學研究。有些對於研究的限制其實是一直都存在的，但是在特定的時間與地點中，這些限制的約束力可能會變得非常強大。在非民主的社會裡，對於社會研究進行嚴格的控制或審查是一種常態、而非例外情況，尤其是對一些具有政治敏感性的主題而言更是如此，其中也包括了對公眾意見所進行的調查。因此，舉例而言，像是在二十世紀晚期時的中國、東歐、南非、以及臺灣等地方，社會研究者都會受到懷疑、被限制只能研究一些「安全的」主題，或者是被強迫去支持官方的政府政策。[18] 而在一些國家裡，當發生過軍事政變之後，社會學研究本身就會因爲被認爲具有顛覆性質而被明令禁止。在一個極端的案例當中，當德國的納粹黨在 1937 年對大學以及研究中心進行「淨化」活動時，有大約 40%的德國科學家因爲政治方面的理由而被解除職務。[19] 在 1950 年代時，在美國有數百位教授以及研究者因爲不願意公開反對共產主義以及與麥卡錫調查行動合作，因而被整肅。在當時，那些反對強

制性忠誠宣誓、支持種族整合、或者是提倡性教育的人，都會被懷疑是心有不軌，並且因而被威脅會失去工作。例如，單單在加州大學裡，就有 25 名教授因爲拒絕簽署一份忠誠宣誓書而失去了他們的工作。[20]

社會研究的兩種限制分別是：（1）控制了接觸到資料或是研究對象的管道的守門員，以及（2）對於如何蒐集官方統計數據的控制權力。守門員可以限制我們所能研究的事物，而且也可能會試圖去保護他們自己或者是他們的組織免於受到批評或者是難堪的情況。守門員通常會限制研究者接觸某些研究對象或是區域的權利，因爲他們與這些研究對象或是區域有一些利害關係存在。例如，在 1997 年時，美國軍方原本打算發放一份關於性騷擾的問卷給 9,000 名士兵去作答。在該份問卷上原本有 153 道題目，但是後來有 6 道題目被刪掉了。而刪除掉這 6 道問題的理由，是「軍方的高階將領擔心所得到的回應會使軍方陷入極度難堪的窘境當中」（Schmitt, 1997）。在該項計畫的顧問當中，有一名社會人類學家以及一名法律教授對此種作法感到不滿，並且提到以該份問卷的早期版本所得到的初步結果顯示出，在軍事基地裡的性騷擾行爲，與那些詢問特定行爲（例如：光顧脫衣舞俱樂部、觀賞限制級影片等等）的問題具有相關性。而身爲守門員的軍方將領們則不想讓這些問題出現在問卷上，因爲所得到的答案可能會讓一些在軍事基地裡普遍存在的現象曝光，並且爲軍方帶來負面的形象。

另外一種限制則牽涉到官方的或者是現存的統計資料，這些資料是由政府或者是其他的大型組織所蒐集的（相關內容已經在第十一章裡討論過了）。某些機構是否決定要蒐集資訊、以及這些機構蒐集資訊的方式，都有可能會影響到研究發現。政治因素經常會決定了某些現象（例如：失業、收入、教育成功、貧窮程度等等）在官方統計資料中的定義，以及此種資料是否會被蒐集。[21]

數以百計的社會科學家會定期地利用由美國人口普查局所蒐集的資料來進行人口統計學的、經濟的、以及其他的研究。進行人口普查的原始目的，是爲了在各州以及行政區之間去分配選舉代表人的名額。之後，人口普查局開始爲了制訂政策、提供社會福利方案、以及依據某個區域的人口數來分配政府經費的目的而蒐集相關的資訊。人口普查局已經變成社會科學中各種相關資訊主要來源，並且是關於許多主題的官方統計資料的情報交換所。在人口普查局的統計資料中出現的嚴重扭曲情況（例如：系統性地多算或是少算某些區域裡的人口總數），會影響到依據這些資訊所申請的研究經費、妨礙了完整的民主代議制度、並且會傷害了社會福利或是基金的公平分配（請參閱在第十一章裡的相關討論內容）。

使用現存的統計資料來進行研究的研究者，通常必須依賴政府來提供資訊或是文件。美國曾於 1980 年時通過了「文書作業簡化法案」（Paperwork Reduction Act），並且依據該法案而設立了「資訊及管制事務局」（Office of Information and Regulatory Affairs）來決定蒐集某些資訊以及保存記錄是否是必要的。該法案導致了由政府贊助的研究出版品的數目逐漸減少。除此之外，該法案曾經「偶爾被用來限制一些不支持行政部門政策目標的資訊」（Shattuck and Spence, 1988: 47）。舉例而言，受到「文書作業簡化法案」的規定所影響，在健康照護的領域內，那些注重環境議題並且間

接地批評企業或是政府政策的研究計畫，就很有可能會被拒絕出版。相較之下，那些焦點在於傳統疾病、並且間接地把責任歸咎於受害者的研究，就比較有可能會被准許出版。以削減成本爲名義，政府機構停止蒐集資訊、不再把資訊公諸大眾、並且把蒐集資訊的工作轉移到私人企業。官員們削減了美國政府的印刷以及出版費用，並且提高了文件的價格。當官僚們決定不去蒐集某些資訊時，會對研究資訊以及政策造成深遠的影響。

受到政客影響而產生的限制　不幸的是，對於某些不屬於科學社群的人而言，當社會研究不符合他們的社會或者是政治價值觀時，他們就會去攻擊社會研究。某個政客或是新聞記者可能聽說了在某個有爭議的領域裡的一項研究計畫、對該項研究作出錯誤的詮釋、然後利用這個機會來炒作新聞以提升自己的知名度。例如，南伊利諾大學的魯賓教授（Professor Harris Rubin）曾經想要研究 THC（一種在大麻中的活性成分）對於性衝動（sexual arousal）的效果。在當時對於這兩者之間的關係只存在著一些相互矛盾的迷思，而且幾乎沒有任何科學方面的證據。魯賓教授非常謹慎地遵循了所有規定的程序以及得到所有必要的許可，而且美國國家心理衛生研究院（National Institute of Mental Health）在經過了科學的同儕審查過程之後，也決定要提供經費來贊助這項研究計畫。然而，有一名保守派的美國國會議員從報紙中得知了這個研究的主題，並且推動了一項修正案來阻止進一步提供經費給這項研究。除此之外，魯賓教授還必須把該項研究計畫已經獲得的所有經費通通都歸還給美國聯邦政府。雖然科學界強烈主張政客不應該干預合法的研究，經費卻仍然被刪減了。政客們不敢公開支持這項社會研究的原因，是因爲他們擔心可能會有其他黨派的候選人告訴選民們，政府正在付錢給一些學生「去服用大麻以及觀看色情影片」。[22] 在 1989 年時，國會議員們曾經阻止美國政府提供經費給一項大型的全國性調查，該項調查主要是在瞭解民眾的性行爲，以對抗愛滋病的蔓延。而國會議員們抵制這項調查的原因，是因爲他們認爲社會研究者去詢問人們的性行爲是不適當的（請同時參閱範例 16.1 美國國會議員質疑研究經費）。[23]

美國的參議院曾經取消過一個由美國國家衛生研究院（National Institutes of Health, NIH）所進行的關於青少年性行爲的研究計畫。這項研究原本預定要調查 24,000 名青少年，並且詢問關於他們的社交活動、家庭生活、以及性行爲方面的問題，以提供瞭解愛滋病以及其他透過性行爲傳染的疾病的相關背景。許多研究者說他們不想針對這個議題公開發表評論，因爲他們擔心會變成一些政治團體攻擊的對象。有些敢於發言的研究者則認爲，某些具有極端政治意識型態的小型少數族群竟然擁有扼殺重要研究的能力，這是一件「可恥的行爲」，而且「令人恐懼」。有一名研究者提到，上述這項研究計畫被取消的原因，並不是因爲該研究的科學品質或是重要性有任何的問題，而是因爲一種來自於意識型態的、認爲「我們不必談論這些事情」的決定。[24]

對於社會研究的攻擊，甚至是連那些不具爭議性、只是被誤解的研究也都受到波及，會對所有的研究者都造成傷害。政客們可能會試圖去阻止那些科學社群認爲

範例 16.1

美國國會議員質疑研究經費

在1998年時，來自於南加州的美國國會議員桑福特（Marshall Sanford）說他想要刪減美國國家科學基金會（National Science Foundation, NSF）對於那些「科學價值」有問題的研究所提供的經費贊助。顯而易見地，他認為自己對於科學價值的判斷力比科學社群更好，而且他還引述了關於自動櫃員機以及撞球的研究。NSF的官員們觀察到，這位議員所提到的研究相關內容雖然縮寫為 ATM，但是所指的卻是「非同步傳輸模式」（asynchronous transfer modes）（這是一種高速的資料傳輸技術），而不是一般人所認為的「自動櫃員機」（automatic teller machines），至於他所提到那個關於「撞球」（billiards）的研究，其實是物理學家們在原子理論中用來指稱一種次原子粒子的詞彙，並不是像他所認為的是一種遊戲。桑福特議員以及另外一位來自於加州的議員，明白地表示他們想要懲罰 NSF，因為他們認為 NSF 竟然把經費投注在一些他們認為不必要的、浪費的研究上面。這些研究包括了那些調查人們為什麼要冒險把他們的資源用在加入社會團體上面、男性與女性的社會行為之間的差異、以及為什麼潛在的政治候選人決定要參加競選活動等等的研究。其他的國會議員則為 NSF 辯護，並且認為造成此種批評的原因，是因為這些政客本身所得到的資訊是錯誤的、草率的，而不是由 NSF 透過同儕審查過程所支持的研究類型所造成的（Lederman, 1998）。

正當合理的研究，或者，政客們也可能會去推動一些他們私下偏好的、但是卻沒有什麼科學價值的研究。申請政府經費補助的研究者們，有時候會以一種毫不吸引人的方式來重新陳述他們的研究計畫。大眾對於研究者的揶揄、或者是研究者申請經費卻被拒絕的情況，也會間接地鼓勵了研究者去進行嚴格的自我審查，以及促使社會大眾對於社會研究抱持著負面的意見與看法（請參閱範例 16.2 政治攻擊對於研究所造成的「寒蟬效應」）。

國家安全以及對於社會研究的限制 在第一次世界大戰以及第二次世界大戰期間，軍事機密以及國家安全在美國成爲重大的議題。大多數此類的議題都牽涉到製造武器的科技，但是有些研究者在研究其他的國家、涉及軍事利益的議題、以及政府本身時，也都會受到限制。美國的安全機構，像是「國家安全局」（National Security Administration）以及「中央情報局」（Central Intelligence Agency, CIA）等等，都會去影響那些探索 1950 年代冷戰時期的社會以及自然科學研究。

在冷戰期間，尤其是從 1940 年代晚期到 1960 年代中期，美國政府嚴格地控管那些探索非西方社會的社會科學研究。情報以及國安機構，與大多數研究中心以及學術機構都有緊密而且秘密的合作關係。在冷戰時期裡，政府裡的國安以及軍事機構、以及一些政治性的基金會，提供了研究經費給大多數關於其他社會的社會研究，而官員們則會嚴密監視研究者們所寫的報告以及陳述，以求符合於政府的政策。那些暗中爲政府機構工作、或者是與政府機構合作的學者，能夠很快地取得研究經費，而且他們的學術生涯也會快速進展。獨立的研究者、或者是那些質疑官方政策的人，就幾乎拿不到補助的經費，而且還會面臨生涯發展方面的限制。至於進行一些違反官方政策的研究，則是幾乎不可能發生的事。[25]

範例　16.2

政治攻擊對於研究所造成的「寒蟬效應」

在二十一世紀最初的幾年裡，一些被稱為「傳統價值聯盟」（Traditional Values Coalition）的美國國會議員，把矛頭對準了一些社會以及醫學方面的研究者，因為這些研究者即將得到「美國國家衛生研究院」（NIH）針對某些主題所提供的經費補助。因此，這些研究者們回報說他們現在正在進行自我審查的過程。坎普娜（Kempner, 2008）在 2005 年以及 2006 年時進行了兩次調查。在其中的一次調查裡，她訪談了一個由 30 名主要研究者（PI）所組成的隨機樣本，而這些研究者都是在上述的爭議當中被點過名的。至於在另外一次的調查當中，她則是寄送了一份問卷給所有被捲入上述爭議裡的主要研究者（其中有 86 人寄回問卷）。坎普娜發現到，其中大多數人會因為政治爭議的緣故而改變了他們的研究主題與過程。在經歷過政治攻擊之後，這些研究者會避免在他們的研究計畫書裡使用某些詞彙，或者是把研究調查的焦點改變為比較不具政治敏感性的主題。

在 1960 年代時，曾經有一個政府的研究計畫掀起了軒然大波。當時美國軍方資助了一個稱作**卡麥洛計畫**（Project Camelot）的研究，在該計畫當中，有數位備受敬重的社會研究者前往智利去研究當地的政治暴動以及動員情況。這項計畫的數個層面都引發了爭議。首先，該項計畫的目標，是應該要如何防止第三世界國家裡的農民或是弱勢團體去採取獨立的政治行動來反對某個獨裁者。此種反叛亂的研究通常都是由 CIA 所進行的。因此，這些研究者就被控訴是在利用他們的技巧以及知識來促進軍方的利益，以對付弱勢的第三世界國家裡的人民。第二，有一些研究者並不清楚研究經費的真正來源。第三，美國官員並未把這項計畫的內容告知智利政府裡的研究參與者。之後，當智利政府發現這件事之後，就馬上要求停止該項計畫，並且要求所有的研究者立刻離開智利。[26]

