

## 質的研究的信度和效度

王文科  
東海大學

王智弘  
國立彰化師範大學

質的研究和量的研究有其差異性和密切關係，有的論者以為量的研究所追求者，質的研究亦應寄予重視，信度和效度即為其中之即使使用名稱雖有不同，但是性質亦頗相近。本文根據可信賴性和瞭解二項品質，探討質的研究的效度，前者又分成可信性和可遷移性。另就涉及研究的客觀性的可確認性，一併予以分析；進而申述提升質的研究效度的可行策略。至於瞭解的品質則分從描述效度、詮釋效度、學理效度、可概括性、和評鑑效度予以分析。最後析述質的研究信度，以可依賴性為主軸探索，並據以闡釋可以提升信度效標的各項可行策略。其中，在提升質的研究效度、信度和客觀性的諸項策略中，三角交叉、同儕或參與者評述、外界審核機制，和延長田野工作時間等策略似具有共同性。

關鍵詞：質的研究、信度、效度、可信賴性、瞭解、可確認性

---

王文科 東海大學教育研究所教授

王智弘 國立彰化師範大學教育研究所副教授



## 壹、緒論

教育領域的研究者與其他學術研究領域的研究者相似，往往在其研究者社群中有他們所根據的假定、概念、價值觀和實踐方式，而形成所謂研究典範(research paradigm)或觀點(perspective)。其中質的研究典範和量的研究典範曾經歷一番典範論戰，迄1980年代早期，量的研究普遍被接納作為教育研究領域的典範(Tashakkori & Teddie, 1998)，儘管如此，某些量的研究者和質的研究者在1980年代依然堅持自己所宗奉的研究典範優於另一種，形成二者壁壘分明且互不相容的論點(incompatibility thesis)(Johnson & Christensen, 2008)。這場論戰迄1990年代，由於有不少研究者拒絕互不相容論點，開始採取實用主義(pragmatism)的觀點，以為量的研究和質的研究各有擅長，俱屬重要，可在同一研究中混合相容，而衍生了混合研究(mixed research)的典範(Creswell & Clark, 2007)。

量的研究向來重視其研究結果的信度(reliability)和效度(validity)問題，但是質的研究的信度和效度問題，卻是長久以來被質疑的議題，Bogdan 和Biklen (2007)在他們的著作中，一直將信、效度視為質的研究的共同問題之一。

Bogdan 和Biklen (2007)對質的研究信度的描述，即是針對質的研究的內在信度(internal reliability)而言，他們指出：「在質的研究，研究者關注的是他們所蒐集得來的資料的準確性和綜合性。質的研究者傾向於將信度視為他們記載的資料與接受研究情境所真正發生的情形之間的適合性，而非指著跨不同觀察而得的刻板的一致性。----兩位研究者研究單一情境，可能提供不同資料與產生不同發現；該兩項研究可能是可信的。假如它們產出矛盾的或不相容的結果，吾人始質疑其中一項或兩項研究的信度。」(p.40)

基於上述的分析，本文嘗試先就量的和質的研究作一比較，以了解二者的分野和關係；然後，專就質的研究的信度和效度及其涉及的相關議題，予以探討。

## 貳、量的和質的研究的比較

先就量的和質的研究二者的不同點加以析述，純正的研究者以為量的和質的研究各自依據的世界假定、研究目標、研究方法、典型研究、研究者角色和研究脈絡的重要性不同(Denzin & Lincoln, 2000；McMillan & Schumacher, 2010)。

有關世界的假定方面，量的研究經常根據邏輯實證論/後實證論(logical positivism /post-positivism)的某些形式，假定具有穩定性的、社會事實的單一實體(single reality)，其與個人的感受和信念有所區隔。質的研究者仰賴詮釋學/建構主義(interpretivism/constructionism)的成分較濃，假定實體是多元的，對相同的情境，透過個別與團體知覺或觀點的方式，進行社會性建構而得(McMillan, 2008)。是以，質的研究認定的實體，係透過個體與其社會的世界互動建構而成。



有關研究目標方面，量的研究針對接受測量的諸項社會事實中，試圖尋找建立彼此之間的關係，以及解釋產生變化的原因之所在。質的研究比較關注從參與者的觀點，了解社會的現象；但是，這種了解，須視研究者擔任研究者角色時，參與接受研究者生活的程度決定。

至於研究的方法和過程方面，在量的研究有一套用以引導研究者的程序和步驟。質的研究策略和過程則較具有彈性，典型上，質的研究者採用緊接著的設計(emergent design)，在研究期間，可以決定修正蒐集資料的策略。恰成對比的是量的研究者，在開始蒐集資料之前選定的方法，是其預先確定設計(pre-established design)的一部份。

在典型研究方面，量的研究者爲了減少誤差、偏見和無關變項的干擾，多使用的典型研究爲實驗設計或相關設計。典型的質的研究者對持續性事件所做質的研究爲人種誌研究(ethnography)，以協助讀者了解接受研究人士對社會情景或社會系統所採取的多元觀點。換句話說，量的研究試圖透過研究設計，以控制偏見；但是，質的研究則尋求考量在資料蒐集和詮釋方面的主體性(subjectivity)。

