



國立臺灣大學理學院心理學系

碩士論文

Department of Psychology

College of Science

National Taiwan University

Master's Thesis

應用古典測驗理論與試題反應理論簡化題目：

以國立臺灣大學身心健康評量系統使用之量表為例

Item Reduction of National Taiwan University

Physical and Mental Health Assessment System through

Classical Test Theory and Item Response Theory

陳彥廷

Yen-Ting Chen

指導教授：姚開屏 博士

Advisor: Kai-Ping Grace Yao, Ph.D.

中華民國 114 年 7 月

July 2025



誌謝



從自己大學二年級開始接觸心理計量領域，直到現在可以完成一份可被接受的學位論文過程中，受到許多人的幫助與關心。一路走來，謝謝大家的陪伴，一同走進我的人生回憶，不論只是點頭之交，或是有著更親密關係的重要他人，在此想和大家說聲：謝謝你們這幾年的陪伴與協助。

對我而言，能夠完成學位論文的撰寫，情緒上的支持與鼓勵是最重要的一環。謝謝白智偉心理師、簡婉曦醫師和戴月明醫師在精神醫療上的支持。自高中一年級一同陪伴進行心理工作，讓我逐漸坦然面對自己的與眾不同，也讓我在碩士階段學習靠近自己的身心狀態，讓我的理想與現實不再是兩條平行線，而是一個可以實際和大家分享的產出。謝謝心理師和醫師在這一路上的建議與鼓勵，讓我從一個懵懂無知的小小孩慢慢變成屬於自己的獨一無二。

進到研究所之後，在學習過程中並不是一路順遂。相較許多同學一路拿高分到畢業，我的成績單更多的是停修和不及格。從碩士班二年級以來，許多老師在課業上給予相對彈性的空間，以輔助度過自己的學習困境。謝謝臺灣大學農藝系蔡政安老師、臺灣大學外語教學中心的張晨老師和臺灣大學農經系的劉鋼老師在修課期間的關心與支持，讓我知道原來我在修課過程中不是只有自己一個人，也謝謝三位老師在我調整藥物的過程中提供不同的學習輔助，終於在碩士班最後一年達成一項我認為不可能的目標：不停修任何一門課和通過學校的及格門檻。自大學以來每學期都在停修和放棄的路上徘徊，直至今日終於體會到堅持到最後一刻的感覺與樣貌，對我而言是一項新的人生成就，開啟人生新頁。

身為一個研究生，對任何事物抱持著打破砂鍋問到底，還問砂鍋在哪裡的精神下，在碩士班三年級時進到桃園機場和星巴克打工。自上次出國參與 IMPS 時遇到航班誤點，對機場運務產生許多的好奇與疑問下，因緣際會和長榮航空地勤人員與長榮航勤地勤人員共事兩個多月。謝謝長榮相關企業的提攜與照顧，儘管曾經都被桃園機場公司進行一些不合理的作業流程修改，但這兩個月讓我學習到許多只有在管制區才會知道的冷知識；諮商完之後常會去星巴克坐

坐，帶著對星巴克本身定價策略與空間安排的疑問進入星巴克打工學習。儘管工作時間不長，學習過程中也充滿著辛苦與挫折，但也謝謝自己曾經進入星巴克企業體系，讓我知道許多星巴克不為人知的小知識與夥伴秘辛。

碩士班四年級開始前，對自己未來的迷惘。當時帶著一顆忐忑的心回去雲林老家一趟。身為家中長子，也快二十年沒回去老家。謝謝當時來星巴克北港門市支援的夥伴于旆，讓我在休學和堅持完成學位論文的交叉路口中選擇繼續努力。雖然只是一杯冷萃，但卻是讓我一路走到現在的精神支柱與契機。這一年間，謝謝北港門市的夥伴芯恩、育譚和育嫻，雖然每個月只會去門市一次，但非常謝謝你一直記得有我這個客人；也謝謝北港門市的眾多夥伴對我的照顧，或許平常沒有太多機會可以和大家聊聊，但自顧客與交流的過程中，大家的熱情跟溫暖都是讓我可以完成碩士學位論文非常無可或缺的關鍵。

走到現在，學位論文能夠完成，除了情感支持外，儘管知識可以透過自我資料查找完成撰寫，但更多的是對一項知識的思考如何進到更深層次的範疇。謝謝我的碩士學位論文指導教授：臺灣大學心理學系姚開屏老師、學位論文口試委員成功大學心理學系陳俊宏老師和臺灣師範大學教育心理與輔導學系陳柏邑老師在本人學位論文上的提點與指教，協助精進對單一議題的思考深度與完整性。謝謝臺灣大學心理系翁儷禎老師和彭昭英老師在學術上的觀念提點，在心理計量領域中的用字遣詞和方法適用情境中給予精闢的見解與提醒。或許以既有的碩士學位論文中仍有許多需進行改進的內容或方法，但相信在各位老師的建議下，碩士學位論文和未來的博士班階段也會秉持各位老師對學術的期待與要求，持續精進自我邏輯思辨與學術知識的文字表達。

最後，謝謝我的家人們在我踏進心理計量領域後對我的支持，雖然他們並不清楚我的論文主題和我在做什麼事情，但他們仍舊守護在我身邊，陪伴著我完成學業與學位論文；謝謝我在東吳的學妹郁潔在學位論文考試當天的協助，讓我可以專心處理口試的內容；也謝謝我帶過的學生們，雖然只有一學期或一學年的互動，但很感謝大家對我教學行政的支持與包容。當然，我也謝謝許多人在這四年對我的指教，儘管在人生中已沒有了過多交集，但仍舊謝謝你們曾經進到我的生命中，讓我的生活變得更加多采多姿。

摘要



2020 年，國立臺灣大學接連發生多起學生自傷或自殺事件。校方輔導體系為了解學生的身心狀態或身心適應程度，自 2020 年起著手「身心健康評量系統建置」計畫。此評量系統建置從 2021 年至 2024 年間收集多份測驗資料，並於每年收集資料後進行測驗與題目評估與增減。儘管歷年已減少部分不適合施測試題，提供學生施測的題目數量仍然偏高下，本研究將針對國立臺灣大學 2024 年大學部部分施測資料進行「正向因應資源」、「負向情緒反應」、「多向度感知社會支持量表」與「自我滿意度量表」試題簡化。除使用描述統計、古典測驗理論（Classical Test Theory [CTT]）及試題反應理論（Item Response Theory [IRT]）分析測驗試題外，亦利用試題訊息量找出適合高情緒風險學生施測試題。本研究發現部分在 CTT 中視為不適合施測試題，在 IRT 中被視為適合高度負向情緒填答者的試題。同時，在綜合兩測驗理論選擇試題結果，並考量身心健康評量系統建置目標：「提供一簡便釐清學生身心狀態工具協助區辨需高度關懷學生」後，建議施測「正向因應資源」十題，「負向情緒反應」十題，「多向度感知社會支持量表」七題，「自我滿意度量表」六題。綜上所述，本研究將提出身心健康評量系統建置之題目簡化建議，並提供第一線測驗實務者有效釐清學生身心狀態工具。

關鍵詞：古典測驗理論、試題反應理論、試題訊息量、試題簡化、身心健康評量系統



Item Reduction Of National Taiwan University

Physical And Mental Health Assessment System Through

Classical Test Theory And Item Response Theory



Yen-Ting Chen

Abstract

In 2020, National Taiwan University experienced a series of student self-harm and suicide incidents. In response, the university's counseling system initiated the “Physical And Mental Health Assessment System” project to better understand students' psychological states. From 2021 to 2024, the assessment system collected extensive test data, with annual evaluations and revisions of the assessment system items. Despite efforts to remove unsuitable questions over the years, the total number of items administered to students remained high. This study aims to reduce items from the "Positive Coping Resources”, “Negative Emotional Responses”, “Multidimensional Scale of Perceived Social Support” and “Self-Satisfaction Scale” by using data from National Taiwan University’s undergraduate students in 2024. In addition to analyzing test items using descriptive statistics, Classical Test Theory (CTT), and Item Response Theory (IRT), the study also conducted the item information function to identify items suitable for students at high emotional risk. The key findings were as follows: (1). Some items deemed unsuitable by CTT standards were identified by IRT as being highly effective for respondents with high levels of

negative emotions. (2). Integrating the results from both theories and considering the expecting goal of the assessment system “to provide a simple tool for clarifying students' mental health status and identifying those in need of high-level intervention”, this study recommends both administering ten items from “Positive Coping Resources”, ten from “Negative Emotional Responses”, seven from “Multidimensional Scale of Perceived Social Support” , and six from the “Self-Satisfaction Scale”. In summary, this study provides recommendations for reducing items within “Physical And Mental Health Assessment System”, offering practitioners an effective tool for assessing students' psychological status.

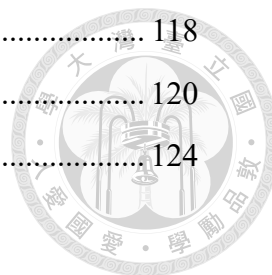
Keywords: Classical Test Theory, Item Response Theory, item information function, Item Reduction, Physical And Mental Health Assessment System

目次



誌謝.....	i
中文摘要.....	iii
英文摘要.....	v
目次.....	vii
圖次.....	ix
表次.....	xi
第一章 緒論.....	1
第二章 文獻回顧.....	3
第一節 身心適應量表相關研究及其問題	3
第二節 國立臺灣大學 2021 至 2024 年身心健康評量系統建置計畫簡介.....	3
第三節 試題簡化策略回顧	12
第四節 古典測驗理論與試題反應理論在臨床實務的研究問題	18
第三章 研究方法.....	21
第一節 計畫參與者	21
第二節 分析量表工具	21
第三節 統計分析方法	23
第四章 結果.....	33
第一節 參與者人口學變項	33
第二節 基於古典測驗理論之試題選擇	37
第三節 基於試題反應理論之試題選擇	51
第四節 綜合兩測驗理論之試題簡化結果	73
第五章 討論.....	89
第一節 試題簡化結果討論：試題層面	89
第二節 試題簡化結果討論：策略層面	94
第三節 研究限制與未來研究方向	100
參考文獻.....	103
附錄一 原有量表施測試題.....	113

附錄二	探索性因素分析因素個數判別圖：平行分析法.....	118
附錄三	本研究使用之程式操作語法.....	120
附錄四	驗證性因素分析標準化參數報表.....	124



圖次



圖 1	題間多向度與題內多向度圖示.....	29
圖 2	正向因應資源之多向度試題鑑別度參數 (MDISC) 分布圖	56
圖 3	正向因應資源之多向度試題閾值參數 (MDIFF) 分布圖	57
圖 4	負向情緒反應之多向度試題鑑別度參數 (MDISC) 分布圖	62
圖 5	負向情緒反應之多向度試題閾值參數 (MDIFF) 分布圖	63
圖 6	多向度感知社會支持量表之多向度試題鑑別度參數 (MDISC) 分布圖	67
圖 7	多向度感知社會支持量表之多向度試題閾值參數 (MDIFF) 分布圖	68
圖 8	自我滿意度之多向度試題鑑別度參數 (MDISC) 分布圖	71
圖 9	自我滿意度之多向度試題閾值參數 (MDIFF) 分布圖	72



表次



表 1	國立臺灣大學身心健康評量系統大學部學生歷年測驗使用測量工具.....	7
表 2	2024 年參與者填答人數（百分比）描述統計.....	34
表 3	2024 年參與者生理男性與生理女性人數（百分比）描述統計.....	35
表 4	2024 年參與者入學管道人數（百分比）描述統計.....	36
表 5	「正向因應資源」因素間相關係數矩陣.....	38
表 6	「正向因應資源」之古典測驗理論分析結果.....	39
表 7	「負向情緒反應」因素間相關係數矩陣.....	42
表 8	「負向情緒反應」之古典測驗理論分析結果.....	43
表 9	「多向度感知社會支持量表」因素間相關係數矩陣.....	46
表 10	「多向度感知社會支持量表」之古典測驗理論分析結果.....	47
表 11	「自我滿意度」因素間相關係數矩陣.....	49
表 12	刪題後測驗模型適配度指標表.....	49
表 13	「自我滿意度」之古典測驗理論分析結果.....	50
表 14	試題反應理論測驗模型適配度指標表.....	52
表 15	「正向因應資源」之試題反應理論分析結果.....	54
表 16	「負向情緒反應」之試題反應理論分析結果.....	59
表 17	「多向度感知社會支持量表」之試題反應理論分析結果.....	65
表 18	「自我滿意度」之試題反應理論分析結果.....	70
表 19	「正向因應資源」簡化試題篩選排序表.....	75
表 20	「負向情緒反應」簡化試題篩選排序表.....	79
表 21	「多向度感知社會支持量表」簡化試題篩選排序表.....	84
表 22	「自我滿意度」簡化試題篩選排序表.....	87
表 23	試題簡化之同時效度報表：以皮爾森相關係數為效度檢驗工具.....	97
表 24	完整量表與簡化量表間相關：以皮爾森相關係數為填答者得分一致性 檢驗工具.....	97
表 25	基於本研究試題簡化策略下的評量系統試題列表.....	98



第一章 緒論



根據臺灣衛生福利部於 2021 年 10 月公告之世界心理健康日衛生福利統計通報（臺灣衛生福利部，2021a），於 2019 年因身心困擾就診約 280 萬人，其中 25 歲以下佔比約 12.5%（約 35 萬人）；且同一年度國人身亡原因資料統計顯示，「自殺」是青壯年族群（15 歲至 24 歲）身亡原因第二名（臺灣衛生福利部，2021b）。以現有資料觀之，其定義之 15 歲至 24 歲涵括台灣高中生及部分持續就學之大專院校學生族群。

針對上述提及統計資料，各大專院校在新生入學時以量表填寫方式觀看學生的身心困擾或狀態。以國立臺灣大學為例，2014 年（含）前的新生身心困擾或狀態評估上，多數採用陳慶餘與吳英璋（1987）發展之「身心生活適應量表」；然而，後因心輔中心主任由精神科醫師擔任，2015 年起新生身心狀態評估改採用心情溫度計（Brief Symptom Rating Scale [BSRS-5]; Lee et al., 2003）的 5+1 題版本作為篩選高關懷學生的量表。此量表雖然簡易使用，但卻無法提供學生更細部資料釐清其身心問題樣貌或根源。

2020 年，國立臺灣大學接連發生多起學生自傷或自殺事件後，校方輔導體系為全面了解學生的身心適應狀態，並希望在學生不同的學習時間點能了解學生的狀況以提供必要的協助下，國立臺灣大學校方委由時任副學務長的心理學系姚開屏教授及心輔中心廖士程主任自 2021 年起著手建置「身心健康評量系統」，每學年採用多種量表讓學生進行網路施測，並依據學生填答資料評估量表是否須納入未來評量系統的建置。

過往在「身心健康評量系統」建置計畫中，透過古典測驗理論（Classical Test Theory [CTT]）之信度分析、探索性因素分析（Exploratory Factor Analysis [EFA]）與驗證性因素分析（Confirmatory Factor Analysis [CFA]）討論量表試題與身心狀態的相關，亦考量學生生理症狀（林青慧，2003）、情緒感受（Lue et al., 2006）、人際互動關係（Segrin et al., 2007）與其身心適應狀態的關係。儘管歷年建置計畫已協助捕捉部分較適合納入身心健康評量系統施測試題，並移除部分不適合提供施測的測驗或試題；然因每年仍有為完成對學生適應狀態研究

而新增新的量表，綜觀下來每年列入施測的試題題數與測驗種類仍然偏多。為便於填答者在身心健康評量系統作答，其一方法為精簡試題題數或測驗種類，但卻可能因減少試題或測驗面向而降低對學生身心狀態不同面向的了解。

為同時取得「身心健康評量系統建置期望目標：『提供校方一簡便釐清學生身心狀態的工具與協助辨別需高度介入關懷之填答者』」、「精簡試題題目以便於作答」與「釐清對學生身心狀態的不同面向」三者取得平衡，本研究除採用 CTT 一體適用所有能力群體的統計方法達成「精簡試題題目」與「釐清學生身心狀態面向」外，亦納入試題反應理論（Item Response Theory [IRT]）的試題訊息量（Item Information）作為題目簡化的指標（Thomas, 2019）。考量多數身心健康量表試題存在試題信度指標（item reliability index）過低、偏態係數（skewness）與峰度係數（kurtosis）過高情況（Reise & Waller, 2009），試題訊息量可描述特定試題在特定能力下選擇特定選項的強弱傾向，可協助身心健康評量系統的試題簡化上具有測驗理論支持與臨床實務應用意義。對此，本研究將基於 CTT 一體適用所有能力群體的統計方法，搭配 IRT 的試題訊息量作為身心健康評量系統試題選擇參考依據。

第二章 文獻回顧



第一節 身心適應量表相關研究及其問題

各大專院校針對學生身心適應狀態上，通常以量表填寫方式觀看學生身心狀態，以協助校方輔導單位進行關懷與介入。以私立東吳大學為例，每學年皆會針對大學部一年級（含進修部一年級學生）進行身心適應普查，採用前東吳大學心理學系朱錦鳳教授編製之身心適應量表進行施測（Ju & Duann, 2002），最終從學生在該量表答題反應進行身心困擾程度歸類、並安排後續心理諮商資源輔導與介入。此普查方法可有效為每位學生進行身心困擾評估，但卻耗費極大的人力與施測時間。除各院系的心理師須親自到各院系進行施測外，該身心適應量表題目長達 200 題，學生填答時無法保證其答題可有效反應其身心狀態。另，該量表自朱錦鳳教授離世後，校方針對該量表未進行再編修與測驗信效度的再檢驗，該測驗反應之身心狀態分數、及現今影響甚鉅之身心困擾變項（如：網路交友等因應科技發展之變項）可能無法忠實反映學生之身心困擾程度。

第二節 國立臺灣大學 2021 至 2024 年身心健康評量系統建置計畫 簡介

西元 2020 年 11 月，國立臺灣大學一星期內陸續出現三名學生自殺事件，後續幾個月中亦有多名學生陸續輕生。事件發生後，除針對學生反應心理輔導資源不足的部分進行補強外，亦希望進一步釐清學生在就學過程中選擇自我傷害的原因，以協助校方預防嚴重自傷事件的發生。

國立臺灣大學在臺灣教育體系中屬於頂尖大學，過往已有研究探討頂尖大學學生的心理壓力及心理健康問題，並說明頂尖大學學生存在嚴重心理健康問題原因與其因應策略（Luthar et al., 2020）。儘管該研究可說明頂尖大學學生自我傷害的部分成因，但如前述所提：協助校方預防嚴重自傷事件的發生上，相較因應策略，更著重能否在嚴重自我傷害事件發生前陪伴學生度過心理困擾。

過往國立臺灣大學 2014 年（含）前的新生身心困擾或釐清學生就學過程中的身心狀態，多數採用陳慶餘與吳英璋（1987）發展之「身心生活適應量表」。2015 年起因心輔中心主任由精神科醫師擔任，新生身心狀態評估改採用心情溫度計（Brief Symptom Rating Scale [BSRS-5]; Lee et al., 2003）作為篩選需高度介入關懷學生的量表工具。儘管心情溫度計具有簡易使用特性，但卻無法提供填答學生更細部資料，進而釐清其身心問題樣貌或根源。

有鑑於此，國立臺灣大學校方為提供一完整而有效對學生身心健康評量的系統，並希望在填答學生不同學習時間點了解學生的狀況以即時介入關懷，國立臺灣大學校方委由時任副學務長的心理學系姚開屏教授及心輔中心廖士程主任自 2021 年起著手建置「身心健康評量系統」，每學年採用多種量表讓學生進行網路施測，並依據學生填答資料評估量表是否須納入未來評量系統的建置。為達成系統建置目標：「提供一完整而有效的身心健康評量系統，並在學生不同的學習時間點能了解學生的狀況以提供必要的協助」，自 2021 年起廣納生理症狀（林青慧，2003）、情緒感受（Lue et al., 2006）、人際互動關係（Segrin et al., 2007）等與身心健康相關的變項進入系統施測外，亦針對陳慶餘與吳英璋（1987）發展的「身心生活適應量表」進行修訂，期望從各個量表所施測的心理變項分數釐清學生的身心健康狀態。

歷年來，除了上述提及對系統試題的增減外，也自歷年填答者資料討論不同身心變項對心理健康的影響關係。不論是從社會性因素層面、情緒性因素層面，或環境因素層面上皆有進行對填答者身心狀態的討論，並以多種統計策略描述多個身心變項之間的關係，協助評量系統建置上能捕捉較適合表徵填答者在學期間的身心狀況試題或測驗（如：高莉，2025；陳伊柔，2024；賴宜琳，2024；張雅筑，2023；陳禹昕，2022）。

壹、國立臺灣大學身心健康評量系統大學部施測測驗種類

本身身心健康評量系統建置第一年，針對大學部學生先行提供以下量表進行測量：「身心生活適應量表」（陳慶餘、吳英璋，1987）、「網路使用習慣量表」（Chen Internet Addiction Scale [CIAS]; Chen et al, 2003）、「家庭功能評估量表」（The McMaster Family Assessment Device [FAD]; Epstein et al., 1983）裡的一般

功能分量表 (The General Functioning 12-item subscale [GF-12])、「多向度感知社會支持量表」(Multidimensional Scale of Perceived Social Support [MSPSS]; Zimet et al., 1988)、「性格量表」(Ten-Item Personality Inventory [TIPI]; Gosling et al., 2003)、「生命意義量表」(Purpose in Life Test [PIL-SF]; Schulenberg et al., 2011)、「情緒調節自我效能量表」(Regularity Emotional Self-Efficacy Scale [RESE]; Caprara, & Gerbino, 2001)、「生活滿意度量表」(Satisfaction With Life Scale [SWLS]; Diener et al., 1985)、「WHO-5 幸福指標量表」(World Health Organization Five-item Well-Being Index [WHO-5]; World Health Organization, 1998)、「簡版完美主義與怕失敗量表」(Frost Multidimensional Perfectionism Scale & The Performance Failure Appraisal Inventory - Revised [FMPS & PFAI-R]; Frost et al., 1990; Conroy et al., 2002)、「大學生身心困擾量表」(Ju & Duann, 2002)。在第二年中，考量新一年度想研究新的議題，在測量題目及量表上有所增減，如：為能更清楚了解學生自殺的風險，在原「負向情緒反應」量表中加入了四題，以區分出不同程度的自殺風險，包括：意念、計畫、行為；另考量第一年 CIAS 中未加入網路使用風險，因此在 CIAS 增加網路使用有關的四個題目：使用網路的時間、遭受網路霸凌的情形，以及使用網路的活動種類。日後若每年持續收集資料，也將增加當年度有興趣、關心的議題，做隨時而有彈性的調整。

除此之外，第三年的建置計畫中亦根據不同量表之心理計量特性進行題目調整，如：「身心生活適應量表」中「我喜歡與朋友一同進行休閒活動。」一題屬於自我肯定對應的題目，但同時測量與朋友關係，測量多重構念，故進行刪除；「寂寞感量表」(UCLA Loneliness Scale [ULS-8]; Russell, 1996) 中，經刪除與大多數題項概念不同題目後保留五題，並維持原量表之單一因素結構。第四年的建置計畫中，除維持第三年提及：根據不同量表之心理計量特性進行題目調整外，建置計畫主持人心理學系姚開屏特聘教授亦詢問經手分析過系統建置資料的夥伴對各個量表的重要性看法。多數夥伴認為最重要的量表或題目需可快速普查填答者身心狀態，可分為「正向情緒相關」：多向度感知社會支持、正向因應資源等社會支持資源，和「負向情緒相關」：自尊量表 (Rosenberg Self-Esteem Scale [RSE]; Rosenberg, 1965)、情緒自評量表 (Depression Anxiety Stress Scales [DASS-21]; Osman et al., 2012)、負向情緒反應等自我情緒感知相關。儘

管量表重要性的排序因人而異，但多數夥伴在題目或測驗選擇上，認為「試題測驗內容重複性高」、「非上述正負向情緒相關的試題測驗」可視為較不重要的測驗或題目。

綜觀過往國立臺灣大學身心健康評量系統建置採用的測驗與題目，測驗與題目選擇上，採納生理症狀（林青慧，2003）、情緒感受（Lue et al., 2006）、人際互動關係（Segrin et al., 2007）等與身心健康相關因素。其在歷年的身心健康評量系統建置中亦有對應量表可釐清填答者在以上三個因素的得分表現。如：生理症狀部分，會在「身心症狀」量表中提供填答者勾選；情緒感受部分，則有「寂寞感」、「完美主義」、「負向情緒反應」等量表提供填答；人際互動關係部分，納入「多向度感知社會支持量表」、「正向因應資源」等量表進行填答者的人際互動狀態評估。

除上述三個因素外，身心健康評量系統建置研究期間於第四年（2024 年）納入「世界衛生組織生活品質量表八題簡版」（EUROHIS-QOL 8-item index; Schmidt et al., 2006）與「自我滿意度」提供評估填答者生活滿意程度的指標，兩測驗皆可說明填答者對生活的總體評估滿意度。

針對上述提及之歷年測驗使用工具可參考表 1。綜觀歷年施測資料，不乏發現本校於身心健康評量系統建置時皆圍繞：「生理症狀」、「情緒感受」、「人際互動關係」與「個人對生活的總體評估滿意度」，期望透過多個量表的施測，盡可能達到對填答者身心適應狀態的理解。

貳、國立臺灣大學身心健康評量系統建置之研究問題

國立臺灣大學身心健康評量系統建置計畫上，自 2021 迄今納入多種量表並採用網路施測方式進行學生身心狀況的評估。此方法存在測驗編製與調整彈性佳的優點外，填答者可自主選擇時段上網進行測驗填寫；然為建置一全面釐清學生身心狀態的評量系統，並隨著有興趣的議題及研究需要而增減不同面向的測驗，每年施測的測驗數量較多，總題目數量皆超過 200 題。儘管每年皆針對多個重要測驗或試題進行增減，但因期望建置一全面的身心健康評量系統，仍會增加新測驗並加以評估該測驗對身心狀態評估的效果。