到了 1960 年代晚期以及 1970 年代時，進行研究的自由逐漸增加，對於跨文化研究者的限制也放寬了，而且被政府歸類爲機密檔案的文件數量也減少了。美國國會在 1966 年通過了「資訊自由法案」（Freedom of Information Act, FOIA），並且在 1974 年時強化了這個法案的內容。根據這個法案的規定，只要學者或是美國社會大眾向適合的機構提出申請，他們就可以取得或是閱覽許多政府文件。這股推動研究開放與自由的趨勢持續了大約 15 年，然後在 1980 年代時出現了逆轉的情況。

到了 1980 年代晚期時，美國聯邦政府擴大了機密文件的範圍，並且減少了可以公開取得的資訊。[27] 美國政府擴大了對於國家安全的定義，擴充了將政府文件列爲機密檔案的系統，並且對那些探討「敏感領域」的研究施加新的限制，即使這些研究並沒有牽涉到任何的政府機構或者是經費時也不例外。在新的規定當中，美國政府可以更加輕易地把原本屬於公共領域的資訊或是文件列爲機密檔案。除此之外，軍事以及國安方面的官員也能夠對外國研究者施加限制，阻止他們參加學術會議或者是參觀美國的教室、圖書館、以及研究中心等等。[28]

卡麥洛計畫　在 1960 年代時，由美國軍方出資、在智利境內所進行的一項具有爭議性的社會研究計畫。該項計畫違反了倫理原則，並且引起了重大的政治議題。

在冷戰時期，CIA 的臥底情報員經常會僞裝成社會研究者，以便於在其他國家裡蒐集資訊。一直到 1986 年爲止，CIA 一直有一條總括性的規定，該規定禁止研究者洩露他們的研究是由 CIA 所贊助的。而到了 1986 年時，這條規定已經鬆綁了，其適用範圍僅限於那些 CIA 認爲洩露背後的實際贊助者會傷害到美國國家利益的研究而已。例如，某位哈佛的教授曾經出版過一本關於美國外交政策的學術書籍，但是這名教授曾經與 CIA 簽訂過一份合約，承諾不會透露正是 CIA 提供經費讓他進行這本書的相關研究。[29]

跨國的研究則會涉及到一些較爲特殊的議題。科學社群譴責利用臥底情報員來僞裝成研究者、以及隱瞞此種研究真正經費來源的作法。關於在其他國家裡進行研究的倫理準則，詳細說明了研究者應該與地主國的官員合作、保護研究參與者、以及必須把資訊留在地主國裡。然而，研究者卻可能發現會有一些來自於他或她自己國家政府的干預，或者，爲了保護在一個非民主社會裡的研究對象的人權，研究者不得不對地主國的政府隱瞞一些敏感的資訊。[30]

在冷戰結束之後，全世界各地的社會研究者擁有更多的獨立性以及學術自由，因而得以研究各式各樣不同的社會。然而，2001 年 9 月 11 日在美國所發生的恐怖攻擊事件，引起了一些政治方面的變化，可能會因而導致美國政府進行更多的監控，而且也會影響到跨國的研究。

透過對於研究經費的控制而產生的間接限制 政治影響社會研究最常見的方式，是透過對於研究經費的控制，而此種情況與受到贊助的研究中所牽涉到的議題相當類似。大規模的研究計畫通常所費不貲，研究的成本甚至可能會高達 1 百萬美金之譜，而研究經費則通常是來自於私人企業或者是政府機構。

大多數的官員們都承認，一個開放而且自主的社會科學社群，是通往公正的、有效的知識的最佳路徑，而同儕審查的過程則是能夠促進自主性的研究。研究者向政府機構提交研究計畫書，以申請經費來進行研究。同儕研究者則爲政府機構來評估研究計畫的價值與優點。雖然政府出資贊助了大多數的基礎研究，但是在美國全國各地許多學院、大學、以及研究中心裡的研究者，也者會進行相關的基礎研究。

與大型企業花費在研究上的數目、或者是與政府贊助其他活動的金額相比之下，社會研究的花費可以說是小巫見大巫。在美國，大部分的研究經費都是由聯邦政府所贊助的，至於那些由大學以及私人基金會所贊助的研究計畫，相對而言在經費數目、規模、以及數量方面就比較有限。因此，對於大型的研究計畫而言，研究者通常還是必須得向政府申請經費才行。

在第二次世界大戰之前，美國境內一些由富有的家族所設立的私人基金會（像是卡內基基金會、福特基金會、洛克斐勒基金會、以及薩吉基金會等等）提供了經費來贊助大部分的社會研究。這些基金會希望能夠獲得一些隨著早期工業發展一起出現的嚴重社會問題的相關資訊。除此之外，這些基金會也想要去防止在激進份子以及社會研究者之間的連結，以及保護既有的社會機構與制度。在經過了幾年之後，「社會科學研究的生產因而變得規律化或是常規化，而社會大眾對於社會科學研究

與其贊助組織之間的關聯性也變得越來越視而不見了」（Seybold, 1987: 197）。私人基金會的經費把社會研究從早期應用的、行動導向的、批判的、以及包含了地區性參與的以社區爲中心的焦點，重新引導到朝向疏離的、專業的、實證主義的、以及學術的研究上面。在第二次世界大戰之後，美國政府擴增了提供給各種研究的補助經費。當聯邦政府所提供的經費數目在 1960 年代超越了私人基金會所提供的經費數目時，私人基金會在決定研究的優先次序方面仍然維持一定的角色。[31] 雖然政府的研究經費已經增加了，但是對於社會科學以及社會學的贊助經費仍然少得可憐。在美國聯邦政府提供給基礎研究的所有經費當中，社會學的相關研究所獲得的研究經費所佔的比例還不到 1%。

在美國，研究者可以向數個聯邦機構申請社會研究的經費，其中包括了「美國國家科學基金會」；國防部門、司法部門、勞工部門、商業部門、住宅與都市發展部門、以及教育部門；「美國國家人文基金會」（National Endowment for the Humanities）；「中小企業管理局」（Small Business Administration）；以及在「衛生以及公眾服務部門」（Department of Health and Human Services）之下的許多機構。聯邦政府本身也雇用了許多研究者，但是大多數的社會研究都是在學院以及大學、或者是在獨立的研究機構當中進行的。

基於政治方面的理由，美國政府裡提供經費給社會研究的主要機構（像是美國國家科學基金會以及美國國家衛生研究院）在剛成立不久時，只有支持基礎的實證主義研究。社會研究者們之所以同意把非實證主義的社會研究以及應用研究排除在外，其目的在於想要獲得自然科學家們的支持、反駁當時認爲社會科學是「軟性學科」的流行看法、以及抵制在意識型態方面的保守份子認爲社會科學具有「左派」性質的指控。除此之外，「美國國家科學基金會」也會避免提供經費給那些主題具有爭議性（例如：性、政治權力）的研究，這是因爲在 1950 年代以及 1960 年代的政治氛圍當中，該基金會擔心針對這些主題所進行的研究有可能會危及大眾對於社會科學研究的支持。

政治過程決定了各種與社會研究有關的機構能夠取得多少經費預算，以及應用研究與基礎研究之間的區隔標準。雖然提出申請的研究計畫書是由各個機構裡的科學審查委員會來評估其科學價值的，但是整體經費的總額度卻是由政府官員來決定的。政客會依據政黨或是意識型態的利益來排定優先順序，而此種情況則會影響到各項研究所能獲得的經費額度（請參閱範例 16.3 政治對於美國的犯罪研究所產生的影響）。

在 1976 年到 1980 年之間，對於社會研究的意識型態批評，導致美國國家科學基金會刪減了 24%對於社會科學研究所提供的經費。雖然研究者們強烈抗議，在 1980 年到 1983 年之間，補助的經費仍然繼續下滑了 17%。此種情況是因爲有些政治領袖相信，研究的結果支持了他們在意識型態上的對手所提出的政策。除此之外，政客們也刪減了應用研究的經費。爲了回應此種情況，數個社會科學領域裡的專業協會聯合起來成立了一個遊說的組織，而該組織的名稱是「社會科學協會聯合會」（Consortium of Social Science Associations, COSSA）。[32]

範例 16.3

政治對於美國的犯罪研究所產生的影響

薩弗斯柏格以及同事（Savelsberg and colleagues, 2002）想要瞭解在 1951 年到 1993 年之間，美國國內的政治壓力是否改變了關於犯罪議題的社會研究的方向。他們閱讀了許多學術期刊的文章，並且提出了下列的問題：在政治上的轉變是否會透過提供經費給研究的方式來影響研究？在學院以及大學裡的學術領域組織方式的變化，是否會影響到研究中所使用的理論（換句話說，犯罪是個人的問題，還是社會力量或不平等現象所造成的），以及研究所檢視的主題（也就是微觀層次的執法情形，或者是鉅觀層次或是對於犯罪行為的瞭解）？他們發現到下列兩種機構都會對研究者所進行的研究類型造成影響：其中一種是試圖推動某種政治議程的政府機構；而另外一種則是一些新的大學系所，而設置這些系所的目的，其實是為了要與政治利益更為緊密合作並且更能迅速作出反應，而不是作為一種獨立的研究社群。然而，雖然經費以及新的組織單位會影響到研究者所研究的主題類型以及他們所考驗的理論，但是這些因素卻沒有影響到最後的資料是否支持這些犯罪的理論。因此，政治力量確實會對研究者所關注以及從事研究的理論、主題、以及觀點造成影響，但是政治因素卻無法影響到研究者設計或者是進行研究的方式、或者是他們所得到的研究結果。

社會研究經費的整體水準，可能在 90 年當中都維持不變的情況。在針對通貨膨脹以及學術專業領域範圍的擴張進行調整之後，在 1920 年代晚期來自於私人的「社會科學研究理事會」（Social Science Research Council）所提供的經費，很可能超過當今的美國國家科學基金會所提供的經費總額。[33]

政治價值觀也可能會限制了研究者能夠研究的問題、以及設定了研究的優先順序。藉由集中焦點在某些研究問題上面、以及限制了其他選擇的方式，一些政治團體試圖去形塑研究者所能進行的研究、以及我們所能取得的各種關於社會的資訊。例如，政客們可能分配經費給某個應用研究，因爲該研究證明了某些法規對於大型企業將帶來許多「難以負擔的」成本。然而，這些政客們卻可能會拒絕補助經費給另外一個研究，因爲該研究的主題在於調查這些法規對於消費者所能帶來的益處。政客們也可能會增加經費以補助研究者們去研究由毒癮犯所犯下的罪行，但是卻拒絕提供經費讓研究者們去研究由企業高層們所犯下的罪行。除此之外，政客們或許會提供經費給那些探討如何促進創業精神的研究，但是卻可能拒絕提供經費給那些探討取消社會福利對於人們會造成何種影響的研究。[34]

在 2001 年到 2008 年之間，政治意識型態對於所有科學研究的介入情況日漸嚴重。在 2004 年時，「科學家關懷社會聯盟」（Union of Concerned Scientists）發表了一份由 8,000 名傑出的科學家所連署背書的聲明，在該份聲明中提到，小布希總統的政府已經把科學政治化到極爲驚人的程度，而此種情況已經嚴重地違反了美國兩黨總統以及行政單位長久以來的作法（Kevlevs, 2006: 761）。

社會研究者所處理的議題，是那些與社會信念、價值觀、以及政策有直接關聯性的議題。政治團體以及充滿意識型態的政客對於這些議題所設定的優先順序，與科學社群所設定的優先順序截然不同。對於社會研究在處理社會需求以及促進知識發展方面的能力而言，此種現象同時具有正面以及負面的影響。首先，此種現象確

保了政客們以及聲勢浩大的政治團體所關注的事情會被優先處理，而且也會有人去研究那些被具有政治影響力的團體定義爲極具重要性的議題。然而，即使科學研究的結果並未支持一個普遍的大眾迷思（例如：死刑具有強烈的嚇阻效應），政客們以及政治團體也會不斷地施壓並且提供經費，以試圖去找出證據來支持他們那些不科學的、民粹的信念。與此同時，那些被科學社群認爲極度重要的議題，卻可能落到沒有經費可以進行研究的下場。

話雖如此，科學社群仍然擁有一些選擇研究問題的自由，但是那些會影響到政治上屬於邊緣地位的社會團體的議題、或者是那些沒有任何組織來進行遊說的議題，仍然只能獲得有限的研究經費補助。此種經費分配不平衡的情況，也會造成關於各種議題的知識出現了不平衡的狀況。到了最後，我們擁有大量知識的議題，都是那些符合強大的政治團體利益的議題，而對於社會中的弱勢團體所抱持的觀點，我們卻可能會一無所知。