有關研究者角色方面，爲了避免偏見，理想上，量的研究者須和研究分離。質的研究者則要「浸淫」於接受研究的情境和現象之中。例如：質的研究者在研究脈絡擔任互動性社會角色時，就要觀察參與者、與參與者訪談，並做成記錄。在蒐集資料時，質的研究者強調要由有技能、且準備就緒人士擔任；而不像量的研究交由「工具」處理。質的研究以Erickson(1973)所指的「受過學科訓練的主體性」(disciplined subjectivity)以及Manson(1996)所提的「反思」(reflexivity)著稱，亦即質的研究強調在整個研究過程中，研究者的角色是要不斷作批判性自我檢查。

至於研究脈絡的重要性方面，多數量的研究試圖建立「普遍不受脈絡約束的通則」(universal context-free generalization)。質的研究者相信人的行動深受引發行動之環境的影響，研究者若對接受研究者用來詮釋其思想、感受及行動所憑依的架構不了解，勢必無從了解人的行爲。質的研究者在資料蒐集和分析期間，需留意該架構或脈絡，質的研究所提出的是受脈絡限制的摘要(context bound summary)。

爲了更深一層了解質的和量的研究的區別，Bogdan 和Biklen (2007)以及Wiersma和Jurs(2009)從二者的特徵著手，使用對比方式呈現，如表1所示，可用以補充上述分析的不足。



表1：質的和量的研究的特徵

質的研究	量的研究
與該研究有關的名詞/術語	
人種誌	實驗
文件	硬性資料
田野工作	外界觀點
軟性資料	經驗
符號互動	實證主義
內在觀點	社會事實
自然主義	統計分析
人種方法論	科學方法
描述	
參與觀察	
現象學	
芝加哥學派	
生活史	
個案研究	
生態學	
敘事描述	
-----	
質的研究	量的研究
詮釋	
與該研究有關的關鍵概念	
意義	變項
常識性了解	操作化
放入括弧	信度
情境定義	假設
每日生活	效度
協商順序	統計顯著性
瞭解	複製
過程	斷定
爲了所有實務的目標	
社會性建構	

表1 續

紮根理論	
理論的根源	
符號互動	結構功能論
人種方法論	實在論，實證論
現象學	行為主義
文化	邏輯經驗論
觀念主義	系統理論
學術根源	
社會學	心理學
歷史	經濟學
人類學	社會學
	政治學
目標	
發展靈敏化概念	測試理論
描述多元實體	建立事實
紮根理論	統計描述
瞭解社會現象	顯示變項間關係
	斷定

質的研究

量的研究

瞭解因果關係

設計

演進的、彈性的、一般的  
作為您如何進行的預感

結構的、預定的、正式的、特定  
細節的運作計畫

書面的研究計畫

簡短的  
思索的  
提出有關聯性的研究領域  
資料蒐集完成之後經常寫成書面  
不做廣泛的實質性文獻探討

廣泛的  
焦點詳細且具體  
程序詳細且具體  
詳細的實質性文獻探討  
在資料蒐集以前寫成書面



表1 續

敘述一般的方法	敘述假設
資料	
描述性	量的
個人文件	可以量化的編碼
田野筆記	計數、測量
相片	操作性變項
人們自己的用語	統計量數
官方文件和其他人工製品	
樣本	
小的	大的
非代表性	分層
理論抽樣	控制組
雪球式抽樣	精確
合目標	隨機選樣
整體探究	控制無關變項
	聚焦個別變項
技術或方法	
觀察	實驗
參與觀察	調查研究
-----	
質的研究	量的研究
閱覽各種文件等	結構性訪談
開放性訪談	準實驗
第一人稱解說	結構性觀察
與研究對象的關係	
同理心	分離的研究者角色
強調信任	短期性
平等主義	疏遠的
研究對象即朋友	研究對象----研究者
密集性接觸	有界限制



表1 續

工具和儀器	
錄音機	量表
轉錄機	問卷
電腦	索引
	電腦
	量尺
	測驗分數
資料分析	
持續性	演繹法
模式、論題、概念	資料蒐集結束進行
歸納法	統計
分析歸納	
恆常比較法	
使用該研究法的問題	
消費時間	控制其他變項
資料化約困難	具體化
信度	強制性
程序未標準化	效度
難以研究大樣本	

註-：取自Qualitative research for education: An introduction to theories and methods (5th ed.) (pp. 44-46), by R. C. Bogdan and S. K. Biklen, Boston: Allyn and Bacon. Copyright 2007 by Allyn and Bacon.. Research methods in education: An introduction (9th ed.) (p.15), by W. Wiersma and S. G. Jurs, . Copyright 2009 by Allyn and Bacon.



最後就質的和量的研究二者的關係言之。揆諸事實，質的和量的研究二者之間的這些區隔，有不少並非絕對。當吾人執行一項研究或閱讀一篇已經完成的研究報告時，當會有如此的感受。尤其當研究者進行單一的評鑑研究或政策分析研究時，就需結合質的和量的兩種研究方法，在這種情況下，質的和量的研究便形成互補的關係，一方面有統計數字作為證據，另一方面也有文字的描述，可以補充其敘述的不足，實用主義論者或許秉持的是這種立場，因而倡導結合二者而成的混合法。同時結合質的和量的研究二者雖然有其頗高的困難度（Tashakori & Teddlie, 1998），但是將二者的差別和關係予以描述，應是有用的安排。