表 1

國立臺灣大學身心健康評量系統大學部學生歷年測驗使用測量工具

量表編號名稱	施測年份與編號	評量向度、尺度 (每題分數範圍)	備註
1 大學生 身心生活 適應	2021 Q1 2022 Q1 2023 Q1 2024 Q1	30 個身心症狀 (是, 否)	
	2021 Q1-31, 32 2022 Q1-31, 32 2023 Q1-31, 32 2024 Q1-31, 32	2 題 (1-5 分)	31 健康狀況與同年 齡人比較 32 健康狀況比半年 前
	2021 Q2 2022 Q2 2023 Q2 (註 1) 2024 Q2 (註 2)	22 題 (1-5 分) 正向因應資源 向度分類： 家庭，自我，師友 資源	1. 2023 年刪除 6, 7, 11, 12 四題 2. 2024 同 2023
	2021 Q3 2022 Q3 (註 1) 2023 Q3 (註 2) 2024 Q3 (註 3)	24 題 (1-5 分) 負向情緒反應 向度分類 1： 負向情緒反應，情緒 焦慮、憂鬱無助 向度分類 2： 身體化症狀，情緒障 礙，神經質症狀，低 信心，無望感	1. 加 3 題自殺題 目與 1 題對未 來無望 2. 刪除 13 並修正 27:我過去曾經 嘗試過自殺 (是, 否) 3. 2024 同 2023
	2021 Q4	12 個感興趣的活動主 題 (勾選)	
	2021 (註 1) 2022 (註 1) 2023 (註 1) 2024 (註 1)	5+1 題 (0-4 分)	1. 由歷年 Q3 內試 題轉換而得
2 類心情溫度計			

表 1 (續)

國立臺灣大學身心健康評量系統大學部學生歷年測驗使用測量工具

量表編號名稱	施測年份與編號	評量向度、尺度 (每題分數範圍)	備註
3 網路使用 習慣	2021 Q5 2022 Q4 (註 1) 2023 Q4 (註 2) 2024 Q4 (註 3)	10 題 (1-4 分) 向度分類： 網路沉迷核心症狀 (含戒斷症狀，耐受 症狀)，網路沉迷相 關問題 (含時間管 理，人際與健康問 題)	1. 加 2 題網路霸 凌、網路使用的 時間與使用網路 的活動種類共 4 題 2. 去除 2 題網路霸 凌與使用網路的 活動種類共 3 題，並調整使用 時間問法 3. 使用時間分兩小 題
4 家庭 一般功能	2021 Q6	12 題 (1-4 分)	
5 多向度感知 社會支持	2021 Q7 2022 Q5 2023 Q5 2024 Q5	12 題 (1-7 分) 向度分類： 重要他人，家庭，朋 友	
6 性格	2021 Q8	10 題 (1-5 分) 向度分類： 外向性，親和性，審 慎嚴謹性，情緒穩定 性，經驗開放性	
7 生命意義	2021 Q9 2024 Q16	4 題 (1-5 分)	
8 情緒調節 自我效能	2021 Q10 2022 Q6 2023 Q6	12 題 (1-5 分) 向度分類： 正向情緒，沮喪痛 苦，憤怒情緒	
9 生活滿意度	2021 Q11 2022 Q7 2023 Q7 2024 Q6	5 題 (1-7 分)	
10 WHO-5 幸福指標	2021 Q11 2022 Q8 2023 Q8	5 題 (0-5 分)	

表 1 (續)

國立臺灣大學身心健康評量系統大學部學生歷年測驗使用測量工具

量表編號名稱	施測年份與編號	評量向度、尺度 (每題分數範圍)	備註
11 完美主義簡版	2021 Q12 2022 Q9 (註 1) 2023 Q14 (註 2) 2024 Q14 (註 3)	7 題 (1-5 分)	1. 置換為 35 題長版 完美主義量表 2. 刪除 6, 7, 12, 13, 16, 17, 20, 22, 27, 30 置換為簡版多層面 完美主義量表 8 題
12 害怕失敗簡版	2021 Q12	7 題 (1-5 分)	
13 大學生身心 生活困擾	2021 Q13	每向度各 8 題 (1-5 分) 向度分類： 時間管理，生涯，學 習，情緒，人際關 係，感情，家庭	
14 心理韌性	2022 Q10 2023 Q9 2024 Q7	10 題 (1-5 分)	
15 寂寞感	2022 Q11 2023 Q10 (註 1) 2024 Q8 (註 2)	8 題 (1-4 分)	1. 刪除 1-3 共 3 題 2. 2024 同 2023
16 拖延	2022 Q12 2023 Q11 (註 1)	16 題 (1-4 分)	1. 刪除 7, 9, 10, 12, 13 共 5 題
17 自尊	2022 Q13 2023 Q12 2024 Q9	10 題 (1-4 分)	
16 拖延	2022 Q12 2023 Q11 (註 1)	16 題 (1-4 分)	1. 刪除 7, 9, 10, 12, 13 共 5 題
17 自尊	2022 Q13 2023 Q12 2024 Q9	10 題 (1-4 分)	
18 克蘭斯冒牌者 現象	2022 Q14	20 題 (1-5 分)	
19 害怕負面 評價	2023 Q13 (註 1) 2024 Q10 (註 2)	12 題 (1-5 分)	1. 刪除 2, 4, 7, 10 共 4 題 2. 2024 同 2023
20 安適幸福感	2023 Q15 2024 Q12	7 題 (1-5 分)	

表 1 (續)

國立臺灣大學身心健康評量系統大學部學生歷年測驗使用測量工具

量表編號名稱	施測年份與編號	評量向度、尺度 (每題分數範圍)	備註
21 情緒自評	2023 Q16 (註 1) 2024 Q13 (註 2)	9 題 (0-3 分) 向度分類： 憂鬱，焦慮，壓力	1. 為 Depression Anxiety Stress Scales [DASS- 21] (Osman et al., 2012)，刪除 1-4, 8, 12, 13, 17- 21 2. 2024 同 2023
22 睡眠品質	2023 Q17	4 題 (1-5 分)	
23 害怕失敗	2023 Q18	3 題 (1-5 分)	
24 自我慈憫	2024 Q14	26 題 (1-5 分) 向度分類： 自我關愛，普同人性，正念覺察，自我批判，自我隔絕，負面經驗過度認同	
25 世界衛生組織 生活品質量表 八題簡版	2024 Q15 (註 1)	8 題 (1-5 分)	1. 為 EUROHIS- QOL 8-item index
26 主客觀 學業成就	2024 Q17 (註 1)	5+1 題 (1-5 分)	1. 含 GPA
27 自我滿意度	2024 Q18 (註 1)	15+1 題 (1-5 分)	1. 第 16 題分成 a, b 小題
28 對心理疾病 的經驗	2024 Q19	10 題 (是，否)	
29 生活壓力指標	2024 Q20	4 題 (0-4 分)	

為捕捉較佳評量學生身心狀態的試題，歷年計劃在試題捕捉上採用古典測驗理論（Classical Test Theory [CTT]）之信度分析、探索性因素分析（Exploratory Factor Analysis [EFA]）與驗證性因素分析（Confirmatory Factor Analysis [CFA]）協助分析與判准試題優劣，如：針對 Cronbach's alpha 過低之測驗或試題、探索性因素分析後因素負荷量過低的試題（如：因素負荷量小於 .4）在評量系統中將被視為不佳試題。反之，其餘試題將會繼續保留施測，持續收集學生填答反應加以分析。

國立臺灣大學身心健康評量系統建置為一個全面評量系統的發展，除提供給一般學生進行自我評估外，亦希望找出身心狀態不穩定的學生加以關懷。為捕捉適合全體學生施測的試題，除透過 CTT 等傳統測驗與統計理論找出適合一般學生自我評估的試題外，也加入試題反應理論（Item Response Theory [IRT]）的試題訊息量（Item Information）作為題目選擇的指標（Thomas, 2019）。試題訊息量數值可表徵一試題在特定能力下的能力估計精確性，為協助找出適合填答者的施測試題工具之一。試題訊息量的數值呈現可協助身心狀態不穩定的同學攫取適合的試題提供施測，如在特定身心狀態下某試題具有最大試題訊息量時，表示某試題在特定身心狀態下較可反映具有特定身心狀態參與者的心身狀態。

除此之外，因應國立臺灣大學健康評量系統建置研究計畫案期望從「生理症狀」、「情緒感受」、「人際互動關係」與「個人對生活的總體評估滿意度」向度評估學生的身心健康狀態，本研究將針對歷年施測且因素結構穩定，但題目多於 10 題的多分題測驗：編號 1「大學生身心生活適應」中的「正向因應資源」分量表、編號 1「大學生身心生活適應」中的「負向情緒反應」分量表、編號 5「多向度感知社會支持量表」與今年新納入之編號 27「自我滿意度」進行分析。

第三節 試題簡化策略回顧

國立臺灣大學身心健康評量系統建置上，採用多個已發展的量表與加入欲研究討論的新試題或測驗進行試題選擇。選擇試題的目的為找尋符合欲測量的身心變項理論架構且可由收集到的答題資料反映最終欲測量的身心變項理論樣貌。因系統建置初期為廣納多個與學生身心健康的已發展量表進入施測，為捕捉更適合提供學生進行簡便施測的試題，則須自納入題庫試題或測驗再進行一次試題選擇，基於已發展量表再進行試題選擇的狀態為試題簡化：自己發展測驗或量表中捕捉更適合提供施測群體的試題。綜合上述所及，國立臺灣大學身心健康評量系統建置上，欲從多個已發展的量表捕捉更適合提供施測群體的試題。相較傳統的試題選擇，系統建置上為多個已發展量表的試題簡化。

儘管試題簡化為從多個已發展的量表捕捉更適合提供施測群體的試題，然因其為將納入題庫試題或測驗再進行一次試題選擇，故試題簡化與試題選擇上多數採取相同方法或策略。換言之，編製測驗後需挑選適合施測試題，為試題選擇；從多個適合施測試題中挑選更能捕捉欲測量變項樣貌的試題，為試題簡化，相較前者設定較寬鬆的試題選擇標準，後者則是將試題選擇的標準提高以取得更適合施測試題。對此，以下文獻回顧中將會著重在「選擇適合施測試題」角度進行回顧，以達成身心健康評量系統建置目標：自既有試題或測驗中捕捉更適合提供施測群體的試題。

壹、以古典測驗理論為基礎之試題選擇

過往心理測驗的編製上，部分測驗以古典測驗理論為基礎進行測驗編製，並於編製過程中提供試題分析的判別標準。Kaplan 與 Saccuzzo (2018) 將身心測驗的編製與試題選擇策略分為演繹策略與實務策略。演繹策略說明試題選擇和編製須具有欲測量的身心變項理論基礎，為理論導向，如一人格測驗採用一人格理論，該人格理論說明人格可分為六大向度時，試題須以六大向度的人格理論為基礎進行選擇和編製；實務策略說明試題選擇和編製是由收集到的答題資料進行試題分析，以評估最終欲測量的身心變項理論樣貌，為資料導向，如欲編製一人格測驗，則會先基於所有參與者在所有試題的答題反應分析，針對

具有欲討論之身心變項的研究群體進行試題選擇後，輔以因素分析討論被選擇試題之所屬因素。

部分測驗編製者合併演繹策略與實務策略進行測驗編製與試題選擇 (Gregory, 2015; Kaplan & Saccuzzo, 2018)。以態度測驗為例，臺灣國中生數學知識信念量表 (Lai & Wu, 2018) 在預試階段採用演繹策略：根據知識信念態度進行試題編寫後，後續採用實務策略進行與試題選擇：校正後試題總分相關係數、決斷值分析、探索性因素分析與 Cronbach's alpha 分析其自編臺灣國中生數學知識信念量表的試題品質；正式量表階段，除採用 Cronbach's alpha 進行信度分析外，亦採用驗證性因素分析、區辨效度、聚斂效度、效標關聯效度與複核效度進行效度驗證。然該研究針對預試階段雖提供完整試題分析數據，但未提供其保留試題標準，僅在效度分析中提供模型適配度的評鑑結果與標準。

自上述例子中，可知在個人態度測驗的試題選擇可採用校正後試題總分相關係數、決斷值分析、探索性因素分析與 Cronbach's alpha 分析，也採用驗證性因素分析與多個效度驗證策略進行效度驗證；然而因其為態度測驗，國立臺灣大學身心健康評量系統屬身心評量測驗，上述提及的試題選擇策略或標準仍需要其餘文獻提供支持，以協助後續身心健康評量系統的試題選擇。

在身心或臨床測驗中，針對身心健康評量系統中提及對填答者的「情緒感受」與「人際互動關係」的測量上，過往在社會情緒感知量表 (Social-Emotional Expertise [SEE]; McBrien et al., 2020) 編製時，採用七點李克特氏量尺。預試階段根據社會情緒感知理論進行試題編寫後，再使用試題平均數、校正後試題總分相關、Cronbach's alpha 與試題間相關進行試題篩選，其採用的方法為 Clark 與 Watson (1995) 的建議。初步刪題完畢後，則依序進行探索性因素分析、區辨效度與聚斂效度檢驗。該研究除試題分析數據外，亦提供其保留試題的判別標準，如：答題平均數大於 5.5 者將會刪除、試題總分相關低於 .3 者會進行刪除。

而在「情緒感受」相關的臨床測驗上，廣泛性焦慮量表 (Generalized Anxiety Disorder [GAD-7]; Spitzer et al., 2006) 編製上，以廣泛型焦慮的定義進行演繹策略編製後，再使用校正後試題總分相關、Cronbach's alpha 與區辨效度確認 7 題的品質。儘管該研究提供證據支持施測七題與施測完整 13 題題庫在接

收者操作特性曲線（receiver operating characteristic [ROC]）曲線下能取得最佳平衡，但未針對校正後試題總分相關、Cronbach's alpha 與區辨效度的判別標準進一步提供說明。

上述兩測驗的試題選擇方法上，除採用前述態度測驗使用之校正後試題總分相關係數、探索性因素分析與、Cronbach's alpha 分析、驗證性因素分析與多個效度驗證策略進行效度驗證外，SEE 加入試題平均數作為題目選擇標準，GAD-7 另提供 ROC 曲線作為題目選擇標準。對此，身心健康評量系統選擇適合施測試題上，可基於上述三種測驗提及選題方法進行試題選擇。Gregory（2015）、Kaplan 與 Saccuzzo（2018）皆在其著作中提到針對身心臨床測驗或態度測驗的編製流程及試題選擇策略上具有共同性（如：皆可採用試題平均數、探索性或驗證性因素分析等方法進行），儘管具有共同策略，但其最終試題選擇的標準上仍需更多實證研究進行輔助。

貳、以試題反應理論為基礎之試題選擇

試題反應理論（Item Response Theory [IRT]）為一個新型測驗理論，相較 CTT 假設測量標準誤（Standard Error of Measurement [SEM]）跨能力間相等，IRT 認為 SEM 隨著填答者的能力不同而有所改變。在身心議題量表上，自 1990 年起即有學者將多向度人格量表（Multidimensional Personality Questionnaire [MPQ]; Tellegen, 1982）使用 IRT 二參數模型進行分析，並指出 IRT 可用來描述二分題型的人格試題反應（Reise & Waller, 1990）。後續亦有針對其餘身心量表使用 IRT 進行量表分析的研究，如：邁爾斯-布里格斯類型指標（Meyers-Briggs Type Indicator [MBTI]; Harvey & Murry, 1994），貝克憂鬱量表（Beck Depression Inventory [BDI]; Santor et al., 1994）。

針對 IRT 二參數模型在臨床測驗的使用上，Gray-Little 等人（1997）使用等級評分模式（Graded Response Model [GRM]; Samejima, 1969）討論羅森伯格自尊量表（Rosenberg Self-Esteem Scale [RSE]; Rosenberg, 1965）的心理計量特性，並得到 GRM 可有效描述填答者在自尊量表上答題反應的結論；然因二參數模式存在鑑別度參數與閾值參數（Threshold）共同影響答題反應，其測驗數值的解釋和群體比較上較單參數模式複雜（Embretson & Reise, 2000）。進一步

而言，單參數模式因只受階難度參數（step-difficulty）影響下，測驗分數上可透過勝算比（Odds Ratio [OR]）進行轉換與比較，二參數模式無法單由勝算比轉換進行轉換與比較。

對身心健康評量系統中提及對填答者的「情緒感受」與「人際互動關係」的測量上，Beesdo-Baum 等人（2012）在焦慮自評量表編製先採用單一因素驗證性因素分析（1-factor confirmatory factor analysis [1-factor CFA]）驗證所有試題的向度性（dimensionality），並使用部分給分模式（Partial Credit Model [PCM]; Muraki, 1990）進行 IRT 試題分析。結果顯示所有試題皆存在鑑別參數大於 1 與階難度參數大於 0 的情況。後續亦針對部分試題的試題反應曲線（Item response curve [IRC]）進行繪製，並發現部分試題的階難度參數出現逆序，反應選項的難易度並非以順序尺度自小到大排列。該量表的編製上，並未提供因素分析的模型適配度標準、試題階難度與鑑別度的判別標準外，在 IRC 的報告上僅使用視覺化方式呈現結果；然該研究提及試題訊息量與測量焦慮的關係：部分試題在極度焦慮個案上存在高訊息量數值，說明部分試題對極度焦慮的測量誤差較低，為少數將試題訊息量納入臨床量表編製的文獻。

Walter 等人（2007）在焦慮電腦化適性測驗編製上，其在測量焦慮的題庫選擇中，除針對所有題目進行 1-factor CFA 外，亦加入試題間殘差相關（residual correlation）檢驗，後續則採用廣義部分給分模式（Generalized Partial Credit Model [GPCM]; Muraki, 1992, 1997）進行 IRT 試題分析。在 1-factor CFA 中，當試題負荷量小於 .4 者將刪除；試題間殘差相關大於 .25 者將刪除；試題鑑別度參數小於 .8 者將刪除。此研究給予向度性檢驗的驗證方式與其標準，亦提供試題鑑別度參數的標準；然而在試題類別曲線（Item category curves [ICCs]）分析中，仍採用視覺化方式觀察試題階難度排序是否出現逆序情況。

Reeve 等人（2007）在健康相關之生活品質量表編製中，針對 IRT 的向度性假設與局部獨立性（Local Independence）假設分別採用 1-factor CFA 與試題間殘差相關檢驗。相較上述兩篇研究僅採用 1-factor CFA 檢驗單一向度，本篇研究在 CFA 模型適配度較差時，會再採用探索性因素分析（Exploratory Factor Analysis [EFA]）分析其向度個數。後續在 IRT 模型使用上，採用二參數 GRM

進行分析，並針對試題與 IRT 模型進行適配度檢驗。最後如前述文獻，ICCs 分析中，採用視覺化方式觀察試題閾值排序是否以順序尺度从小到大排列。儘管本篇文獻在 IRT 分析中提供較其餘兩文獻完整的流程，但卻未提供完整的標準參考，未來以 IRT 進行試題檢驗或簡化時可參考本篇文獻提供流程，但在試題檢驗標準上仍需輔以其他文獻支持。

參、綜合兩測驗理論之試題選擇

近年身心或臨床量表編製研究中，相較單由因素分析進行的試題分析方法，部分研究者會將 CTT 與 IRT 共同加入進行量表選題或編製（Thomas, 2019）。

簡版慢性疾病生活品質量表（Leng & Yao, 2024）中，自原本 26 題簡化成四個範疇、每個範疇有三題的簡版測驗。除採用古典測驗理論之校正後試題總分相關進行初步篩選外，在試題反應理論階段採用多向度部分給分模式（Multidimensional Partial Credit Model [MPCM]; Kelderman, 1996）討論試題與其適配性。最終計算每一題在每一向度上的試題訊息量總和，以協助找尋測量精確性較高的試題。其中，校正後試題總分相關小於 .4、模型適配度指標 *infit* 或 *outfit* 超出 .6 至 1.4 的範圍、試題訊息量偏低者將會被視為不適合納入量表的試題。本篇研究在 IRT 模型適配度上採用 *infit* 與 *outfit* 指標，前者為試題訊息量加權適合度均方：填答者的能力與試題階難度匹配時數值較大；後者為未加權適合度均方：填答者能力與試題階難度不匹配時數值較大。同時，本篇研究為少數納入試題訊息量進行題目選擇的研究，其結果亦反映出最終保留的試題具有高訊息量數值，說明試題訊息量在題目簡化或篩選流程下的可參考性。

Li 等人（2025）在新版護理學生文化容忍與敏感度測驗中，其簡化測驗的核心為先採用古典測驗理論後再行使用試題反應理論。古典測驗理論中，鑑別度考量得分前後各 27% 參與者進行獨立樣本 *t* 檢定，內部一致性信度採用 Cronbach's α ，最後採用一因素驗證性因素分析檢驗試題與文化容忍與敏感度間的因素結構。試題反應理論中，針對局部獨立性假設使用試題間殘差相關，IRT 模型適配度上，採用 *infit* 與 *outfit* 指標驗證試題與單向度單參數模型的適配程度，並呈現試題選項曲線進行視覺化判別。在鑑別度指標 CR 小於 3、

Cronbach's alpha 小於 .8、一因素驗證性因素分析的因素負荷量小於 .5、試題間殘差相關小於 .3、模型適配度指標 infit 或 outfit 超出 .6 至 1.4 的範圍者將會進行刪除。相較前篇文獻，其納入試題鑑別度與驗證試題與因素間的關係，且加入試題選項曲線作為題目選擇參考。



肆、測驗試題的選擇與簡化

上述文獻回顧為針對測驗中較差的試題進行選擇並排除，採用的標準多為過往相同領域的測驗編製標準；然而，當一份測驗已存在良好的測驗理論基礎（如：具有穩定的因素結構）下，需要進行題目選取以簡化測驗時，需將題目選擇的標準提高，以利找出測驗中更好的試題。

以憂鬱—焦慮—壓力量表（Depression Anxiety Stress Scales [DASS]; Lovibond & Lovibond, 1995）為例，原始版本為 42 題。Osman 等人（2012）在 DASS 量表簡化至 21 題的策略上採用保留試題總分相關較高者作為簡化版本，並命名其為 DASS-21。其研究結果亦發現 21 題版本存在較 42 題版本高的測量因素解釋力。Lee 等人（2019）在 21 題版本簡化至 12 題版本的策略上，與傳統試題選擇方法相同（如：利用因素負荷量做為選題標準），但其在試題總分相關、鑑別度分析與 Cronbach's alpha 皆採用較高標準。儘管其研究發現兩版本皆具有高測驗信度與效度，但簡化的 12 題版本因題數較短，臨床使用上較 21 題版本便捷。

綜合上述例子與說明，不論是古典測驗理論或試題反應理論的試題選擇標準上，在試題簡化情境下會需要提高題目選擇標準。當一測驗中研究者需要簡化的題目數量越多，題目選擇標準會越嚴格。

儘管如此，本研究在進行試題簡化之前，仍會先觀察欲分析的測驗是否具有有良好的測驗理論基礎，故會先採用傳統試題選擇的策略標準來進行評估。其後的簡化程序上，則會以試題訊息量進行試題排序。當今天評量系統需要較多測驗試題時，即可放寬試題訊息量的數值大小標準，以提供較多試題進行施測。

第四節 古典測驗理論與試題反應理論在臨床實務的研究問題

傳統古典測驗理論與試題反應理論中，在試題的優劣評估上會以能區辨個體、且保留難易適中的試題進行施測；然在臨床實務中，以身心狀態量表為例，多數填答試題的參與者在填答選項上偏向選擇最高點數或最低點數（Embretson & Reise, 2000; Gregory, 2015; Kaplan & Saccuzzo, 2018; Reise & Waller, 2009; Thomas, 2019），如：「我有自殺的想法」中，存在自殺意念或具有自殺行為的個案，傾向選擇「非常同意」，然對於一般身心狀態較穩定的群體而言，則傾向選擇「非常不同意」。此時在參與者答題分布上存在高度偏態，以傳統古典測驗理論或試題反應理論的試題保留或優劣的評判標準中，該題將不會納入最後的量表施測。

Gregory (2015) 提到身心行為量表的編製上需考量臨床治療實務者需求：以一快速且簡便的方式協助臨床治療者與個案評估其身心狀態；同時，身心行為量表或系統的建置上，相較傳統測量工具需因應測驗理論的基本假設及因素結構，更著重在簡易且直觀辨別填答者的身心狀態，以輔助臨床治療實務者後續的治療計劃調整。Clark (1988) 亦提到部分身心行為量表儘管符合測驗理論的基本假設及穩定的因素結構，但卻未提出相對應的同時效度和區辨效度支持。

Reise 與 Waller (2009) 建議身心健康量表編製上，除利用 CTT 與 IRT 協助觀察參與者在試題上的答題情況外，也需加入臨床治療實務者需求。以上述：「我有自殺的想法」為例，該題具有指標意義，可初步捕捉需初步介入關懷的個案學生，儘管該試題在古典測驗理論或試題反應理論上不具有其統計學上的證據支持其保留在題庫中，但卻可有效協助臨床實務者達成提早介入關懷的測驗目的。同時，作者也提到試題訊息量可做為找出試題適合高能力者與低能力者施測的工具，進一步而言，當試題具有高試題訊息量數值時，說明試題可有效反映特定能力填答者的狀態，如：診斷為重度憂鬱個案在「我有自殺的想法」試題中具有較非憂鬱個案高的試題訊息量數值，其反應此題在重度憂鬱個案上較能捕捉其憂鬱程度；儘管此題以傳統選擇試題優劣的標準上為高偏態試題，但考量其臨床實務的應用上仍有其必要保留性。

