

中華傳統術數之量化指標發展與信效度分析

林聖軒

國立陽明交通大學統計學研究所副教授

國立陽明交通大學數據科學與工程研究所 合聘副教授

劉政猷

輔仁大學宗教研究所博士生

鄭志明

輔仁大學宗教研究所教授

摘要

中華傳統術數源遠流長，目前已經發產出龐大的思想體系，也和中華文化與精神生活緊密結合。相對於這些完善的「質性」理論與實踐基礎，「量化」研究目前正處於發展中的階段。本文提出未來術數量化研究的三個重點方向：(1) 命運測量問卷發展(2)信度研究以及(3)效度研究。問卷發展能夠針對傳統術數可預測且可量化的項目，建立標準量化指標。信度研究目的是估計命理師之間對於論命結果的共識程度。效度研究目的是估計「命理師預測命主的命運」是否和「命主真實發生的命運」的吻合程度。然而研究者在進行術數量化研究時，也要同時注意到量化方法的諸多限制，且應在研究中對於傳統數術在量化之外所代表之民族精神生活、處世哲學、以及背後深層的文化價值，抱持著崇敬的態度。

關鍵詞：傳統術數、紫微斗數、量化指標、問卷設計、信度研究、效度研究、統計科學。

壹、前言

量化研究是所有社會科學發展的趨勢，例如當今語言學開始著重音聲的物理性質音頻分析，教育學針對新發展的教學方法在施教前後使用問卷量測受教者的認知是否有改變，經濟學開始以計量經濟為主流等等。除了社會科學量化方法蓬勃發展以外，臨床醫學也是另外一個例子：傳統西方醫學非常重視臨床經驗豐富的資深醫師之見解，但如今實證醫學當道，舉世皆奉臨床試驗的結果為黃金標準；相較於實證研究的結果，沒有量化數據的專家意見在證據等級上反而排在較低的位階。

中華傳統術數(本文中簡稱「術數」、或是「命理學」)源遠流長，目前已經發產出龐大的思想體系，多元的論命方法與派別，以及深邃的理論基礎；同時也和中華文化、精神生活、哲學思想、以及天人合一的世界觀都緊密的結合在一起。相對於這些完善的「質性」理論與實踐基礎，「量化」研究目前正處於發展中的階段。

在〈命理學量化研究之模型與方法論〉¹一文中，已經有提出針對傳統術數預測準確度研究的框架。該文探討了傳統術數的統計模型，以及「命運」的科學解釋。然而在該文中，有三項問題尚未被觸及：第一、命運的測量並沒有明確的操作型定義。第二、在測量準確度的同時，參與實驗的命理師是否具有代表性？或是更進一步的探討(1)命理師之間對於論命理論與實務是否有共識？(2)應該如何定義與測量命理師間對於論命結果的共識程度？第三、過去關於命理準確度的量化研究，所定義的命運乃是針對命主的「性格」，並且是使用概略性印象分數(General Impression)，如此一來，不但結果僅限於性格相關的預測準確度，其他各方面，諸如婚姻、與親友鄰居的關係、子女、事業、收入等等的預測項目，完全沒有觸及。

因此針對以上三個問題，本文提出未來傳統術數量化研究的三個重點方向：(1)測量問卷的發展(2)信度研究以及(3)效度研究。問卷發展能夠針對傳統術數可預測、可量化的項目，建立量化指標。信度研究和效度研究是心理學、醫學、公共衛生學與社會科學常見的研究之一。信度研究主要目的是研究不同測量是否有一致的結果。這個概念可以用來測量命理師之間對於論命結果的共識程度。效度研究主要是看特定測量方法是否能夠準確的測量到黃金標準的測量。就應用在命理學上來說，本研究提出將命理師預測命主的命運當做「特定測量方法的測量」，

¹ 命理學量化研究之模型與方法論。林聖軒、孫保羅(民 107)。第七屆【中華傳統術數文化學術研討會】論文集。

而將命主真實發生的命運當作「黃金標準的測量」，接著可用與效度研究同樣的研究設計以及統計方法，比較兩者的接近程度。

本文架構如下：第二章探討量化指標建立的意義以及問卷發展的重要性；第三章探討命理學的信度研究；第四章探討命理學的效度研究；第五章介紹命理量化研究中常見的誤解與限制。