## 參、質的研究的效度

Dey(1993)以為為了檢核蒐集得來之資料的品質或效度，經常發問以下六個問題：

- 資料係根據自己的觀察而得，或出自道聽塗說？
- 所蒐集的資料是否有經過他人進一步的確證？
- 在什麼條件之下，進行的觀察或提出的報告？
- 提供資料的人士的可靠性如何？
- 什麼動機影響參與者提出的報告？
- 什麼偏見影響了觀察和提出的報告？

這些問題似乎涉及資料是否可信和有效的問題，質的研究者採取與量的研究者不同的術語，來探討信、效度問題。雖然Smith(1983)主張，質的與量的研究二者的基本假定，如前一節的分析，可知並不相容，因此信、效度的概念應該拋棄，然而多數質的研究者並不接納該種說法。

有關蒐集得來之資料的品質所涉及的概念，質的和量的研究各自使用的術語並不一致，Anfara, Brown和Mangione(2002)將之做了以下的歸納，可供參考。本文所使用的信、效度術語，即依據他們的分析而來。

- 可信性〈credibility〉=內在效度〈internal validity〉
- 可遷移性〈transferability〉=外在效度〈external validity〉
- 可依賴性〈dependability〉=信度〈reliability〉
- 可確認性〈confirmability〉=客觀性〈objectivity〉

傳統所指量的研究效度，係指一種測驗能夠測量出其所要測量的東西的程度。質的研究所指的效度，與其說法不同，專指研究者蒐集得來的質的資料，能夠準確地判斷吾人想要估計的東西的程度，亦即關注研究發現的準確性(accuracy)和真實性(truthfulness)。從歷史演變的角度觀之，一開始效度是和數量本位的量的研究連結在一起的。惟自1970年代末和1980年代初，質的研究深受重視之後，質的研究者開始感覺到有一股壓力，需要去證明其研究具有合理性，以及確保研究的準確性和可信性。





質的研究者用來描述效度的專門術語有二，一是Guba(1981)以及Lincoln和Guba(1985)所主張的可信賴性(trustworthiness)；另一為Maxwell(1992)採用的瞭解(understanding)概念。

先就可信賴性分析，Guba將之包括可信性(credibility)、可遷移性(transferability)、可依賴性(dependability)，和可確認性(confirmability)。Gay, Mills和Airasian(2009)以及Mills(2010)於分析質的研究效度時，即據此而來，惟有的學者(如：Ary, Jacobs, & Sorensen, 2010; Marshall & Rossman, 2006)則認為，可依賴性應屬於信度的範疇；可確認性則與客觀性、中立性較有關聯。本文係根據後者的見解，將可信賴性移往信度分析，可確認性隨之移往客觀性探討。

## 一、可信賴性：可信性、可遷移性

評估可信賴性效度的標準之一為可信性。當質的研究者在評估研究者的觀察、詮釋和結論是否具有可信性(credibility/ believability)時，可信性即可取來和量的研究中的效度規準(包括測量的效度和內在效度)並行探討。

質的研究追求的可信性之目標在於：探索問題，或者描述場景、過程、社會群體或互動組型。亦即研究者能夠對呈現在研究當中的一切複雜性予以通盤考量，同時處理不容易被解說的問題。是以可信性以獲知參與者對情景或事件的知覺，和了解研究者在研究報告中對它們這些所做的描繪是否匹配。換言之，可信性乃在於探求研究者是否已經能夠準確表述參與者的所思、所感和所為，以及影響他們的思維、感受和行動的過程。

質的研究者爲了了解可信性，通常要發問以下的問題：研究者以重複的、延長時間的和實質的投入田野工作的方式如何？研究者耗於場景的時間足夠嗎？研究者以有意義的方式和參與者保持互動嗎？出現的觀察者效應減輕，觀察者深受參與者信賴，並與其建立密切關係嗎？研究者所使用的方法，是否易於對研究的場景和參與者提出準確的和深度的圖像？

爲了回答上述問題，並確保可信性，研究者可採取以下的策略(Ary, Jacobs, & Sorensen, 2010; Guba, 1981; Gay, Mills, & Airasian, 2009; Johnson & Christensen, 2008; Marshall & Rossman, 2006; Richard, 2005)：

- (一) 延長在研究場所參與的時間：俾讓研究者發展出一種適合於解釋資料且能發揮防衛作用的學理，即所謂獲致學理適當性本位的證據(evidence based theoretical adequacy)，以克服曲解和偏見。
- (二) 進行持續性觀察：以確認研究者所蒐集之資料，具有普遍性和非典型的特徵。
- (三) 作同僚探詢(peer debriefing)：係指透過同儕間專業性互動，以增進洞察力，也就是就有能力的他人相互之間，對於描述、詮釋、與確定的主題，有一致性共識，以獲致一致本位的證據(evidence based on consensus)。
- (四) 平常檢查田野筆記並向研究者的同事發問問題：以協助研究者再度檢查假定，並考量查看資料的替代方式。在許多行動研究中，與其他實務工作者合作與定期性會議，可作爲