在美國，政客可以否決那些已經通過嚴謹的同儕審查、而且被認爲具有科學價值的研究計畫書，即使這些政客並未閱讀過這些計畫書、並且只是因爲政治意識型態的理由而不喜歡其研究主題而已。[35] 例如，在 2005 年時，美國的眾議院就拒絕提供經費給兩個在美國國家心理衛生研究院（NIMH）中已經通過同儕審查的研究計畫。在這兩個研究計畫中，其中一個計畫所探討的是鴿子的視覺能力，而另外一個則是檢視了心理特質如何有助於維繫成功的婚姻。阻止補助經費給這兩項研究的，是一名來自於德州地區的眾議院議員，他本身是德州地區的不動產開發商，而且沒有任何科學方面的專門知識。此種政治方面的干預，會對同儕審查的過程造成傷害。事實上，這位議員也曾經試圖阻止美國國家心理衛生研究院提供經費給另外兩項研究（其研究主題分別是人們的自我表達以及價值體系），不過幸好他並沒有成功。[36]

指定用途或者是「政治分肥」的研究經費　從 1990 年代開始，在分配美國政府對於研究所提供的經費時，有越來越多的政客刻意地去規避科學的同儕審查過程。這些政客們會爲經費「指定用途」（“earmarked”），也就是刻意把研究經費分配給特定的大學或是研究機構裡的特定研究計畫。這些政客們在分配研究經費時，所依據的標準是政治上的偏好，並不是研究計畫書在研究品質以及價值方面的公平競爭結果，而研究計畫的品質以及價值則是由科學社群中經驗豐富的成員所評定的。表面上看來，此種「政治分肥」（“pork barrel”）的作法，已經擴散到了對於研究經費的分配過程當中了。所謂的「政治分肥」，指的是政客在決定該分配多少預算給重大的政府工程計畫時，其考量的標準並不是這些工程計畫的重要性或是高度優先性，而是這些計畫能不能爲自己選區內的企業以及支持者帶來財務上的利益。

因此，越來越常見的情況是，在某些州或是選區裡的研究者通常可以獲得鉅額的研究經費補助，而在其他地區裡的研究者則幾乎完全申請不到經費，其中的區別就在於政治方面的人脈與關係，而不是科學方面的價值。例如，紐約州立大學水牛城分校（State University of New York at Buffalo）曾經獲得一筆 1,200 百萬美元的經費，讓該校去進行一項關於交通傷害的研究。這筆經費並非是經過良性競爭而獲得

的，而只是來自於一個政治的決策。從 1989 年到 1993 年之間，在政治方面指定用途的研究經費總額增加了一倍，而且在之後的幾年當中，其總額都維持著相當穩定的狀態。而自從 1996 年開始，指定用途的研究經費總額則是快速成長了五倍，目前是每年高達大約 20 億美金左右。

此種政府補助的研究經費被政治化的現象，迫使許多大學以及研究機構必須去討好一些具有影響力的政客。舉例而言，在 1995 年時，新罕布夏州沒有獲得任何指定用途的研究經費。四年之後，當新罕布夏州的共和黨參議員葛瑞格（Judd Gregg）擔任參議院裡的預算委員會的主席時，新罕布夏州裡的研究者們就因而受惠，而當時該州在獲得最多政府研究經費補助的州裡面排名第七。當參議員麥肯（John McCain）於 2001 年試圖終止此種研究經費的政治分肥現象時，美國的參議院以 87 比 12 的票數否決了他的提議。許多政客對於此種政治分肥的情況相當自豪，並且還會去誇耀他們依據政治偏好（而非科學價值）爲自己的選區所爭取到的各種研究經費。因此，爲了獲得研究經費，大學以及研究機構不得不花費更多的時間與努力來討好政客以及進行遊說，而不是去鼓勵研究者們寫出在以價值爲主的同儕審查過程中具有競爭力的研究計畫書。[37]

正如同在《高等教育內幕》（*Inside Higher Ed*）這本雜誌於 2010 年 4 月 29 日那一期裡所報導的：「毫不令人意外地，在學術領域內獲得最多指定用途補助經費的學校，都是位於那些在國會裡一些最有權力的議員所代表的州。有四個密西西比州的研究機構……是在獲得最多研究經費的前 25 名之內，而此種情況大部分都是參議員寇可蘭（Sen. Thad Cochran）的功勞，因爲這名密西西比州的資深參議員在參議院的預算審議委員會裡是最具影響力的共和黨員」（請參閱表 16.1）。有一項研究認爲，對於獲得經費補助的大學而言，指定用途的研究經費或許會增加這些大學裡的研究人員所發表的報告數量，然而，同時卻可能會降低了新增加的研究報告的整體品質（Payne, 2002）。

許多研究機構以及大學已經開始轉向私人捐贈者（富裕的個人、企業、或者是基金會）申請研究經費。然而，來自於私人捐贈者的經費通常都會附帶一些條件或者是限制。例如，曾經有一名私人捐贈者在提供經費給某位研究者之後，卻收回了 45 萬美金，因爲這名研究者曾經公開批評過這個捐贈者所支持的一個政策。[38] 有些捐贈者會毫無附帶條件地支持獨立的研究者，但是卻有許多其他的捐贈者利用捐出去的經費來製造出一些微妙的壓力，以提倡某個他們偏好的政策立場、意識型態、或者是政治議程。雖然大學以及研究機構會盡力去避免來自於私人捐贈者對研究經費所設下的限制，但是它們仍然必須在下列兩者之間找到一個平衡點：其中一方是來自於捐贈者的、而且是大學與研究機構所迫切需要的現金；而另一方則是抽象的理想，像是研究者有自由去進行以及發表任何能夠促進知識發展的研究等等。然而，有些大學或是研究機構爲了解決要去歸還或是拒絕來自於捐贈者的經費的窘境，只好同意對公開、自由的研究設下一些限制。

對於散佈知識的限制 科學社群的重大規範之一，就是要公開地散播知識。強而有力的團體或者是機構，則會試著以限制資訊流通、限制發表研究結果、或者是強

表 16.1　2010 年獲得最多美國國會研究經費的 15 個高等教育機構

高等教育機構	獲得的經費總額
1. 阿拉巴馬大學吐斯卡魯薩校區（University of Alabama at Tuscaloosa）	$ 58,755,000
2. 密西西比州立大學（Mississippi State University）	$ 47,919,000
3. 德州農工大學（Texas A&M University）	$ 40,150,000
4. 北達科塔大學（University of North Dakota）	$ 39,660,000
5. 北達科塔州立大學（North Dakota State University）	$ 37,040,000
6. 密西西比大學（University of Mississippi）	$ 33,655,000
7. 夏威夷大學（University of Hawaii）	$ 33,503,000
8. 麻州大學波士頓校區（University of Massachusetts at Boston）	$ 33,002,000
9. 猶他州立大學（Utah State University）	$ 27,190,000
10. 新墨西哥礦業與科技學院（New Mexico Institute of Mining and Technology）	$ 27,000,000
11. 路易斯安娜州立大學（Louisiana State University）	$ 26,650,000
12. 南密西西比大學（university of Southern Mississippi）	$ 22,590,000
13. 西維吉尼亞大學（West Virginia University）	$ 21,920,000
14. 路易斯維爾大學（University of Louisville）	$ 20,150,000
15. 肯塔基大學（University of Kentucky）	$ 19,709,000

資料來源："The Academic Pork Barrel, 2010" from *Inside Higher Ed*, April 29, 2010; http://www.insidehighed.com/news/2010/04/29/earmark（本資料擷取於 2010 年 5 月 16 日）

迫研究者保持沉默的方式，來侵犯或者是打擊社會研究。

有一則 1997 年的新聞報導，可以用來證明外界的機構或是團體打壓研究發現的嚴重情況。[39] 有一家製藥公司生產了一種廣泛被用來對抗甲狀腺問題的藥物，然而，該製藥公司卻曾經試圖阻止一支大學的研究團隊發表其研究結果，因爲該研究結果顯示出上述的藥物其實根本不具有任何療效。爲了交換研究經費，研究者們之前就已經簽署過一份合約，而這份合約賦予該製藥公司阻止他們發表研究結果的權利。其他的研究則顯示出，在這家製藥公司所提供經費補助的研究裡，其發表的研究發現當中有 98%都顯示出上述的藥物是具有療效的。然而，當研究的經費並不是來自於這家製藥公司時，認爲上述藥物具有療效的研究發現的數目，就會低很多。有些人相信，當牽涉到某家公司數百萬美金的利潤時，關於新產品的負面發現就會被隱瞞起來。研究者可能會獲得股票或者是財務方面的誘因，以誘使他們發表正面的發現、或者是延後發表負面發現的時間。在接受過來自於製藥公司或者是生技公司經費贊助的大學研究者當中，有超過一半以上的人都提到，這些私人捐贈者會對他們進行研究的方式施加壓力。

反公共參與的策略性訴訟　一種訴訟的類型，一些富裕的、有權力的組織利用此種訴訟來威脅研究者，並且阻止研究者公開地表達自己的想法或是揭露一些資訊。

在醫藥或是生物科技方面的研究，並不是唯一一個私人利益會與散布研究發現產生衝突的領域。在 1997 年時，一名康乃爾大學的教授在一場鎮民大會中作證發言 10 分鐘，內容是關於全美國最大的療養院公司「比佛利企業」（Beverly Enterprises）（該企業在全美國總共經營了 700 家療養院）對待勞工的不當作法。這名教授的證詞是來自於多年的研究結果，並且也被國會報告、新聞報導、法庭記錄、訪談、以及其他學者所詳細記錄下來。到了 1998 年時，這家公司控告這名教授毀謗，求償美金 22 萬 5 千美元，並且要求索回多年的研究文件以及筆記。此種訟訴案件被稱爲**反公共參與的策略性訴訟**（Strategic Lawsuits against Public Participation, SLAPP），而其目的是在於阻止人們作出對某些企業或是組織不利的公開證詞。

此種作法肇始於 1970 年代，當時有一些公司提起了「策略性訴訟」，以消除在一些爭議性議題上面的反對聲浪。

在一項關於企業罪行的研究中，由於接受過訪談的管理者們威脅要對研究者提起訴訟，因而使得發表研究的時程延後，並且也迫使研究者修改研究結果。有一項針對寄宿學校所進行的研究，也因爲學校行政人員威脅要提出訴訟，因而阻止了研究結果的發表。學校裡的行政人員想要更改他們在訪談中所說過的話，並且作出其他的更正，因爲他們並不認同研究者所作出的結論。在另外一個例子裡，一名可疑的研究者在被控告有利益衝突的問題之後，則是藉由提出訴訟的威脅方式，迫使其他同儕研究者針對他們所發表的一篇研究文章作出一些更改。[40]

在 2002 年到 2004 年之間，曾經有過一些對於政府的嚴重指控，當時美國的聯邦政府限制了某些科學資訊的發表，因爲這些資訊違反了、或者是未能支持行政團隊的政策立場。這些對於政府干預的指控內容，包括了嚴格審查保險套效果的資料、阻止研究者發表一些顯示了禁慾沒有比性教育更有效果的證據、以及政府官員指示「美國國家癌症研究院」（National Cancer Institute）在其網站上發布一項聲明，宣稱墮胎會增加罹患乳癌的機會（雖然某個大型研究已經顯示出這兩者之間沒有任何關聯性）。除此之外，政府官員也因爲某些政治顧問的抗議、而非科學家的意見，而禁止散布任何關於全球暖化的報告。美國環保署（Environmental Protection Agency）也宣布，該機構不會去分析任何違反官方政策的關於污染的研究。一位美國農業部門的研究者曾經研究過如何透過飲食來減少豬圈裡的臭味，而他也同時發展出能夠偵測空氣污染物的相關應用方法。該項研究在無意之間透露出大規模的養豬場經常違反了聯邦的污染法規，並且製造了能夠抵抗抗生素的細菌。養豬業裡的一位成員得知了這項研究的內容，並且與這名研究者的上司取得聯繫，之後，這位上司就禁止這名研究者在研究會議中發表研究發現、或者是把研究內容投稿到科學期刊上。[41] 事實上，上述這些行動遠不如伊朗政府所曾經做過的事情那麼激烈。伊朗政府曾經把一些調查研究者關入監獄當中，只因爲這些研究者的研究發現顯示出，絕大多數的伊朗人民都希望改善與美國之間的關係，而此種結果與伊朗政府的官方政策是完全背道而馳的（請參閱重點摘要 16.1 政治議題摘要）。

重點摘要　16.1

政治議題摘要

對於研究的直接限制

1. 政府（或者是有權力的政治團體）會去禁止、開除、囚禁、或者是威脅某些教授或是研究者，因為這些教授或研究者研究了一些不受歡迎的主題、公開討論一些「禁忌的」話題、或者是作出一些政府反對的言論。
2. 在政府機構或是大型組織裡的官員或人員，阻止研究者去取得官方文件或統計資料，或者是試圖去限制官方資料的蒐集方式或公開的方式。
3. 政客或是高層官員批評、攻擊、或是公開施加壓力，意圖阻止合法正當的社會研究，因為他們基於個人的、宗教的、或者是意識型態的理由而反對這些研究。
4. 官員們試圖去阻止或是嚴格審查某些研究，因為他們相信這些研究可能會妨害了國家安全。除此之外，官員們也可能暗中試圖去控制社會研究，以達到其本身在軍事或是秘密蒐集情資方面的目的。