貳、量化指標的建立與問卷發展

一、如何使用現有學科的量化指標來量化命理學

在進行量化的時候，首先要問：「到底想要量化什麼項目？」本文以紫微斗數為例，其他傳統術數方法可以使用同樣的方法學來建立量化指標以及問卷。紫微斗數主要使用命主出生時的八字，依照古書的算法排成命盤，再依照命盤中星斗所處的位置以及相互之間的關係，預測命主一生的命運。本文沿用〈命理學量化研究之模型與方法論〉一文的名詞並且稍作修改。首先定義「真實的命運」(或稱「命理事實」)為命主一生中各方面實際發生各種可測量的主觀與客觀項目，例如感情狀況、婚姻狀態、健康指標、存活狀態、收入、學歷、與親友子女之間的關係與互動。命理學家使用傳統術數預測命主的命運則定義為「預測的命運」，或稱「命理跡象」。不論是「真實的命運」、「命理事實」、「預測的命運」、或是「命理跡象」在此概括統稱為「命運」。

用這個定義來說，命理跡象有以下四種可能的情況：(1)真實命主已經發生的事情(2)真實命主原本應該發生不過因為某些因素而未發生的事情(3)真實命主正在發生或是尚未發生的事情(4)虛構命主(例如隨意給定一個八字以及性別所排出的命盤)。

使用這個名詞定義，「信度」就是相同的命主(真實或虛構皆可)，使用不同命理學家或是不同術數方法所預測的「命理跡象」相符合的程度。「效度」則是相同命主的「命理跡象」以及「命理事實」相符合的程度。信效度會在後章詳細介紹，在本章我們針對命理跡象和命理事實的定義與測量議題詳加闡述。

在紫微斗數中有兩個相對可測量的東西：輸入的「時辰」，以及輸出的「命理跡象」(以及相對應的「命理事實」)。前者相對無爭議，後者則是本文要探討的重點。命運可以分別對應到「心理學/精神醫學」、「醫學/公共衛生學」、以及「社會學」這三門學科常見的測量變項。以下章節將詳述如何使用現有學科測量命運，以及相關的研究與議題。

二、出生時辰的量化

出生時辰本身就是一個高度量化的東西。古人對於時間的量化測量已經有非常完整的理論與實踐體系，然而在本節簡單列舉三個用來論命時常見的量化議題：(1)出生時的時區：應該用中原標準時間，還是用各地的時區？如果是用時區，和太陽相對的關係，在同一個時區的兩個交會點應該算同一時間，卻要把在不同時區但非常鄰近的兩個點算同一個時區，這樣是否合理？(2)雙胞胎甚至多胞胎

應該如何取命盤較為合理？(3)因醫療理由、以及因非醫療理由、進行剖腹產的時辰，是否需要校正？這些問題各派都有自己的理論解釋。這和堪輿學常見的問題也很類似：「應該使用地球北極還是磁場的北極當作正北方？」²

三、命運的量化可以分成三個學門：心理學、社會學、與醫學。

紫微斗數有十二個宮位³，其中命宮是整個命盤的基礎，代表了整體格局以及基本盤的走勢，然而在單獨預測方面，命宮最主要代表的是命主的性格，人品，以及心理狀態，相對應的科學學門是心理學(或精神醫學)的各種心理計量以及診斷。在疾厄宮預測命主的身體健康狀態，相對應的科學學門是醫學以及公共衛生學所定義的健康狀況。其他宮位，諸如官祿宮預測命主的學歷或事業，財帛宮預測命主的收入，田宅宮預測命主的不動產或理財狀況，六親宮預測命主的感情、婚姻、與至親朋友的關係，以上這些都屬於社會學的測量範疇。