同儕報告的功能。

- (五) 確立有關項目的適當性：即蒐集文件、錄影帶、錄音紀錄、人工製品、其他原始的或生活切片的資料項目等，可作為研究的一部份，以供分析和詮釋之用。
- (六) 成員查核(member checks亦稱參與者回饋)：訪談稿或研究者研究結論摘要，要送給參與者表示評述意見，以探求接受研究者是否同意研究者對他們所說到的一切。此外，研究者藉著記錄反思的田野筆記，以及記錄他們的思維報告，以持續調控他們自己主觀的觀點和偏見。此種準確描繪與參與者有關聯性、且由研究者所研究的意義；以及參與者的觀點、思維、感受、意向和經驗被準確了解與描繪的程度，亦謂之獲得指涉或詮釋適當性本位的證據(evidence based on referential or interpretive adequacy)。
- (七) 實施三角交叉(triangulation)：指的是所有質的研究者使用多種方法、資料蒐集策略和資料來源的過程，以獲得接受研究之對象的完整圖像以及作複核資料之用，確定他們對研究的地方和人員有廣泛的陳述。由這些不同出處的資料，透過三角交叉來做比較，以確證研究者所作的結論。如果資料來源提供的是衝突的資訊，研究者應為這種衝突，探討出可能的理由。三角交叉可分成以下五種：1. 資料三角交叉法，在研究中採用各種來源(程序或工具蒐集而得)的資料，以協助了解現象，是一種用以取得結構確證本位證據(evidence based on structural corroboration)的手段，透過該種手段，將多種類型的資料彼此關連起來，以支持或否定對某事務狀態所做的詮釋與評鑑。2. 方法的三角交叉法，使用多種研究方法，如人種誌法和文件分析法，來研究現象，亦為取得結構確證本位的證據的一種手段。3. 研究者的三角交叉法，運用好幾個不同的研究者或評鑑者來蒐集、分析和詮釋資料，亦是用以取得結構確證本位的證據的手段之一。4. 理論的三角交叉法，對單一資料運用多種觀點來詮釋和解說資料，以獲致學理適當性本位的證據(evidence based theoretical adequacy)。5. 科際三角交叉法，使用其他學科推論過程以及了解研究發現，亦為獲致學理適當性本位的證據的一種手段。由於三角交叉在運用上的慎重其事，步步為營，所以據以提出的研究報告，使用多種低推論的描述 (low-inference descriptors) 以及豐厚的描述(thick ,rich description)方式，前者如逐字稿或直接引述，以協助讀者體驗參與者的世界。後者乃在於協助研究傳遞對研究脈絡的了解，即在於取得指涉或詮釋適當性本位的證據。
- (八) 注意聲音(attention to voice)：採取解放架構的研究者堅持，評鑑一項研究要採用此策略。問題有如：誰替誰說話？研究者找出社會中最無權力的人，讓他們參與活動嗎？這些聲音提供細微末節，且位於研究的最佳位置嗎？採取解放觀點的研究者也堅持從學理架構如女性主義理論，來檢查資料。
- (九) 建立結構性的確認或連貫性：以確保研究報告內部沒有存在衝突和矛盾之處。
- (十) 自我反思(self-reflexivity)：研究者積極地批判、反思自己可能具有的偏見，並予以確認。在資料分析過程中指出記載的反思，可使研究者變得更自知，且試圖去調整和控



制自己的偏見，以獲致控制偏見本位的證據(evidence based on control of bias)。在研究計畫中，可載有一節「研究者的偏見」，以討論研究者的個人背景、該背景如何影響研究，以及研究者使用什麼策略來探討可能發生的問題。

(十一) 分析負向的個案(negative case analysis)：這也是研究者用來減弱研究者偏見的一種策略。研究中所列出的假設，須要提供證據來予以支持。負向個案分析涉及檢查與假設矛盾或不一致的資料。當負向的例子獲得確認之後，研究者就要修正假設，或者提供解說該個案未適合的理由，亦在於獲致控制偏見本位的證據。

(十二) 建立外部查核機制(external audit trail)：即由一位獨立的研究者針對以下的問題，檢查研究過程中所蒐集而得的全部資料，1.研究發現紮根於所蒐集的資料嗎？每項研究發現與資料中的某些部分有清楚的連結嗎？2.提出的論題和資料的適當性如何？所有的詮釋和結論得到資料的支持嗎？3.研究者的偏見獲得控制嗎？

有關於上述位確保可信性所採取之策略與取得各項證據效度的對應，可參閱如表2所載。

表2: 提升可信性的策略簡表

證據效度	策略
結構確證本位的證據	方法的三角交叉 資料的三角交叉
一致本位的證據	同僚探詢/評論 研究者的三角交叉
指涉或詮釋適當性本位的證據	成員查核/參與者回饋 低推論描述/豐厚的描述
學理適當性本位的證據	延長在研究場所參與的時間 學理的三角交叉 科際的三角交叉

接下頁



表2: (續)

證據效度	策略
控制偏見本位的證據	組型匹配(涉及根據理論預測一系列結果會形成組型，然後決定真正結果適合預測組型的程度)
	反思 負向個案抽樣(分析)

註：修正自Introduction to research in education(8th ed.)(p.501), by D. Ary, L. C. Jacobs, & C. Sorensen, Copyright 2010 by Wadsworth, Cengage Learning. Methods in education research: From theory to practice(p.276), by M. G., Lodico, D. T. Spaulding, & K. H. Voegtle, Copyright 2006 by Jossey-Bass.