透過經費補助而對研究產生間接限制

1. 研究經費的限制或是刪減，將會妨礙了新知識的產生，而這些新知識或許能夠挑戰固有的意識型態信念或者是政治觀點。
2. 對於接受經費補助的研究在主題或是議題方面的控制，會改變了產生新知識的方向，如此一來，這些研究的結果將會支持特定的政策立場。
3. 政客們在經費補助方面的政治分肥作法，會迴避了同儕審查的過程，而且在分配研究經費時，所依據的標準不是研究的科學價值，而是政治上的偏好、或者是圖利於政客自己選區裡的選民們。
4. 為了達到政治目標，而對研究者所能使用的技術、工具、或是服務設下重重的限制，這些限制不但與科學研究的過程毫無關係，而且還會增加進行研究的成本、時間、或者是複雜程度。

在散布研究結果方面的限制或是偏見

1. 如果研究者在一個公共的論壇中自由地發言、或者是公開地發表他們的研究結果，將會受到法律訴訟或是罰款的威脅。
2. 在研究者被允許去與科學社群或是大眾分享研究發現之前，一些非科學家的人（也就是企業管理者或者是政府官員等等）會要求先進行審閱以及過濾。
3. 政府官員以及其他具有影響力的人們，會去宣傳一些特定的研究發現，雖然這些研究發現被科學社群認為是具有嚴重缺陷、薄弱而且無說服力、或者是不恰當的，但是卻可能對上述這些人的政治議程有所幫助。

散播研究發現

實證主義的研究者們認爲，價值觀可以合理地在下列的兩個領域裡發揮作用。首先，研究者可以選擇一個主題領域或者是研究問題。雖然在主題領域當中有一些尚待進一步探索的「科學新領域」（“scientific frontiers”），但是研究者仍然可以根據個人的喜好來選擇研究問題。[42] 第二，在完成一項研究之後，研究者的價值觀會影響到他們散播或是發表研究發現的管道。科學社群認爲研究者應該要公開地發表研究發現，而贊助研究經費的機構也會要求一份研究報告，但是除了上述這些要求之

外，決定要如何以及在何處發表研究發現，其實是由個別的研究者自己決定的。

關聯模式 在完成了一項研究之後，可能會產生一種盧爾（Rule, 1978a, 1978b）稱爲**關聯模式**（models of relevance）的倫理政治議題。盧爾檢閱了社會研究者對於自己的研究以及其用途所採取的立場，並且主張研究者們所採取的立場可以合併成爲五種基本的類型（請參閱資訊箱 16.6 關聯模式）。

關聯模式是一種關於社會研究者所應該採取的立場的理想型。例如，研究者是否爲一種技術人員，專門生產關於社會運作方式的有效、可靠的資訊，以供其他人去使用？或者，研究者是否屬於一種獨立的專業人士社群，對於提出何種研究問題以及研究結果的運用方式具有發言的權利？在一條連續線上，其中的一個極端就是不具道德意識的研究者，此種研究者對於研究以及研究結果的使用方式既不關心、而且也沒有控制的權力。此種研究者只供應其他人所要求的知識，然後對於其他的事情全都漠不關心。此種情況就是許多在德國被納粹統治時的科學家所採取的立場，並且以此種立場來爲自己與納粹黨合作的事情作出辯護。當時這些科學家所做的事在之後被歸類爲「危害人類的罪行」（"crimes against humanity"），然而，這些科學家卻認爲他們「只是遵守命令」以及「做該做的工作」，但是「沒有提出任何問題」。而在另外一個極端，則是那些能夠全權掌控研究以及研究結果的使用方式的研究者。

我們在第四章中所討論過的各種社會科學取向，與各種不同的關聯模式以及政治觀點之間具有相關性。[43] 實證主義者傾向於依循「直接而且正面的效果」或是「特定的支持對象：政府」這兩種關聯模式。詮釋取向的研究者則是依循著「沒有淨效果」或是「特定的支持對象：弱勢團體」這兩種關聯模式。至於批判取向的社會研究者，則是依循著「特定的支持對象：勞工階級」或是「特定的支持對象：弱勢團體」這兩種關聯模式。

特定的研究者或者是研究計畫可能會橫跨不同的模式。例如，懷特（Whyte, 1986）就曾經描述過，關於員工所有權的研究，就橫跨了三種不同的特定支持對象（勞工階級、弱勢團體、以及政府），而且具有直接而且正面的效果。

自從盧爾提出了關聯模式的概念之後，在美國境內隨著一種非政府的私人**智庫**（think tank）在數量以及規模方面大幅度地成長，已經出現了一種新的關聯模式。而這第六種模式就是*特定的支持對象：富裕的個人以及企業*。此種模式認爲，社會研究能夠反映出研究者的政治價值觀，並且促進那些想要維持或是擴張其權力的富裕團體的政治目標。智庫是由富裕的個人、企業、以及政治團體所支持贊助的研究以及公共組織。例如，像是「曼哈頓研究院」（Manhattan Institute）、「卡托研究院」（Cato Institute）、「傳統基金會」（Heritage Foundation）、以及「美國企業研究院」（American Enterprise Institute）等等智庫，在 1980 年代到 1990 年代之間有著戲劇性的急速成長。

關聯模式 一組理想型，內容是關於社會研究者對於進行研究的目的、以及研究結果用途的理解方式。

智庫 一種組織（通常是非營利的、非政府的組織），在其中有一個或是多個研究者、作者、記者、以及其他人去發展、修正、詳細闡述、以及宣傳一些關於政策議題的想法。

資訊箱 16.6　關聯模式

1. *沒有淨效果*　社會科學的發現沒有為社會帶來更大的益處。有數個知名的社會科學家主張此種觀點，其中包括了孫末楠（William Graham Sumner）、巴烈圖（Vilfredo Pareto）、史賓塞（Herbert Spencer）、班菲爾德（Edward Banfield）、以及威爾森（James Wilson）等等。這些保守派的社會科學家認為，研究的產物能夠被用於促進任何人的自我利益，而且他們也相信，從長遠來看，社會科學產生的知識所帶來的害處，和好處是相等的。
2. *直接而且正面的效果*　社會科學的知識能夠為所有人帶來益處。採取此種立場的自由派社會科學家，像是莫頓（Robert Merton）等等，認為關於各種社會關係的知識，有助於創造一個更為理性的社會。關於社會問題的研究結果，能夠幫助我們更加瞭解社會世界，並且使我們有能力知道，我們可以用什麼方法使社會世界變得更好。舉例而言，林德布洛姆以及柯漢（Lindblom and Cohen, 1979）就曾經強烈建議，社會科學應該要改變方向，朝向一種他們認為是解決社會問題的方向發展前進。
3. *特定的支持對象：勞工階級*　社會科學應該被用來提升勞工階級的利益以及地位。這是關於社會研究適當用途的馬克思主義模式。根據此種模式，所有的社會研究都可以被分為三種類別：無價值的、幫助中產階級的、以及協助勞工階級的。與批判科學取向一致的是，此種模式認為，研究發現應該要能夠揭發及對抗對於勞工的剝削、壓迫、不公、以及迫害等等情況，藉由此種方式提倡以及捍衛勞工階級的利益，並且為勞工們提供協助。
4. *特定的支持對象：弱勢團體*　社會科學應該被用來幫助社會中任何一個弱勢的，或者是未能享有任何權益的團體。此種模式通常會與曼海姆（Karl Mannheim）以及米爾斯（C. Wright Mills）聯想在一起，而且也比馬克思主義者的立場更為廣泛。此種模式認為，在社會中有許多社會團體是缺乏力量的（像是女性、消費者、少數種族團體、同性戀者、以及窮人們等等），並且主張這些弱勢團體受到了社會中其他握有權力的團體所壓迫。因為後者擁有獲得教育、財富、以及知識的管道。社會研究者應該捍衛那些在社會上沒有發言權、而且被有權勢者所操縱的人們。有權勢者可以利用或收買社會科學研究，以達到他們的目的。由於社會研究者在社會中扮演了獨特的角色、並且處於能夠瞭解社會所有領域的位置上，因此，社會研究者有義務幫助弱勢者，並且與他們分享知識。
5. *特定的支持對象：政府*　社會科學的適當角色，就是協助社會中的決策制訂者，尤其是政府官員們。此種模式是由美國參議員莫依尼漢（Senator Daniel Moynihan）所表達的，而且也出現在美國國家科學基金會的官方政策報告當中。除此之外，此種模式在非民主的社會當中也相當普遍。此種模式類似於第二種模式（直接而正面的效果），但是其中增加了一項假定，也就是政府最有資格運用社會研究的發現，並也會完全地投入於消除社會的問題。此種模式也類似於第一種模式（沒有淨效果），但是卻暗示了在對國家忠誠的限制之下，「販售」或是把研究發現提供給出價最高的人。此種模式也假定，政府是以對所有人都最有利的方式運作的，而且研究者具有一種愛國的義務，必須把他們所得知的資訊交給擁有政治權力的官員。

這些智庫都具有特定的政治意識型態觀點，而且會利用社會研究或者是僞科學來推展其特定的議程。智庫會支付薪資給研究者、贊助研究報告、並且使大眾注意到那些支持其特定觀點的研究結果（請參閱範例 16.4 倫理、政治、以及調查研究的濫用）。

由智庫所進行的研究，在品質方面有很大的落差。智庫所進行的研究通常沒有經過同儕審查的過程、而且也缺乏堅實的證據，但是卻有冗長的建議事項。這些研究的讀者群並不是科學社群，而是政客、記者、以及一般大眾。這些研究的主要目標並不是在推動知識的進展，而是在於提倡一種意識型態的觀點，以形塑大眾的思想或者是影響政治辯論。智庫通常會獲得大量的媒體宣傳、名望、以及財富。同時，那些只能靠著微薄的經費來進行研究、而且缺乏與大眾媒體的關係的傳統社會科學家們，則會發現到，大眾以及記者們完全忽視了他們針對同樣的公共議題所進行的嚴謹而仔細的研究。這是因爲大眾以及記者們的注意力已經完全地被智庫所發表的研究結果給吸引過去了。

在發表研究結果之後 科學社群裡的成果共享主義規範，要求科學家們必須公開他們的研究發現。一旦研究發現成爲公共領域中的一部分，研究者就失去了對於這些發現的掌控權力，而此種情況意味著其他人可以把這些發現用於他們自己的目的上。雖然研究者或許是根據自己的價值觀來選擇研究主題，然而一旦研究發現被發表之後，其他人就可以利用這些發現來推展對立的價值觀。

例如，讓我們假定你想要提升某個美國原住民部落的政治權利。你研究了該部落的社會習俗，其中包括了那些使該部落無法在社區裡取得更大政治權力的障礙。一旦你發表了研究發現之後，雖然部落裡的成員們可以運用研究結果來打破這些障礙，但是反對者也可以利用相同的研究發現來限制該部落的權力，以及強化這些障礙。

會影響未來行為的研究發現 你是否曾經因爲讀到了某個研究發現，所以採取了和過去不同的方式來做某件事？如果是的話，那麼你並不是一個特例。有些時候，研究發現的散播可能會對社會行爲產生影響。此種情況的範例之一，就是政治民調結果所產生的效果。民意調查會影響到選民的政治偏好，換句話說，有部分的人會因爲民意調查的結果而改變了他們的觀點。[44]

其他的社會研究發現也會對行爲產生影響。事實上，研究發現的散播可能會以一種抵消或是改變原來發現的方式而影響了行爲。例如，有一項研究發現到，專業人士傾向於給他們的孩子關於學業成就方面的大量壓力，也因此而造成他們的孩子極度焦慮而且很不快樂。如果專業人士們讀到了這項發現，他們可能會改變他們養育孩子的行爲方式。而在數年之後，另外一項研究或許會發現到，與其他的團體相較之下，專業人士對於他們孩子在課業成就方面的要求並沒有比較高。

對於那些會影響社會行爲的研究發現，研究者們可能有下列幾種反應：

1. 此種研究發現破壞了人類社會行爲的可預測性以及規律性，因而破壞了複製的可能性。
2. 此種研究發現所改變的只是瑣碎的行爲，所以此種情況是只有在非常狹隘的應用

範例　16.4

倫理、政治、以及調查研究的濫用

在一次極為不尋常的事件當中，具有領導地位的專業民意組織「美國民意研究學會」（American Association for Public Opinion Research, AAPOR）尖銳地批評了兩個組織，因為這兩個組織在調查研究中明目張膽地從事了一些不道德的行為，以推動其狹隘的政治目標。在 1997 年時，美國民意研究學會發現到，「倫茲研究公司」（Luntz Research Corporation）裡的倫茲（Frank Luntz）「不斷地拒絕公開關於他的研究裡的關鍵因素。」倫茲的調查顯示出民眾強烈支持一項由共和黨於 1994 年 11 月提出的、稱為「與美國的合約」（"Contract with America"）的提案，但是其他的研究者並未發現此種現象。倫茲廣為宣傳他的調查發現，但是卻拒絕揭露基本的方法論資訊，而此種資訊在符合倫理的調查中是必須要揭露的。