綜上所述，本研究為協助臨床實務者達成提早介入關懷的目的、並期望以客觀測驗理論指標協助簡化試題，後續將以 CTT 輔以 IRT 的試題訊息量協助試題簡化，透過前述眾多文獻提及之選擇題目的標準與策略，進一步提供身心健康評量系統中試題簡化的建議與評判標準，以期簡化後試題具有測驗理論基礎支持及臨床實務的可應用性。



第三章 研究方法



第一節 計畫參與者

本研究將使用國立臺灣大學 2024 年大學部身心健康評量系統建置計畫（行為與社會科學倫委會審查案號：202106HS019）之參與者填寫資料，其篩選參與者的條件如下：年齡介於 17 歲至 25 歲、需為國立臺灣大學大學部在學學生、所有測驗試題並未存在遺漏值。符合上述條件之參與者共計 3838 位年齡介於 17 歲至 25 歲間的大學部在學學生。為保障測驗參與者於測驗填寫前、中、後的研究參與權益，所有參與者進行測驗填答前皆須瀏覽知情同意書，完成簽署同意後方進行國立臺灣大學 2024 年大學部身心健康評量系統的資料填寫。

第二節 分析量表工具

國立臺灣大學 2024 年大學部身心健康評量系統的填寫資料中，本研究將採用的量表分述如下：

壹、「大學生身心生活適應量表」(陳慶餘、吳英璋，1987)

全量表原先包含三大部份，依序為「身心症狀」、「自我與社會資源」及「心理症狀」。目前採用之編修量表為 2009 年修訂版本，分別為「身心症狀」、「正向因應資源」及「負向情緒反應」。兩個版本皆以李克特式五點量尺：「從未如此」、「很少如此」、「有時如此」、「經常如此」及「總是如此」進行施測。本研究中，將針對「正向因應資源」及「負向情緒反應」進行討論。

一、「正向因應資源」

原先量表共 28 題，包含三個因素：「師友資源」、「家庭資源」及「自我資源」。本研究自 2021 年至 2023 年身心健康評量系統建置計畫相關研究進行題目編修與刪減共計六題後，現今施測版本為 22 題。其題目陳述如：「我滿意老師、同學或朋友關懷我情緒的方式」，「當我嘗試新事物時，我會感到家人的接受與支持」，完整題目可參考附錄 1 表 1-1。

二、「負向情緒反應」

原先量表共 26 題，包含五個因素：「情緒障礙」、「低自信」、「身體化症狀」、「神經質症狀」及「無望感」。除於 2022 年身心健康評量系統建置計畫相關研究額外加入與自殺相關題目與對未來無望感題目，並自 2021 至 2023 年間進行編修與刪減後，現今施測版本為 24 題版本，包含「負向情緒反應」、「情緒焦慮、憂鬱無助」兩大因素。其題目陳述如：「我會有一段時間心情低落、憂鬱、鬱悶」，「我覺得人生很無望，對事物皆失去興趣」，完整題目可參考附錄 1 表 1-2。

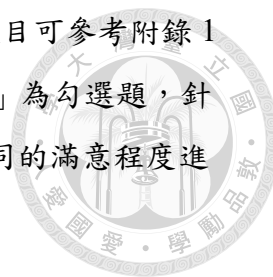
貳、多向度感知社會支持量表 (Multidimensional Scale of Perceived Social Support [MSPSS]; Zimet et al., 1988)

用於測量從家人、朋友及重要他人身上感知到的社會支持程度，並以大學生為施測對象。共計 12 題，以李克特氏七點量尺：「非常強烈不同意」、「強烈不同意」、「有些不同意」、「無意見」、「有些同意」、「強烈同意」、「非常強烈同意」進行測量。其題目陳述如：「我生命中有個特別的人會關心我的感受」，「我的家人真的會試圖幫助我」，完整題目可參考附錄 1 表 1-3。根據 Zimet 等人 (1988) 研究指出，此量表總分的內部一致性信度為 .88，重要他人、家人及朋友三個向度的內部一致性信度分別為 .91、.87 及 .85。在建構效度上，MSPSS 的三個向度與身心症狀自評量表 (Hopkins Symptom Checklist [HSCL]; Syed et al., 2008) 的憂鬱、焦慮有顯著的相關性，家人的社會支持與憂鬱 ($r = -.24, p < .01$)、焦慮 ($r = -.18, p < .01$) 呈負相關；朋友的社會支持與憂鬱 ($r = -.24, p < .01$) 呈負相關；重要他人的支持與憂鬱 ($r = -.13, p < .05$)、焦慮 ($r = -.25, p < .01$) 呈負相關。

參、自我滿意度

為 2024 身心健康評量系統建置計畫新加入的量表。由過往 2021 年至 2023 年提供施測的身心健康評量系統題庫針對不同可能影響「對填答者自己的滿意度」的變項進行詢問，共計 16 題以李克特氏五點量尺：「非常不滿意」、「不滿意」、「普通」、「滿意」、「非常滿意」進行測量。其題目陳述如：「我對於自己的

個人特質感到...」，「我對於自己的身心健康感到...」，完整題目可參考附錄 1 表 1-4。其中第十六題分為兩個子題：「我傾向認同的性別為：」為勾選題，針對填答者的性別認同進行詢問外，也針對環境支持自身性別認同的滿意程度進行詢問：「我對於環境對我自己性別認同的支持感到...」。



第三節 統計分析方法

本研究目的為利用古典測驗理論 (Classical Test Theory [CTT]) 與試題反應理論 (Item Response Theory [IRT]) 進行國立臺灣大學 2024 年大學部身心健康評量系統的題目選擇。分析過程中，將使用 R 語言的 lavaan (Rosseel, 2012)、psych (Revelle, 2025)、psychometric (Fletcher, 2023)、mirt (Chalmers, 2012)、WrightMap (Irribarra & Freund, 2014) 套件與自行撰寫程式碼進行。詳細程式碼可參考附錄三。試題篩選的統計分析方法與步驟如下：

壹、 人口學變項描述統計

起初，針對現有填答人數進行人口學變項的人數分布與其比例，如：生理性別人數與比例、學生之所屬學院人數與比例、學生在不同入學管道的人數與比例。針對上述統計資料，將以 2024 年參與者填答人數比例和 2023 年國立臺灣大學統計學報 (National Taiwan University, 2023) 之學生統計資料進行比較，以觀察填答人數比例和全校學生的人數比例是否存在較大落差。

貳、 試題描述統計

試題描述統計上，在古典測驗理論中，將會呈現試題平均數、試題標準差、試題偏態與峰度係數。其中，試題平均數定義為試題難度；在試題反應理論中，會根據測量模型具有的參數進行試題各向度鑑別度參數與試題閾值（或階難度）參數。其中，閾值（或階難度）參數為填答者自特定選項和前一選項的能力交界值。

參、古典測驗理論 (Classical Test Theory, CTT)

針對 CTT，依序將編號 1「大學生身心生活適應」中的「正向因應資源」分量表、編號 1「大學生身心生活適應」中的「負向情緒反應」分量表、編號 5「多向度感知社會支持量表」與今年新納入之編號 27「自我滿意度」進行試題鑑別度分析、探索性因素分析、內部一致性信度分析與驗證性因素分析。

一、試題鑑別度分析

本研究將此相關係數列在第一步驟進行試題評估。試題鑑別度分析上採用校正後試題總分相關係數：為排除特定試題後各試題與其所有試題得分加總間的皮爾森相關係數協助進行鑑別度分析，其數理表示如下：

$$\text{cor}\left(\left(\sum TS_{(X-i)} - TS_i\right), TS_i\right) \quad (1)$$

其中， $\sum TS_{(X-i)}$ 代表向度總分 X 扣除第 i 題的試題總分， TS_i 代表第 i 題的試題得分。Murphy 與 Davidshofer (2004) 在其著作說明部分測驗發展者使用特定的理論體系引導撰寫題目，將該理論中的觀點轉譯成測驗試題。對此，因測驗試題編寫上依照欲測量變項背後之理論體系進行撰寫，以利試題測量所屬的向度，故此相關係數的計算會以「編寫測驗試題預先設定試題所屬的向度總分」進行討論。

該相關係數可表示單一試題與測驗總分之間的相關程度，如該測驗總分為一個有效之測驗效標，該試題的得分將與測驗總分存在高度相關；反之，該試題的得分高低與測驗總分存在負相關或無相關時，除說明該試題與總分之間的相關性外，也因得分高低與欲測量變項的分數存在負相關或無相關，該試題並無法協助鑑別欲測量變項的分數高低。Piedmont (2014) 指出，當校正後試題總分相關低於 .4 時，試題將被視為不適合施測試題，其標準較 McBrien 等人 (2018) 與 Lai 等人 (2018) 使用的 .3 嚴苛；然因本研究目標為試題簡化，欲選擇較適合填答者的試題上，採用相對嚴格的標準可捕捉較佳填答試題。

二、探索性因素分析 (Exploratory Factor Analysis, EFA)

因素分析 (Factor Analysis) 為討論可觀察變項 (如：試題) 與研究者所欲測量心理變項的關係。本研究將進行探索式因素分析，觀察測驗中所有試題的因素負荷量，以辨別試題對特定欲測量因素的關係強弱。EFA 部分，為採用最大概似法 (Maximum likelihood) 進行因素選取，平行分析 (Parallel Analysis [PA]; Horn, 1965) 為因素選擇個數策略，並搭配 PROMAX 斜交轉軸進行最終因素負荷量的呈現。其結果判准上，Gregory (2015) 指出因素負荷量低於 .4 時，試題將被視為不適合施測試題，其標準與 McBrien 等人 (2018) 相同，但低於 Spitzer 等人 (2006) 的 .58；然而 Spitzer 等人 (2006) 的因素負荷量為試題選擇完畢後得到的因素負荷量，並非選擇試題使用的標準，故本研究仍維持 Gregory (2015) 與 McBrien 等人 (2018) 建議使用大於 .4 作為試題保留標準。

三、內部一致性信度分析

本階段為判別各測驗經過探索性因素分析後，各測驗與其試題的內部一致性。對此，除了上述提及校正後試題總分相關表示單一試題與測驗總分之間的相關程度外，亦提供 Cronbach's alpha 做為內部一致性信度指標。Cronbach's alpha 為所有折半測驗可能性之信度 (Split-half reliability) 平均數，其數學關係式如下：

$$\text{Cronbach's } \alpha = \left(\frac{N}{N-1} \right) \left(\frac{\sigma^2 - \sum \sigma_i^2}{\sigma^2} \right) \quad (2)$$

其中， N 為測驗所具有的試題總數， σ^2 為測驗總分的變異數， $\sum \sigma_i^2$ 則為試題各自得分的變異數總和。相較計算單一折半測驗信度，Cronbach's alpha 提供全面的信度指標，避免單一折半測驗信度指標造成錯誤判別測驗信度高低。針對 Cronbach's alpha，Cortina (1993) 在其著作中說明當 Cronbach's alpha 大於 .7 時為可接受之內部一致性信度，且文獻回顧之研究中的試題在內部一致性信度係數上皆大於 .7，本研究在試題評估選擇上將以需大於 .7 作為試題保留的標準。

另本研究亦提供排除特定試題後的 Cronbach's alpha，其數值表明單一試題對內部一致性係數的貢獻。具體而言，當排除特定試題之內部一致性係數高於未排除之內部一致性係數時，說明該試題移除後可增加測驗中特定向度的施測試題一致性；反之，如未高於未排除之內部一致性係數，則該試題不需進行移除。

四、 驗證性因素分析 (Confirmatory Factor Analysis, CFA)

CFA 部分，則為進一步確認試題與潛在變項間存在的模型適配程度。在模型參數估計與適配指標計算上，將使用 50% 的樣本進行參數估計、50% 的樣本進行適配指標計算。本研究採用的模型適配度指標與其標準分別為：卡方統計量 (Chi-square) 需未達統計上顯著、標準化殘差均方根 (Standardized Root Mean Square Residual [SRMR]) 小於 .08、可比級適配度指標 (Comparative Fit Index [CFI]) 大於 .95、近似誤差均方根 (Root Mean Square Error of Approximation [RMSEA]) 小於 .06 與非規範擬合指標 (Tucker-Lewis index [TLI]) 大於 .95 (Hu & Bentler, 1999)；然而，McDonald 與 Ho (2002) 認為在實務研究上，RMSEA 在 .08 內可視為模型良好適配，Hu 與 Bentler (1999) 亦提供實務研究上的擬合指標寬鬆標準：CFI 在 .9 以上視為模型良好適配。

除此之外，多向度試題反應理論分析方法上同連續變項之驗證性因素分析 (McDonald, 1999)，本研究將在 EFA 確認試題所屬向度後，利用參數估計使用之樣本進行試題反應理論試題鑑別度 (Multidimensional Discrimination index [MDISC]) 與試題閾值 (或階難度) 參數 (Multidimensional Difficulty index [MDIFF]) 計算，當試題無法確認其所屬向度時，將無法基於刪題後的結果進行參數估計。

肆、 試題反應理論 (Item Response Theory, IRT)

針對 IRT，依序將編號 1「大學生身心生活適應」中的「正向因應資源」分量表、編號 1「大學生身心生活適應」中的「負向情緒反應」分量表、編號 5「多向度感知社會支持量表」與今年新納入之編號 27「自我滿意度」進行測量模型擬合與試題訊息量分析。

一、 試題反應理論之測量模型擬合

試題反應理論 (Item Response Theory [IRT]) 為一個新型測驗理論，相較 CTT 在測量標準誤 (Standard Error of Measurement [SEM]) 跨試題間為相等，IRT 則認為 SEM 應隨著填答者的能力不同而在不同試題上具有不同的 SEM。儘管 IRT 具有多個優於 CTT 的優勢，但 IRT 多年來被認為是一具有強力假設前提的測驗理論：需滿足測驗向度性假設。對此，本研究基於 IRT 進行試題分析時，首先將確定測驗本身的向度個數為單一向度或多個向度。因四份量表皆採用李克特氏量尺，屬順序尺度、且每試題在已知答題反應會受到鑑別參數與閾值參數影響下，故採用部份給分模式 (Partial Credit Model [PCM]; Muraki, 1990) 與等級評分模式 (Graded Response Model [GRM]; Samejima, 1969) 進行細部測量模型擬合。PCM 數學關係式如下：

$$P_{ix}(\theta) = \frac{\exp[\sum_{j=0}^x (\theta - \delta_{ij})]}{\sum_{r=0}^{m_i} [\exp(\sum_{j=0}^r (\theta - \delta_{ij}))]} \quad (3)$$

而 GRM 的數學關係式如下：

$$P_{ix}(\theta) = P_{ix}^*(\theta) - P_{i(x+1)}^*(\theta) \quad (4)$$

$$P_{ix}^*(\theta) = \frac{\exp[\alpha_i(\theta - \beta_{ij})]}{1 + \exp[\alpha_i(\theta - \beta_{ij})]} \quad (5)$$

其中， $P_{ix}(\theta)$ 為特定能力 θ 達到第 i 題之第 x 個閾值 (或階難度) 的機率， δ_{ij} 為第 i 題中第 j 個階難度參數， α_i 為第 i 題試題鑑別度指標， β_{ij} 為第 i 題中第

j 個閾值參數。其次基於向度個數與測量模式進行模型適配度檢驗。本研究採用之適配度統計量上，如為 PCM，則採用 infit 與 outfit 指標，infit 值為觀察填答者基於測驗模型下預期得分結果，outfit 值為觀察基於測驗模型下極端非預期得分結果，Wright 與 Linacre (1994) 指出當一試題的 infit 與 outfit 皆落在 .6 至 1.4 區間時，該試題將會被視為基於測量模式下可進行施測的適當試題，其標準與 Leng 與 Yao (2024) 及 Li 等人 (2025) 進行題目選擇的標準相同；如為 GRM，將採用 Orlando 與 Thissen (2003) 之符號卡方檢定 (signed chi-squared test [$S - \chi^2$]) 進行適配度分析。如試題在符號卡方檢定未達顯著，或基於符號卡方檢定之均方根誤差 (RMSEA $S - \chi^2$) 小於 .06 時，則視為試題適配於 GRM 測量模式。

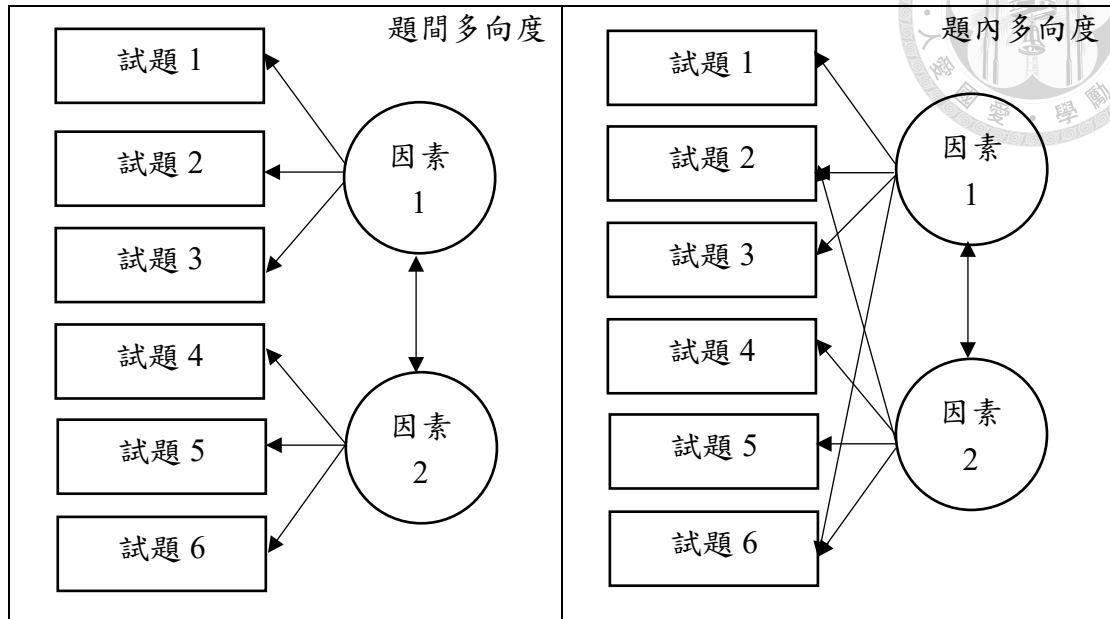
二、 試題反應理論之統計量分析

在上階段確認試題適配的 IRT 測量模式後，本研究將計算所有試題之試題訊息量 (item information)。當總體訊息量數值大時，試題具有較強的能力反應填答者各向度真實表現，且表達能力估計正確性的貢獻量大小。訊息量最大值對應的能力水準，即代表該試題所能最精確測量或估計到的能力參數估計值：估計標準誤為試題訊息量開根號的倒數。對此，試題訊息量可協助判斷一試題適合測量何種能力者。如試題本身為單一向度，僅需考量單一能力區間特定試題反應填答者在對應測量向度之能力估計正確性；如為多個向度，則需要考量不同向度上的能力區間特定試題反應填答者在對應測量向度之能力估計正確性。

過往使用試題訊息量時，因試題反應理論之基本假定：單向度假設存在，故在試題訊息量的計算上不需考量因素間的相關係數矩陣。當測驗建基在多向度試題反應理論時，如測驗試題屬題間多向度 (between-item multidimension) 時，則需加入因素間的相關係數矩陣考量單一因素對能力估計的貢獻程度；然而在測驗試題屬題內多向度 (within-item multidimension) 時，為加入因素間的相關係數矩陣考量多個因素對能力估計的貢獻程度。有關題間多向度與題內多向度測驗圖示如圖 1 所示。

圖 1

題間多向度與題內多向度測驗圖示



對此，本研究採用較常使用的最小能力估計變異法作為試題訊息量計算（D-Optimality Fisher Information; Segall, 1996）。基於多向度試題反應理論之試題訊息量，第 i 題、且未知能力 θ 時試題訊息量的數學關係式如下：

$$I_i(\theta) = -E \left(\frac{\partial^2}{\partial \theta \partial \theta'} \log f(x_i | \theta) | \theta \right) \quad (6)$$

當採用 D-Optimality 最佳化方法時，第 i 題、且未知能力 θ 時試題訊息量會依照以下算則進行：

$$I_{D-Optimality(i)}(\theta) = \det((\Sigma_0^{-1}) + I_i(\theta)) \quad (7)$$

Σ_0^{-1} 為先備變異數共變異數矩陣，本研究則採用既有填答資料進行變異數共變異數矩陣計算。Wang 與 Chang (2011) 指出該試題訊息量計算方法在多向度測驗情境下可提供具有最小能力估計偏誤的試題，且在最終能力估計時亦可維持良

好的均方差 (Mean Square Error [MSE])；然而上述試題訊息量計算需掌握填答者的能力狀態才能計算，但在無法提前取得填答者的能力狀態下，本研究採用 Leng 與 Yao (2024) 的計算取向：分向度能力值 -4 到 4，以 .1 為單位的等差級數所形成的能力向量進行討論。當一試題具有兩個向度時，其能力向量 $[\theta_1, \theta_2]$ 如下：

$$[-4, -4], [-4, -3.9], \dots, [0, 0], \dots, [4, 3.9], [4, 4]$$

並在最後計算試題在難度中等 ($-.5 \leq \theta_1 \leq .5, -.5 \leq \theta_2 \leq .5$)、能力偏高 ($\theta_1 > .5, \theta_2 > .5$) 與能力偏低 ($\theta_1 < -.5, \theta_2 < -.5$) 各自的試題訊息量。然而，因對參與者的身心健康狀態 θ 未知，Embretson 與 Reise (2000) 指出在參與者能力未知下，基於 IRT 進行初步試題選擇上以難度適中 ($-.5 \leq \theta \leq .5$) 為主，故後續分析上將分為難度適中的試題 ($-.5 \leq \theta \leq .5$)、 $\theta > .5$ 與 $\theta < -.5$ 呈現訊息量，觀察試題本身傾向提供高或低能力者訊息量的多寡。當訊息量數值越大者，越適合納入身心健康評量系統進行施測。

另考量上述提及試題訊息量計算區間範圍不同，在未知填答者的能力下，將以標準常態分配之機率密度 $\Phi(\theta)$ 進行試題訊息量的加權後再行討論訊息量多寡，其數學關係式如下：

$$I_{D-Optimality(i)}^{weighted}(\theta) = \int_{-\infty}^{\infty} I_{D-Optimality(i)}(\theta) \Phi(\theta) d\theta \quad (8)$$

藉由訊息量加權，可反映不同族群的能力特性，以利評估測驗試題對不同群體能力估計精確性的貢獻量。

伍、基於兩測量理論的分析結果

根據上述兩測驗理論的試題選擇結果，本研究將分別討論以上測驗的試題選擇排序，並將所有測驗內的試題分為：「適合所有群體施測試題」、「適合焦點群體施測試題」、「不建議提供施測試題」。其中，焦點群體是指基於身心健康評量系統建置目的：欲透過特定試題找尋需儘早介入關懷的同學，如：身心狀態較差的填答者。

對此，本研究會先以古典測驗理論的分析項目：試題難度、試題偏態係數與峰度係數、校正後試題總分相關、Cronbach's alpha、探索性因素分析因素負荷量進行討論。試題難度與試題鑑別度為第一標準（並以試題偏態係數與峰度係數作為輔助參考工具），後續再以校正後試題總分相關與探索性因素分析因素負荷量分別作為第二標準與第三標準。後續以試題反應理論之試題訊息量協助確認試題對特定焦點群體是否具有較小的能力估計偏誤。當特定試題對焦點群體具有較小的能力估計偏誤、但被古典測驗理論視為不建議提供施測試題時，如果試題對應之因素結構符合國立臺灣大學身心健康評量系統建置目的，基於測驗目的應予以保留。

在古典測驗理論上，針對探索性因素分析之因素負荷量協助試題簡化的部分，Comrey 與 Lee（2013）在其著作認為經由探索性因素分析後的因素負荷量可分為以下五個標準：當因素負荷量大於 .71 為極度良好試題（excellent），介於 .63 到 .71（含）為良好試題（very good），介於 .55 到 .63（含）為好試題（good），介於 .45 到 .55（含）為可接受試題（fair），介於 .32 到 .45（含）為不佳試題（poor）。Tabachnick 與 Fidell（2019）認為試題的因素負荷量最少需達 .32 才可判斷試題和因素之間的顯著相關與否。對此，以試題選擇上，多數研究者會採用大於 .40 的標準，雖然大於 Tabachnick 與 Fidell（2019）所認為的最小因素負荷量數值；然而當進到試題簡化時，因需採用較高的試題選擇標準，本研究將採用 Comrey 與 Lee（2013）判斷極度良好試題的標準：因素負荷量大於 .71 的標準作為試題簡化的標準。

在試題反應理論上，針對訊息量協助試題簡化的部分，除了觀看試題數值判斷能力估計偏誤外，因對參與者的身心健康狀態未知，Embretson 與 Reise（2000）指出在參與者能力未知下，基於 IRT 進行初步試題選擇上以難度適中

($-.5 \leq \theta \leq .5$) 為主。基於此建議下，為提升所選試題在施測上的能力估計精確性，故在「適合所有群體施測試題」上優先選擇難度適中 ($-.5 \leq \theta \leq .5$) 時試題訊息量大或對所有能力群體 ($\theta > .5$ 與 $\theta < -.5$) 具有相近數值的估計精確性；「適合焦點群體施測試題」中則優先選擇在特定能力群體 ($\theta > .5$ 或 $\theta < -.5$) 估計精確性較佳者，且為協助所有群體的能力估計，亦搭配 $-.5 \leq \theta \leq .5$ 時訊息量進行題目選擇。綜上所述，基於訊息量的試題保留上，「適合所有群體施測試題」為基於全體能力進行選擇：針對難度適中 $-.5 \leq \theta \leq .5$ 與考量對全體能力估計上具有較一致的估計精確性，「適合焦點群體施測試題」為基於測驗目標群體進行選擇：針對情緒狀態不佳者捕捉具有較高能力估計精確性的試題為主。