評估可信賴性效度的標準之二為可遷移性。可遷移性原由Lincoln & Guba(1985)提出，係指研究者相信每件事情是受脈絡限制的；因而主張，研究發現對於有與研究問題或實務問題類似的情境或其他團體，將有可應用性或概括性的程度。是以一套研究發現可否遷移至其它脈絡，仰賴的是研究脈絡与其它脈絡之間的相似性(similarity)與適配性(goodness of fit)而定。

至於這種研究場域與其他場域之間是否存在著相似性，係由讀者去做判斷，讀者要作如是的判斷，需有對研究場景和參與者、學校、資源、政策、文化和其他特徵間的差異，提供豐厚且詳盡的描述才行，此即涉及描述的適當性(descriptive adequacy)。所以可遷移性要問的，不是樣本是否具有代表性的問題；而是指該項研究，讓讀者們透過深入了解在研究場域發生的過程之後，決定在他們自己社區中，是否會要運作類似的過程。這種可遷移性的評估與量的研究外在效度的差異，在於後者係由研究者作推廣概括；前者則由可能的使用者作推廣概括的決定(Marshall & Rossman, 2006)。

影響可遷移性的因素，舉隅言之，約有：1.選樣效應(selection effect)，事實是接受研究的構念，對單一的群體來說，具有獨特性，因此以其為對象進行的研究結果，遷移的可能性如何，就會產生爭議。2.場景效應(setting effect)，事實是研究結果為接受研究之特定場景的函數，以其進行研究的結論，欲遷移至另一情境，也會有討論的餘地。3.臨時事故效應(history



effect)，事實是參與者獨特的臨時事故經驗，可能影響比較。4.反應性(研究本身的效應)(reactivity: the effect of the reach itself): 在質的研究中，當研究者為關鍵的搜集資料工具時，欲消除其影響力，是不可能的事；但是研究者透過反思的陳述，來描述他自己的偏見，和提供詳細描述觀察策略與訪談問題等事項，可協助讀者了解可能的影響力。在採用訪談技術的研究，反應性是一項嚴格要求的影響因素(Ary, Jacobs, & Sorensen, 2010)。基於上述的分析，為了提升質的研究之可遷移性，可採取如下的策略 (Marshall & Rossman, 2006；Mills,2010)：

1. 蒐集詳細的描述性資料：即所謂描述適當性(descriptive adequacy)，將蒐集而得的資料，用來比較已知之情境(班級或學校)和可能可被遷移至的其他情境二者之間的關係，以判斷研究結果可遷移至其他可能之脈絡的適宜性。
2. 做跨個案比較(cross-case comparison)：研究者研究兩個以上的個案，如果研究發現一致，將增加研究發現可遷移至其它情景或脈絡的可能性。
3. 三角交叉多元資料：多元資料可以用來確證、精密或闡明研究中的問題；多個案、多位資訊提供者、或多種資料蒐集方法的研究設計，將會強化該研究在其他情境中的適用性 (Rossman,& Wilson, 1994)。

茲參考Ary, Jacobs, & Sorensen,(2010)以及Lodico, Spaulding, 和 Voegtler(2006)的觀點將提升可遷移性的策略繪製如表3所示：

表3：提升可遷移性的方法簡表

效標	策略
描述的適當性	豐厚的描述參與者、互動、文化、政策、 場景與背景
相似性	跨個案比較 文獻比較 描述限制
限制反應性	反思性陳述 詳細描述方法

註：參考自Introduction to research in education(8th ed.)(p.502), by D. Ary, L. C. Jacobs, & C. Sorensen, Copyright 2010 by Wadsworth, Cenage Learning.. Methods in education research: From theory to practice(p.276), by M. G., Lodico, D. T. Spaulding, & K. H. Voegtler, Copyright 2006 by Jossey-Bass.



## 二、瞭解：描述效度、詮釋效度、學理效度、可概括性和評鑑效度

最後就Maxwell(1992)所認定比效度還要基本的瞭解(understanding)概念(p.281)來做分析。其根據瞭解而建立的類型學包括描述效度(descriptive validity)、詮釋效度(interpretive validity)、學理效度(theoretical validity)、可概括性(generalizability),和評鑑效度(evaluative validity)。

描述效度係指研究者在研究報告中所做的解說，具有事實上的準確性，沒有將個人所見、所聞予以組裝、或歪曲。描述效度有其重要性，因為描述幾乎是所有質的研究的主要目標和基本的要求(Maxwell, 1992, pp.285-286)。根據該項標準，研究者可藉著和資訊提供者(informants)一起檢核、或查對訪談錄音帶或錄影帶內容，來查核在解說中所複製的引用內容(包括對事件、物件、行為、人員、情景、時間、地點等的描述)是否正確。

一般探討描述效度的關鍵問題，有如：研究者所報告的與接受研究的群體中真正發生的一樣嗎？研究者準確報告所見、所聞嗎？

爲了回答上述這些問題，且能獲致描述效度的有效的策略之一爲，研究者使用三角交叉法。該策略涉及使用多位觀察者所做的紀錄，以及描述參與者的行為和所處的脈絡。或使用多位觀察者複核觀察所得的資料，以確認研究者所同意發生的事件。如果經多位研究者確證(同意)的觀察結果，外界評述者質疑事件是否發生的情況，就可能減少。結果是該研究會更具有可信性，且易於獲得支持。

詮釋效度係指準確描繪研究者所研究的附著於參與者行為的意義；亦即指涉由質的研究者準確了解參與者的觀點(participants' perspective)、思維、感受、意向和經驗，並在報告中描繪出來的程度。了解研究參與者的「內在在世界」(主觀世界)是質的研究的重要部分，且詮釋效度乃在於指涉準確提出這些內在世界的程度。誠如Maxwell(1992)所言：「詮釋性解說築基於接受研究者的語言，且儘可能根據他們的用字遣詞和概念來做詮釋，再則，問題不在於用來解說這些概念的適當性與否，而在於應用於解說的個人觀點是否具有準確性。」(p.289)