四、心理學常見的量表以及過去的研究

依照命宮所預測出來命主的人格、個性、脾氣，主要是心理學的測量範疇。然而心理學可以測量的問卷有非常多，常見的有五大人格量表、三向度人格特質量表(Tridimensional Personality Questionnaire, 簡稱 TPQ)、自殺危險程度量表(Scale of Suicidal Risk)、貝氏憂鬱量表、貝氏焦慮量表，以及世界衛生組織生活品質量表(WHO Quality of Life)中「心理健康」的問卷。以上不同的問卷分別測量不同的心理狀態或是人格向度。舉例來說，五大人格量表將人格分成經驗開放性、盡責性、外向性、親和性、情緒不穩定性等五個向度。三向度人格特質量表則是將人格分成「追求新鮮感」(Novelty Seeking)，「躲避傷害」(Harm Avoidance)及「報酬依賴性」(Reward Dependence)三型，這個量表所測量的三個向度分別和三種神經傳導物質有相對應的關係，相對受到生醫學者研究的歡迎。而貝氏焦慮/憂鬱量表則如其名，是臨床醫學研究最常用來測量焦慮以及憂鬱程度的標準量表之一。

目前大部分的命理量化研究主要著重在心理學變項這部分，在 1983 年以及 2006 年的準確度研究⁴，就是命理學家使用西洋占星學預測的人格特質，以及心

²三元玄空起星與磁偏角的對應。劉政猷(民 108)。第八屆【中華傳統術數文化學術研討會】論文集。

³十二個宮位分別為：命宮、兄弟宮、夫妻宮、子女宮、財帛宮、疾厄宮、遷移宮、僕役宮、官祿宮、田宅宮、福德宮、父母宮。

⁴ 詳見以下兩篇研究論文：(1) Carlson S. A double-blind test of astrology. *Nature* 1985;**318**(6045):419-425. (2) Wyman AJ, Vyse S. Science versus the stars: A double-blind test of the validity of the NEO five-factor inventory and computer-generated astrological natal charts. *The Journal of general psychology* 2008;**135**(3):287-300.

理學家使用五大人格量表預測的人格特質，當作「特定測量方法的測量」，與命主/受試者自行評估自己人格的相似性為「黃金標準的測量」，進行效度研究。但以上的人格測量都是質性的敘述，目前尚未有研究拿心理學的量表來量測命理針對性格的預測準確度。

五、醫學、公共衛生學與傳統術數量化研究

疾厄宮代表命主的健康狀態的命運，對應到醫學以及公共衛生學對於疾病以及事故傷害的諸多量測。紫微斗數興起於古代中國，因此預測的健康狀況大多是以中醫學的辯證方法來進行定義。中醫學已經是一個很完整的醫學體系；西醫的診斷有嚴格的定義以及編碼系統⁵，因此在量化上最大的問題是「如何將中醫診斷與西醫診斷進行配對」。因為健保的普及，台灣有十分完整的資料庫，然而這些都是西醫診斷。台灣其他各種資料庫也都非常的齊全，例如事故傷害資料庫，癌症登記檔、死亡登記檔等等。然而具有個人資料的健康數據庫，在使用上都有非常嚴格的規定，諸如去識別化等。因此理論上最多只有可能得知其出生年月日，要得到完整的八字或是姓名則相當的困難。因此如果要使用此一數據庫，只能使用不需要「時」甚至「日」的算法。例如用生肖犯太歲是否會增加該年度的死亡率、事故發生率，就是一個可以驗證的題目。

另外命理學家常有「斷生死」的技術，由於台灣的死亡登記檔建立得非常完整，因此在實務上可以在收案時登記命主的身分證，用此串聯死亡登記檔得到存活狀態的資料。存活分析是很成熟的統計方法，不過需要取得命主的身分證，能否通過倫理審查是一大難關。

六、社會學的變項、社會變遷調查量表以及過去的研究

除了命身宮以及疾厄宮以外的其他宮位，調查的幾乎都是社會學的測量範疇。舉例來說：官祿宮所預測的命運對應到命主的學歷、職業相關的狀態；財帛宮所預測的命運對應到命主的收入多寡以及收入來源；田宅宮預測的命運對應到命主的不動產狀況或理財態度與狀況；六親宮預測的命運則對應到命主的感情狀況、婚姻狀態、和父母、兄弟姊妹、子女、周遭鄰居的關係與狀態等等。

目前社會學的調查百家爭鳴，不過目前最著名的社會學研究資料庫當屬中央研究院社會學研究所主持的「台灣社會變遷基本調查」(以下簡稱「變遷調查」)問卷以及資料庫。該問卷可以被直接拿來使用當作與社會學相關命運變項的測量