三年之後，美國民意研究學會則是批評「競選活動電話公司」（Campaign Tel, Ltd.）嚴重地違反了保密原則。競選活動電話公司使用了一份名單，名單中列出了威斯康辛州已經登記註冊的選民名字以及電話號碼，而且該公司還宣稱是正在進行一項調查。事實上，該公司後來把關於調查結果的詳盡資訊以及電話號碼交給了威斯康辛州的共和黨分部。美國民意研究學會提到，該學會「強烈譴責任何以除了合法的調查研究之外的目的而偽裝成調查研究、並且從受訪者取得資訊的作法。」競選活動電話公司刻意地隱瞞了該公司以及上述調查研究的真正本質。而當美國民意研究學會發現並且公布了上述的不道德行為時，競選活動電話公司早已經不存在了。

*資料來源：*請參閱 AAPOR 的網站：www.aapor.org/main1.html

領域裡工作的研究者才會關心的議題。

3. 此種研究發現改變了人類行爲，因爲在人類行爲中幾乎沒有什麼不能改變的法則，而且人們將會運用在公共領域裡的知識來改變自己的生活。

不論在任何情況下，社會研究都無法去揭露人類關係以及行爲的完整複雜性。即使社會研究真的能夠做到這一點，而且此種知識也能完整而且正確地散播給所有人知道，社會研究者仍然必須去研究哪些人類行爲發生了變化、以及發生變化的過程爲何。

學術自由　大多數的學生都曾經聽說過學術自由，但是卻沒有什麼人真正瞭解學術自由的真正涵意。**學術自由**（academic freedom）所指的是一種開放而且幾乎不受限制的氣氛，以促進意見以及資訊的自由交流。在開放的民主社會當中，許多人會珍惜學術自由，並且也堅決相信應該要提供學者們免於受到干預的自由。此種觀念是建立在下列這個信念之上的：基礎的民主機構、公正知識的發展、以及言論自由等，都需要意見以及資訊的自由流通才能達成。學術自由與研究自主性有關。關於研究主題的新想法、對於研究發現的詮釋、理論或是假設的發展、以及觀念的公開討論，都必須以學術自由爲前提才行。

在學院、大學、以及研究機構裡的學術自由，提供了進行科學研究不可或缺的自由討論以及意見交流的環境。爲了促進知識的發展，研究者、教授、

學術自由　一種概念，也就是研究者以及／或是教師能夠自由地檢視所有的主題以及討論所有的觀念，不必遭受到來自於教師、學者、以及科學家社群之外的任何人或是權威所施加的任何限制、威脅、或者是干預。

以及學生需要一個自由的環境，讓他們可以在其中去自由地提出或是辯論各種不同的、有時可能是不受歡迎的意見或是立場。換句話說，在此種自由的環境中，人們可以無所畏懼地在公開的討論中、在教室裡、在公開的談話中、以及在出版品裡去盡情地探索各種想法與觀念的完整內容。

你如果想要瞭解學術自由的重要性，只要去看看在沒有學術自由的地方裡，社會研究的數量有多麼稀少，就可以略知一二了。那些想要限制討論或者是強加某種觀點在他人身上的社會政治後援團體以及政府官員，都可能會威脅到學術自由。對於學術自由的限制，將會妨礙了與社會有關的知識的成長，而且也會傷害到研究過程的完整性。

在十九世紀晚期時，由於當時社會科學正在大學裡逐漸制度化，學術自由才成爲一個具有重要意義的議題。在早期時，教授們可能會因爲政府官員或是經濟菁英不喜歡他們在課堂裡或是在文章裡所表達的意見而失去工作。像是韋伯倫（Thorsten Veblen）之類的美國早期社會科學領域裡的知名學者，曾經因爲他們在課堂中的言論、或者是他們在文章裡所表達的想法，而被大學裡的行政人員強行解職。之後的終身職位（tenure）制度的發展雖然促進了學術自由，但是仍然無法完全保障學術自由。所謂的終身職位制度，就是當教職人員在經過了一段長時間的試用期（通常是六年）之後，除非有非常充分的理由，否則不能被任意地解雇。終身職位的制度大幅度地減少了教授以及研究者因爲發表不受歡迎的想法而被大學行政人員強迫解雇的情況。[45]

對於社會科學的政治性攻擊，早已經不是什麼新鮮事了。此種攻擊正好可以顯現出下列二者之間的衝突：其中之一是對於知識的獨立追求、而另外一方則是想要強加其信念於他人身上的政治團體所抱持的觀點。這些對於社會科學的攻擊，引發了下面這個問題：在面對較大文化的價值觀時，社會科學應該要擁有多大的自主性？社會研究所得到的發現，經常會與那些來自於非科學的知識系統（像是宗教或者是政治意識型態）的社會信念相互抵觸。在大約 400 年前，當自然科學尚未被大眾所接受時，伽利略就面對過同樣的問題。他依據自由思想的科學所得到的天文學方面的發現，與當時教會的官方教條互相抵觸。在遭受到折磨的威脅之下，伽利略被迫公開地撤回自己的發現。這件事使得科學的進展被拖慢了大約一個世代之久。而對於演化論的質疑與挑戰，則再一次證明了存在於科學知識與普遍信念之間的衝突與矛盾。

對於良好的研究而言，學術自由是不可或缺的。科學研究所牽涉到的不只是瞭解技術層面的資訊而已（例如，如何得到隨機樣本），還必須要具備自由以及公開討論的精神、依據科學價值而不是個人價值觀所提出的批評、以及對於社會生活所有領域的探究過程。當學術自由受到限制時，上述這些價值也都會受到威脅。

客觀性以及價值中立

有些人主張社會科學必須和自然科學一樣既客觀又公正，而其他人則認爲，社會科學不可能達到價值中立以及客觀的標準。造成此種混淆狀況的部分原因，是因

爲上述的兩個詞彙都有至少兩種不同的定義，而且有些時候兩個不同的詞彙還會擁有相同的定義（請參閱圖表 16.2）。

實證主義的取向認爲科學是價值中立的、公正的、以及客觀的，也就是說，科學就等於是同時包含上述這三種定義。邏輯演繹、正式理論、以及把事實和以價值爲基礎的概念區隔開來的作法，確保了科學的價值中立性。科學社群不應該受到任何偏見所影響，並且是由自由而公開的討論所主導的。在完全價值中立以及客觀的情況下，科學就能夠揭露出單一且唯一的、一致的、清楚明白的真理。

在討論社會科學在社會中所扮演的角色時，韋伯（Max Weber）、高德納（Alvin Gouldner）、以及曼海姆（Karl Mannheim）是三種重要的非實證主義社會思想家。韋伯（Weber, 1949）主張，在社會科學當中，事實與價值之間的區隔並非是清楚明確的。他認爲所謂的社會事實或者是具有社會意義的行動，正是由充滿價值的理論所定義的。因此，社會理論必然會包含一些以價值爲基礎的概念，因爲特定文化當中的成員會創造出關於社會世界的所有概念。我們無法把社會概念從文化內容當中清除掉，而且具有社會意義的行動只有在文化脈絡之中才能產生意義。例如，當我們在研究種族團體時，這些團體對於種族之間的生物性差異其實並不感興趣。種族只是一種社會概念，而我們之所以會去研究這個概念，是因爲某個文化裡的成員爲不同種族的外表附加了特定的社會意義。如果人們沒有爲可以觀察到的種族差異附加任何意義的話，種族這個概念就會變得毫無意義可言。

其他的社會研究者則以韋伯的想法爲基礎，並加以擴展。例如，摩爾（Moore, 1973）想要知道多數族群（例如：英語系美國人、白人）的研究者是否能夠以「外來者」的身分來正確地研究少數民族的族群，因爲這些研究者的研究問題、假定、以及興趣，都是源自於一種優勢的、非少數族群的觀點。佔優勢的白人文化中的文化、價值觀、以及信念體系，是否適合於提出關於少數民族次文化的重要問題？是否能夠真正地瞭解這些少數民族的次文化？類似的狀況也出現在與性別有關的研究當中。[46] 雖然來自於一個不同的文化這件事，或許並不會對於研究另外一個族群造成妨礙，但是研究者本身卻必須要更加謹慎小心、而且也要更加敏感。

韋伯（Weber, 1949）也曾經主張過，社會科學家無法避免在他們所研究的社會議題上面去選擇一個特定的立場。當在運用一些已經被接受的研究技術時，研究者*必須*要保持公正（也就是中立的，而且不能有個人的意見以及沒有證據支持的觀點），並且把焦點集中在社會世界運作的方式或者是機制上面，而不是把焦點放在目的、價值觀、或者是規範性的目標上面。研究者本身的價值觀必須與研究發現區隔開來，而且研究者應該只有在以個別公民的身分發言時，才能夠去宣揚自己在特定議題上面所抱持的立場。

高德納（Gouldner, 1976）則是攻擊了價值中立的、客觀的社會科學這個概念。他主張，價值中立性在過去只是被當成掩飾價值立場的工具而已。事實上，價值中立性本身就是一種價值，也就是一種偏好「不帶任何價值」的立場。高德納曾經提到過，完全的價值中立是不可能的，而且科學家以及其他的專業人士只是利用這個詞彙來隱瞞他們本身的價值觀而已。他建議研究者應該要公開表達自己的價值觀。

圖表 16.2 客觀的、價值中立的、以及公正的

1. *客觀的：*
   a. 「主觀的」相反詞；外部的、可觀察的、事實的、精確的、量化的
   b. 合乎邏輯的；由明確的理性程序所創造出來的；沒有任何個人的或是任意的決定；依循著明確的預定規則
2. *價值中立的：*
   a. 沒有任何後設的價值觀或者是假定；沒有任何先驗的哲學成分；與道德無關的
   b. 沒有受到個人偏見或者是文化價值觀的影響；沒有個人的意見；不能接受沒有根據的看法；中立的
3. *公正的：*
   a. 排除了非隨機的錯誤；不存在任何系統性的誤差；在技術方面是正確無誤的
   b. 不受到任何個人偏見或者是文化價值觀的影響；沒有個人的意見；不能接受沒有根據的看法；中立的

關係位置 一種由曼海姆所提出的概念，也就是專業的學術研究者以及知識份子佔據了一種獨特的社會位置，而且也與社會中的主流團體保持一定的距離。此種情況使得這些研究者或是知識份子處於一種發展出公正知識的最佳位置上。

研究者去進行研究的動機，不只是以一種不帶情感的方式來研究這個世界而已。當一位研究者擁有想要造成某些改變的強烈道德動機時，此種情況並不必然會使得該研究者所進行的研究的品質變差。

曼海姆（Mannheim, 1936）也質疑了*價值中立性*以及*客觀性*這兩個觀念。他認爲社會中的知識份子，尤其是那些進行社會研究的知識分子，扮演了一個獨特的社會角色。一個人在社會中的位置，會形塑了他或是她的想法以及觀點，然而，社會研究者與大多數其他的人是有所不同的。社會研究者所處的社會位置，對他們所造成的影響較小，因爲這些研究者可以利用進行研究的方式來瞭解其他人的觀點，並且與社會中所有領域內的人感同身受。與大多數的其他人相較之下，社會研究者比較不需要去迎合握有權力的菁英階級，而且也比較不會受到民眾的意見、短暫的風尚、以及一時的狂熱所影響。因此，社會研究者能夠、而且也應該採取一種**關係位置**（relational position），也就是一種與任何其他特定的社會團體保持距離、但是卻又與所有的團體都有接觸的位置。換句話說，社會研究者應該在社會中處於一種疏離或是邊緣的位置，但是仍然要與社會中的所有部分都保持聯繫，即使是那些經常被忽視或是被隱藏起來的部分亦然。

結論

在科學事業當中，傳達研究結果和研究的倫理與政治具有同等的重要地位。關於社會研究的政治議題，你可能面臨的「解決之道」總共有三個層次。首先，你必須注意到此種議題的存在、察覺到潛在的危險性、並且對社會政治環境採取切合實際的看法，而不是一廂情願地以天真的看法來看待社會研究。其次，你應該與其他人一起合作，以提倡社會研究應該擁有一種免於受到外在壓力脅迫的獨立性。第三，

你應該要教育社會大眾以及大型機構的領導者關於獨立的社會研究所具有的價值以及重要性。

最後，在本章結束之前，對於你這位社會研究的消費者或者是新手社會研究者，我強烈鼓勵你去發展出一種自我察覺的能力。你應該要去察覺到社會研究在社會中所處的地位，以及社會研究本身所處的社會脈絡。社會研究者能夠爲較大的社會帶來一種獨一無二的觀點。我們對於我們自己、科學社群、以及整個社會都負有責任，而且我們也必須去瞭解社會科學達到今日在社會中的地位的演變過程。

問題回顧

1. 請討論在撰寫研究報告過程中的事前準備、自由寫作、重寫、編輯、以及撰寫等等各個階段之間的關係。
2. 在量化以及質化研究報告的組織方式之間的主要差異是什麼？
3. 研究計畫書與最後的研究報告之間有何相似以及相異之處？
4. 政客的行動會對社會研究造成什麼種類的限制？
5. 對於研究經費的控制會以何種方式而影響了研究所探討的議題類型？
6. 提供研究經費的政府機構或是私人捐贈者所使用的標準、與科學社群的同儕所使用的標準之間可能會有什麼樣的差異性？
7. 在過去的 20 年之間，美國政府對於研究經費補助的趨勢爲何？而這些趨勢可能會對現在所進行的研究造成什麼樣的影響？
8. 盧爾所提出的關聯模式的來源爲何？而關聯模式的用途又是什麼？
9. 學術自由會以何種方式來支持或者是抵觸關係位置？
10. 進行客觀而且價值中立的研究的意義是什麼？

注解

1. See "Plagiarism Case Documented," American Sociological Association *Footnotes*, 17(2), p. 2 or "Noted Harvard Psychiatrist Resigns Post after Faculty Group Finds He Plagiarized," *Chronicle of Higher Education*, 35(15), p. 1.
2. From Sociology Writing Group (1991).
3. For suggestions on writing, see Donald et al. (1983) and Leggett et al. (1965).
4. From Sociology Writing Group (1991:40).
5. See Fine (1988) for this and other suggestions on writing.
6. See Mullins (1977:11–30) for a discussion of outlines and the organization of quantitative research reports. Also see Williams and Wolfe (1979:85–116) for good hints on how to organize ideas in a paper.
7. Grosof and Sardy (1985:386–389) have provided suggestions on how to explain quantitative findings.
8. Lofland (1974) inductively discovered what he identifies as five major writing styles for reporting field research (generic, novel, elaborated, eventful, and interpenetrated) and discusses how they are evaluated.
9. The error of segregation is discussed in Lofland and Lofland (1984:146).
10. See Becker and Geer (1982:244) and Schatzman and Strauss (1973:130) for a discussion of this and

related issues.