有鑑於上述的觀點，研究者常使用多種低推論的描述方式行之，如逐字稿便是所有推論描述成份最低者，與其以直接引用來提供參與者正確的用語。另外爲了達成詮釋效度的最重要策略爲參與者回饋(或稱成員查核)，藉之可以澄清溝通上的誤差。

學理效度涉及由研究報告發展出來對現象所採取的學理性解說，與資料或被研究之現象適合的程度(Maxwell, 1992)，且這種解說是可信的和可以辯護的。例如：敘述有關發生在學校中的教育變革過程，以及對在該過程中的性別角色和權力所提供的解說，可用來協助讀者了解，在該過程中涉及的學理構念以及該理論在學校社群其他方面的應用情形。

提升學理效度的基本策略爲延伸田野工作的期間，以便有足夠時間去蒐集資料、產生和考驗所做的詮釋，則學理性的解說會更深入且詳細。另亦可採用理論三角交叉法、同儕評述、組型配對等策略進行之。

當研究者發展出學理的解說時，他應根據理論去做某些預測，並就這些預測的準確性，予



以考驗，此時研究者即進行組型配對。此項策略係指研究者要立即提出數個預測，如果發生的全部的預測如原先預測的一樣，研究者即獲得支持解說的證據。又當研究者發展學理的解說時，他也必須使用負向個案抽樣策略，以找尋與其解說不配合的例子或個案，因而研究者不僅僅去發現支持其所發展之理論的資料而已。

可概括性，依照Maxwell(1992)的說法，包括內在可概括性和外在可概括性兩種，前者係指對研究社區內部的可概括性；後者則指對非為研究者所研究之情景的可概括性。內在可概括性比外在可概括性重要，因為質的研究者很少主張研究的外在可概括性。

評鑑效度係指研究者是否盡可能以無偏見的方式、客觀報告資料；而非對資料作判斷或評鑑。

對多數質的教育研究者而言，使用可信賴性與瞭解之術語及其涉及的相關概念，來探討研究的效度，有其取向的根據。並賦予研究者來檢證自己的研究是否滿足專業的標準，這既是一個機會，也是一種挑戰。

## 肆、質的研究的可確認性

評估質的研究品質的另一項規準為可確認性(confirmability)。可確認性涉及掌握傳統上所指的客觀性概念，強調以蒐集得到之資料具有中立性和客觀性而言。

Lincoln 與 Guba (1985)主張質的研究中所做的邏輯推論和詮釋，能讓他人容易了解，才能增加論點的強度。是以質的研究需要發問：研究發現是否能為另一項研究發現所確認？如此作為，即在於了解：研究者所做出的邏輯推論和詮釋，他人是否易於了解？讀者或諍友了解推論是如何做成的嗎？這些推論容易被了解嗎？或研究者在程序和研究結果的詮釋中，不受偏見約束的程度如何？

由於在質的研究中，想要達成像量的研究那樣程度的客觀性，常不是一件容易的事，質的研究者關注的是他們蒐集的資料和從中獲致的結論，是否為探究相同情境的其他研究者所確認。所以在質的研究中，焦點是從研究者的中立性轉移到對資料和詮釋的確認。

至於提升可確認性的方法，建立完備的外在審核機制頗為重要，藉此，研究者可使另一位研究者就所賦予的相同資料和脈絡中，可以獲致或未能獲致相同的結論。其他如：方法三角交叉、同儕評述、反思等，均為可行的運用途徑。Ary, Jacobs, 與Sorensen( 2010, p.504)對於提升可確認性的方法，以表4示之：



表4：提升質的研究可確認性的方法

效標	策略
檔案文件	外在審核機制
確證	三角交叉 同儕評述
控制偏見	反思

又在上述各種涉及質的研究效度相關的品質概念時，分別提出各自提升效度的策略，固然有其著眼點，但是Wolcott(1994)以及McMillan 和Schumacher(2010)所提出提升質的研究效度策略的綜合見解，亦有其價值，茲一併引述於後。Wolcott(1994)提出確保質的研究效度的實用策略包括：多聽少講、準確紀錄觀察、提早開始撰寫、讓讀者充分見到自己的資料、充分報告、要坦白、尋求回饋、準確撰寫(pp.348-370.)。另McMillan 和Schumacher(2010)提出確保質的研究設計效度的策略有：延長與持續在田野工作的期間、運用多種方法的策略、逐字紀錄參與者使用的語言、作低推論描述、使用多位研究者、以機械記錄資料、使用參與研究者提供的資料、由成員參與查核、使用參與者評述、尋求負向的或不一致的資料。這些見解，當可做為綜合前述提升質的研究各項策略的觀點。

## 伍、質的研究的信度

信度泛指研究者測量任何事物所得之資料一致的程度。在量的研究的方面，信度通常用來指涉工具或測驗具有的特質；惟質的研究者雖然也在他們的研究中使用信度，卻不去評鑑工具或測驗的信度，而去考量他們所使用的技術搜集得來之資料的信度。例如：一位質的研究者檢查自己的研究結果時，如果使用相同技術有一段時間，他們應當考量搜集得到的資料是否有一致性。然而信度和效度不同，一種有效的測驗測量其想要測量的事物，在一段期間具有一致性；至於可信的測驗，可能測量錯誤的事務所具有的一致性。