⁵ 國際疾病傷害及死因分類標準，The international statistical classification of diseases and related health problems, 簡稱 ICD code，目前最新的是第十版

標準，也可以以此加以修改成更適合命理學命運的預測。

在過去的研究中已經有初步使用變遷調查資料庫來進行效度研究⁶；而在本期研討會論文集〈八字論命的量化問卷發展〉⁷使用八字論命為例，根據變遷調查問卷改編為針對「婚姻」、「收入」、以及「子女」三大項的問卷。該問卷相較於原本的變遷調查，更加接近八字論命中對於「婚姻」、「收入」、以及「子女」相關命運的概念。

值得一提的是，心理學研究也可以如同社會變遷調查一樣，先拿傳統問卷，例如五大人格量表，逐一與術數專家討論是否為命理可以預測之項目，如果傳統心理測量和命理學能夠預測的個性、人格有出入的話，可以發展一個針對於命理學的量化問卷。

⁶ 以社會變遷調查資料庫進行統計命理學之實證研究。林聖軒、劉政猷(民 109)。2020 台灣科技與社會研究學會年會。

⁷ 八字論命的量化問卷發展。陳彥宏、林聖軒(民 110)。第十屆【中華傳統術數文化學術研討會】論文集。

參、命理學家是否有共識——信度研究

一、信度研究的目的是量化命理師間的共識程度

信度研究是心理學、醫學、公共衛生學與社會科學最常見的研究之一，主要目的是看不同測量間的相關性(也就是「信度」)有多高。其中信度有分成許多種的信度，在此簡單列舉兩種最常使用的信度：如果是同一種測量方法或是施測者，針對同一事件/狀態重複進行測量之間叫做「再測信度(test-retest reliability)」；不同施測者(rater)測量的相關性叫做「觀察者間信度(inter-rater reliability)」。除此之外還有許多種信度，不過共同點都是比較不同測量的相關性的強弱，而非與黃金測量的結果(或是真實的結果)相比較。

本文中，我們使用「信度」這個概念，以及相對應的研究設計和統計分析，來量化命理師之間的共識程度。「信度」我們在這裡定義為相同的命主(真實或虛構皆可)，使用不同命理學家或是不同術數方法所預測的「命理跡象」相符合的程度。在這裡我們統稱為「信度」，而不再細分是哪一種細項的信度分類。

二、信度測量的研究設計

信度的研究根據命理師的多寡可以分成兩個階段來進行。在第一階段屬於先驅實驗，在此階段我們聘請二至三位命理師，分別預測命主的命運(命理跡象)，在進行統計分析得到經過統計量化後的共識度。信度研究的統計分析將在下節介紹。經過第一階段的研究之後，即可粗略得到命理師對於各方面命理跡象共識度的估計。

以本期研討會〈八字論命的量化問卷發展〉⁸所發展的八字論命問卷作為例子。該問卷中第三部分即是八字論命「命理跡象」的量化問卷。在第一階段的八字研究，兩位命理師獲得十二位虛構命主的命盤，根據這些命盤，預測十二位虛構命主的命理跡象，再根據這些命理跡象填寫這些問卷的第三部分。接著使用統計分析，分析這兩位命主在問卷第三部分的每個項目的相關性高低，如此一來，就可以知道使用八字論命下，在該問卷「婚姻」、「收入」、以及「子女」三大項目中命理師論命結果的共識程度有多高。

然而這兩位命理師是否具有代表性？共識度是否會根據不同門派，不同師承，或是不同算法而有不同的結果？如果門派相同且繼承同一師門，是否有學到老師的精髓，學習時間是否夠久？即便學習時間相似，職業時間的長短，影響到經驗

⁸ 八字論命的量化問卷發展。陳彥宏、林聖軒(民 110)。第十屆【中華傳統術數文化學術研討會】論文集。

值，是否也會影響到共識程度？即便以上都一樣，每個人論命的風格不同，著重的地方不一樣，是否也會影響到共識程度呢？以上這些問題，都可以從第二階段的信度研究來得到回答。