第四章 結果



第一節 參與者人口學變項

針對參與者人口學變項部分，因本學年度的線上問卷填答系統設定每題皆為必須填答，所有測驗試題並未存在遺漏值。後續本研究聚焦在年齡介於 17 歲至 25 歲、需為國立臺灣大學大學部在學學生，共計 3838 位。表 2 呈現各學院的填答人數與性別人數，括號為該人數佔總體人數的比例（性別的括號則為該性別佔總體學院人數的比例）。根據 2024 年國立臺灣大學線上問卷系統的設定，性別部分因其可以透過參與者手動進行更改至多元性別，儘管多數學生並未做更改，但由表 3 可發現文學院、社會科學院、生物資源暨農學院與法律學院有部分同學進行修改。後續分析上，因本研究目標為提供學生群體進行少數優良試題進行施測，故未來分析上並未將性別做更動的同學進行答題反應的刪除。

表 2 呈現 2024 年參與者填答人數與 2023 年大學部在學人數，括號為該人數佔總體人數的比例。結果顯示，所有學院或學位學程的填答人數與實際在校人數比例接近。表 3 中呈現 2024 年參與者各生理性別人數與 2023 年大學部在學學生生理性別人數，括號為該性別佔總體學院人數的比例。結果顯示在 2024 年參與者中，理工學院（如：理學院、工學院、電機資訊學院）的男性比例居多，且以電機資訊學院的性別比例差距最為懸殊；文學院、社會科學院、管理學院則以女性比例居多。醫學院男女性別比例接近，而生命科學院則以女性居多。自表 3 可知跨學院間皆存在生理男性填答比例較生理女性低的現象；在學院內亦可發現生理男性填答比例較真實在學生理男性比例低、生理女性填答比例則較真實在學生理女性比例高，反應不論所屬學院為何，生理女性皆較願意完成今年所有身心健康評量系統建置試題。

有關最初填答者進入國立臺灣大學入學管道如表 4 所示，多數學生採用個人申請方式入學，其次為考試分發。其中，他國學生（僑生、外籍生）、希望入學學生、離島學生、原住民學生共計佔全校入學管道的 11.41%，身障學生佔 .55%。進一步與過往入學管道人數比例比較，繁星與僑生入學者在今年填答

比例上升、考試分發者填答比例下降，然個人申請填答比例與往年入學管道人數比例差異較小。



表 2

2024 年參與者填答人數（百分比）描述統計

學院	參與者人數（百分比）大學部在學學生數（百分比）	
	2024	2023
文學院	497 (12.95)	2164 (12.49)
理學院	433 (10.50)	1519 (8.77)
社會科學院	419 (10.92)	1820 (10.51)
醫學院	478 (12.45)	2154 (12.43)
工學院	410 (10.68)	2091 (12.07)
生物資源暨農學院	519 (13.52)	2411 (13.92)
管理學院	438 (11.41)	2211 (12.76)
公共衛生學院	35 (0.91)	161 (0.93)
電機資訊學院	321 (8.36)	1469 (8.48)
法律學院	188 (4.90)	817 (4.72)
生命科學院	122 (3.18)	465 (2.68)
共同教學中心學位學程	8 (0.21)	41 (0.24)
全校大學部在學學生數	3838 ^a	17323

^a 為全校大學部在學學生參與者人數總和。

表 3

2024 年參與者生理男性與生理女性人數（百分比）描述統計



學院		參與者生理性別人數 (百分比)	大學部在學 學生生理性別數 (百分比)
			2023
文學院	男	156 (31.39)	757 (34.98)
	女	340 (68.41)	1407 (65.02)
理學院	男	255 (63.28)	1049 (69.06)
	女	178 (44.17)	470 (30.94)
社會科學院	男	159 (37.95)	892 (49.01)
	女	259 (61.81)	928 (50.99)
醫學院	男	224 (46.86)	1179 (54.74)
	女	254 (53.14)	975 (45.26)
工學院	男	278 (67.80)	1591 (76.09)
	女	132 (32.20)	500 (23.91)
生物資源暨農學院	男	183 (35.26)	1187 (49.23)
	女	335 (64.55)	1224 (50.77)
管理學院	男	176 (40.18)	1057 (47.81)
	女	262 (59.82)	1154 (52.19)
公共衛生學院	男	14 (40.00)	70 (43.48)
	女	21 (60.00)	91 (56.52)
電機資訊學院	男	256 (79.75)	1260 (85.77)
	女	65 (20.25)	209 (14.23)
法律學院	男	81 (43.09)	367 (44.92)
	女	104 (55.32)	450 (55.08)
生命科學院	男	45 (36.89)	223 (47.96)
	女	77 (63.11)	242 (52.04)
共同教學中心學位學程	男	2 (25.00)	20 (48.78)
	女	6 (75.00)	21 (51.22)
全校	男	1829 (47.66)	9652 (55.72)
	女	2033 (52.97)	7671 (44.28)

註：因線上問卷系統的設定，將生理性別調整為多元性別者將不列入生理性別人數計算。

表 4

參與者入學管道人數（百分比）描述統計



入學管道	2024（百分比）	2023（百分比）
個人申請	1502 (39.13)	1678 (38.95)
考試分發	1223 (31.87)	1424 (33.05)
繁星	432 (11.26)	436 (10.12)
僑生	225 (5.86)	262 (6.08)
非上述入學管道	488 (13.06)	459 (10.65)
轉學	112 (2.92)	154 (3.57)
希望入學	108 (2.81)	74 (1.72)
外籍生	57 (1.49)	95 (2.21)
離島生	21 (0.55)	17 (0.39)
特殊選才	37 (0.96)	34 (0.79)
奧林匹亞	29 (0.76)	33 (0.77)
身障	21 (0.55)	19 (0.44)
原住民	27 (0.70)	21 (0.49)
其他 ^a	76 (1.98)	12 (0.28)
總人數	3838	4308

註：表中呈現資料含不同年份的入學學生。

^a 含醫事人員、外交子女等入學管道。

第二節 基於古典測驗理論之試題選擇



壹、正向因應資源

表 6 中呈現「正向因應資源」的試題描述統計值，如：試題平均數、試題標準差、試題偏態係數、試題峰度係數與校正後試題總分相關。表格結果顯示，所有試題難度介於 2.77 至 3.80 間、標準差介於 .78 至 1.09 間、偏態係數與峰度係數分別介於 -.46 至 .40 與 -.69 至 .26 間。多數試題的校正後試題總分相關大於 .40，僅第七題「在別人面前，我不怕說出自己的缺點或過去的失敗」為 .39。

探索性因素分析結果呈現在表 5 與表 6。在表 5 中，Kaiser 總體合適性指標為 .93，說明現有未刪除試題下適合進行探索性因素分析。同時表 5 亦顯示各因素間的相關係數介於 .47 至 .53 間，其因素萃取個數判別圖可參閱附錄二圖 2-1。後續分析中，顯示多數試題皆與原先試題所屬的向度相同，其中第一題「家人對我的生活方式感到滿意」、第二題「家人對我的學習與學業目標感到滿意」、第三題「我的生活目標同時符合我的與父母的需求」在「自我資源」與「家庭資源」上具有相近的因素負荷量，與原先試題僅歸屬「家庭資源」向度不符；第七題「在別人面前，我不怕說出自己的缺點或過去的失敗」在「自我資源」因素負荷量為 .35，對「自我資源」向度的解釋能力較低。以上四題基於探索性因素分析的簡化試題標準：「因素負荷量大於 .4 或單一試題未在兩個或以上因素具有數值接近之因素負荷量」下，將被列為優先簡化試題。

內部一致性信度指標：標準化 Cronbach's alpha 中，排除經探索性因素分析提及之優先排除試題後，結果如表 6 所示。三個因素的內部一致性信度分別為：「家庭資源」.91、「自我資源」.83、「師友資源」.89，顯示優先排除試題移除後，各因素內試題對因素測量上具有高度一致性。

驗證性因素分析結果中，排除上述優先簡化試題後，先採用 50% 樣本數進行模型參數估計，再利用剩餘 50% 的樣本進行模型適配度指標的計算。模型參數估計報表可參閱附錄 4 表 4-1-1、表 4-1-2 與表 4-1-3。五種常見的模式適配度指標（Goodness-of Fit indices）：卡方統計量（Chi-square）及其顯著性、標準化殘差均方根（Standardized Root Mean Square Residual [SRMR]）、可比級適配度

指標 (Comparative Fit Index [CFI])、近似誤差均方根 (Root Mean Square Error of Approximation [RMSEA]) 與非規範擬合指標 (Tucker-Lewis Index [TLI])，除卡方統計量外皆落在合理區間中，其結果如表 12 所示。

綜上所述，基於古典測驗理論之簡化試題流程中，在正向因應資源內會將第一題「家人對我的生活方式感到滿意」、第二題「家人對我的學習與學業目標感到滿意」、第三題「我的生活目標同時符合我的與父母的需求」與第七題「在別人面前，我不怕說出自己的缺點或過去的失敗」列為優先簡化試題。

表 5

「正向因應資源」因素間相關係數矩陣^a

	家庭資源 (F_1)	自我資源 (F_2)	師友資源 (F_3)
家庭資源 (F_1)	—		
自我資源 (F_2)	.52	—	
師友資源 (F_3)	.47	.53	—

註：所有試題之校正後試題總分相關皆大於 .4，故未刪除題目計算相關係數。

^a Kaiser's Overall MSA = .93.

表 6

「正向因應資源」之古典測驗理論分析結果

編號	試題內容	平均數	標準差	偏態	峰度	It.coW	探索性因素分析				α
家庭資源 (F_1)											
12	當我有麻煩和煩惱時，我可以從家庭得到幫助	3.47	1.08	-0.31	-0.69	.73	F_1	.82	-.02	.04	.89
13	我滿意家人對我的情緒（喜、怒、哀、樂）表示關心和愛護的方式	3.33	1.09	-0.22	-0.64	.79	F_1	.91	-.06	.01	.87
14	我和家人可以共度愉快的時光	3.70	0.95	-0.44	-0.20	.76	F_1	.87	-.07	.03	.88
15	當我嘗試新事物時，我會感到家人的接受與支持	3.60	1.03	-0.35	-0.56	.75	F_1	.79	-.05	.10	.90
自我資源 (F_2)											
4	我能夠從容地應付我的課業	2.77	0.94	0.08	-0.26	.54	F_2	.01	.63	-.04	.81
5	我知道我為何要讀書	3.34	1.07	-0.21	-0.61	.48	F_2	-.03	.53	.05	.82
6	我是個心情愉快的人	3.34	0.88	-0.11	-0.22	.68	F_2	.03	.66	.12	.79
8	早上起床時，我覺得精神良好	2.57	0.91	0.40	0.02	.51	F_2	.06	.53	-.02	.82
9	我對於自己休閒活動的安排感到滿意	3.15	1.00	0.04	-0.61	.56	F_2	.01	.56	.07	.81
10	碰到困難時，我會冷靜下來去尋找解決困難的方法	3.56	0.81	-0.17	-0.11	.57	F_2	-.02	.58	.08	.81
11	一般而言，我喜歡自己，滿意自己	3.28	1.02	-0.17	-0.50	.70	F_2	.08	.72	.03	.79



表 6 (續)

「正向因應資源」之古典測驗理論分析結果

編號	試題內容	平均數	標準差	偏態	峰度	It.coW	探索性因素分析			α
師友資源 (F_3)							原屬向度	F_1	F_2	F_3
16	當我遇到麻煩，能從老師、同學或朋友獲得協助	3.67	0.87	-0.38	-0.06	.78	F_3	.08	.02	.81
17	我滿意老師、同學或朋友關懷我情緒的方式	3.57	0.86	-0.24	-0.17	.81	F_3	.02	.03	.86
18	我能和老師、同學或朋友共度愉快的時光	3.80	0.78	-0.41	0.26	.75	F_3	-.01	.13	.74
無法歸屬特定向度										
1	家人對我的生活方式感到滿意	3.41	0.90	-0.19	-0.35	.65	F_1	.46	.36	-.08
2	家人對我的學習與學業目標感到滿意	3.73	0.92	-0.46	-0.16	.62	F_1	.42	.33	-.04
3	我的生活目標同時符合我的與父母的需求	3.41	0.96	-0.22	-0.41	.68	F_1	.47	.36	-.05
7	在別人面前，我不怕說出自己的缺點或過去的失敗	3.13	1.06	-0.01	-0.73	.39	F_2	-.04	.35	.16

註：It.coW (Item-Total correlation With item adjustment)：校正後試題總分相關（以編寫測驗試題預先設定試題所屬的向度總分進行計算）； α ：排除特定試題 Cronbach's alpha。

^a 表示特定向度之 Cronbach's alpha。



貳、 負向情緒反應

表 8 中呈現「負向情緒反應」的試題描述統計值，其中原屬「自傷想法行為」的試題難度小於 2，試題標準差小於 1；原屬「精神病狀」的試題難度亦小於 2，試題標準差亦小於 1；其餘試題難度與標準差分別介於 2.01 至 3.66 間與 .96 至 1.17 間。其中，第六題「我會有一段時間（至少持續一週）情緒高昂，容易興奮」是唯一非屬「自傷想法行為」或「精神病狀」向度、且標準差為 .96 的試題。偏態係數與峰度係數部分，「自傷想法行為」與「精神病狀」所有試題皆超出正負 1，說明填答者在兩向度試題的選項選擇分布不均，其餘試題在兩係數上皆落在正負 1 之間。校正後試題總分相關中，第六題「我會有一段時間（至少持續一週）情緒高昂，容易興奮」為 .40。

探索性因素分析結果呈現於表 7 與表 8。在表 7 中，Kaiser 總體合適性指標為 .95，說明現有試題適合進行探索性因素分析。同時表 7 亦顯示各因素間的相關係數介於 .35 至 .64 間，其因素萃取個數判別圖可參閱附錄二圖 2-2。後續分析中，顯示多數試題皆與原先試題所屬的向度相同，其中第八題「我對一般人不感到害怕的環境（如：高處、人群、封閉的地方或出家門），我卻感到極度害怕，以至於想要離開」與第十三題「我有『將要生病』的感覺」在「因應情緒行為」與「精神病狀」上具有相近的因素負荷量，與原先試題僅歸屬「精神病狀」向度不符、第十四題「我覺得心情不好，沒辦法應付周圍發生的事」與第十六題「我覺得處在壓力中」在「因應情緒行為」與「無望感」上具有相近的因素負荷量，與原先試題僅歸屬「無望感」向度不符。以上四題基於探索性因素分析的簡化試題標準：「因素負荷量大於 .4 或單一試題未在兩個或以上因素具有數值接近之因素負荷量」下，將被列為優先簡化試題。

內部一致性信度指標：標準化 Cronbach's alpha 中，排除經探索性因素分析提及之優先排除試題後，結果如表 8 所示。四個因素的內部一致性信度分別為：「因應情緒行為」.88、「無望感」.89、「自傷想法行為」.79、「精神病狀」.73，前三者顯示優先排除試題移除後，因素內試題對因素測量上具有高度一致性。

驗證性因素分析結果中，排除上述優先簡化試題後，先採用 50% 樣本數進行模型參數估計，再利用剩餘 50% 的樣本進行模型適配度指標的計算。模型參數估計報表可參閱附錄 4 表 4-2-1、表 4-2-2 與表 4-2-3。表 12 的結果中顯示 TLI 皆未達標準；然而，綜合其餘指標評估下，發現 RMSEA、SRMR、CFI 皆落在合理區間中，可說明刪題後量表具有良好的模型適配度。

綜上所述，基於古典測驗理論之簡化試題流程中，負向情緒反應中第八題「我對一般人不感到害怕的環境（如：高處、人群、封閉的地方或出家門），我卻感到極度害怕，以至於想要離開」、第十三題「我有「將要生病」的感覺」、第十四題「我覺得心情不好，沒辦法應付周圍發生的事」與第十六題「我覺得處在壓力中」列為優先簡化試題。

表 7

「負向情緒反應」之因素間相關係數矩陣^a

	因應情緒行為 (F_1)	無望感 (F_2)	自傷想法行為 (F_3)	精神病狀 (F_4)
因應情緒行為 (F_1)	—			
無望感 (F_2)	.64	—		
自傷想法行為 (F_3)	.44	.50	—	
精神病狀 (F_4)	.44	.35	.47	—

註：僅一試題之校正後試題總分相關為 .39，其餘皆大於 .4，故未刪除題目計算相關係數。

^a Kaiser's Overall MSA = .95.

表 8

「負向情緒反應」之古典測驗理論分析結果

編號	試題內容	平均數	標準差	偏態	峰度	It.coW	探索性因素分析					α
							原屬向度	F_1	F_2	F_3	F_4	
因應情緒行為 (F_1)												
1	我會因為不安與生氣，造成我的身體不舒服（如：拉肚子、頭痛、皮膚起疹子、手腳麻木等）	2.54	1.09	0.30	-0.63	.59	F_1	.55	.16	.10	.19	.88 ^a
2	我會覺得自己很神經質，很容易緊張，對事情都想得太多，考慮太多	3.15	1.17	-0.13	-0.85	.67	F_1	.62	.32	.04	.16	.87
3	我會沒有理由地感到害怕、驚慌、不安	2.46	1.07	0.46	-0.47	.73	F_1	.62	.34	.16	.25	.86
4	我會有一段時間心情低落、憂鬱、鬱悶	3.06	1.02	-0.09	-0.39	.75	F_1	.75	.36	.18	.02	.86
5	我會有一段時間心情不好、容易苦惱、動怒、生氣	2.88	1.02	0.04	-0.45	.76	F_1	.77	.31	.19	.03	.86
6	我會有一段時間（至少持續一週）情緒高昂，容易興奮	2.18	0.94	0.70	0.25	.40	F_1	.42	.01	.06	.22	.89
7	日常生活中，我情緒會不可預測地忽高忽低，讓我感到困擾	2.40	1.06	0.53	-0.33	.72	F_1	.63	.27	.21	.22	.86
9	我曾因生活上的變化（如親人離別、與人吵架等）導致情緒不穩與痛苦，進而影響工作與人際關係等	2.31	1.01	0.49	-0.29	.59	F_1	.49	.21	.18	.30	.87
15	我睡眠有困難，如：難以入睡、易醒來或早醒	2.41	1.13	0.48	-0.54	.54	F_1	.44	.21	.19	.21	.88

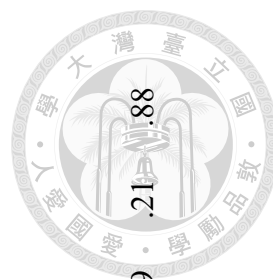


表 8 (續)

「負向情緒反應」之古典測驗理論分析結果

編號	試題內容	平均數	標準差	偏態	峰度	It.coW	探索性因素分析				α	
無望感 (F_2)												
17	我覺得自己比不過別人	3.36	1.12	-0.21	-0.67	.74	F_2	.30	.77	.03	.02	.86 ^a
18	我不滿意自己的學業或工作表現	3.19	1.04	0.03	-0.51	.73	F_2	.27	.74	.09	.04	.86
19	我覺得自己無法做決定	2.75	1.08	0.24	-0.57	.59	F_2	.25	.55	.10	.12	.88
20	我覺得自己是一文不值的人	2.14	1.08	0.78	-0.07	.76	F_2	.21	.72	.30	.19	.85
21	我覺得人生很無望，對事物皆失去興趣	2.06	1.02	0.83	0.18	.72	F_2	.24	.63	.39	.19	.86
24	我對未來不抱甚麼期待	2.17	1.05	0.66	-0.15	.68	F_2	.20	.61	.38	.16	.87
自傷想法行為 (F_3)												
22	我有尋死（自殺）的想法	1.49	0.78	1.74	2.99	.68	F_3	.21	.28	.78	.08	.70
23	我有想傷害別人的念頭或衝動	1.37	0.69	2.10	4.85	.46	F_3	.14	.10	.45	.22	.81
25	我有尋死的計畫	1.23	0.58	2.96	10.21	.69	F_3	.10	.12	.75	.17	.71
26	我有做過傷害自己的事情	1.28	0.63	2.71	8.32	.60	F_3	.15	.11	.59	.22	.75
精神病狀 (F_4)												
10	我曾因不必要的重複動作（如：洗手、關門、關瓦斯）或因停不下來的想法（如：一直以為是髒的，或一直以為可能會傷害他人，但我又明知那是不可能發生的事情）而心理感到不安與痛苦	1.75	0.98	1.33	1.25	.60	F_4	.29	.14	.16	.60	.66
11	我相信一些事情（如：有人跟蹤我，或有人要害我等等）但是別人認為那是不存在的，或是不能發生的	1.55	0.82	1.68	2.81	.60	F_4	.19	.10	.15	.69	.58



表 8 (續)

「負向情緒反應」之古典測驗理論分析結果

編號	試題內容	平均數	標準差	偏態	峰度	It.coW	探索性因素分析				α	
精神病狀 (F_4)											.73 ^a	
12	我會聽到一些別人聽不到的聲音，看到別人看不到的事物。	1.29	0.59	2.30	6.17	.50	F_4	.11	.01	.26	.59	.70
無法歸屬特定向度												
8	我對一般人不感到害怕的環境（如：高處、人群、封閉的地方或出家門），我卻感到極度害怕，以至於想要離開	2.08	1.01	0.83	0.23	.47	原屬向度 F_4	.36	.27	.20	F_4	.40
13	我有「將要生病」的感覺	2.03	0.99	0.70	-0.15	.57	F_1	.38	.28	.24	F_1	.36
14	我覺得心情不好，沒辦法應付周圍發生的事	2.43	1.00	0.37	-0.30	.73	F_1	.56	.45	.22	F_1	.23
16	我覺得處在壓力中	3.46	1.00	-0.30	-0.24	.59	F_2	.49	.50	.06	F_2	.05

註：It.coW (Item-Total correlation With item adjustment)：校正後試題總分相關（以編寫測驗試題預先設定試題所屬的向度總分進行計算）； α ：排除特定試題 Cronbach's alpha。

^a 表示特定向度之 Cronbach's alpha。



參、多向度感知社會支持量表

表 10 中呈現「多向度感知社會支持量表」的試題描述統計值，其報表顯示試題難度介於 4.74 至 5.36 間、標準差介於 1.23 至 1.76 間、偏態係數與峰度係數落在正負 1 之間。校正後試題總分相關上，所有試題皆大於 .70。

探索性因素分析結果呈現於表 9 與表 10。在表 9 中，Kaiser 總體合適性指標為 .90，說明現有試題適合進行探索性因素分析。同時表 9 亦顯示各因素間的相關係數介於 .33 至 .45 間，其因素萃取個數判別圖可參閱附錄二圖 2-3。後續分析中，所有試題皆與原先試題所屬的向度相同，且所有試題皆未需進行簡化：所有試題的因素負荷量大於 .4、且單一試題未在兩個或以上因素具有數值接近之因素負荷量。

內部一致性信度指標：標準化 Cronbach's alpha 結果如表 10 所示。三個因素的內部一致性信度分別為：「重要他人」.94、「家人」.93、「朋友」.92，顯示各因素內試題對因素測量上具有高度一致性。

驗證性因素分析結果中，先採用 50% 樣本數進行模型參數估計，再利用剩餘 50% 的樣本進行模型適配度指標的計算。模型參數估計報表可參閱附錄 4 表 4-3-1、表 4-3-2 與表 4-3-3。表 12 顯示四項適配度指標：SRMR、RMSEA、CFI 與 TLI 皆落在合理標準中，僅卡方適配度檢定數值存在高度顯著。

表 9

「多向度感知社會支持量表」之因素間相關係數矩陣^a

	重要他人 (F_1)	家人 (F_2)	朋友 (F_3)
重要他人 (F_1)	—		
家人 (F_2)	.33	—	
朋友 (F_3)	.45	.33	—

註：所有試題校正後試題總分相關皆大於 .4，故未刪除題目計算相關係數。

^a Kaiser's Overall MSA = .90.