可依賴性(dependability)雖然指涉研究者探討資料蒐集、詮釋的程序和過程，是否具有穩定性，可供他人追隨而言；所以也有人將之列入信度考量(Marshall & Rossman, 2006)。是以一項良好的質的研究會對資料蒐集與分析方式，提供鉅細靡遺的解說。





可依賴性經常是指在於只有摘述研究者提出之經驗性的結論報告，與包括對方法採取完整解說在內的實證性、研究本位的質的研究報告二者之間存在著差異。雖然在質的研究結果裏，不可能包括所有資料，但是許多質的研究者會將他們的資料提供給其他研究者，俾供隨時檢閱之用(Gay, Mills, & Airasian, 2009; Mills,2010)。爲了探討提升研究者搜集得來之資料的可依賴性問題，Guba(1981)建議可以採取如下步驟：

- (一)重疊法(overlapping methods，類似三角交叉法的過程)：研究者使用兩種以上方法，使一種較弱的方法爲較強的方法所補償。如訪談法可用來補償觀察法的不足。
- (二)外界的審核機制(external audit trail)：在該機制中，由他人決定做決定的方式以及情境具有何種獨特性，以及記載研究進行方式的文件，包括做什麼、何時做和爲什麼如此做等在內。審核機制涵蓋在訪談和觀察蒐集而得的原始資料、記錄研究者決定要訪談的對象或觀察的標的或觀察什麼和要做如此觀察的理由、彙存從原始資料中發展工作假設的方式以及後續的修正和考驗、研究發現等記載。此一步驟可以由外界的審核者(可能是諍友、校長或研究生)來檢查執行資料蒐集、分析和詮釋的過程，即審核機制檢查執行研究的方式，包括在什麼時候，做了什麼，理由爲何等。因此，研究者必須完整保持筭記和活動紀錄，以及保管做好組織的資料，以便隨時可供檢索之需。研究者也須提供研究樣本、選樣過程、脈絡描述、資料蒐集方法、詳細田野筭記、錄音帶、錄影帶以及可供他人檢核的描述性資料等。使用檢核機制是運用一個指導者，他是屬於獨立的第三團體的審核者，來檢查研究者所執行的研究，以證實所採用之程序的可依賴性，以及檢查研究發現是否可以證實的，亦即這些發現是否合乎邏輯地衍自且紮根於所蒐集的資料。完整提出的程序和結果，能給讀者在自然脈絡限制的範圍內，判斷該研究是否可以複製。
- (三)複製邏輯(replication logic)：爲了說明可依賴性，可就研究者在跨多元場景或跨多位人士進行的研究，所顯示出來研究發現的一致性表之。複製邏輯涉及在多個場所或以多個團體執行研究，以決定該項研究是否具有可依賴性，俾提供建議。依照該項邏輯，以多個時間、於不同場景或對不同的人員爲對象進行研究，如果研究發現爲真的次數愈多，就愈具有可依賴性，研究者據以做成的結論，就愈有信心。換句話說：如果研究發現可以以不同的人在不同場所做複製研究，則該項證據建議此一研究發現可以廣泛應用。
- (四)逐步複製(stepwise replication)：將資料交給兩位探究者，各自獨立進行區分和分析，然後將結果予以比較，該結果如有一致性，即作爲具有可依賴性的證據(Johnson & Christensen,2008)。
- (五)資料或方法的三角交叉，以確證得到類似的結果。
- (六)編碼---再編碼/評定者內部一致(code-recode/intra-rater agreement)：係指研究者將資料編碼，然後騰出一段時間來分析資料，接著回來重新將原來已經編碼的資料再次做編碼，進而比較該兩組經過編碼的資料，是否有一致性。



(七)評定者間/觀察者間一致(inter-rater agreement /inter-observer agreement)：由於質的研究常使用多位探查者，而使研究者爲了評估可依賴性，主張使用該項策略。如研究者隨機選取一份謄本，交給一位同僚，由他依照研究者確定的編碼系統，進行編碼；另一位編碼者則可自由地添加自己認定的編碼系統去做編碼工作，當後者編碼完成之後，將之取來和原先的編碼相比較，以了解兩位編碼者所編的成分是否相同。  
Ary等(2010)將上述提升質的研究策略和欲達成之效標並列如表5所示。

表5：提升質的研究可信性的途徑

效標	策略
檔案文件	外界的審核機制
一致的發現	複製邏輯 逐步複製
編碼一致性	編碼---再編碼/評定者內一致 評定者間或觀察者間一致
確證	重疊法 資料三角交叉 方法三角交叉

註：參考自：參考自Introduction to research in education(8th ed.)(p.504), by D. Ary, L. C. Jacobs, & C. Sorensen, Copyright 2010 by Wadsworth, Cenage Learning.  
“Criteria for assessing the trustworthiness of naturalistic inquires,” by E.G. Guba,1981, Educational Communication and Technology,29(2),75-91.