在第二階段，會由多位命理師預測命理跡象以外，還會請命理師填寫諸多自身的資料，以分析影響「共識度」的因子。為了確保命理師的隱私以及其個資安全，這些共識度的因子僅限於論命相關的資料，例如師承、門派、學習術數時間長短，以及著重的論命理論。與論命背景無關的問題只調查最基本的人口學變項：性別、年齡、以及教育程度三項。以八字論命為例，可參照就本期研討會〈八字論命的量化問卷發展〉⁹的第一部分與第二部分。

三、信度研究的統計方法

信度研究是非常廣泛使用的研究，不論是研究設計或是統計方法早已發展成熟¹⁰。有興趣的讀者可以參閱相關書籍。本章簡略敘述信度研究幾個最常見的統計方法，並且使用幾種假設的命理研究案例作為示範。

信度最常使用的量測是 Kappa 統計量、Spearman 相關係數以及皮爾森相關係數三個統計方法。這三個統計方法分別是針對類別資料、有序資料或是非常態分佈的連續型資料以及連續型資料這三種資料型態。為求簡潔易懂，我們使用「共識度」來代表「信度」，不論是用這三種統計方法的哪一種得到的。共識度的值是介於-100%到 100%之間的數字。如果是 0%的話，代表命理師之間毫無關連；如果是-100%代表命理師之間的預測結果完全相左。100%則代表命理師之間的預測結果完全相同。

以下我們列舉三個假設的案例，簡單示範如何使用這些統計量來解釋命主在婚姻方面的命理跡象之間的共識程度。在這裡我們假設結婚的命理跡象有三種分法：(1)分成二元變量：早婚與晚婚。(2)分成三元有序變量：早婚、平均時間結婚、晚婚。(3)確切的結婚年齡。以下的 Kappa 統計量使用網路上的免費計算軟體¹¹，可以在網頁直接輸入數據進行計算。

案例一、二元變量(使用 Kappa 統計量計算)

假設有十二位虛構命主以及兩位命理師使用八字論命方法預測命運，第一位命理師認為六位早婚，六位晚婚；如果第二位命理師也是判定同樣的結果，如此

⁹ 八字論命的量化問卷發展。陳彥宏、林聖軒(民 110)。第十屆【中華傳統術數文化學術研討會】論文集。

¹⁰ 王榮德。(2009)。流行病學方法論。

¹¹ 本文中的 Kappa 統計量網路計算機網址：<https://www.graphpad.com/quickcalcs/kappal.cfm>

一來使用 Kappa 統計方法，會得到 Kappa 統計量，也就是共識度，為 100%，代表兩位命理師間八字論命對於結婚狀態的結果完全有共識。

如果對於第一位命理師預測六位早婚的虛構命主，第二位命理師卻預測三位早婚、三位晚婚；對於第一位命理師預測六位晚婚的虛構命主，第二位命理師也是預測三位早婚、三位晚婚(如同表一所示)。如此一來 Kappa 統計量共識就會是 0%。

表一、共識度=0%的信度研究案例。其中列位項目以及欄位項目分別代表第一位預測的命理跡象/第二位預測的命理跡象。

	早婚	晚婚	總人數
早婚	3	3	6
晚婚	3	3	6
總人數	6	6	12

如果第二位命理師和第一位命理師認定的完全相反，十二位都是晚婚的話；這樣子在最後統計結果共識就會是-100%。

表二、共識度=-100%的信度研究案例。其中列位項目以及欄位項目分別代表第一位預測的命理跡象/第二位預測的命理跡象。

	早婚	晚婚	總人數
早婚	0	6	6
晚婚	6	0	6
總人數	6	6	12

以上的共識度並不一定是全有全無律，也有可能介於 0%到 100%(或是 0%到 -100%)的中間值。舉例來說，如果第一位命理師預測命理跡象皆為早婚的六位虛

構命主中，第二位命理師認為有四位是早婚，兩位晚婚；而第一位命理師預測命理跡象皆為晚婚的六位虛構命主中，第二位命理師認為有兩位是早婚，四位晚婚(如表三所述)，這樣子共識值則是 33.3%。