表 10

「多向度感知社會支持量表」之古典測驗理論分析結果

編號	試題內容	平均數	標準差	偏態	峰度	It.coW	原屬向度	探索性因素分析	α
重要他人 (F_1)									
1	每當我有需要時，有個特別的人會在我身邊幫助我	4.74	1.60	-0.47	-0.56	.84	F_1	F_1	.92
2	有個特別的人能讓我分享快樂與悲傷	4.96	1.66	-0.67	-0.40	.88	F_1	F_1	.91
5	有位特別的人是我慰藉的主要來源	4.73	1.76	-0.49	-0.68	.84	F_1	F_1	.92
10	我生命中有個特別的人會關心我的感受	4.99	1.68	-0.62	-0.43	.85	F_1	F_1	.92
家人 (F_2)									
3	我的家人真的會試圖幫助我	5.30	1.46	-0.79	0.12	.84	F_2	F_2	.90
4	我能從家人得到所需的情緒上的協助與支持	4.86	1.64	-0.59	-0.38	.87	F_2	F_2	.89
8	我可以告訴家人我的煩惱	4.58	1.72	-0.41	-0.72	.81	F_2	F_2	.91
11	我的家人願意協助我做決定	5.16	1.45	-0.78	0.29	.79	F_2	F_2	.92
朋友 (F_3)									
6	我的朋友真的會試圖幫助我	5.36	1.23	-0.80	0.84	.82	F_3	F_3	.90
7	當萬事不順遂時，我可以依賴我的朋友	5.02	1.43	-0.65	0.08	.83	F_3	F_3	.90
9	我有朋友可以分享我的快樂與悲傷	5.46	1.30	-1.04	1.29	.82	F_3	F_3	.90
12	我可以告訴我的朋友我的煩惱	5.38	1.34	-0.97	0.98	.82	F_3	F_3	.90

註：It.coW (Item-Total correlation With item adjustment)：校正後試題總分相關（以編寫測驗試題預先設定試題所屬的向度總分進行計算）； α ：排除特定試題 Cronbach's alpha。

^a 表示特定向度之 Cronbach's alpha。

綜上所述，基於古典測驗理論之簡化試題流程中，多向度感知社會支持量表未存在需優先排除試題。



肆、自我滿意度

表 13 中呈現「自我滿意度」的試題描述統計值。報表顯示試題難度與標準差分別介於 2.52 至 4.15 間與 .85 至 1.10 間、偏態係數與峰度係數皆落在正負 1 之間。校正後試題總分相關上，因本測驗為由過往 2021 年至 2023 年提供施測的身心健康評量系統題庫針對不同可能影響「對填答者自己的滿意度」的變項進行詢問，未事先指定試題所屬向度下，本研究將採用測驗總分（總體滿意度）與各自試題進行校正後試題總分相關計算。基於上述計算設定下，第十六題「我對於環境對我自己性別認同的支持感到」之校正後試題總分相關為 .29，第一題「我對於自己所就讀的科系感到」則為 .39。

探索性因素分析結果呈現於表 11 與表 13。表 11 中，Kaiser 總體合適性指標為 .91，說明現有試題適合進行探索性因素分析。同時表 11 亦顯示因素間的相關係數為 .58，其因素萃取個數判別圖可參閱附錄二圖 2-4。後續分析中，顯示多數試題可分類至兩大向度，並將其命名為：「身心狀態滿意度」與「日常生活滿意度」，其中第五題「我對於自己的時間管理能力感到」與第十六題「我對於環境對我自己性別認同的支持感到」在「身心狀態滿意度」與「日常生活滿意度」上具有相近的因素負荷量。以上兩題基於探索性因素分析的簡化試題標準：「因素負荷量大於 .4 或單一試題未在兩個或以上因素具有數值接近之因素負荷量」下，將被列為優先簡化試題。

內部一致性信度指標：標準化 Cronbach's alpha 中，排除經探索性因素分析提及之優先排除試題後，結果如表 13 所示。兩個因素的內部一致性信度分別為：「身心狀態滿意度」.86、「日常生活滿意度」.83，顯示優先排除試題移除後，各因素內試題對因素測量上具有高度一致性。

驗證性因素分析結果中，排除上述優先簡化試題後，先採用 50% 樣本數進行模型參數估計，再利用剩餘 50% 的樣本進行模型適配度指標的計算。模型參數估計報表可參閱附錄 4 表 4-4-1、表 4-4-2 與表 4-4-3。表 12 結果中顯示 TLI 未達標準；然而，綜合評估其他指標後，發現 SRMR、RMSEA 皆落在合理區

間，CFI 則落在 .89，其接近標準 .90 情況下，顯示刪題後量表具有良好模型適配度。

綜上所述，基於古典測驗理論之簡化試題流程中，在自我滿意度中第五題「我對於自己的時間管理能力感到」與第十六題「我對於環境對我自己性別認同的支持感到」列為優先簡化試題。

表 11

「自我滿意度」之因素間相關係數矩陣^a

	身心狀態滿意度 (F_1)	日常生活滿意度 (F_2)
身心狀態滿意度 (F_1)	—	
日常生活滿意度 (F_2)	.58	—

註：僅一試題之校正後試題總分相關為 .36，其餘皆大於 .4，故未刪除題目計算相關係數。

^a Kaiser's Overall MSA = .91.

表 12

刪題後測驗模型適配度指標

	χ^2	SRMR	RMSEA	CFI	TLI
正向因應資源	1306.91*	0.03	0.05	0.97	0.96
負向情緒反應	5486.30*	0.05	0.07	0.90	0.88 ^b
多向度感知社會支持量表	1321.66*	0.03	0.08	0.97	0.96
自我滿意度量表	2896.03*	0.05	0.09 ^a	0.89	0.87 ^b

註：卡方統計量 (Chi-square [χ^2])，標準化殘差均方根 (Standardized Root Mean Square Residual [SRMR])，近似誤差均方根 (Root Mean Square Error of Approximation [RMSEA])，可比級適配度指標 (Comparative Fit Index [CFI])，非規範擬合指標 (Tucker-Lewis index [TLI])。

^a RMSEA > .08 表示模型適配度不佳。

^b TLI < .95 表示模型適配度不佳。

* $p < .0001$.

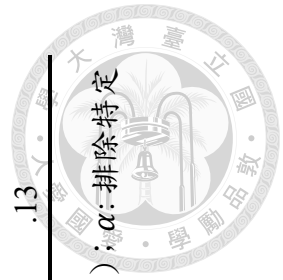
表 13

「自我滿意度」之古典測驗理論分析結果

編號	試題內容	平均數	標準差	偏態	峰度	It.coW	探索性因素分析		α
身心狀態滿意度 (F_1)									
							歸屬向度	F_1	F_2
6	我對於自己的人際關係感到	3.27	0.99	-0.33	-0.40	.55	F_1	.50	.28
7	我對於自己的情感關係感到	3.07	1.10	-0.10	-0.64	.47	F_1	.45	.22
8	我對於自己的情緒穩定度感到	3.23	1.10	-0.23	-0.66	.61	F_1	.68	.20
9	我對於自己的個人特質感到	3.39	0.97	-0.33	-0.29	.67	F_1	.65	.33
10	我對於自己的身心健康感到	3.16	1.04	-0.16	-0.59	.71	F_1	.76	.25
11	我對於自己的自我掌控能力感到	3.08	1.08	-0.14	-0.74	.69	F_1	.68	.33
12	我對於自己的家庭狀況感到	3.70	1.05	-0.63	-0.13	.41	F_1	.42	.14
13	我對於自己的休閒活動安排感到	3.42	0.98	-0.32	-0.35	.63	F_1	.62	.24
14	我對於自己的運動習慣感到	2.92	1.19	0.15	-0.91	.46	F_1	.47	.17
15	我對於自己的睡眠狀況感到	2.78	1.09	0.17	-0.71	.47	F_1	.50	.16
.83 ^a									
日常生活滿意度 (F_2)									
1	我對於自己所就讀的科系感到	3.51	0.99	-0.44	-0.18	.39	F_2	.17	.45
2	我對於自己的學業或工作表現感到	2.99	0.99	-0.10	-0.57	.62	F_2	.40	.56
3	我對於自己的生涯規劃感到	2.97	0.99	-0.07	-0.47	.62	F_2	.21	.90
4	我對於自己的生活與人生目標感到	3.12	1.01	-0.16	-0.53	.68	F_2	.33	.81
.74									
無法歸屬特定向度									
5	我對於自己的時間管理能力感到	2.52	1.09	0.34	-0.65	.57		.45	.40
16	我對於環境對我自己性別認同的支持感到	4.15	0.85	-0.75	0.18	.29		.26	.13

註：It.coW (Item-Total correlation With item adjustment)：校正後試題總分相關 (測驗總分與各自試題進行計算)； α ：排除特定試題 Cronbach's alpha。

^a 表示特定向度之 Cronbach's alpha。



第三節 基於試題反應理論之試題選擇



壹、 試題反應理論之測量模型擬合：向度個數與測量模式的選擇

起初，基於試題反應理論（Item Response Theory [IRT]）下，本研究所採用的四份量表皆會進行單向度與多向度的模型適配度檢定，且採用樣本為。四份量表於向度模型適配度檢定上，單向度試題反應理論模式則使用單參數羅吉斯迴歸模式，多向度試題反應理論模式採用二參數羅吉斯迴歸模式。表 14 呈現測驗分析的試題反應理論測驗模型適配度指標結果。「正向因應資源」分量表、「負向情緒反應」分量表、「多向度感知社會支持量表」與今年新納入「自我滿意度」儘管在卡方統計量上存在高度顯著（ $p < .0001$ ），但所有分析量表與測驗 AIC（Akaike Information Criterion）、BIC（Bayesian Information Criterion）與 SABIC（Sample-Size-Adjusted Bayesian Information Criterion）指標在多向度試題反應理論模式較單向度試題反應理論模式低，且對數轉換概似函數數值（log-Likelihood [logLik]）在多向度試題反應理論模式較單向度試題反應理論模式高下，本研究基於綜合評估結果得知多向度模型較單向度模型適配既有的測驗試題。

同時，因四份量表欲討論學生的身心適應狀態，相比成就測驗存在猜測情況，填答者在身心適應狀態的填寫上傾向反應自身身心樣貌，故本研究僅考量具有兩參數的試題反應理論測量模式。根據表 14 顯示，「正向因應資源」分量表、「負向情緒反應」分量表、「多向度感知社會支持量表」與「自我滿意度」測驗試題因 AIC、BIC、SABIC 指標與對數轉換概似函數數值皆在多向度 GRM 下較多向度 PCM 小，且對數轉換概似函數數值皆在多向度 GRM 下較多向度 PCM 大，表明多向度 GRM 模式較多向度 PCM 適配既有的測驗試題。綜合以上向度個數與測量模式的分析結果，本研究在後續試題反應理論分析上將採用多向度 GRM 進行討論。

表 14

試題反應理論測驗模型適配度指標表

正向因應資源						
	AIC	SABIC	BIC	logLik	χ^2	df
單向度	124738.8	124954.1	125176.5	-62299.4		
多向度	116177.3	116401.7	116633.7	-58015.6	8567.59*	3
PCM	116269.1	116561.3	116863.1	-58039.6		
GRM	115812.2	116104.3	116406.2	-57811.1	456.93	0
負向情緒反應						
	AIC	SABIC	BIC	logLik	χ^2	df
單向度	178437.8	178776.0	179125.6	-89108.9		
多向度	173449.2	173805.9	174174.5	-86608.6	5000.59*	6
PCM	170416.7	170939.5	171479.7	-85038.4		
GRM	169150.1	169672.9	170213.1	-84405.1	1266.59	0
多向度感知社會支持量表						
	AIC	SABIC	BIC	logLik	χ^2	df
單向度	133811.5	134069.8	134336.7	-66821.8		
多向度	118205.2	118472.8	118749.2	-59015.6	15612.30*	3
PCM	119649.9	119972.8	120306.4	-59719.9		
GRM	117798.3	118121.2	118454.8	-58794.1	1851.58	0
自我滿意度						
	AIC	SABIC	BIC	logLik	χ^2	df
單向度	134515.2	134730.4	134952.9	-67187.6		
多向度	131427.9	131646.2	131871.8	-65642.9	3089.30*	1
PCM	131953.2	132208.5	132472.2	-65893.6		
GRM	131063.0	131318.3	131582.0	-65448.5	890.17	0

註：單向度為單參數羅吉斯迴歸模式，多向度為二參數羅吉斯迴歸模式，PCM 為部份給分模式（Partial Credit Model），GRM 為等級評分模式（Graded Response Model），AIC 為 Akaike Information Criterion，SABIC 為 Sample-Size-Adjusted Bayesian Information Criterion，BIC 為 Bayesian Information Criterion，logLik 為對數轉換概似函數數值（log-Likelihood）。

* $p < .0001$.

貳、 等級評分模式下的試題分析：正向因應資源

測驗向度性部分，表 6 探索性因素分析顯示測驗試題可歸類至三個向度，且表 12 驗證性因素分析結果顯示該測驗在三個因素下具有良好的模型適配度。後續表 14 呈現經單向度與多向度模型假設考驗後亦支持測驗具有多個向度。

表 15 呈現「正向因應資源」中剩餘可歸類至特定向度的試題在等級評分模式下的擬合表現、試題參數與試題訊息量。擬合表現上，本研究採用 Orlando 與 Thissen (2003) 之符號卡方檢定 (signed chi-squared test [$S - \chi^2$]) 進行適配度分析，其結果顯示所有試題之符號卡方檢定存在高度顯著，說明試題與測量模式的不適配情況；然而，因卡方檢定易受樣本數大小影響，故測量模式擬合評估可藉由觀察符號卡方檢定下的 RMSEA 表現，其 RMSEA 結果顯示所有試題皆落在合理區間，表明試題與測量模式間的良好適配情況。試題參數上，為基於驗證性因素分析的結果而得。多向度試題鑑別度參數 (MDISC) 與多向度試題閾值參數 (MDIFF) 分別如圖 2 與圖 3 所示。自圖 2 可知多向度試題鑑別度參數普遍落在 $\text{Logits} > 0$ 的範圍，圖 3 可知多向度試題閾值參數普遍落在 $\text{Logits} < 0$ 的範圍，且所有試題皆未出現逆序現象，反應選項的閾值難易度以順序尺度自小到大批列。

同時，表 15 亦呈現「正向因應資源」中剩餘可歸類至特定向度試題之試題訊息量，其經過標準常態分配進行加權後，可知六題在正向因應資源 $-0.5 \leq \theta \leq 0.5$ 中具有最高試題訊息量數值，說明此六題在未知填答者的能力下較適合提供填答者填寫。部分試題在 $\theta > 0.5$ 上具有較高的試題訊息量數值，則說明此些試題對具有高度正向因應資源的填答者上存在較小的測量誤差；反之部分試題在 $\theta < -0.5$ 上對具有低度正向因應資源的填答者具有較小測量誤差。

表 15

「正向因應資源」之試題反應理論分析結果

編號	試題內容	$S - \chi^2$	RMSEA, $S - \chi^2$	α_M	MDIFF	試題訊息量		
						θ_M^a	θ_H^a	θ_L^a
4	我能夠從容地應付我的課業	210.98*	.014	1.34	-0.59	1.25	3.13	403.34
5	我知道我為何要讀書	148.90*	.007	1.13	-1.45	0.10	1.83	330.40
6	我是個心情愉快的人	147.09*	.012	2.60	-1.28	0.13	1.62	621.71
8	早上起床時，我覺得精神良好	194.35*	.016	1.27	-0.10	1.72	3.39	344.96
9	我對於自己休閒活動的安排感到滿意	195.39*	.014	1.52	-0.97	0.42	1.98	394.20
10	碰到困難時，我會冷靜下來去尋找解決困難的方法	144.63*	.011	1.57	-2.14	-0.21	1.79	399.32
11	一般而言，我喜歡自己，滿意自己	156.43*	.012	2.82	-0.99	0.12	1.36	806.94
12	當我有麻煩和煩惱時，我可以從家庭得到幫助	193.52*	.015	3.08	-1.15	-0.28	0.89	804.20
13	我滿意家人對我的情緒（喜、怒、哀、樂）表示關心和愛護的方式	191.70*	.016	4.10	-1.00	-0.08	1.02	1226.90
14	我和家人可以共度愉快的時光	156.22*	.013	3.47	-1.60	-0.51	0.76	859.30
15	當我嘗試新事物時，我會感到家人的接受與支持	178.20*	.014	2.79	-1.39	-0.37	0.81	780.64
16	當我遇到麻煩，能從老師、同學或朋友獲得協助	132.17*	.010	3.75	-1.55	-0.42	1.08	637.83
							238.85	203.79
							309.48	182.62
							253.15	164.41
							273.28	



表 15 (續)

「正向因應資源」之試題反應理論分析結果

編號	試題內容	$S - \chi^2$	RMSEA. $S - \chi^2$	α_M	MDIFF	試題訊息量					
						θ_M^a	θ_H^a	θ_L^a			
17	我滿意老師、同學或朋友關懷我情緒的方式	172.61 [*]	.015	4.59	-2.53	-1.43	-0.24	1.20	832.11	272.56	231.94
18	我能和老師、同學或朋友共度愉快的時光	165.57 [*]	.015	3.18	-2.98	-1.96	-0.66	1.06	472.45	198.33	172.05

註： $S - \chi^2$ ：符號卡方檢定統計量，RMSEA. $S - \chi^2$ ：基於符號卡方檢定之 RMSEA 統計量， α_M ：多向度試題鑑別度參數，MDIFF：多向度試題閾值參數。

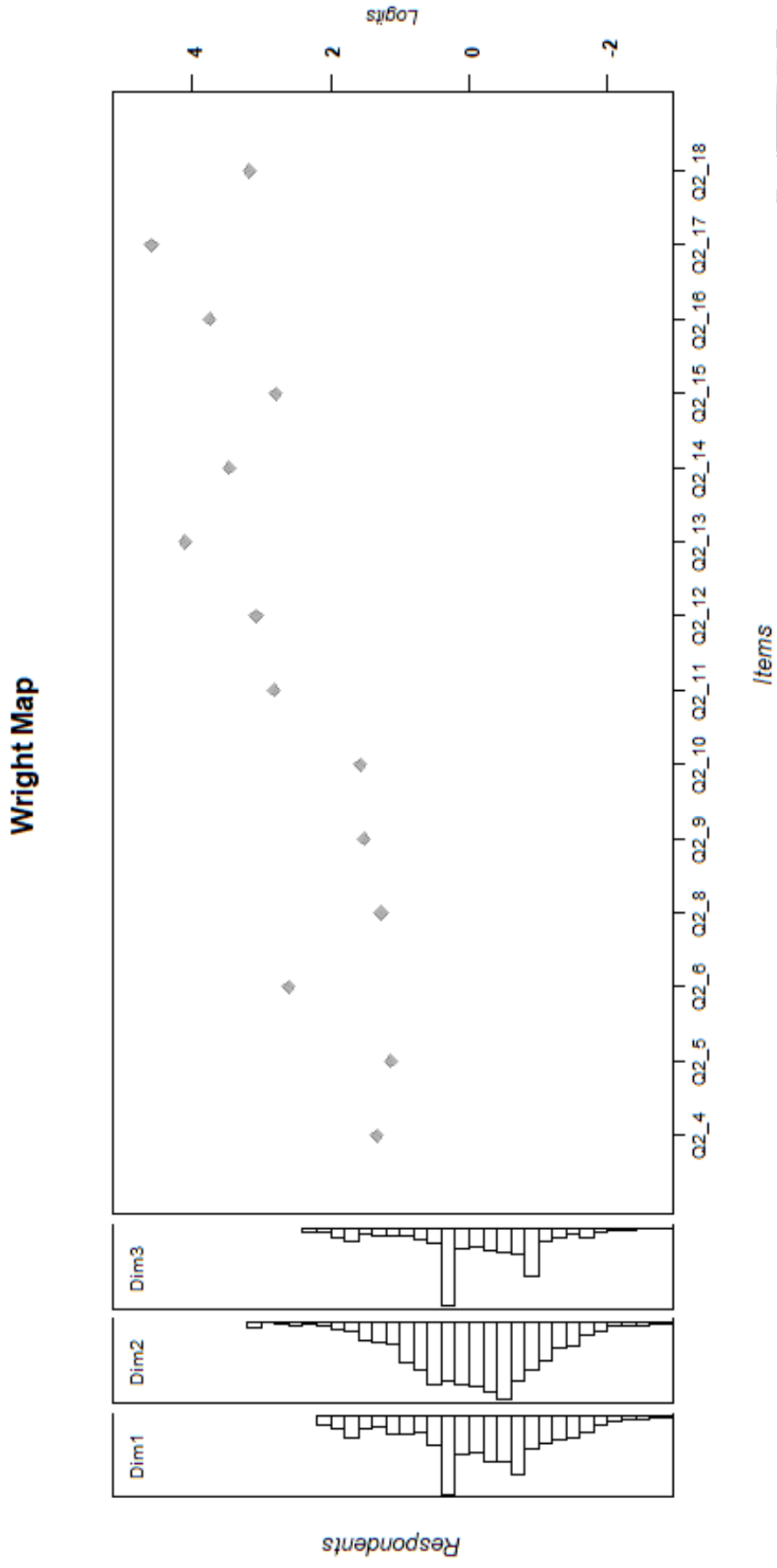
^a θ_M 為 $-.5 \leq \theta \leq .5$ 試題訊息量加權總和， θ_H 為 $\theta > .5$ 試題訊息量加權總和， θ_L 為 $\theta < -.5$ 訊息試題量加權總和。

* $p < .001$.



圖 2

正向因應資源之多向度試題鑑別度參數 (MDISC) 分布圖

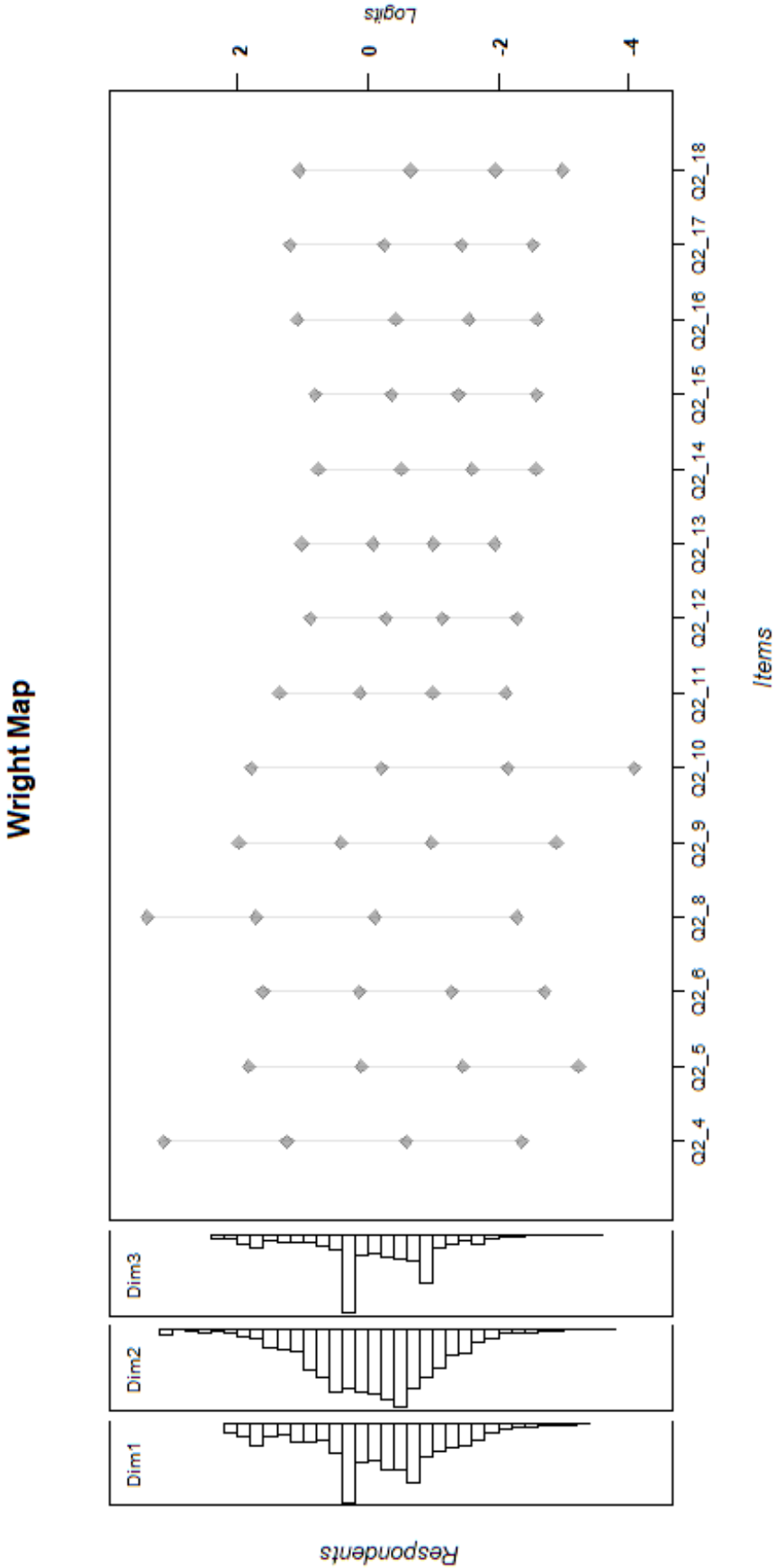


註：Dim1 為家庭資源向度的參與者能力分布，Dim2 為師友資源向度的參與者能力分布，Dim3 為自我資源向度的參與者能力分布，Q2 為正向因應資源於 2024 年施測測驗編號。



圖 3

正向因應資源之多向度試題閾值參數 (MDIFF) 分布圖



註：Dim1 為家庭資源向度的參與者能力分布，Dim2 為師友資源向度的參與者能力分布，Dim3 為自我資源向度的參與者能力分布，Q2 為正向因應資源於 2024 年施測測驗編號。



參、等級評分模式下的試題分析：負向情緒反應

測驗向度性部分，表 8 探索性因素分析顯示測驗試題可歸類至四個向度、且表 12 驗證性因素分析結果顯示該測驗在四個因素下具有良好的模型適配度。後續表 14 呈現經單向度與多向度模型假設考驗後亦支持測驗具有多個向度。

表 16 呈現「負向情緒反應」中剩餘可歸類至特定向度的試題在等級評分模式下的擬合表現、試題參數與試題訊息量。擬合表現上，採用 Orlando 與 Thissen (2003) 之符號卡方檢定 (signed chi-squared test [$S - \chi^2$]) 顯示所有試題之符號卡方統計量存在高度顯著，說明試題與測量模式的不適配情況；然而，RMSEA 結果顯示所有試題皆落在合理區間，表明試題與測量模式間的良好適配情況。試題參數上，為基於驗證性因素分析的結果而得。多向度試題鑑別度參數 (MDISC) 與多向度試題閾值參數 (MDIFF) 分別如圖 4 與圖 5 所示。自圖 4 可知鑑別度參數普遍落在 $\text{Logits} > 1$ 的範圍。圖 5 可知閾值參數部分落在 $\text{Logits} > 0$ 的範圍，說明填答者存在低負向情緒狀態時，儘管已填答「完全不同意」(第一點) 的最低選向下，仍可能在部分試題上被認為存在偏高之負向情緒狀態 (i.e., $\text{Logits} > 0$)；同時，自圖 5 可知所有試題皆未出現逆序現象，反應選項的閾值難易度以順序尺度自小到大排列。