## 陸、結論

質的研究的效度，可從研究的品質切入分析，針對可依賴性與瞭解爲基準進行探討，前者又細分成可信性和可遷移性、可確認性和可依賴性四方面，本文除了將可信性和可遷移性納入探討質的研究效度的範疇，另將可確認性單獨處理，用以探討質的研究的客觀性或中立性。至



於可依賴性則視為質的研究的信度領域予以析述，此外，並針對可以用來提升這些品質的可行策略，逐項分析。至於瞭解此一品質的探究，則分成描述效度、詮釋效度、學理效度、可概括性和評鑑效度五項，分別檢測研究者蒐集的資料，能否準確判斷所要測量的東西。惟這裡所探討的效度涉及的判斷基準，係屬於概括性的。如涉及較專門性的質的研究效度，又會增列其較獨特的要求標準。如個案研究(case study)偏重詮釋效度(interpretive validity) (Gall, Gall, & Borg, 2007)，行動研究者(action researcher)亦倡導適用於該類研究的民主效度、成果效度、過程效度、觸媒效度和對話效度等基準(引自Mills, 2010)。

另質的研究的可確認性，即在於追求研究結果的客觀性，本文根據Ary 等(2010)的見解，也提出可以迎合提升可確認性效標的可行策略。

至於質的研究的信度，則以可依賴性表之。本文除了敘述其意義之外，並採用Guba所提提升信度可行的諸項策略予以分析，以供參考。

又在提升質的研究效度諸項策略中，依序以三角交叉、同儕或參與者評述、外界審核機制、延伸田野工作期間等被提出次數較多；另三角交叉與外界審核機制則為提升質的研究之信、效度共同採取的策略。揆諸事實，在提升質的研究信度使用的評定者內部一致、觀察者間一致等策略，與提升質的研究效度中的採用多位觀察者、同儕評述、成員查核等策略，似有相似之處。甚至提升質的和量的研究的信度策略，即使做法不同，但運用的術語卻有雷同之處，或許更因而凸顯需探討質的研究之信、效度的必要性。



## 參考文獻

- Ary, D., Jacobs, C., & Sorensen, C. (2010). *Introduction to research in education*(8th ed.). Belmont, CA: Wadsworth, Cengage Learning.
- Anfara, V. A., Jr., Brown, K. M., & Mangione, T. L. (2002). Qualitative analysis on stage: Making the research process more public, *Educational Research*, 31(7), 28-38.
- Bogdan , R. C., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theories and methods* (5th ed.). Boston: Allyn and Bacon.
- Creswell, J. W., & Clark, V. L. P. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Dey, I. (1993). *Qualitative data analysis: A user-friendly guide for social scientists*. London: Routledge.
- Denzin, N. K., & Lincoln, Y. S. (Eds.) ( 2000). *Handbook of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.
- Erickson, E. (1973). What makes school ethnography “ethnographic?” *Anthropology and Education Quarterly*, 9, 58-69.
- Gay, L. R., Mills, G. E., & Airasian, P. (2009). *Educational research: Competencies for analysis and applications* ( 9th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Boston: Allyn and Bacon.
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology*, 29(2), 75-91.
- Johnson, B., & Christensen, L. ( 2008). *Educational research: Quantitative, qualitative, and mixed approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Lincoln, Y. S., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Manson, J. (1996). *Qualitative researching*. Thousand Oaks, CA: Sage.
- Maxwell, J. H. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, 62(3), 279-300.
- McMillan, J. H. (2008). *Educational research: Fundamental for the consumer* (5th ed.). Boston: Allyn and Bacon.
- McMillan, J. H., & Schumacher, S. (2010). *Research in education: Evidence-based inquiry* (8th



ed.). Boston: Pearson Education.

Marshall, C., & Marshall, C. (2006). *Designing qualitative research* (4th ed.). Thousand Oaks, CA: Sage.

Richards, L. (2005). *Handling qualitative data: A practical guide*. Thousand Oaks, CA: Sage.

Rosman, G. B., & Wilson, B. L. (1994). Number and words revisited: Being shamelessly eclectic. *Quality and Quantity*, 28, 315-327.

Smith, J. K. (1983, March). Quantitative versus qualitative research: An attempt to clarify for issue. *Educational Research*, 12(3), 6-13.

Tashakkori, A., & Teddie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.



# Validity and Reliability of Qualitative Research in Education

Wen-Ke Wang  
Tunghai University

Chih-Hung Wang  
National Changhua University of Education

The article explored the difference and relationship between qualitative and quantitative research paradigm, using trustworthiness and understanding as criteria, to assess validity of qualitative research and confirmability to confirm objectivity of qualitative research, and dependability of qualitative research to measure reliability. To analyze validity of qualitative research, trustworthiness was classified into topics of credibility and transferability, and understanding was included descriptive, interpretive, theoretical, and evaluative validity and generalization. Then, approaches for enhancing each validity were proposed. The procedure of processing objectivity and reliability is the same as validity. Triangulation, peer briefing, participant review, external audit trail, and prolonged on-site work were more important ways among strategies for enhancing validity, reliability, and objectivity of qualitative research. The nature of enhancing reliability strategies of qualitative research may vary, but the terms, such as intra-rater agreement and intra-rater agreement /inter-observer agreement are similar to those in quantitative research, and the same as strategies of multiple observations, member checking and participant review for validity of qualitative research.

*Keyword* : qualitative research validity reliability trustworthiness understanding confirmability

---

Wen-Ke Wang, Professor, Graduate Institute of Education, Tunghai University

Chih-Hung Wang, Associate Professor, Graduate Institute of Education National Changhua University of Education