表三、共識度=33.3%的信度研究案例。其中列位項目以及欄位項目分別代表第一位預測的命理跡象/第二位預測的命理跡象。

	早婚	晚婚	總人數
早婚	4	2	6
晚婚	2	4	6
總人數	6	6	12

其中值得注意的是，統計學會把隨機(白話一點來說就是亂猜或丟銅板)的狀況進行校正。就像以上的例子，如果命理師完全相反的情況，共識度會是-100%，而不是 0%。而一半相同一半相反的情況，和丟銅板決定的事一樣的，如此一來會被認定是 0%而不是 50%。

案例二、有序變量——可以使用 Kappa 檢定量或是 Spearman 相關係數

如果結婚狀態有三種：早婚、平均時間結婚，以及晚婚這三種選項，而第一位命理師判定 1-4 號四位會早婚、5-8 號四位正常時間，9-12 號四位晚婚。如果第二位命理師也是做出一樣的判斷的話，那共識度就是 100%。如果全部都猜錯，共識度並不會變成 0%，而會根據猜錯的程度而做調整。例如早婚猜成平均時間結婚，也會獲得部分的共識度。

案例三、連續變量

八字論命量化問卷中，在進行婚姻的量化之時，是請命理師猜測命主「大約幾歲結婚」。這樣的連續變量和案例二一樣，歲數完全一樣就會獲得全部的共識度，但是如果猜錯，相差的歲數越少的話，就會越高的共識度的分數。至於共識度的算法則是採用皮爾森相關係數。這個係數和 Kappa 統計量一樣，也是-100%到 100%之間的數值，在解釋上也同樣代表「共識度」。

四、量化命理師間的共識度的必要性

在進行效度分析的時候，參與研究的命理師很容易被質疑是否夠資格代表命理界？是否為合格、正派的命理師？因此在執行效度分析之前，信度分析的結果相當必要。在高度量化的學門，各種變項都有標準的量測原則。以醫學來當作例子。各種疾病的診斷，如果持有專科執照醫師經過標準的診斷，再確認該名醫師有一定時間的執業資歷、執業過程沒有重大爭議案件發生過、人品無重大瑕疵、且無利益衝突之後，基本上就不會有人質疑該名醫師的診斷結果。但命理學家的教育養成相對是口授心傳，學派理論百家爭鳴，因此到底誰的論命理論與結果具有代表性，誰又具有資格做裁決，相對於醫學等有經過量化實證洗禮的學科，還是比較主觀的。

經過信度測量之後，如果不論是門派、年資、師承、性別、年齡等等的因素，命理師的共識度都很高的話，那代表性相對就沒有太大的問題。另外，如果「學習時間」是一個影響共識度的因子，例如學習時間大於十年的命理師之間共識度相當高，不過和小於十年的命理師之間共識度低，這樣子就可以得到「八字論命的學習時間至少要十年才能夠出師」的結論。

信度研究經過嚴謹的命理專家審查之後，還有機會可以做為「命理師認證考試」之用。若是以當代公認的大師預測結果為正確答案，如果一位剛出道命理師論命結果與正確答案有高度共識的話，則代表有資格。

值得一提的是，相對於需要真實命主的效度研究，信度研究因為是採用虛構命盤，在倫理議題上單純許多。同時因為不需要真實命主，因此在研究的成本大為降低，在選取命盤也有相對大的自由空間。

肆、命理預測和事實吻合的程度——效度研究

效度研究主要是看「特定的測量方法」是否能夠準確的測量到「黃金標準的測量」。將這個概念用在命理學時，會把命理師預測命主的命運(也就是「命理跡象」)當做「特定測量方法的測量」，而把命主真實發生的命運(也就是「命理事實」)當作「黃金標準的測量」，然後使用同樣的研究設計以及統計方法，比較兩者的接近程度。總而言之，「效度研究」就是想要測量相同命主的「命理跡象」以及「命理事實」相符合的程度，或是換成白話一點來說，就是「命理準不準？」因此先前「命理學準確度」研究，就是本文所提的效度研究。

目前大部分的命理量化研究都是效度研究，主要著重在心理學變項，在 1983 年以及 2006 年的研究¹²，就是命理學家使用西洋占星學預測的人格，以及心理學家使用五大人格量表預測的人格，當作「特定測量方法的測量」，與命主/受試者自行評估與自己人格的相似性為「黃金標準的測量」，進行效度研究。