11.See Hammersley and Atkinson (1983) and Van Maanen (1988).

12.Discussed in Spradley (1970:162–167).

13.See Van Maanen (1988:13).

14.See Dabbs (1982) and Jackson (1978).

15.For a discussion of ethical concerns in writing field research reports, see Becker (1969), Punch (1986), and Wax (1971).

16.See Barzun and Graff (1970) and Shafer (1980) for excellent suggestions on writing about historical research.

17.For more on writing proposals to fund research projects, see Bauer (1988), Locke et al. (1987), and Quarles (1986). A somewhat dated but useful short introduction to proposal writing is found in Krathwohl (1965).

18.For Russian social science research, see Keller (1988, 1989) and Swafford (1987). Also see "Soviet Sociologist Calls Attention for Her Science," American Sociological Association *Footnotes* (April 1987), p. 2.

19.See Greenberg (1967:71).

20.For more on the decade of the 1950s and its effect on social reseachers, see Caute (1978:403–430), Goldstein (1978:360–369), and Schrecker (1986).

21.See Block and Burns (1986) and Starr (1987).

22.See Bermant (1982:138). Nelkin (1982a) provided a general discussion of "forbidden" topics in social science research.

23."Sex Survey Is Dealt a Setback," *New York Times* (July 26, 1989), p. 7.

24.See Stephen Burd, "Scientists Fear Rise of Intrusion in Work Supported by NIH," *Chronicle of Higher Education* (October 2, 1991), p. Alff.

25.See Cumings (1997), Sanders (1979), and Simpson (1993) on U.S. government influence on area studies and internationally related academic research during the Cold War era.

26.Project Camelot is described in Horowitz (1965).

27.See Dickson (1984), Nelkin (1982b), and Shattuck and Spence (1988:2).

28.See Shattuck and Spence (1988) and Josephson (1988). Also see "Librarians Charge Plan Would Cut Flow of Data," *New York Times* (February 21, 1989).

29.For more on the CIA and social researchers, see Shattuck and Spence (1988:39–40) and Stephenson (1978).

30.For sensitive situations involving cross-national research, see Fuller (1988) and Van den Berge (1967).

31.For discussion, see Bannister (1987), Blumer (1991b), D'Antonio (1992), Hyman (1991), Ross (1991), and Seybold (1987).

32.See Dynes (1984) on COSSA.

33.The SSRC spent $20 million for the social sciences from 1924 to 1928 (Gieger, 1986:152) compared to $136 million allocated in 1989 by the NSF for the social sciences (D'Antonio, 1992). In the late 1920s, the number of academic social scientists was about one-tenth and a dollar purchased over six times more than it does today. The number of social science doctorates—including psychology, teaching, or conducting basic research—in 1986 was about 129,000 (Science and Engineering Personnel: A National Overview, Document NSF 90-310). The number of higher education faculty in all academic fields in 1930 was less than 83,000 (Historical Statistics of the United States, 1970, Table H696). The $20 million over 4 years in the 1920s, or $5 million per year, would be equivalent to roughly $300 million in 1990. The median family income before taxes in 1929 was $2,335 (Historical Statistics, Table G308).

34.For more on the effects of politics and funding cuts on social research in the 1980s, see Cummings (1984), Himmelstein and Zald (1984), and Zuiches (1984). For more general discussion of the effect of funding on research, see Galliher and McCartney (1973) and Dickson (1984).

35.See "NIH FY 1991 Budget Rescinded by $3.1 Million, Congress Objects to 31 Research Projects Funded

by NSF," *The Blue Sheet* (F-D-C Reports, Inc.) (May 27, 1992), p. 3.

36.Nature Neuroscience Editorial (2005).

37.See Brainard and Borrego (2003), Brainard and Southwick (2001), Cordes (1998), Payne (2003a, 2003b), and Savage (2001) on rapidly increasing pork barrel academic spending.

38.See Golder (1996).

39.See Punch (1986:18–19; 49–69).

40.Lawrence Altman, "Experts See Bias in Drug Data," *New York Times* (April 29, 1997), and Sheryl Gay Stolberg, "Gifts to Science Researchers Have Strings, Study Finds," *New York Times* (April 1, 1998). On the nursing home "slap suit," see Greenhouse, "Cornell Professor Fights a Slander Suit," *New York Times* (April 1, 1998), and news report of Morning Edition, National Public Radio (April 27, 1998).

41.See Block (2003), Clymer (2002), Krider (2004), Lee (2003), and Union of Concerned Scientists (2004).

42.For more discussion on how researchers select research questions or problems, see Gieryn (1978) and Zuckerman (1978).

43.See Brym (1980) on role of intellectuals in society.

44.Marsh (1984), Noelle-Neumann (1974, 1984), and Price (1989) discussed the effects of research results on subsequent public behavior and opinion.

45.Bartiz (1960), Schrecker (1986), Schwendinger and Schwendinger (1974), and Silva and Slaughter (1980) discuss the history of social researchers in society.

46.Committees on the Status of Women in Sociology (1986).

五位數字亂數表

10819	85717	64540	95692	44985	88504	50298	20830	67124	20557
28459	13687	50699	62110	49307	84465	66518	08290	96957	45050
19105	52686	51336	53101	81842	20323	71091	78598	60969	74898
35376	72734	13951	27528	36140	42195	25942	70835	45825	49277
93818	84972	66048	83361	56465	65449	87748	95405	98712	97183
35859	82675	87301	71211	78007	99316	25591	63995	40577	78894
66241	89679	04843	96407	01970	06913	19259	72929	82868	50457
44222	37633	85262	65308	03252	36770	51640	18333	33971	49352
54966	75662	80544	48943	87983	62759	55698	41068	35558	60870
43351	15285	38157	45261	50114	35934	05950	11735	51769	07389
11208	80818	78325	14807	19325	41500	01263	09211	56005	44250
71379	53517	15553	04774	63452	50294	06332	69926	20592	06305
63162	41154	78345	23645	74235	72054	84152	27889	76881	58652
17457	68490	19878	04981	83667	00053	12003	84614	14842	29462
28042	42748	55801	94527	21926	07901	89855	21070	80320	91153
32240	24201	24202	45025	07664	11503	97375	83178	26731	45568
87288	22996	67529	38344	29757	74161	16834	40238	48789	99995
39052	23696	42858	85695	50783	51790	80882	97015	81331	76819
71528	74553	32294	86652	15224	07119	45327	69072	64572	07658
76921	04502	78240	89519	02621	40829	88841	66178	01266	10906
45889	22839	77794	94068	85709	96902	19646	40614	03169	45434
10486	79308	75231	33615	42194	49397	91324	79553	66976	83861
42051	14719	80056	74811	58453	04526	90724	36151	09168	04291
47919	11314	80282	09297	02824	59530	31237	26311	62168	46591
19634	40589	28985	40577	33213	52852	17556	85342	66881	18944
10265	45549	38771	38740	48104	63990	73234	19398	33740	97345
74975	33526	36190	25201	19239	06254	02198	99109	01005	20983
37677	76778	15736	57675	81153	59651	69262	89250	75156	59164
18774	15979	26466	80236	65400	24272	02088	09307	33426	11230
93728	14965	85141	27821	53791	38728	66369	29415	55330	99228
34212	15590	41336	23614	26153	19466	44176	80885	00015	40077
81984	54478	45226	97338	14064	45768	13538	49093	05691	69720
72755	15743	00552	89374	85400	37392	26598	71917	64275	16125
13162	57044	75982	15819	23385	40860	51585	44542	39656	91139
64686	62224	34124	79171	73909	26196	54057	63264	72089	06658
00157	64594	03178	75774	32315	34443	37224	85593	55251	42666
84194	83591	82152	24311	22414	43244	81542	31491	42075	17275
05776	60399	65218	89299	20273	30071	53077	18853	56652	63896
33365	18314	81074	49433	10884	75467	56085	14731	98085	60895

67928	38976	38480	59980	23156	72335	33489	59420	67819	51874
64394	45154	81851	54228	73095	97217	16908	90242	92869	17311
73000	20948	57065	70195	87563	41590	85047	71743	94916	50534
63555	03388	96638	16591	13641	73342	59131	63144	63587	62084
84005	02035	08182	16395	44928	08897	44750	71378	67522	20180
42593	35102	14577	38102	60403	04540	53992	27069	69574	76682
49519	49517	88147	83375	87045	57466	91259	06680	45586	36257
42149	01579	83056	19423	28165	25620	68035	17919	09120	59078
66192	98427	10152	96970	89990	34604	49632	46533	63362	43151
16124	88620	87074	37851	77131	73855	03740	10306	63858	04349
35492	47334	57189	26465	70078	14477	00881	00929	86907	73764
54503	40155	94734	20689	32475	62851	13216	21419	95502	36783
88063	53451	15642	67345	06935	70644	68570	79176	31975	83082
83689	14426	40357	34906	56282	96104	83796	57663	88627	17521
40393	72810	00681	15351	28858	72086	99090	39741	17914	27385
76648	61322	06817	64674	50317	52373	78223	84222	14021	43432
42091	27088	37686	88033	68007	71009	24018	49568	64351	94130
78925	41509	14319	92389	85492	40880	01487	85509	48316	62618
61915	98081	87996	53798	51485	38912	85858	43392	64678	44458
29504	66960	42645	54547	20615	77035	79942	33972	46112	78290
90170	97643	46284	34591	42692	72933	66166	98389	37460	14545
96439	06806	76714	80084	57685	37447	44901	64699	89142	64657
98365	28725	84376	50634	79289	31106	71351	10533	57545	27399
74794	91013	89791	54236	02369	35317	31103	82481	52256	94510
37499	85907	16293	17673	13373	06599	50138	19860	46716	36928
77530	25960	33671	54383	25144	82627	99266	75134	96539	47242
67990	35106	05214	82928	39824	11128	31390	76293	52809	54881
07355	29187	09357	94498	69697	92515	89812	90794	44738	46806
40716	05787	68975	38937	44033	50064	25582	09428	10220	42455
97748	64395	13937	60406	99182	92720	80805	26242	81943	40341
83682	18775	60095	78600	03994	30313	21418	58563	47258	75582
73506	30672	18213	37887	26698	87700	75784	86878	74004	88636
36274	02333	43132	93725	87912	90341	74601	77001	30717	60002
73508	00852	94044	98474	12621	91655	55258	85551	76122	68052
06488	12362	60020	66902	90734	73689	22382	40896	09028	72925
20201	31560	98885	32275	46818	76114	07959	65639	33267	98595
49947	13114	06773	06454	95070	26564	08974	11640	76202	86105
79928	50600	06586	72129	37233	02564	83265	32579	21234	83535
76360	86412	36240	20210	17692	80482	67007	15474	23198	74250
54601	84643	66759	57661	16434	61708	93185	75957	61056	90678
23441	63863	95238	59665	55789	26180	12566	58645	15125	76707
47093	90509	48767	09874	23363	84954	09789	30178	28804	93294
93603	11580	94163	85561	71328	88735	69859	84563	25579	52858
68812	15299	99296	45906	37303	49507	70680	74412	96425	38134
69023	84343	36736	52659	90751	20115	89920	44995	17109	96613
76913	03158	83461	27842	03903	34683	89761	80564	45806	88009
99426	99643	00749	79376	44910	27490	59668	93907	73112	46365
59429	08121	06954	28120	17606	22482	91924	00401	16459	15570
38121	05358	01205	00662	73934	97834	56917	64058	05148	87599
97781	32170	99914	75565	79802	38905	17167	08196	46043	72094
79068	21760	78832	93795	67798	54968	87328	46494	74338	89805
46601	04015	00484	39366	56233	22622	90706	02327	60807	39009

漢英索引

Z

Z 分數（z-scores），515

一劃

一致性（consistency），279
一致性檢查（consistency checking），507
一致法（method of agreement），684, 691
一般性解釋（ordinary explanation），100
一般的實務工作者(generalist practitioner), 51

二劃

人工製品（artifacts），386
人文科學（Geisteswissenschaft），138
人性能動作用（human agency），143
人類的機械模式（mechanical model of man），132
人類關係區域檔案(Human Relations Area Files, HRAF），647