同時，表 16 亦呈現「負向情緒反應」中剩餘可歸類至特定向度試題之試題訊息量，其經過標準常態分配進行加權後，可知二十六題中有十題在負向情緒反應中 $\theta < -.5$ 上對具有低度負向情緒反應的填答者具有較小測量誤差，其中「自殺行為與想法」和「精神病狀」向度上，所有試題對低度負向情緒反應的填答者具有較小測量誤差，說明這些向度上的試題較可反應出情緒狀態較佳填答者的真實情緒狀態。

表 16

「負向情緒反應」之試題反應理論分析結果

編號	試題內容	$S - \chi^2$	RMSEA. $S - \chi^2$	α_M	MDIFF	試題訊息量				
						θ_M^a	θ_H^a	θ_L^a		
1	我會因為不安與生氣，造成我的身體不舒服（如：拉肚子、頭痛、皮膚起疹子、手腳麻木等）	217.76	.008	1.14	0.16	1.77	3.56	337.39	112.12	114.14
2	我會覺得自己很神經質，很容易緊張，對事情都想得太多，考慮太多	258.17*	.013	1.47	-0.69	0.50	2.05	373.01	125.65	124.99
3	我會沒有理由地感到害怕、驚慌、不安	233.97*	.012	1.97	0.35	1.51	2.82	474.12	153.60	164.89
4	我會有一段時間心情低落、憂鬱、鬱悶	241.18*	.015	2.05	-0.75	0.78	2.29	477.39	163.42	164.10
5	我會有一段時間心情不好、容易苦惱、動怒、生氣	232.87*	.015	2.14	-0.42	1.06	2.48	504.22	172.84	171.73
6	我會有一段時間（至少持續一週）情緒高昂，容易興奮	277.70*	.011	0.74	1.40	3.59	5.96	296.05	98.29	101.83
7	日常生活中，我情緒會不可預測地忽高忽低，讓我感到困擾	230.02*	.011	1.71	0.48	1.73	3.03	469.36	150.77	164.92
9	我曾因生活上的變化（如親人離別、與人吵架等）導致情緒不穩與痛苦，進而影響工作與人際關係等	237.28	.010	1.15	0.62	2.35	4.06	377.02	121.30	128.71



表 16 (續)

「負向情緒反應」之試題反應理論分析結果

編號	試題內容	S - χ^2	RMSEA. S - χ^2	α_M	MDIFF			試題訊息量			
					θ_M^a	θ_H^a	θ_L^a				
10	我曾因不必要的重複動作（如：洗手、關門、關瓦斯）或因停不下來的想法（如：一直以為手是髒的，或一直以為可能會傷害他人，但我又明知那是不可能發生的事情）而心理感到不安與痛苦	286.57*	.014	1.76	0.27	1.53	2.41	3.29	452.69	92.84	168.41
11	我相信一些事情（如：有人跟蹤我，或有人要害我等等）但是別人認為那是不存在的，或是不可能發生的	265.04*	.020	2.31	0.52	1.82	2.61	3.40	452.07	83.77	181.69
12	我會聽到一些別人聽不到的聲音，看到別人看不到的事物。	258.96*	.008	2.01	1.21	2.55	3.56	4.42	407.98	75.23	225.02
15	我睡眠有困難，如：難以入睡、易醒來或早醒	238.41	.015	0.97	-1.36	0.50	2.11	3.82	321.41	105.30	109.20
17	我覺得自己比不過別人	271.39*	.013	1.60	-2.61	-1.21	0.27	1.60	551.99	181.77	184.75
18	我不滿意自己的學業或工作表現	254.74*	.011	1.56	-2.82	-1.06	0.66	2.02	443.22	153.80	149.69
19	我覺得自己無法做決定	265.16*	.010	1.15	-2.22	-0.21	1.43	3.07	279.14	94.33	94.19
20	我覺得自己是一文不值的人	194.38	.013	2.53	-0.46	0.69	1.82	2.73	471.31	153.52	158.92
21	我覺得人生很無望，對事物皆失去興趣	247.44*	.013	2.15	-0.47	0.88	2.07	3.06	447.98	141.20	153.06
22	我有尋死（自殺）的想法	202.87*	.014	2.70	0.54	1.77	2.83	3.66	398.53	86.62	170.09

表 16 (續)

「負向情緒反應」之試題反應理論分析結果

編號	試題內容	$S - \chi^2$	RMSEA. $S - \chi^2$	α_M	MDIFF	試題訊息量		
						θ_M^a	θ_H^a	θ_L^a
23	我有想傷害別人的念頭或衝動	213.72*	.012	1.25	1.14	2.76	4.21	5.39
24	我對未來不抱甚麼期待	208.90	.009	0.00	-0.68	0.68	2.10	3.20
25	我有尋死的計畫	245.36*	.021	2.85	1.40	2.36	3.40	3.99
26	我有做過傷害自己的事情	229.55*	.016	1.72	1.38	2.59	3.70	4.58
						294.81	76.84	141.63

註： $S - \chi^2$ ：符號卡方檢定統計量，RMSEA. $S - \chi^2$ ：基於符號卡方檢定之 RMSEA 統計量， α_M ：多向度試題鑑別度參數，

MDIFF：多向度試題閾值參數。

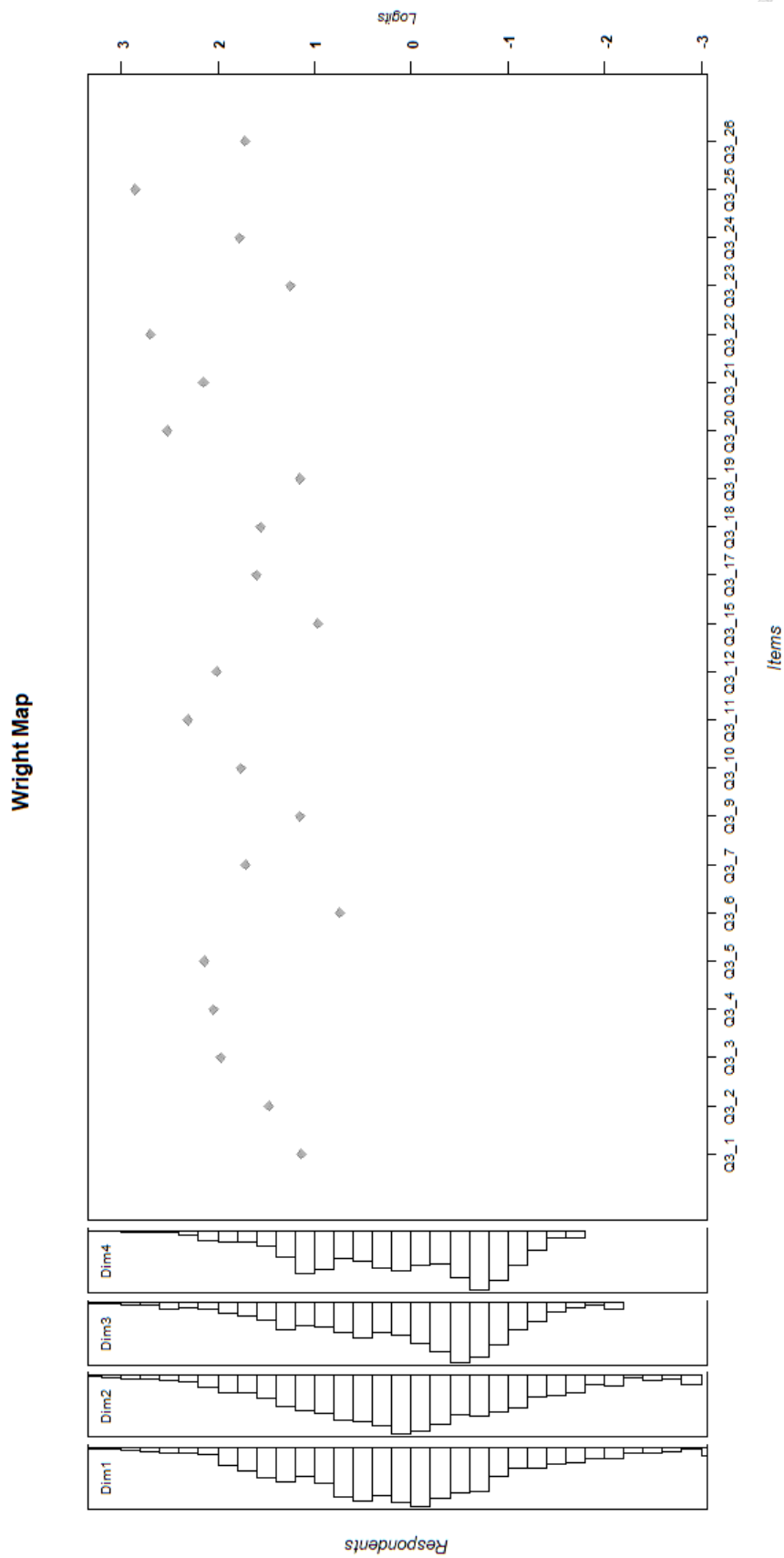
^a θ_M 為 $-.5 \leq \theta \leq .5$ 試題訊息量加權總和， θ_H 為 $\theta > .5$ 試題訊息量加權總和， θ_L 為 $\theta < -.5$ 訊息試題量加權總和。

* $p < .001$.



圖 4

負向情緒反應之多向度試題鑑別度參數 (MDISC) 分布圖

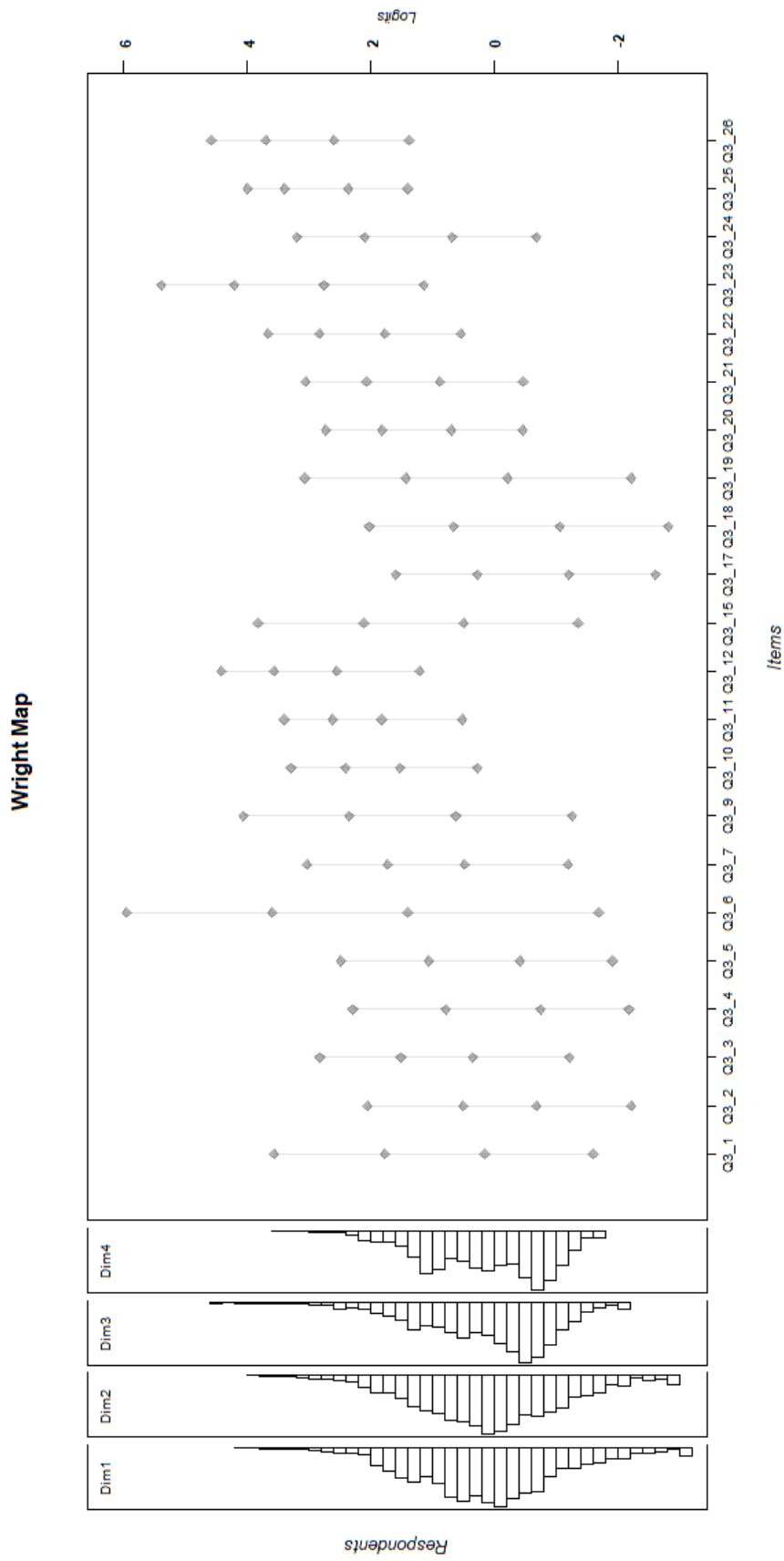


註：Dim1 為因應情緒行為向度的參與者能力分布，Dim2 為無望感向度的參與者能力分布，Dim3 為自傷想法行為向度的參與者能力分布，Dim4 為精神病狀向度的參與者能力分布，Q3 為負向情緒反應於 2024 年施測測驗編號。



圖 5

負向情緒反應之多向度試題閾值參數 (MDIFF) 分布圖



註：Dim1 為因應情緒行為向度的參與者能力分布，Dim2 為無望感向度的參與者能力分布，Dim3 為自傷想法行為向度的參與者能力分布，Dim4 為精神病狀向度的參與者能力分布，Q3 為負向情緒反應於 2024 年施測測驗編號。



肆、等級評分模式下的試題分析：多向度感知社會支持量表

測驗向度性部分，表 10 探索性因素分析顯示測驗試題可歸類至三個向度、且表 12 驗證性因素分析結果顯示該測驗在三個因素下具有良好的模型適配度。後續表 14 呈現經單向度與多向度模型假設考驗後亦支持測驗具有多個向度。

表 17 呈現「多向度感知社會支持量表」中剩餘可歸類至特定向度的試題在等級評分模式下的擬合表現、試題參數與試題訊息量。擬合表現上，採用 Orlando 與 Thissen (2003) 之符號卡方檢定 (signed chi-squared test [$S - \chi^2$]) 顯示所有試題之符號卡方統計量存在高度顯著，說明試題與測量模式的不適配情況；然而，RMSEA 結果顯示所有試題皆落在合理區間，表明試題與測量模式間的良好適配情況。試題參數上，為基於驗證性因素分析的結果而得。多向度試題鑑別度參數 (MDISC) 與多向度試題閾值參數 (MDIFF) 分別如圖 6 與圖 7 所示。自圖 6 可知鑑別度參數普遍落在 $\text{Logits} > 2$ 的範圍，較其他測驗的鑑別度參數高。圖 7 可知閾值參數普遍落在 $\text{Logits} < 0$ 的範圍，且發現自「沒意見」(第四點) 到「有些同意」(第五點) 的閾值參數落在 $\text{Logits} < 0$ ，說明填答者在多向度感知社會支持程度偏低下，仍有可能選擇「有些同意」的選項；同時，自圖 7 可知所有試題皆未出現逆序現象，反應選項的閾值難易度以順序尺度自小到大批列。

同時，表 17 亦呈現「多向度感知社會支持量表」中剩餘可歸類至特定向度試題之試題訊息量，其經過標準常態分配進行加權後，四題在社會支持程度 $-0.5 \leq \theta \leq 0.5$ 中具有最高試題訊息量數值，說明此四題在未知填答者的能力下較適合提供填答者填寫。另外八題中，有七題在 $\theta > 0.5$ 上具有較高的試題訊息量數值，則說明此些試題對具有高度社會支持程度的填答者上存在較小的測量誤差。

表 17

編號 7「多向度感知社會支持量表」之試題反應理論分析結果

編號	試題內容	$S - \chi^2$	RMSEA. $S - \chi^2$	α_M	MDIFF	試題訊息量		
						θ_M^a	θ_H^a	θ_L^a
1	每當我有需要時，有個特別的人會在我身邊幫助我	337.90*	.013	3.43	-2.20 -1.46 -0.81 -0.30 0.58 1.49	480.64	164.90	160.31
2	有個特別的人能讓我分享快樂與悲傷	336.55*	.014	4.38	-2.05 -1.41 -0.88 -0.46 0.27 1.17	681.80	232.60	228.47
3	我的家人真的會試圖幫助我	338.93*	.013	3.76	-2.47 -1.83 -1.18 -0.66 0.12 0.99	1400.25	480.98	189.56
4	我能從家人得到所需的情緒上的協助與支持	350.09*	.013	4.26	-1.90 -1.31 -0.83 -0.32 0.46 1.22	1892.31	576.40	318.34
5	有位特別的人是我慰藉的主要來源	416.12*	.016	3.12	-1.92 -1.29 -0.77 -0.24 0.49 1.26	461.67	155.15	155.58
6	我的朋友真的會試圖幫助我	288.60*	.013	3.47	-2.74 -2.14 -1.48 -0.79 0.24 1.30	691.74	260.49	207.41
7	當萬事不順遂時，我可以依賴我的朋友	342.16*	.014	3.69	-2.28 -1.67 -1.00 -0.44 0.51 1.42	782.64	293.03	243.45
8	我可以告訴家人我的煩惱	378.59*	.014	2.92	-1.84 -1.20 -0.61 -0.14 0.66 1.46	1060.86	300.13	240.64
9	我有朋友可以分享我的快樂與悲傷	321.24*	.013	3.42	-2.54 -1.96 -1.44 -0.95 0.13 1.11	649.52	241.35	194.51
10	我生命中有個特別的人會關心我的感受	345.85*	.014	3.37	-2.10 -1.50 -0.98 -0.40 0.30 1.07	499.09	169.14	166.15



表 17 (續)

編號 7「多向度感知社會支持量表」之試題反應理論分析結果

編號	試題內容	$S - \chi^2$	RMSEA. $S - \chi^2$	α_M	MDIFF				試題訊息量				
					θ_M^a	θ_H^a	θ_L^a	θ_M^a	θ_H^a	θ_L^a			
11	我的家人願意協助我做決定	359.14 [*]	.012	2.85	-2.43	-1.78	-1.22	-0.63	0.30	1.25	905.93	315.71	196.06
12	我可以告訴我的朋友我的煩惱	339.04 [*]	.013	3.11	-2.47	-1.92	-1.37	-0.86	0.19	1.16	641.92	239.88	190.50

註： $S - \chi^2$ ：符號卡方檢定統計量，RMSEA. $S - \chi^2$ ：基於符號卡方檢定之 RMSEA 統計量， α_M ：多向度試題鑑別度參數，

MDIFF：多向度試題閾值參數。

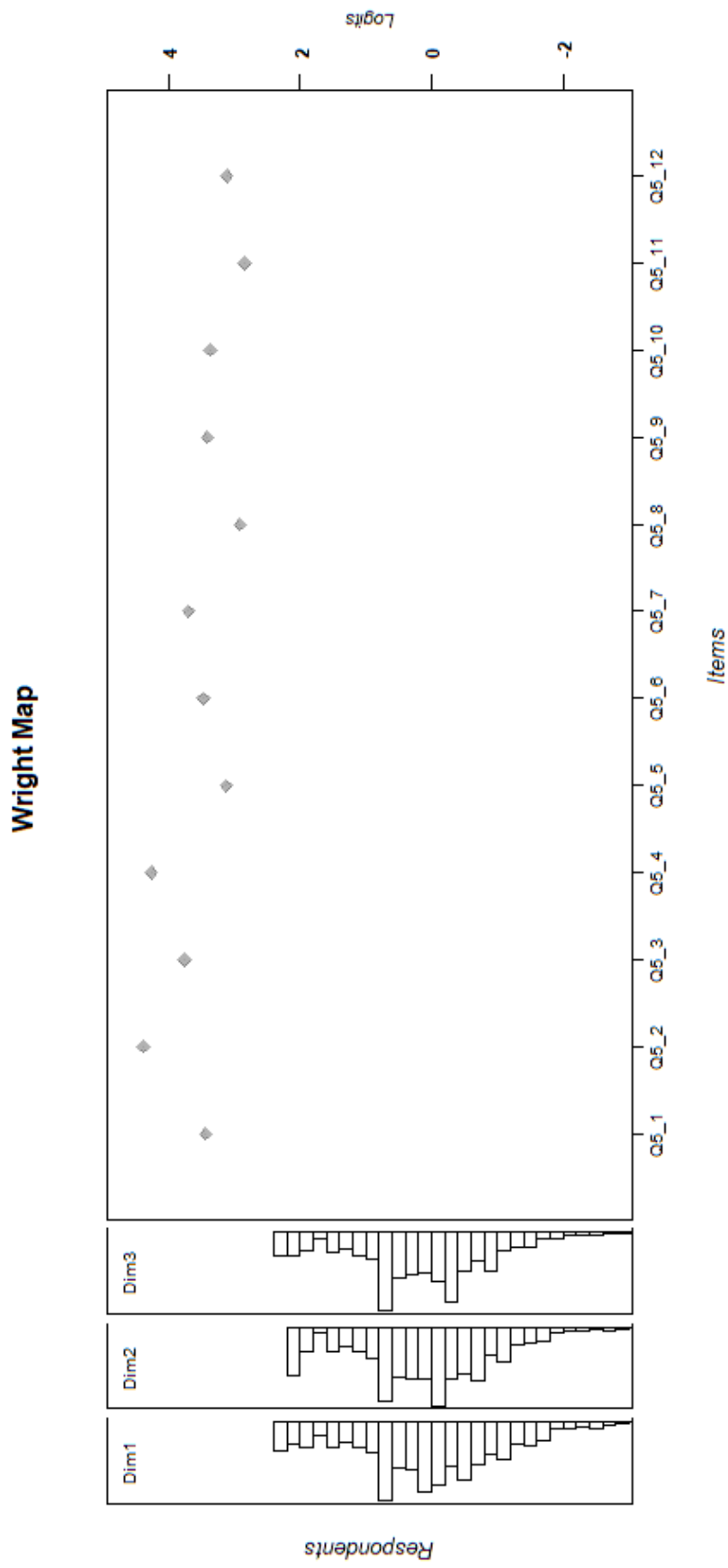
^a θ_M 為 $-.5 \leq \theta \leq .5$ 試題訊息量加權總和， θ_H 為 $\theta > .5$ 試題訊息量加權總和， θ_L 為 $\theta < -.5$ 訊息試題量加權總和。

* $p < .001$.