目前本團隊已經開始使用社會變遷調查問卷來測試紫微斗數的準確度。研究設計以及統計分析可見初步研究的成果摘要¹³。以社會變遷調查問卷為藍本進行修改的問卷，則正在發展並進行信度研究中，等待完成後才會進行效度研究。

¹² 詳見以下兩篇研究論文：(1) Carlson S. A double-blind test of astrology. *Nature* 1985;**318**(6045):419-425. (2) Wyman AJ, Vyse S. Science versus the stars: A double-blind test of the validity of the NEO five-factor inventory and computer-generated astrological natal charts. *The Journal of general psychology* 2008;**135**(3):287-300.

¹³ 以社會變遷調查資料庫進行統計命理學之實證研究。林聖軒、劉政猷(民 109)。2020 台灣科技與社會研究學會年會。

伍、術數量化研究中常見的誤解與限制

一、量化是準確度的基礎，但項目更為廣泛

值得注意的是，很多人把量化研究和準確度(也就是本文的「效度」)畫上等號。然而本文中量化研究包含了三個部分，效度只是最後一個步驟。跳過了前兩個步驟直接進行效度研究，會有諸多問題：首先如果沒有經過量化指標的建立，怎麼知道彼此討論的「命運」是不是在講同一個概念？舉例來說，在 1983 年 Carlson 在 Nature 發表西洋占星學沒有顯著預測能力的研究，在本研究提出的「量化指標」的角度來看，只是針對心理學中「個人性格描述」的項目，且以命主對於自己的評分當作黃金標準。不但測量準確度未經確認，也較為缺乏統計上的檢定力；其二，在信度沒有確認之前，如何得知參與該研究的占星師具有代表性？如果換較為資深的占星師，或是找不同門派的占星師參與，會不會有不一樣的結果？最後，「效度」很容易被誤以為在挑戰傳統術數的真實性，相對會引起不必要的衝突與誤會。在本文提到的研究框架之下，傳統術數的量化研究將會更為全面而嚴謹。

二、不能量化的東西不代表不存在

雖然量化有諸多科學上的重要性，不過也要記得能夠量化的項目只是其中的一小部分。舉例來說，「父母對子女的愛」這件事情就非常難以量化。在研究親緣關係可能可以使用「每天陪孩子時間的長度」來量化，但是陪伴的品質、內容、互動關係，就屬於非常難以量化的項目。除此之外，還必須考慮到量化後測量誤差或是錯誤分類(measurement error or misclassification)的議題，命理亦然。因此研究者要非常清楚量化研究的極限，謙虛謹慎的解讀闡釋量化後的結果。

三、準度研究的論命過程完全不同於平日的論命經驗

命理師在進行論命準度研究時，會有和平日完全不同的論命經驗。傳統論命時，除了看命盤中的線索以外，還可以藉由觀察命主本身的氣色、談吐、舉止行為，以及言語互動中，得到更正確的推論。不過用問卷進行信效度測量中，完全沒有與命主的互動，雖然是排除「命理跡象」以外干擾的必要之惡，但如何把「命理跡象」與「與命主互動」的加成效應計算進去，這也是未來設計研究要達成的重要目標之一。

四、科學家常遭受的科學界以及命理界的誤解

在進行命理科學研究的科學家，會遇到兩面不討好的窘境。對科學研究生態不熟悉的命理專家，可能會誤以為科學家不是前來踢館，就是想用霸權的心態想要「破除迷信」。另一方面，進行命理科學研究又往往會被科學界認定為不入流的偽科學研究。這些科學與命理交會時的衝擊和宗教學界的「宗教對談」相當類似¹⁴；所遇到的困境以及和解共生之道亦可以取為借鏡；細節可參閱上期研討會論文集〈命理與科學的衝突和解與共生——從「宗教對談」出發尋求解決之道〉一文¹⁵。在此略為補充投稿國際期刊的實務層面：在「否證」的科學原則之下，「命理跡象無法預測命理事實(命理學不準)」是預設值，也就是統計檢定上的「虛無假說」；而「命理跡象能夠預測命理事實(命理學準確)」，也就是統計檢定上的「對立假說」。進行統計檢定會有「不顯著」以及「顯著」兩種結果，這兩種結果代表著「接受虛無假說」以及「拒絕虛無假說」兩種結論。為求白話，我們把這兩種統計檢定的結果稱做「虛無假說為真」以及「對立假說為真」(不過請注意在統計的語言這樣說並不完全正確)。如果套用到命理學，統計結果顯著以及不顯著對應到的是「命理學不準」以及「命理學準確」兩種結果。