三劃

三角檢證法（triangulation），224
口述歷史（oral history），633
女性主義研究（feminist research），160
小資產階級（petite bourgeoisie），92
工具（instrument），389
工具性知識（instrumental knowledge），51
工具性導向（instrumental orientation），136

四劃

中介變項（intervening variable），244
中央極限定理（central limit theorem），332
中位數（median），511
中距理論（middle-range theory），115
中觀層次理論（meso-level theory），98
互爲主體性（intersubjectivity），135
互動語音回覆(interactive voice response, IVR), 446
內在效度（internal validity），292, 386
內含詞彙（included terms），688
內容分析（content analysis），69
內容效度（content validity），286
內部一致性（internal consistency），598
內部評論（internal criticism），640
內團體（in-groups），368
公平偏好（equity preference），394
分析比較（analytic comparison），690
分析敘事（analytic narrative），695
分析備忘錄（analytic memos），587
分析單位（units of analysis），95
分析領域（analytic domain），689
分散情況（dispersion），512
分層抽樣（stratified sampling），337
分類學（taxonomy），90
化約論（reductionism），254
反公共參與的策略性訴訟（Strategic Lawsuits against Public Participation, SLAPP），750
反身—辯證導向（reflexive-dialectic orientation），157
反身性知識（reflexive knowledge），51
反面個案方法（negative case method），701
反應心向（response set），307, 432
反應式的（reactive），69, 470
反應風格（response style），307
反應偏頗（response bias），307
反應順序（reactive sequence），697
天真的假定模型（naïve assumption model），450
引用（citation），178
引導式問題（leading question），415
心理狀態（mentality），625
支持性證據（supporting evidence），629

文化知識（cultural knowledge），554
文化脈絡研究（cultural-context research），643
文化領域（cultural domain），688
文本（text），139, 473
文本檢索（text retrieval），707
文獻（literature），170
方法論（methodology），2
方法論的實證主義（methodological positivism），130
月暈效應（halo effect, 6
比例機率（probability proportionate to size, PPS），344
比率（rate），302
比率尺度測量（ratio-level measurement），294
比較判斷法則（law of comparative judgment），309

五劃

世代研究（cohort study），65
世界體系理論（world system theory），644
主要研究者（principal investigator, PI），737
主要效果（main effects），383
主軸式編碼（axial coding），677
以文本爲基礎的管理程式（text-based managers），708
以編碼爲基礎的理論建構程式（code-based theory builders），709
他觀的（etic），658
他觀概念（etic concepts），625
代表性信度（representative reliability），281
出生世代（birth cohorts），65
出局者（freeze-outs），577
加法量表（additive scales），305
加權（weighting），300
加權的指數（weighted index），300
功能理論（functional theory），113
包氏社會距離量表（Bogardus social distance scale），310
半過濾式問題（quasi-filter question），430
卡麥洛計畫（Project Camelot），743
去除焦點（defocusing），559
可比較性誤差（comparability errors），412
可信度的階層（hierarchy of credibility），603
可能性檢查（wild code checking），507
可能編碼式清理（possible code cleaning），507
可得性抽樣（availability sampling），322
可被接受的無能者（acceptable incompetent），578
可靠性（dependability），279
可證僞性（falsifiability），636
古典實驗設計（classical experimental design），378
古特曼量表指數（Guttman scaling index），313
史學（historiography），630
外在效度（external validity），292, 393
外部一致性（external consistency），598
外部評論（external criticism），640
外團體（out-groups），368
平行示範（parallel demonstration），687
平均數（mean），511
平均數值（averages），509
本質主義（essentialist），132
本體論（ontology），125
未加權的指數（unweighted index），300
未被注意到的觀察（unnoticed observation），470
正向關係（positive relationship），107
母群（universe），236, 325
母數（parameter），329
母體（population），320
民俗領域（folk domain），688
民族性（nationhood），88
民族誌（ethnography），554
生命史訪談（life history interview），594
生活真實性（mundane realism），395
生態效度（ecological validity），599
生態謬誤（ecological fallacy），253
田野研究（field research），70, 550
田野場域（field site），560
田野實驗（field experiment），397
申請補助費技巧（grantsmanship），736
目的論（teleology），252
目標母體（target population），326
立意抽樣（purposive sampling），353

六劃

交叉列表（cross-tabulation），521
交叉設計（crossover designs），208
交互作用的因果關係（reciprocal-effect causal relations），103

交互作用效果（interaction effects），250, 383
任意修改（bowdlerization），640
光學掃瞄（optical scan），505
全容（exhaustive），417
全容的特性（exhaustive attribute），296
全距（range），332, 512
全過濾式問題（full-filter question），430
共同發生的情況（conjuncture），637
共同變異（concomitant variation），103
共享主義規範（norm of communalism），717
共變（covariation），517
列聯式清理（contingency cleaning），507
列聯表（contingency table），521
印象管理（impression management），394
同時效度（concurrent validity），286
同等性時間序列（equivalent time-series design），380
同義（equivalence），642, 659
同樣本多次研究（panel study），65
同謀（confederate），377
名目尺度測量（nominal-level measurement），294
名目比較（nominal comparison），691
合作的邂逅模型（collaborative encounter model），459
因果法則（causal laws），131
因果假設（causal hypothesis），245
因果解釋（causal explanation），101
因果機制（causal mechanism），105
因素（factors），383
回憶（recollections），633
回譯法（back translation），652
多因子設計（factorial design），382
多重指標（multiple indicators），281
多國研究（transnational research），644
多國單位（multinational units），644
多變項（multivariate），509
字彙的同義（lexicon equivalence），659
守門員（gatekeepers），562
安慰劑效應（placebo effect），392
年鑑學派（Annales School），625
成本效益分析（cost-benefit analysis），48
成員（members），551
成員確認（member validation），600
成熟效應（maturation effect），388
曲線關係（curvilinear relationship），520
有限的自主（bounded autonomy），153
有意義的社會行動（meaningful social action），140
次母群分析（subpopulation analysis），281
次序分析（sequence analysis），695
次序理論（sequential theory），110
次級分析（secondary analysis），470
次級資料（secondary sources），632
次級資料分析（secondary data analysis），69
次數分配（frequency distribution），509
次數分配多邊圖（frequency polygon），509
百分位數（percentiles），513
自由寫作（freewriting），723
自我評估（self-assessment），29
自我歸類理論（self-categorization），368
自然主義（naturalism），130, 556
自然科學（Naturwissenschaft），138
自然推論性（naturalistic generalization），395
自然態度（natural attitude），144
自然歷史（natural history），600
自願同意原則（principle of voluntary consent），203
自變項（independent variable），105, 244
自觀的（emic），658
行爲主義（behaviorism），130, 369
行動主義者（activists），51
行動研究（Action research），42
行話（argot），581
西方文化的偏誤（Western cultural bias），654

七劃

伴隨發生的事件（contingency），637
克里本多夫的阿爾法係數（Krippendorff's α），480
判斷抽樣（judgmental sampling），353
否定假設的邏輯（logic of disconfirming hypotheses），248
告密者（whistle-blower），214
均衡（balanced），417
完形（Gestalt），597
局部表格（partials），534
形式理論（formal theory），99
形成性評估研究（formative evaluation），42
抄襲（plagiarism），197

技術官僚觀點（technocratic perspective），136
批判取向社會科學（critical social science, CSS），129, 148
批判結構主義（critical structuralism），148
批判實在論者（critical realist），126
折半法（split-half method），281
抑制變項型態（suppressor variable pattern），538
改寫（paraphrasing），722
李克特量表（Likert scales），305
決定性實驗（crucial experiments），247
決定論（determinism），132
沒有偏誤的（unbiased），373
系統抽樣（systematic sampling），334

八劃

事件結構分析（Event-structure analysis, ESA），710
事前準備（prewriting），723
事後（post facto），636
依變項（dependent variable），105, 244
具有統計顯著性的結果（statistically significant results），541
具體性錯置的謬誤（fallacy of misplaced concreteness），493
典型（typifications），144
典範（paradigm），29, 115, 129
受訪者（respondents），406
受訪者驅動抽樣（respondent-driven sampling），355
受試者（subjects），375
受試者內實驗（within-subject experiment），371
命題（proposition），94
垃圾科學（junk science），16
委託研究（commissioned research），52
彼此互斥（mutually exclusive），417
所羅門四組設計（Solomon four-group design），381
承載體（carrier），321
拉丁方塊設計（Latin square design），381
抽象解釋（Erklärung 或 abstract explanation），138
抽象層次（level of abstraction），87
抽樣元素（sampling element），324
抽樣分配（sampling distribution），331
抽樣比（sampling ratio），327
抽樣架構（sampling frame），326
抽樣間距（sampling interval），334
抽樣誤差（sampling error），329, 412
明白的偽裝（explicit cover），205
物化（reification），152
直方圖（histogram），509
直接輸入法（direct-entry method），505
直接觀察的筆記（direct observation notes），585
知情的同意（informed consent），204
社會交換理論（social exchange theory），442
社會地圖（social map），588
社會安全（social security），649
社會指標（social indicator），486
社會崩潰（social breakdown），574
社會理論（social theory），13
社會期望（social desirability），420
社會期望偏誤（social desirability bias），423
社會認同理論（social identity theory），368
社會影響評估（social impact assessment），44
社會整合（social integration），244
社會關係圖（sociogram），355
空盒子（empty boxes），686
初級資料（primary sources），630
表面上的興趣（appearance of interest），578
表面事實（outcroppings），682
表面效度（face validity），285
表面詞彙（cover term），688
表現型（phenotype），88
長期（longue durée），625
長期影響（long-term effects），62
附加測量（accretion measure），471
非介入性研究（unobtrusive research），69
非反應式研究（nonreactive research），69, 470
非來自於資料的知識（nonsource-based knowledge），640
非侵擾性測量（unobtrusive measures），470
非循環（nonrecursive），103
非虛假關係（no spuriousness），104
非線性研究路徑（nonlinear research path），229

九劃

信賴區間（confidence intervals），332

便利抽樣（convenience sampling），322
保密性（confidentiality），209, 601
促發（priming），26, 368
俗民方法論（ethnomethodology），555
前測（pretest），376
前實驗設計（preexperimental design），378
前導測試（pilot test），282, 376
客觀主義（objectivist），132
封面故事（cover story），377
封面頁（face sheet），437, 587
封閉式問題（closed-ended question），426
建構主義導向（constructionist orientation），141
律則的（nomothetic），134
後現代研究（postmodern research），162
後設分析（meta-analysis），172, 442
後設敘事（metanarratives），693
後測（posttest），376
後實證主義（postpositivism），130
拼湊（bricolage），240
指數（index），298, 299
政治機會結構（political opportunity structure, POS），61
流失（mortality），389
相互排除的特性（mutually exclusive attribute），296
相對主義（relativism），147
相關係數（correlation coefficient），103, 528
研究法（methods），2
研究計畫主持人（project director），737
研究計畫書徵求訊息（requests for proposals, RFPs），737
研究造假（research fraud），197
科學（science），11
科學方法（scientific method），128
科學的不當行爲（scientific misconduct），197
科學社群（scientific community），17
科學社群規範（norms of the scientific community），19
科學素養（scientific literacy），15
背景假定（background assumption），84
要求特徵（demand characteristic），391
負向關係（negative relationship），107
重建邏輯（reconstructed logic），228
限額抽樣（quota sampling），323
陌生的態度（attitude of strangeness），573

十劃

個案化（casing），279, 674
個案研究（case-study research），58, 241
個案研究的比較研究（case-study comparative research），643
個案澄清（case clarification），687
個殊的（idiographic），144
修改（revising），725
倫理規範（code of ethics），212
套套邏輯（tautology），250
差異法（method of difference），692
捕捉複雜性與追蹤過程的能力（Ability to capture complexity and trace processes），60
效標效度（criterion validity），286
時代錯誤（anachronism），628
時近效應（recency effect），431
時期劃分（periodization），697
時間地圖（temporal map），588
時間序列研究（time-series research），63
時間預算調查（time budget survey），448
校準（Calibration），60
格狀問題（grid question），438
消減誤差比例（proportionate reduction in error），528
特殊族群（special populations），207
狹義的實務工作者（narrow practitioner），51
真實性（truthfulness），279
矩陣式問題（matrix question），438
破壞性實驗（breaching experiment），555
純研究（pure research），37
紐倫堡規範（Nuremberg code），212
耗損測量（erosion measure），471
脈絡的同義（contextual equivalence），660
脈絡效應（context effects），435
能夠勝任的圈內人表現（competent insider performance），601
逆推法（abduction），154
高爾頓問題（Galton's problem），645