圖 6

多向度感知社會支持量表之多向度試題鑑別度參數 (MDISC) 分布圖

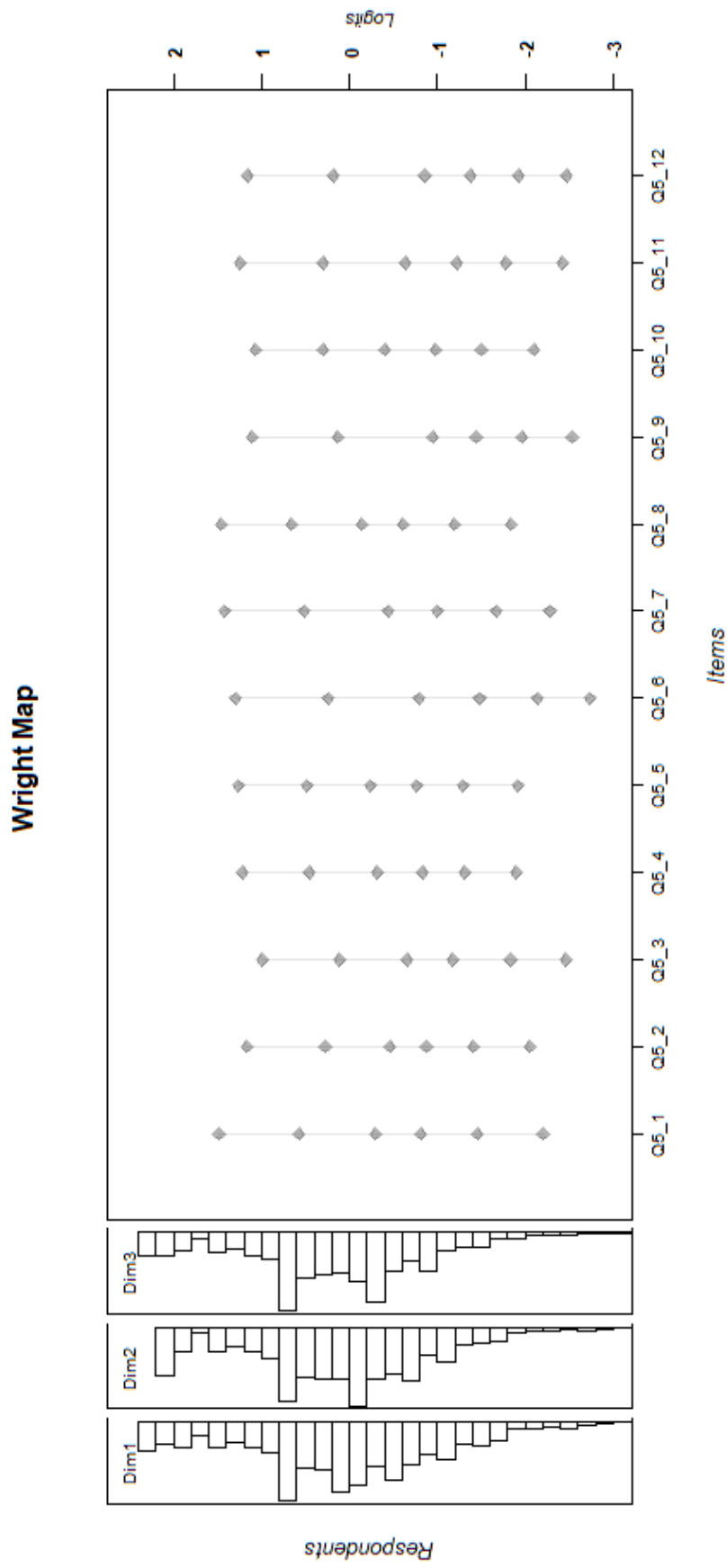


註：Dim1 為重要他人向度的參與者能力分布，Dim2 為家人向度的參與者能力分布，Dim3 為朋友向度的參與者能力分布，Q5 為多向度感知社會支持量表於 2024 年施測測驗編號。



圖 7

多向度感知社會支持量表之多向度試題閾值參數 (MDIFF) 分布圖



註：Dim1 為重要他人向度的參與者能力分布，Dim2 為家人向度的參與者能力分布，Dim3 為朋友向度的參與者能力分布，Q5 為多向度感知社會支持量表於 2024 年施測測驗編號。



伍、等級評分模式下的試題分析：自我滿意度

表 13 測驗向度性部分，探索性因素分析顯示測驗試題可歸類至兩個向度、且表 12 驗證性因素分析結果顯示該測驗在兩個因素下具有良好的模型適配度。後續表 14 呈現經單向度與多向度模型假設考驗後亦支持測驗具有多個向度。

表 18 呈現「自我滿意度」中剩餘可歸類至特定向度的試題在等級評分模式下的擬合表現、試題參數與試題訊息量。擬合表現上，採用 Orlando 與 Thissen (2003) 之符號卡方檢定 (signed chi-squared test [$S - \chi^2$]) 顯示所有試題之符號卡方統計量存在高度顯著，說明試題與測量模式的不適配情況；然而，RMSEA 結果顯示所有試題皆落在合理區間，表明試題與測量模式間的良好適配情況。試題參數上，為基於驗證性因素分析的結果而得。多向度試題鑑別度參數 (MDISC) 與多向度試題閾值參數 (MDIFF) 分別如圖 8 與圖 9 所示。自圖 8 可知鑑別度參數普遍落在 $\text{Logits} > 0$ 的範圍，除第三題與第四題的鑑別度參數偏高外，其餘試題大致落在 $0 < \text{Logits} < 2$ 的範圍中。圖 9 可知閾值參數普遍以 $\text{Logits} = 0$ 為切分點，高點數之閾值參數落在 $\text{Logits} > 0$ ，低點數則落在 $\text{Logits} < 0$ ；同時，自圖 9 可知所有試題皆未出現逆序現象，反應選項的閾值難易度以順序尺度自小到大排列。

同時，表 18 亦呈現「自我滿意度」中剩餘可歸類至特定向度試題之試題訊息量，其經過標準常態分配進行加權後，十六題中僅兩題在自我滿意程度 $-0.5 \leq \theta \leq 0.5$ 中具有最高試題訊息量數值，表示未知填答者的能力下存在兩題較適合提供填答者填寫。其他試題的試題訊息量數值不論在 $\theta > 0.5$ 與 $\theta < -0.5$ 、或在試題訊息量數值總和上皆無存在明顯的數值大小傾向，意即試題訊息量數值在 $\theta > 0.5$ 與 $\theta < -0.5$ 上近乎相等，且題目間的試題訊息量數值間較其他測驗而言落差較小。

表 18

「自我滿意度」之試題反應理論分析結果

編號	試題內容	$S - \chi^2$	RMSEA. $S - \chi^2$	α_M	MDIFF	試題訊息量		
						θ_M^a	θ_H^a	θ_L^a
1	我對於自己所就讀的科系感到	180.11	.008	1.16	-1.83	-0.22	1.84	26.73
2	我對於自己的學業或工作表現感到	242.85*	.016	1.90	-0.65	0.56	2.35	41.86
3	我對於自己的生涯規劃感到	232.61*	.018	4.33	-0.52	0.51	1.80	141.18
4	我對於自己的生活與人生目標感到	178.01*	.013	5.25	-0.63	0.29	1.52	266.77
6	我對於自己的人際關係感到	207.20*	.012	1.45	-1.16	0.18	2.18	25.96
7	我對於自己的情感關係感到	227.74*	.013	1.18	-0.90	0.61	2.30	46.26
8	我對於自己的情緒穩定度感到	171.42	.011	2.10	-0.81	0.18	1.51	64.90
9	我對於自己的個人特質感到	178.31*	.013	2.41	-1.10	0.03	1.50	78.57
10	我對於自己的身心健康感到	174.77*	.014	2.88	-0.72	0.26	1.54	92.33
11	我對於自己的自我掌控能力感到	182.28*	.013	2.33	-0.62	0.30	1.74	89.19
12	我對於自己的家庭狀況感到	257.40*	.014	0.98	-3.77	-2.20	-0.65	41.21
13	我對於自己的休閒活動安排感到	188.77*	.012	1.86	-2.59	-1.25	-0.02	60.21
14	我對於自己的運動習慣感到	181.14	.009	1.14	-2.16	-0.42	0.74	21.36
15	我對於自己的睡眠狀況感到	211.04*	.012	1.17	-2.07	-0.30	1.03	21.50
							2.83	44.78
								21.67

註： $S - \chi^2$ ：符號卡方檢定統計量，RMSEA. $S - \chi^2$ ：基於符號卡方檢定之 RMSEA 統計量， α_M ：多向度試題鑑別度參數，

MDIFF：多向度試題閾值參數。

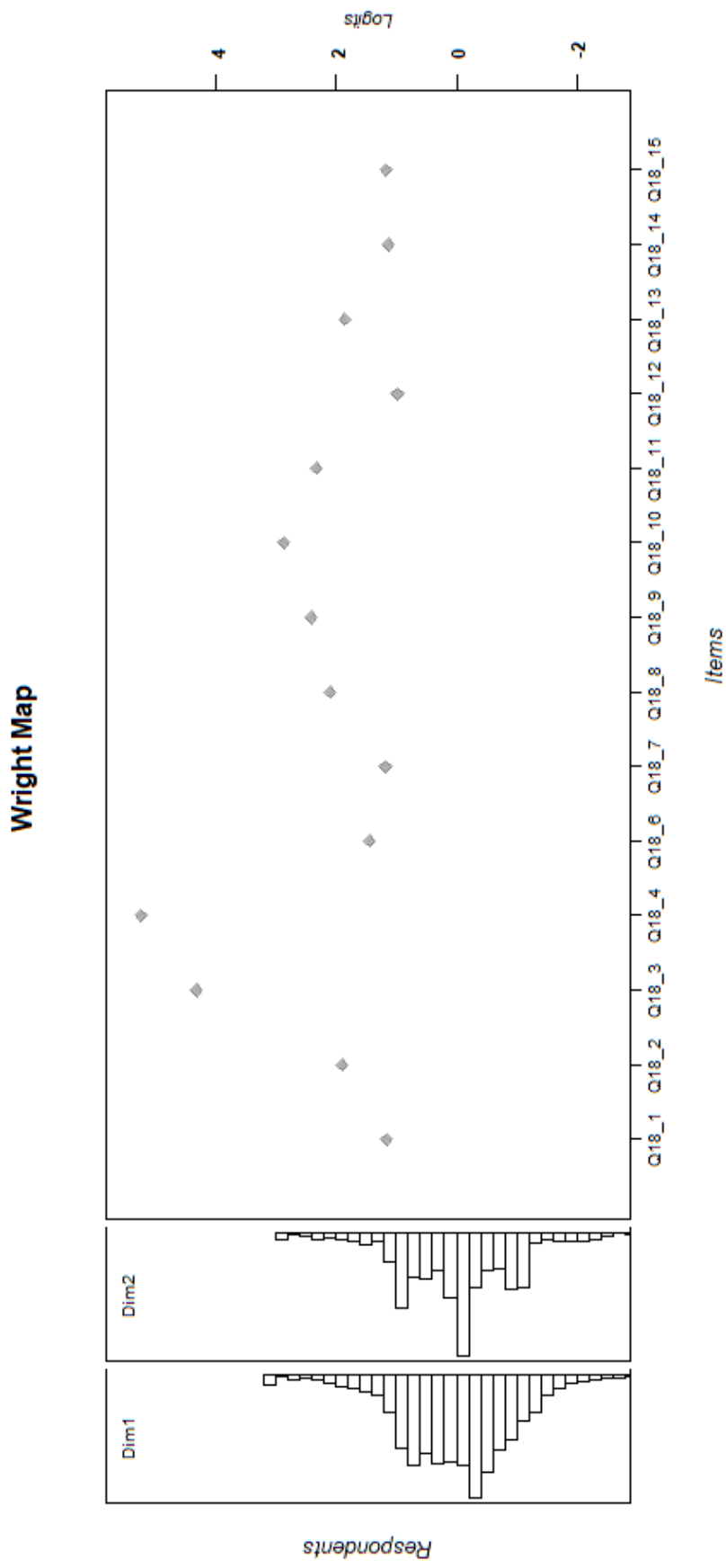
^a θ_M 為 $-0.5 \leq \theta \leq 0.5$ 試題訊息量加權總和， θ_H 為 $\theta > 0.5$ 試題訊息量加權總和， θ_L 為 $\theta < -0.5$ 訊息試題量加權總和。

* $p < .001$.



圖 8

自我滿意度之多向度試題鑑別度參數 (MDISC) 分布圖



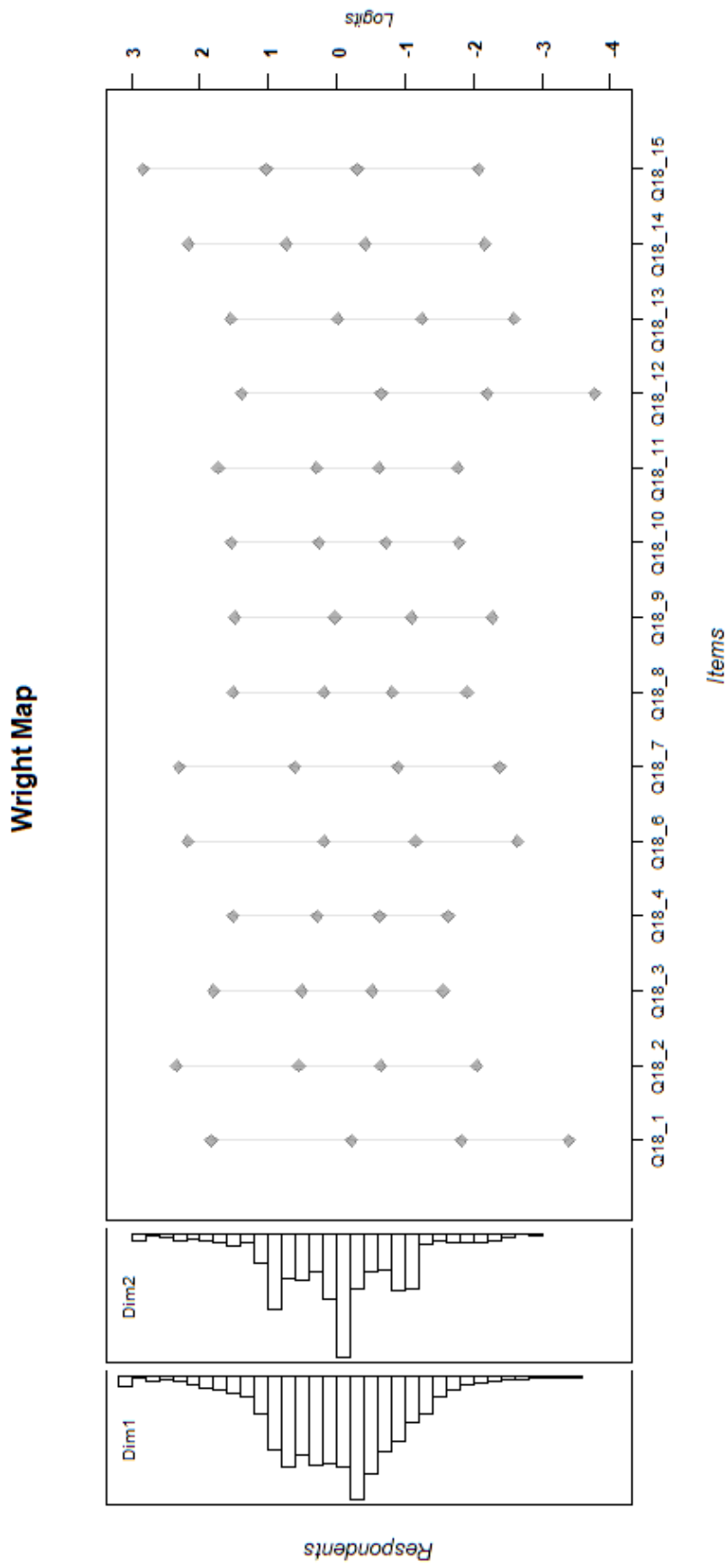
註：Dim1 為身心狀態滿意度向度的參與者能力分布，Dim2 為日常生活滿意度向度的參與者能力分布，Q18 為自我滿意度於

2024 年施測測驗編號。



圖 9

自我滿意度之多向度試題閾值參數 (MDIFF) 分布圖



註：Dim1 為身心狀態滿意度向度的參與者能力分布，Dim2 為日常生活滿意度向度的參與者能力分布，Q18 為自我滿意度於 2024 年施測測驗編號。



第四節 綜合兩測驗理論與實務之試題簡化結果

根據上述兩測驗理論的試題選擇結果，本研究將分別討論以上測驗的試題選擇排序，並將所有測驗內的試題分為：「適合所有群體施測試題」、「適合焦點群體施測試題」、「不建議提供施測試題」。其中，焦點群體是指基於身心健康評量系統建置目的：欲透過特定試題找尋需儘早介入關懷的同學，如：身心狀態較差的填答者。

壹、 正向因應資源

一、 適合所有群體施測試題

試題排序後結果如表 19 所示。「適合所有群體施測試題」總計六題，含括家庭資源（F1）與自我資源（F2）兩向度試題。基於古典測驗理論分析項目，試題難度適中、偏態係數與峰度係數未超出正負 1、探索性因素分析因素負荷量大於 .71；基於試題反應理論分析項目，在未知填答者真實正向因應資源能力下，此六題為訊息量在 $-.5 \leq \theta \leq .5$ 較大或未偏好較高或較低正向因應資源能力。

二、 適合焦點群體施測試題

基於測驗理論評估適合焦點群體施測試題上，編號 16「當我遇到麻煩，能從老師、同學或朋友獲得到協助」與編號 18「我能和老師、同學或朋友共度愉快的時光」兩者在「師友資源」向度上除具有高因素負荷量（皆大於 .71）外，為釐清焦點群體在「師友資源」向度的狀態，考量單一因素需具有三道試題較能完整進行估計，故將上述兩題列入「適合焦點群體施測試題」。編號 6「我是個心情愉快的人」與編號 4「我能夠從容地應付我的課業」中，則以試題反應理論的分析結果為主。前者為訊息量在 $-.5 \leq \theta \leq .5$ 較大，後者為在 θ_L 上訊息量較 θ_H 高，表明此題對捕捉「正向因應資源能力低」填答者的狀態佳。為釐清焦點群體在「自我資源」向度的狀態，亦考量單一因素需具有三道試題較能完整進行估計。綜上基於試題反應理論之訊息量評估後，將上述兩題列入「適合焦點群體施測試題」中。

基於測驗實務觀點評估焦點群體施測試題上，編號4「我能夠從容地應付我的課業」雖然存在較低的因素負荷量（因素負荷量為.63），但過往Luthar等人（2020）在其研究中表明頂尖大學學生在學業方面較其他大學存在高度困擾，如：學業成績表現不如預期，學業表現與個人付出不成正比等狀態較易存在位處頂尖大學的學生中。對此，考量填答者在學業方面的應對可協助校方輔導體系釐清填答者在自我與學業之間的關係下，本研究將此題列入「適合焦點群體施測試題」中。編號6「我是個心情愉快的人」則因為身心健康評量系統為評量填答者之身心狀態，此題目敘述可直接反映填答者的情緒狀態下，本研究將此題列入「適合焦點群體施測試題」中。

三、 不建議提供施測試題

在「不建議提供施測試題」中，四道題目為具有交叉因素負荷量者，無法確切釐清此四題對應之測量因素為何。剩餘四題在「自我資源」向度上具有較低因素負荷量，且在試題訊息量表現上較其餘適合所有群體和焦點群體施測試題差，故將此八題列為不建議提供施測試題。

表 19

「正向因應資源」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA. $S-\chi^2$	θ_M^a	θ_H^a	θ_L^a
適合所有群體施測試題												
13	我滿意家人對我的情緒 (喜、怒、哀、樂)表示關 心和愛護的方式	3.33	1.09	-0.22	-0.64	.87	F1	.91	.016	1226.90	356.64	273.28
14	我和家人可以共度愉快的時 光	3.70	0.95	-0.44	-0.20	.88	F1	.87	.013	859.30	309.48	182.62
17	我滿意老師、同學或朋友關 懷我情緒的方式	3.57	0.86	-0.24	-0.17	.81	F3	.86	.015	832.11	272.56	231.94
11	一般而言，我喜歡自己，滿 意自己	3.28	1.02	-0.17	-0.50	.79	F2	.72	.012	806.94	247.44	199.55
12	當我有麻煩和煩惱時，我可 以從家庭得到幫助	3.47	1.08	-0.31	-0.69	.89	F1	.82	.015	804.20	248.09	176.82
15	當我嘗試新事物時，我會感 到家人的接受與支持	3.60	1.03	-0.35	-0.56	.90	F1	.79	.014	780.64	253.15	164.41

適合焦點群體施測試題

16	當我遇到麻煩，能從老師、 同學或朋友獲得協助	3.67	0.87	-0.38	-0.06	.84	F3	.81	.010	637.83	238.85	203.79
6	我是個心情愉快的人	3.34	0.88	-0.11	-0.22	.79	F2	.66	.012	621.71	204.78	187.81
18	我能和老師、同學或朋友共 度愉快的時光	3.80	0.78	-0.41	0.26	.87	F3	.74	.015	472.45	198.33	172.05
4	我能夠從容地應付我的課業	2.77	0.94	0.08	-0.26	.81	F2	.63	.014	403.34	131.37	134.93



表 19 (續)

「正向因應資源」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA. $S-\chi^2$	θ_M^a	θ_H^a	θ_L^a
不建議提供施測試題												
1	家人對我的生活方式感到滿意	3.41	0.90	-0.19	-0.35		CL					
2	家人對我的學習與學業目標感到滿意	3.73	0.92	-0.46	-0.16		CL					
3	我的生活目標同時符合我的與父母的需求	3.41	0.96	-0.22	-0.41		CL					
5	我知道我為何要讀書	3.34	1.07	-0.21	-0.61	.82	F2	.53	.007	330.40	110.30	104.88
7	在別人面前，我不怕說出自己的缺點或過去的失敗	3.13	1.06	-0.01	-0.73		CL					
8	早上起床時，我覺得精神良好	2.57	0.91	0.40	0.02	.82	F2	.53	.016	344.96	111.45	118.12
9	我對於自己休閒活動的安排感到滿意	3.15	1.00	0.04	-0.61	.81	F2	.56	.014	394.20	125.27	126.24
10	碰到困難時，我會冷靜下來去尋找解決困難的方法	3.56	0.81	-0.17	-0.11	.81	F2	.58	.011	399.32	133.71	127.54

註： α ：排除特定試題 Cronbach's alpha；F：所屬因素；FL：所屬因素之因素負荷量；RMSEA. $S - \chi^2$ ：基於符號卡方檢定之

RMSEA 統計量；CL (Cross Loading)：跨向度間具有近似因素負荷量。

^a θ_M 為 $-.5 \leq \theta \leq .5$ 試題訊息量加權總和， θ_H 為 $\theta > .5$ 試題訊息量加權總和， θ_L 為 $\theta < -.5$ 訊息試題量加權總和。





貳、 負向情緒反應

一、 適合所有群體施測試題

試題排序後如表 20 所示。「適合所有群體施測試題」總計五題，包括因應情緒行為 (F1) 與無望感 (F2) 兩向度試題。基於古典測驗理論之分析項目，試題難度適中、偏態係數與峰度係數未超出正負 1、探索性因素分析因素負荷量大於 .71；基於試題反應理論分析項目，在未知填答者真實負向情緒狀態下，此五題為訊息量在 $-.5 \leq \theta \leq .5$ 較大、或未偏好較高或較低負向情緒狀態。同時，在心情溫度計中有三題被列在此歸類中，顯示心情溫度計的部分題目亦可提供給一般同學評估自身的情緒狀態。

二、 適合焦點群體施測試題

基於測驗理論評估適合焦點群體施測試題中，編號 3「我會沒有理由地感到害怕、驚慌、不安」在試題反應理論中顯示其試題訊息量總和偏高，比部分「適合所有群體施測試題」適合捕捉填答者之負向情緒狀態，故將此題列入「適合焦點群體施測試題」中。編號 22「我有尋死（自殺）的想法」、編號 25「我有尋死的計畫」與編號 26「我有做過傷害自己的事情」中，因對個體自傷想法行為的貢獻性強（因素負荷量皆大於 .71），故將此三題列入「適合焦點群體施測試題」中。

測驗實務觀點評估焦點群體施測試題上，編號 15「我睡眠有困難，如：難以入睡、易醒來或早醒」為心情溫度計試題，儘管在因素負荷量或試題訊息量上表現較其他分類在「適合所有群體施測試題」之心情溫度計試題差，但因心情溫度計屬臨床區辨高風險學生的有效實用工具，對此本研究仍將編號 15 列入「適合焦點群體施測試題」中。編號 26「我有做過傷害自己的事情」儘管對高負向情緒能力者的捕捉上較編號 23「我有想傷害別人的念頭或衝動」弱，然考量本身心理健康評量系統欲找尋存在具有「自我」傷害風險高的填答者，著重在自己傷害上，故本研究將編號 26 列入「適合焦點群體施測試題」中。

三、 不建議提供施測試題

在「不建議提供施測試題」中，四道題目為具有交叉因素負荷量者，無法確切釐清此四題對應之測量因素為何，故將此四題列為不建議提供施測試題。

「因應情緒行為」向度五題、「無望感」向度三題、「自傷想法行為」一題具有較低因素負荷量，儘管部分試題具有較佳的試題訊息量表現，但經由上述臨床實務需求考量後，仍將此九題列為不建議提供施測試題。

另針對精神病狀向度上的三題，因身心健康評量系統建置為提供校方簡便捕捉負向情緒狀態較高者進行關懷、並提供填答者快速釐清目前自身身心狀態的工具，考量本向度上試題屬臨床特定病況描述，難以適用對非特定病況且具有高度負向情緒狀態填答者，該向度三試題與身心健康評量系統建置目的不符而列為不適合提供施測試題。

表 20

「負向情緒反應」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA	θ_M^a	θ_H^a	θ_L^a
適合焦點群體施測試題												
17 ^b	我覺得自己比不過別人	3.36	1.12	-0.21	-0.67	.86	F2	.77	.015	551.99	181.77	184.75
5 ^b	我會有一段時間心情不好、容易苦惱、動怒、生氣	2.88	1.02	0.04	-0.45	.86	F1	.77	.015	504.22	172.84	171.73
4 ^b	我會有一段時間心情低落、憂鬱、鬱悶	3.06	1.02	-0.09	-0.39	.86	F1	.75	.015	477.39	163.42	164.10
20	我覺得自己是一文不值的人	2.14	1.08	0.78	-0.07	.85	F2	.72	.009	471.31	153.52	158.92
18	我不滿意自己的學業或工作表現	3.19	1.04	0.03	-0.51	.86	F2	.74	.013	443.22	153.80	149.69
適合焦點群體施測試題												
3 ^b	我會沒有理由地感到害怕、驚慌、不安	2.46	1.07	0.46	-0.47	.86	F1	.62	.012	474.12	153.60	164.89
15 ^b	我睡眠有困難，如：難以入睡、易醒來或早醒	2.41	1.13	0.48	-0.54	.88	F1	.44	.008	321.41	105.30	109.20
22 ^b	我有尋死（自殺）的想法	1.49	0.78	1.74	2.99	.70	F3	.78	.013	398.53	86.62	170.09
25	我有尋死的計畫	1.23	0.58	2.96	10.21	.71	F3	.75	.021	314.41	73.79	233.42
26	我有做過傷害自己的事情	1.28	0.63	2.71	8.32	.75	F3	.59	.016	294.81	76.84	141.63



表 20 (續)

「負向情緒反應」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA, $S - \chi^2$	θ_M^a	θ_H^a	θ_L^a
不建議提供施測試題												
1	我會因為不安與生氣，造成我的身體不舒服（如：拉肚子、頭痛、皮膚起疹子、手腳麻木等）	2.54	1.09	0.30	-0.63	.87	F1	.55	.008	337.39	112.12	114.14
2	我會覺得自己很神經質，很容易緊張，對事情都想得太多，考慮太多	3.15	1.17	-0.13	-0.85	.87	F1	.62	.013	373.01	125.65	124.99
6	我會有一段時間（至少持續一週）情緒高昂，容易興奮	2.18	0.94	0.70	0.25	.89	F1	.42	.011	296.05	98.29	101.83
7	日常生活中，我情緒會不可預測地忽高忽低，讓我感到困擾	2.40	1.06	0.53	-0.33	.86	F1	.63	.011	469.36	150.77	164.92
8	我對一般人不感到害怕的環境（如：高處、人群、封閉的地方或出家門），我卻感到極度害怕，以至於想要離開	2.08	1.01	0.83	0.23		CL					
9	我曾因生活上的變化（如親人離別、與人吵架等）導致情緒不穩與痛苦，進而影響工作與人際關係等	2.31	1.01	0.49	-0.29	.87	F1	.49	.010	377.02	121.30	128.71