在科學上，絕大多數的發表只會發表顯著的結果，因此在命理學的統計研究中，「命理學不準」的結果是預設值，這類被命理學家認為是「踢館」的研究結果，根本無法獲得期刊刊登發表。

即便是得到「命理學準確」的結果，在期刊刊登的過程也是困難重重。首先命理學界並沒有相對應的科學學門，因此目前的相關研究最好的是刊登在 1985 年《自然(Nature)》期刊。除此之外，近年來的命理研究只有數篇關於「人格特質」相關的刊登於心理學期刊，以及一篇使用易經分類候選人與當選人的姓名刊登於綜合型期刊「Quality and Quantity」¹⁶；大部分的研究集中在針對「術數行為」的研究，而非術數理論的量化與信效度分析。

五、量化科學的限制與反思

科學家或是數據學家，在進入傳統術數領域進行研究時，務必要認清科學或是量化的限制。舉例來說，人格雖然已經有許多心理量表來進行人格的測量，不過目前對於人格是否能夠測量還是有極大的爭論。亦有文獻指出著名的心理學量化研究目前只有三成的再現性¹⁷。因此在量化科學上，除了更加仔細地設計量表

¹⁴ 房志榮, 武金正, 莊宏誼, 陳敏齡, 陳德光, & 黃懷秋. (2000). 宗教交談: 理論與實踐.

¹⁵ 命理與科學的衝突和解與共生 - 從「宗教對談」出發尋求解決之道。林聖軒、劉政猷、鄭志明(民 109)。第九屆【中華傳統術數文化學術研討會】論文集。

¹⁶ Ho, L. S., & Chen, H. Y. (2019). A name classification method using the I Ching name configurations for Taiwan local elections. *Quality & Quantity*, 53(2), 641-651.

¹⁷ Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*,

以外，還要向術數學家請益確認「『量表所測量的命運』以及『術數學家論命時所說的命運』是否為相同的概念」；以上兩者都確認無誤之後，再多次進行精準且嚴謹的信效度研究後，才能夠謹慎保守的發表自己的研究成果。

由於統計科學的思維是以「排除虛無假說」或是「否證」的科學哲學下進行的，不應該在粗略的測量以及研究證據下直接否定傳統術數。在頂尖醫學期刊《英國醫學期刊》就曾經有一篇標題為〈Absence of Evidence is not Evidence of Absence〉短評¹⁸。如果這句話套用在命理效度研究結果的中文翻譯是：沒有科學證據支持「命理是準確的」，不代表有科學證據支持「命理不準確」。

最後，傳統術數除了預測命運以外，還有許多重要的意義，諸如民俗學、社會學，同時也代表了一個民族的精神文化。不論傳統術數量化的結果如何，研究者都應對於傳統數術在量化之外所代表之民族精神生活、處世哲學、以及背後深層的文化價值，抱持著崇敬的態度。

349(6251).

¹⁸ Altman, D. G., & Bland, J. M. (1995). Statistics notes: Absence of evidence is not evidence of absence. *Bmj*, 311(7003), 485.

參考書目

- 第七屆【中華傳統術數文化學術研討會】論文集。
- 第八屆【中華傳統術數文化學術研討會】論文集。
- 王榮德。(2009)。流行病學方法論。

參考期刊論文

- Carlson S. A double-blind test of astrology. *Nature* 1985;318(6045):419-425.
- Wyman AJ, Vyse S. Science versus the stars: A double-blind test of the validity of the NEO five-factor inventory and computer-generated astrological natal charts. *The Journal of general psychology* 2008;135(3):287-300.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251).
- Altman, D. G., & Bland, J. M. (1995). Statistics notes: Absence of evidence is not evidence of absence. *Bmj*, 311(7003), 485.