十一劃

僞科學（pseudoscience），14
僞調查（pseudosurvey），462
僞證據（pseudoproof），628

假定（assumptions），84
假設（hypotheses），23
假設（hypothesis），94
偶然性抽樣（accidental sampling），322
偏差行爲脫序理論（anomie theory of deviance），90
偏態分配（skewed distribution），512
動機框架（motivational frame），61
匿名性（anonymity），209
匿名審查（blind review），21
區別效度（discriminant validity），287
區間尺度測量（interval-level measurement），294
參與式行動研究（participatory action research），43
參與研究者（participatory researcher），53
唯名論者（nominalist），125
唯後測非同等群組設計（posttest-only nonequivalent group design），379
唯意志論（voluntarism），143
基模（schemes），687
基礎研究（basic research），37
執行摘要（executive summary），727
培根式謬誤（Baconian fallacy），626
專業術語（jargon），87
專題研究報告（research monographs），21
專題著作（monographs），173
將社會研究正常化（normalize social research），569
常態分配（normal distribution），309, 511
控制組（control group），376
控制變項（control variable），532
探索性研究（exploratory research），53
探問（probe），428, 454
接觸假說（contact hypothesis），57, 96, 399
措辭效應（wording effects），434
掩護（fronts），600
推論統計（inferential statistics），351, 540
推論獨立（separation of inference），587
啓發式影響（Heuristic impact），59
敘事（narrative），693
敘事分析（narrative analysis），695
敘事文本（narrative text），694
敘事呈現（narrative presentation），695
敘事探究（narrative inquiry），694
敘事結構分析（narrative structural analysis），695
敘事解釋（narrative explanation），695
敘事實踐（narrative practice），694
敘事歷史（narrative history），637
敘述統計（descriptive statistics），509
族群性（ethnicity），88
族群界線（ethnic boundary），88
條件式問題（contingency question），425
條碼（bar code），506
混合領域（mixed domain），689
混淆變項（confounding variables），365
涵蓋律則模式（covering law model），130, 134
涵蓋誤差（coverage errors），412
深厚描述（thick description），554
深度僞裝（deep cover），205
淨效果（net effect），534
理想型（ideal type），90, 684
理論（theory），76
理論系統（theoretical systems），115
理論性解釋（theoretical explanation），100
理論抽樣（theoretical sampling），357
理論架構（theoretical frameworks），115
理論推論性（theoretical generalization），397
理論概念（theoretical concept），85
現代主義（modernism），162
現存的統計數據（existing statistics），470
現存統計資料研究（existing statistics research），69
異例抽樣（deviant case sampling），356
異常個案分析（deviant case analysis），701
眾數（mode），509
第一級詮釋（first-order interpretation），242
第一類型錯誤（Type I error），544
第二級詮釋（second-order interpretation），242
第二類型錯誤（Type II error），544
第三級詮釋（third-order interpretation），242
統計（statistics），507
統計回歸效應（statistical regression effect），389
統計值（statistic），329
統計效度（statistical validity），292
統計獨立（statistical independence），517
統計關係（statistical relationship），517
統計顯著水準（level of statistical significance），

542
統計顯著性（statistical significance），541
紮根理論（grounded theory），97, 238, 621
組內實驗（within-group experiment），372
組合解釋（combinational explanation），623
組間實驗（between-group experiment），372
累積量表（cumulative scaling），313
處遇（treatment），68, 375
處遇擴散（diffusion of treatment），390
規範性陳述（normative statement），81
訪談大綱（interview schedule），410
設計、計畫、預算制度（Planning, Programming, and Budgeting System, PPBS），42
設計記號（design notation），384
軟性資料（soft data），226
連鎖轉介抽樣（chain referral sampling），355
連續記錄（running records），633
連續變項（continuous variables），293
逐次抽樣（sequential sampling），356
部分開放式問題（partially open questions），428
都市衰敗（urban decay），89

十二劃

博士論文（dissertation thesis），181
單向度性（unidimensionality），297
單次個案研究設計（one-shot case-study design），378
單組前測後測設計（one-group pretest-posttest design），379
單組唯後測設計（one-group posttest-only design），378
單變項（univariate），509
單變項統計（univariate statistics），517
循環（recursive），103
描述性研究（descriptive research），54
散佈圖（scattergram，或者也被稱爲 scatterplot），518
智庫（think tank），752
欺瞞（deception），377
游移者（floaters），431
減損效應（attrition），389
測量尺度（level of measurement），293
測量的三角檢證法（triangulation of measure），224
測量的同義（measurement equivalence），662
測量信度（measurement reliability），280
測量效度（measurement validity），284
測謊題（sleeper question），425
測驗效應（testing effect），388
焦點團體（focus group），603
無回應誤差（nonresponse errors），412
無意識地（subliminally），26
硬性資料（hard data），226
等值信度（equivalence reliability），281
等距區間法（method of equal-appearing intervals），309
結構式觀察（structured observation），476
結構性解釋（structural explanation），109
虛假關係（spuriousness），257
虛無假設（null hypothesis），249
評估研究（Evaluation research），40
超驗（transcend），147
超驗觀點（transcendent perspective），147
進入的階梯（access ladder），564
進行推論（infer），670
量化（quantification），369
量化的（quantitative），14
量身打造（tailoring），442
量表（scale），298, 305
量度（scaling），305
量度圖分析（scalogram analysis），314
開放式問題（open-ended question），426
開放式編碼（open coding），674
間斷性時間序列設計（interrupted time-series design），380
間斷變項（discrete variables），293
階級分析（class analysis），148
集中趨勢量數（measures of central tendency），509
順序尺度測量（ordinal-level measurement），294
順序效應（order effects），435

十三劃

亂數表（random-number table），330
剽竊（plagiarism），722
微觀層次理論（micro-level theory），98
意識型態（ideology），80
極化效應（polarization effect, 605
極端個案抽樣（extreme case sampling），356

概念分類（concept classification），90
概念化（conceptualization），271
概念性定義（conceptual definition），271
概念性假設（conceptual hypothesis），275
概念的同義（conceptual equivalence），661
概念效度（Conceptual validity），59
概念網絡建構程式（conceptual network builders），710
概念叢（concept clusters），89
準實驗設計（quasi-experimental designs），379
罪惡的知識（guilty knowledge），602
補償行爲（compensatory behavior），390
解釋（explanation），100
解釋性批評（explanatory critique），154
解釋性研究（explanatory research），56
解釋型態（explanation pattern），538
詳析典範（elaboration paradigm），535
詮釋（interpret），242
詮釋取向社會科學（interpretive social science, ISS），129, 138
詮釋性解釋（interpretive explanation），114
詮釋學（hermeneutics），139
資料（data），13
資料記錄（data records），505
資料欄位（data field），505
跨個案研究（cross-case），241
跨國研究（cross-national research），643
路徑相依（path dependency），696
跳躍式（skip），425
過早下結論（premature closure），6
過度推論（overgeneralization），5
過程分析（process analysis），623
過程追蹤（process tracing），623
鉅觀層次理論（macro-level theory），98
隔離的錯誤（error of segregation），729
電腦輔助自填問卷訪談（computer-assisted self-administered interviewing, CASAI），422
電腦輔助面訪（computer-assisted personal interviewing, CAPI），422
電腦輔助電話訪談（computer-assisted telephone interviewing, CATI），444
零變異（zero variation），512
預先編碼（precode），504
預測（prediction），100
預測效度（predictive validity），286

十四劃

實用性導向（practical orientation），147
實在論者（realist），125
實務中的邏輯（logic in practice），228
實際成本評估（actual cost evaluation），49
實際的（practical），648
實徵的（empirical），13
實徵假設（empirical hypothesis），275
實徵推論（empirical generalization），115
實徵實在論（empirical realist），132
實質的（substantive），648
實質理論（substantive theory），99
實踐（praxis），155
實證主義（positivism），129
實證主義社會科學（positivist social science, PSS），129-130
實驗反應（reactivity），391, 396
實驗者期望（experimenter expectancy），390
實驗室實驗（laboratory experiment），397
實驗研究（experimental research），68
實驗真實性（experimental realism），397
實驗組（experimental group），376
實驗設計（experimental design），377
實驗對象流失（experimental mortality），389
對立假設（alternative hypothesis），249
對照脈絡（contrast contexts），684
對話式訪談（conversational interview），428, 450
對應規則（rules of correspondence），275
摘要（abstracts），178
槓桿顯著性理論（leverage saliency theory），442
構念效度（construct validity），287
演繹取向（deductive direction），96
滾雪球抽樣（snowball sampling），355
漏斗順序（funnel sequence），435
滿意（satisficing），429
漸次接近（successive approximation），678, 686
種族（race），88
算術平均數（arithmetic average），511
綜合（synthesis），629
網絡抽樣（network sampling），355

網絡理論（network theory），112
與當地人同化（go native），575
語意分析（semantic analysis），478
語意差別（semantic differential），312
語意關係（semantic relationship），688
認知人類學（cognitive anthropology），687
認知訪談（cognitive interviewing），461
認識論（epistemology），127
輕度偽裝（shallow cover），205
遠征研究（safari research），653
需求評估（needs assessment），47
領域分析（domain analysis），596, 688

十五劃

價值中立的科學（value-free science），136
層（strata），337
數盲（innumeracy），15
樣本（sample），320
樣本調查（sample survey），464
標明型態（specification pattern），535
標記（markers），593
標準化（standardization），302
標準化分數（standardized score），515
標準形式問題（standard-format question），430
標準差（standard deviation），513
模式效應（mode effects），412
模式配對（pattern matching），687
模稜兩可的問題（double-barreled question），414
確認因果機制（Causal mechanisms identification），59
編碼以及檢索程式（code-and-retrieve programs），708
編碼系統（coding system），476
編碼者間信度（intercoder reliability），479
編碼單（code sheet），505
編碼程序（coding procedure），504
編碼簿（codebook），504
編輯（editing），725
線性研究路徑（linear research path），229
線性關係（linear relationship），520
複合次數分配（compound frequency distribution, CFD），521
複製型態（replication pattern），535
複雜概念（complex concepts），89
調查後誤差（postsurvey errors），412
調查研究（survey research），68
質化比較分析（Qualitative comparative analysis, QCA），709
質化的（qualitative），14
適當性（adequacy），670
適當性假定（postulate of adequacy），145
適應性抽樣（adaptive sampling），357

十六劃

凝聚力（cohesion），112
學術自由（academic freedom），755
學術研究（academic research），37
學術期刊文章（scholarly journal articles），21
操作化（operationalization），274
操作性定義（operational definition），274
操弄確認（manipulation check），392
整體性闡述（Holistic elaboration），60
整體設計方法（total design method），442
橫斷性研究（cross-sectional research），62
機率抽樣（probability sampling），320
機構審查委員會（institutional review board, IRB），204
歷史（history），630
歷史比較研究（historical-comparative research），70
歷史效應（history effect），388
歷史偶然性（historical contingency），698
篩選（screen），425
輻合效度（convergent validity），287
選擇性編碼（selective coding），679
選擇性觀察（selective observation），5
選擇偏誤（selection bias），386
遺失資料（missing data），301
錯誤共識（false consensus），6
錯誤的解釋（false explanations），250
錯誤意識（false consciousness），153
隨手筆記（jotted notes），583
隨意抽樣（haphazard sampling），322
隨機（random），329
隨機分派（random assignment），372
隨機回應技術（randomized response technique, RRT），423
隨機數碼撥號（random-digit dialing, RDD），

344
隨機樣本（random samples），329
霍桑效應（Hawthorne effect），396
靜態群組比較（static group comparison），379

十七劃

壓縮回答（telescope），418
應用研究（applied research），38
檔案館（archive），630
總合評分（summated-rating），305
總結性評估研究（summative evaluation），42
縱貫性研究（longitudinal research），62
聲望抽樣（reputational sampling），355
聲望的偏誤（prestige bias），414
舉例法（illustrative method），686
賽斯通量表法（Thurstone scaling），309
隱性知識（tacit knowledge），554
隱性編碼（latent coding），478
隱密研究（covert research），601
隱藏性母體（hidden populations），357

十八劃

叢集（cluster），340
叢集抽樣（cluster sampling），340
歸納取向（inductive direction），96
禮貌偏誤（courtesy bias），651
簡要（parsimony），78
簡單概念（simple concepts），89
簡單隨機樣本（simple random sample），330
簡報（debrief），399
職業隔離（occupation segregation），28
謬誤（fallacies），250
轉型（transform），157
轉型的觀點（transformative perspective），157
離散程度（spread），512
雙盲實驗（double-blind experiment），391
雙重性別的名字（androgynous first name），491
雙組唯後測設計（two-group posttest-only design），380
雙載假設（double-barreled hypothesis），250
雙變項（bivariate），509
雙變項統計（bivariate statistics），517

十九劃

簽約研究（contract research），215
羅森伯格自尊量表（Rosenberg self-esteem scale），650
贊助者（sponsorship），442
關係（correlation），103
關係位置（relational position），81, 758
關聯性（association），103
關聯性（correlational），410
關聯量數（measure of association），528
關聯模式（models of relevance），752
關鍵指標（critical indicator），629
關鍵時刻（critical juncture），698
類型學（typology），90

二十劃

懸置法（bracketing），146
鐘形曲線（bell-shaped curve），511
闡明型態（interpretation pattern），538

二十一劃

屬性（attributes），243
欄外值（marginal），523
辯證（dialectic），151
辯證唯物論（dialectical materialism），148

二十二劃

權變評估（contingency evaluation），49
權變關係（contingent relationships），94

二十三劃

變異性（variability），512
變焦鏡頭（zoom lens），732
變項（variable），89, 243
變項概念（variable concept），89
邏輯實徵主義（logical empiricism），130
顯性知識（explicit knowledge），554
顯性編碼（manifest coding），477
體悟（verstehen），138