表 20 (續)

「負向情緒反應」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA, $S-\chi^2$	θ_M^a	θ_H^a	θ_L^a
10	我曾因不必要的重複動作 （如：洗手、關門、關瓦斯）或因停不下來的想法 （如：一直以為手是髒的，或一直以為可能會傷害他人，但我又明知那是不可能發生的事情）而心理感到不安與痛苦	1.75	0.98	1.33	1.25	.66	F4	.60	.014	452.69	92.84	168.41
11	我相信一些事情（如：有人跟蹤我，或有人要害我等）但是別人認為那是不能存在的，或是不可能發生的	1.55	0.82	1.68	2.81	.58	F4	.69	.014	452.07	83.77	181.69
12	我會聽到一些別人聽不到的聲音，看到別人看不到的事物。	1.29	0.59	2.30	6.17	.70	F4	.59	.020	407.98	75.23	225.02
13	我有「將要生病」的感覺	2.03	0.99	0.70	-0.15		CL					
14	我覺得心情不好，沒辦法應付周圍發生的事	2.43	1.00	0.37	-0.30		CL					
16	我覺得處在壓力中	3.46	1.00	-0.30	-0.24		CL					
19	我覺得自己無法做決定	2.75	1.08	0.24	-0.57	.88	F2	.55	.011	279.14	94.33	94.19



表 20 (續)

「負向情緒反應」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA, $S-\chi^2$	θ_M^a	θ_H^a	θ_L^a
21	我覺得人生很無望，對事物皆失去興趣	2.06	1.02	0.83	0.18	.86	F2	.63	.013	447.98	141.20	153.06
23	我有想傷害別人的念頭或衝動	1.37	0.69	2.10	4.85	.81	F3	.45	.012	313.87	79.51	129.95
24	我對未來不抱甚麼期待	2.17	1.05	0.66	-0.15	.87	F2	.61	.009	384.41	124.98	130.21

註： α ：排除特定試題 Cronbach's alpha；F：所屬因素；FL：所屬因素之因素負荷量；RMSEA.S- χ^2 ：基於符號卡方檢定之 RMSEA 統計量；CL (Cross Loading)：跨向度間具有近似因素負荷量。

^a θ_M 為 $-.5 \leq \theta \leq .5$ 試題訊息量加權總和， θ_H 為 $\theta > .5$ 試題訊息量加權總和， θ_L 為 $\theta < -.5$ 訊息試題量加權總和。

^b 為類心情溫度計試題。



參、多向度感知社會支持量表



一、適合所有群體施測試題

試題排序後如表 21 所示。「適合所有群體施測試題」共計四題，皆屬家庭資源向度 (F2) 基於古典測驗理論之分析項目，試題難度適中、試題偏態係數與峰度係數未超出正負 1、探索性因素分析因素負荷量大於 .71；基於試題反應理論分析項目，此四題為訊息量在 $-0.5 \leq \theta \leq 0.5$ 較大，且未偏好較高或較低社會支持程度。

二、適合焦點群體施測試題

基於測驗理論評估適合焦點群體施測試題中，編號 6「我的朋友真的會試圖幫助我」、編號 7「當萬事不順遂時，我可以依賴我的朋友」與編號 12「我可以告訴我的朋友我的煩惱」除對朋友向度的貢獻性強（因素負荷量皆大於 .71）外，在試題反應理論下的試題訊息量表現與「適合所有群體施測試題」於 θ_L 區間接近，表明這三題可與「適合所有群體施測試題」共同捕捉對朋友支持向度較低之填答者狀態，故將此三題列入「適合焦點群體施測試題」中。

測驗實務觀點評估焦點群體施測試題中，編號 9「我有朋友可以分享我的快樂與悲傷」與編號 12「我可以告訴我的朋友我的煩惱」上，前者著重在內在正向與負向情緒感受，後者則涵蓋負向情緒與對填答者日常生活造成困擾事項，考量身心健康評量系統著重在負向情緒感受的測量，故本研究將編號 12 納入「適合焦點群體施測試題」中。

三、不建議提供施測試題

在「不建議提供施測試題」中，四題為「重要他人」向度；然而因每位填答者對重要他人的定義不同，未來校方進行身心輔導時較難掌握填答者所指對象下，儘管其因素負荷量偏高，本研究經實務考量後仍列為不建議提供施測試題。另外一題為上述提及的編號 9「我有朋友可以分享我的快樂與悲傷」，考量身心健康評量系統著重在負向情緒感受的測量，故僅納入編號 12 進行施測。

表 21

「多向度感知社會支持量表」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA, $S-\chi^2$	θ_M^a	θ_H^a	θ_L^a
適合所有群體施測試題												
4	我能從家人得到所需的情緒上的協助與支持	4.86	1.64	-0.59	-0.38	.89	F2	.91	.013	1892.31	576.40	318.34
3	我的家人真的會試圖幫助我	5.30	1.46	-0.79	0.12	.90	F2	.86	.013	1400.25	480.98	189.56
8	我可以告訴家人我的煩惱	4.58	1.72	-0.41	-0.72	.91	F2	.82	.014	1060.86	300.13	240.64
11	我的家人願意協助我做決定	5.16	1.45	-0.78	0.29	.92	F2	.78	.012	905.93	315.71	196.06
適合焦點群體施測試題												
7	當萬事不順遂時,我可以依賴我的朋友	5.02	1.43	-0.65	0.08	.90	F3	.81	.014	782.64	293.03	243.45
6	我的朋友真的會試圖幫助我	5.36	1.23	-0.80	0.84	.90	F3	.81	.013	691.74	260.49	207.41
12	我可以告訴我的朋友我的煩惱	5.38	1.34	-0.97	0.98	.90	F3	.81	.013	641.92	239.88	190.50
不建議提供施測試題												
1	每當我有需要時,有個特別的人會在我身邊幫助我	4.74	1.60	-0.47	-0.56	.92	F1	.77	.013	480.64	164.90	160.31
2	有個特別的人能讓我分享快樂與悲傷	4.96	1.66	-0.67	-0.40	.91	F1	.85	.014	681.80	232.60	228.47
5	有位特別的人是我慰藉的主要來源	4.73	1.76	-0.49	-0.68	.92	F1	.85	.016	461.67	155.15	155.58



表 21 (續)

「多向度感知社會支持量表」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA, $S - \chi^2$	θ_M^a	θ_H^a	θ_L^a
9	我有朋友可以分享我的快樂與悲傷	5.46	1.30	-1.04	1.29	.90	F3	.78	.013	649.52	241.35	194.51
10	我生命中有個特別的人會關心我的感受	4.99	1.68	-0.62	-0.43	.92	F1	.82	.014	499.09	169.14	166.15

註： α ：排除特定試題 Cronbach's alpha；F：所屬因素；FL：所屬因素之因素負荷量；RMSEA.S - χ^2 ：基於符號卡方檢定之 RMSEA 統計量；CL (Cross Loading)：跨向度間具有近似因素負荷量。

^a θ_M 為 $-0.5 \leq \theta \leq 0.5$ 試題訊息量加權總和， θ_H 為 $\theta > 0.5$ 試題訊息量加權總和， θ_L 為 $\theta < -0.5$ 訊息試題量加權總和。



肆、自我滿意度

一、適合所有群體施測試題

試題排序後如表 22 所示，兩題被列為「適合所有群體施測試題」。基於古典測驗理論之分析項目，試題難度適中、試題偏態係數與峰度係數未超出正負 1、探索性因素分析因素負荷量大於 .71；基於試題反應理論分析項目，此兩題為訊息量在 $-0.5 \leq \theta \leq 0.5$ 較大，且未偏好較高或較低自我滿意程度。

二、適合焦點群體施測試題

基於測驗理論評估適合焦點群體施測試題中，編號 10「我對於自己的身心健康感到」在因素負荷量上達到 .76 外，其在總體試題訊息量上為總測驗為第三高數值的試題，且對低自我滿意度的填答者具有較佳的能力捕捉程度下本研究將其納入適合焦點群體施測試題。

基於測驗實務觀點評估焦點群體施測試題中，編號 2「我對於自己的學業或工作表現感到」為 Luthar 等人（2020）所提及位處頂尖大學學生的壓力來源之一，編號 13「我對於自己的休閒活動安排感到」與世界衛生組織對生活品質的定義：「生活品質是指個人在所生活的文化價值體系中，對於自己的目標、期望、標準、關心等方面的感受程度」（Costanza et al., 2007）之個人目標安排與期望相同，編號 15「我對於自己的睡眠狀況感到」則與心情溫度計試題近似下，儘管上述三題皆存在因素負荷量偏低的情況，但基於過往研究與測驗目標群體考量後，將這三題納入適合焦點群體施測試題。

三、不建議提供施測試題

在「不建議提供施測試題」中，兩題為具有交叉因素負荷量者，其餘試題為因素負荷量過低，基於古典測驗理論將列為不建議提供施測試題。另在編號 8「我對於自己的情緒穩定度感到」上，考量本評量系統已有負向情緒反應進行完整評估，故將其列為不建議提供施測試題。



表 22

「自我滿意度」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要						
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA. $S - \chi^2$	θ_M^a	θ_H^a	θ_L^a	
適合所有群體施測試題													
4	我對於自己的生活與人生目標感到	3.12	1.01	-0.16	-0.53	.74	F2	.81	.013	266.77	104.91	96.52	
3	我對於自己的生涯規劃感到	2.97	0.99	-0.07	-0.47	.73	F2	.90	.018	357.06	141.18	131.16	
適合焦點群體施測試題													
10	我對於自己的身心健康感到	3.16	1.04	-0.16	-0.59	.83	F1	.76	.014	92.33	44.13	40.64	
2	我對於自己的學業或工作表現感到	2.99	0.99	-0.10	-0.57	.79	F2	.56	.016	91.54	41.86	39.36	
13	我對於自己的休閒活動安排感到	3.42	0.98	-0.32	-0.35	.84	F1	.62	.012	60.21	29.47	27.96	
15	我對於自己的睡眠狀況感到	2.78	1.09	0.17	-0.71	.86	F1	.50	.012	44.78	21.50	21.67	
不建議提供施測試題													
1	我對於自己所就讀的科系感到	3.51	0.99	-0.44	-0.18	.85	F2	.45	.008	54.39	26.73	25.51	
5	我對於自己的時間管理能力感到	2.52	1.09	0.34	-0.65		CL						
6	我對於自己的人際關係感到	3.27	0.99	-0.33	-0.40	.85	F1	.50	.012	53.21	25.96	24.82	
7	我對於自己的情感關係感到	3.07	1.10	-0.10	-0.64	.85	F1	.45	.013	46.26	22.41	22.29	
8	我對於自己的情緒穩定度感到	3.23	1.10	-0.23	-0.66	.84	F1	.68	.011	64.90	31.41	30.00	
9	我對於自己的個人特質感到	3.39	0.97	-0.33	-0.29	.84	F1	.65	.013	78.57	38.54	34.94	
11	我對於自己的自我掌控能力感到	3.08	1.08	-0.14	-0.74	.84	F1	.68	.013	89.19	41.28	38.33	
12	我對於自己的家庭狀況感到	3.70	1.05	-0.63	-0.13	.86	F1	.42	.014	41.21	20.11	19.77	

表 22 (續)

「自我滿意度」簡化試題篩選排序表

編號	試題內容	古典測驗理論結果摘要					試題反應理論結果摘要					
		平均數	標準差	偏態	峰度	α	F	FL	RMSEA. $S-\chi^2$	θ_M^a	θ_H^a	θ_L^a
14	我對於自己的運動習慣感到	2.92	1.19	0.15	-0.91	.86	F1	.47	.009	44.50	21.36	21.52
16	我對於環境對我自己性別認同的支持感到	4.15	0.85	-0.75	0.18		CL					

註： α ：排除特定試題 Cronbach's alpha；F：所屬因素；FL：所屬因素之因素負荷量；RMSEA. $S - \chi^2$ ：基於符號卡方檢定之 RMSEA 統計量；CL (Cross Loading)：跨向度間具有近似因素負荷量。

^a θ_M 為 $-.5 \leq \theta \leq .5$ 試題訊息量加權總和， θ_H 為 $\theta > .5$ 試題訊息量加權總和， θ_L 為 $\theta < -.5$ 訊息試題量加權總和。



第五章 討論



本研究為探討古典測驗理論與試題反應理論在國立臺灣大學身心健康評量系統建置計畫（行為與社會科學倫委會審查案號：202106HS019）中試題篩選與簡化策略，並基於既有納入建置計畫的四份測驗分量表：「正向因應資源」、「負向情緒反應」、「多向度社會支持量表」與「自我滿意度」討論不同測驗理論下的簡化策略與簡化結果。最後針對上述四種量表中的試題分為「適合所有群體施測試題」、「適合焦點群體施測試題」與「不建議施測試題」作為最終簡化結果呈現。本研究經由古典測驗理論與試題反應理論策略簡化後，在「正向因應資源」、「負向情緒反應」、「多向度社會支持量表」與「自我滿意度」中「適合所有群體施測試題」分別為六題、五題、四題與兩題。相較原先測驗總題數分別為十八題、二十六題、十二題與十六題（共計七十二題），經考量評量系統建置目的：捕捉負向情緒程度較高的填答者介入關懷，並基於測驗理論與實務後分別保留十題、十題、七題與六題（共計三十三題）。本研究除協助遴選出適合納入最終評量系統的試題外，亦針對既有測驗試題進行簡化：自所有試題中保留適合所有身心狀態的參與者與協助捕捉負向情緒程度較高的參與者的題目。

第一節 試題簡化結果討論：試題層面

壹、「正向因應資源」

「正向因應資源」中，以古典測驗理論進行題目簡化時，因題庫內試題本身已具有完整因素結構，且試題平均數與標準差等描述統計量各試題間數值表現穩定，故較難依此進行本測驗的題目簡化；然而，考量 Comrey 與 Lee（2013）對極度優良試題的標準需符合因素負荷量高於 .71 情況下，部分試題因低於 .71 的標準而被排除。後續採用試題反應理論進行題目簡化時，因著重在對填答者的能力估計誤差：以試題訊息量為工具，除可協助判斷試題對參與者不同正向因應資源程度的測量精確性外，可觀察到在正向因應資源中「適合

焦點群體施測試題」的試題對偏高正向因應資源程度具有較良好的測量精確性。由於身心健康評量系統的目的為完整捕捉填答者之正向因應資源程度，故本研究建議將「適合所有群體施測試題」與「適合焦點群體施測試題」納入系統評量。綜上所述，在正向因應資源中，由古典測驗理論可協助釐清試題的因素結構及其試題平均數標準差的表現，當採用較嚴格之因素負荷量標準下仍發現試題表現皆良好穩定時，可採用試題反應理論的試題訊息量評估對不同能力程度下的能力估計精確性，作為未來測驗試題簡化的工具。

貳、「負向情緒反應」

「負向情緒反應」為本研究中與評量系統建置目標最為相關的測驗。當以古典測驗理論進行題目簡化時，儘管試題具有穩定的因素結構，但可發現填答者在高程度負向情緒狀態因素：「自我傷害行為」與「精神病狀」中，試題平均數偏低，且試題偏態與峰度係數偏高下，表明多數填答者選擇「非常不同意」或「不同意」。對此，基於古典測驗理論的試題選擇標準下，這些因素下的試題會被視為不適合施測試題。

這兩項因素中，「自我傷害行為」內存在廣為使用評估填答者心理情緒狀態的「心情溫度計」試題。心情溫度計（Brief Symptom Rating Scale [BSRS-5]; Lee et al., 2003）為五題加一題自殺意念試題的設計，臨床上主要做為第一線醫師篩檢精神疾患狀態的工具。因該測驗工具具有穩健的信效度指標，在臨床實務中儘管背離測驗理論的試題選擇標準，但仍有其保留的必要性。除此之外，基於試題反應理論的試題選擇上，當採用試題訊息量評估試題對不同程度負向情緒反應的測量精確性時，可發現負向情緒反應中「適合焦點群體施測試題」的試題對所有負向情緒反應程度群體具有較良好的測量精確性，反映此因素的試題亦可有效反映所有填答者之負向情緒反應能力狀態。進一步觀察上述提及的「自我傷害行為」內存在之「心情溫度計」試題，可發現以試題反應理論為簡化基礎下的「適合焦點群體施測試題」群中，「自我傷害行為」內存在之「心情溫度計」試題為「適合焦點群體施測試題」之候選試題，此發現亦佐證心情溫度計的試題設計本身可做為第一線醫師捕捉精神疾患身心狀態的工具；然而，因「自我傷害行為」因素試題較多填答者填寫非常不同意與不同意，採用

試題訊息量進行試題選擇時則會出現試題對低度負向情緒反應填答者具有高度測量精確性，其反應填答選項對試題訊息量選題造成的盲點：無法透過試題訊息量捕捉適合高度負向情緒反應者填答的試題。儘管如此，本研究仍透過總體試題訊息量捕捉到適合所有填答者施測、且屬「自我傷害行為」向度的試題，表明總體訊息量可做為找尋高度負向情緒反應者適合試題的統計工具之一。

「精神病狀」上，在古典測驗理論與試題反應理論架構下的試題表現同「自我傷害行為」，說明對高程度負向情緒狀態的測量精確性較高；然而因身心健康評量系統主要設計為給一般學生群體進行施測填寫，且針對高風險群體進行快速評估下，相較上述提到的心情溫度計，評量系統建置上會以區分具有自我傷害之高風險群體與非具有自我傷害之高風險群體為主，且因臨床精神症狀較為複雜，以既有負向情緒反應量表的試題評量難以捕捉真實填答者之情緒狀態，如：當參與者本身並未具有類強迫症症狀，但具有輕度憂鬱風險時，其在原測驗中第十題「我曾因不必要的重複動作（如：洗手、關門、關瓦斯）或因停不下來的想法（如：一直以為手是髒的，或一直以為可能會傷害他人，但我又明知那是不可能發生的事情）而心理感到不安與痛苦」傾向回答「非常不同意」或「不同意」，但在原測驗中第二十一題「我覺得人生很無望，對事物皆失去興趣」上可能傾向回答「同意」或「非常同意」。對此，相較利用臨床精神症狀進行負向情緒的評估，本研究基於評量系統建置以區分自我傷害高風險群體為目的，並考量負向情緒在臨床精神症狀上的表現難以少部分進行評估下，建議不列入「精神病狀」的試題進入身心健康評量系統。

綜上所述，在負向情緒反應上，基於古典測驗理論簡化中，適合高度負向情緒程度填答者的試題因試題平均數與試題答題傾向而被視為不適合施測試題；然而在試題反應理論的試題簡化程序中，當進一步考量試題對高度負向情緒程度填答者的測量精確性時，發現僅觀察 $\theta > .5$ （高度負向情緒狀態）無法協助臨床實務者捕捉適合該能力群體施測的試題，需綜觀所有能力群體之總體訊息量較可捕捉適合高度負向情緒狀態的試題；同時，既有題庫中部份「心情溫度計」試題為「適合焦點群體施測試題」，表明基於試題反應理論的試題訊息量可協助捕捉適合高度負向情緒程度填答者的試題，且捕捉的試題結果亦與過往臨床所使用評估精神疾患身心狀態的工具相符。本研究建議未來在臨床測驗

的簡化程序上加入試題訊息量進行評估，作為捕捉適合施測族群的試題簡化工具。



參、「多向度感知社會支持量表」

「多向度感知社會支持量表」中，為一長期在建置過程中施測且具有穩定因素結構的測驗，儘管其在古典測驗理論的探索性因素分析與過往分析結果一致，但卻存在試題平均數偏高的情況（為七點量尺，多數試題平均數大於五點）。Zimet（2016）針對該測驗試題平均點數進行測驗群體的社會支持程度進行討論，當試題平均點數大於 5.1 時，該群體將被視為具有高度社會支持。此外，該研究亦表明採用試題平均點數進行群體區分時（如：區分成高度社會支持、中度社會支持、低度社會支持），可增加該測驗的測驗效度。綜合上述內容，透過多向度社會支持量表進行施測時，2024 年填答的學生被視為具有高度社會支持的情況。

上述測驗在試題反應理論架構下進行簡化時，保留三題與朋友支持相關的試題，其與前述「正向因應資源」的師友資源近似，試題文字與語意上仍與「正向因應資源」保留的題目存在差異下，此三題因具有高度捕捉所有施測群體的社會支持程度，本研究建議保留此測驗。因測驗題目較少，如需細緻釐清填答者的家庭支持程度與朋友支持程度時，Zimet（2016）表明本測驗可利用分向度試題分數平均數觀察填答者家庭支持程度或朋友支持程度傾向：填答平均分數大於 5.1 者為高度社會支持，填答分數介於 3 至 5 分則為中度社會支持，填答分數介於 1 至 2.9 分時為低度社會支持；同時，其研究亦表明可透過所有試題分數平均數觀察填答者整體的感知社會支持程度，其分數判准同上方描述：填答平均分數大於 5.1 者為高度社會支持，填答分數介於 3 至 5 分則為中度社會支持，填答分數介於 1 至 2.9 分時為低度社會支持。

肆、「自我滿意度」

「自我滿意度」中，為今年新加入施測的測驗。在古典測驗理論架構進行試題簡化下，多數試題接落入可保留施測試題。其中，可發現第十六題：「我對於環境對我自己性別認同的支持感到」在測驗的兩個因素上具有近似的因素負

荷量外，且存在校正後相關係數偏低。根據測驗本身所具有的因素：「身心狀態滿意度」與「日常生活滿意度」觀之，此試題在「環境」描述上屬日常生活，而「自己性別認同的支持」描述屬身心狀態，故題目描述可能為此試題在測驗的兩個因素上具有近似的因素負荷量。進一步觀察此題目的試題平均數，可發現其平均數為 4.15（本量表為五點量尺），顯示多數填答者在此題填答上落在「非常滿意」或「滿意」，難以對不同自我滿意度程度之試題總分區辨，可能為其校正後相關係數偏低之原因。

在試題反應理論架構下進行試題簡化時，相較前三種測驗藉由試題訊息量容易區分試題至「適合所有群體施測試題」、「適合焦點群體施測試題」與「不建議施測試題」作為最終簡化結果，本測驗中的試題在高自我滿意度程度和在低自我滿意程度上具有近似的能力估計偏誤，反映多數試題可協助捕捉不同情緒狀態程度下填答者的自我滿意程度。本研究經考量 Comrey 與 Lee（2013）對極度優良試題的標準需符合因素負荷量高於 .71 作為簡化條件時，僅捕捉個體對「身心狀態滿意度」的試題兩題；後續經由對低自我滿意度存在較高的試題訊息量觀之，發現「我對於自己的身心健康感到」、「我對於自己的學業或工作表現感到」、「我對於自己的休閒活動安排感到」與「我對於自己的睡眠狀況感到」存在對低自我滿意度填答者的偏高強度能力捕捉。對於本研究在此量表上的簡化試題結果，與世界衛生組織對生活品質的定義：「生活品質是指個人在所生活的文化價值體系中，對於自己的目標、期望、標準、關心等方面的感受程度」（Costanza et al., 2007）含括個體對「身心健康滿意度」與「日常生活滿意度」相同（『適合所有群體施測試題』兩題，『適合焦點群體施測試題』四題，兩者加總為六題）。對此，日後納入身心健康評量系統的建議施測試題。本研究建議採用上述提及的六題版本進行施測。

總結而言，填答者在回應對自我的滿意度時，採用古典測驗理論可捕捉含有測量多重概念的試題（如：第十六題），但在基於試題反應理論下，試題訊息量在自我滿意度的試題選擇或簡化上較無法反映試題適合高或低自我滿意度程度的填答者，但仍可從試題反應理論之試題訊息量捕捉適合施測的試題，為利於後續試題的簡化策略之一。

第二節 試題簡化結果討論：策略層面

本研究在試題簡化採用的標準上，為針對試題選擇策略與標準進行回顧。試題簡化上，為針對既有已發展測驗選擇更適合提供施測群體的試題，其策略近似試題選擇，但會採取較高的標準，以利捕捉較佳提供施測的試題。以古典測驗理論或試題反應理論為基礎的試題選擇上，多數文獻採用試題難度、試題鑑別度、校正後試題總分相關等統計量進行試題選擇；然而，多份有關測驗編製或回顧性文章表明上述統計量在身心相關測驗選擇較佳試題或簡化時難以以過往的試題優劣判別標準或指標進行題目選擇（Reise & Waller, 2009; Gregory, 2015; Thomas, 2019）。對此，本研究在試題簡化上，並未將上述三者統計量作為絕對刪除或保留試題的統計策略，而是會同步考量測驗試題於臨床實務的使用需求後再行判斷是否有保留的必要性。簡言之，試題簡化上除了「量」性考量外，也需進行「質」性考量。上述質性與量性同步考量的題目簡化策略在 Thomas（2019）回顧性文章中被視為較佳的臨床身心評量測驗或系統選擇試題的方式，儘管可能造成違反部分統計理論或測驗理論假設，然當試題具有臨床使用必要性時，加入施測可有效區辨高低風險填答者。

在基於試題反應理論的簡化上，為先驗證測量模型的適配度，而後搭配試題訊息量進行選題的方式。針對驗證測驗模型的適配度上，本研究在測驗模型使用上採用多向度等級評分模式進行討論，並輔以符號卡方檢定評估試題對測量模型的適配程度。在試題反應理論中具有多種測量模式，這些測量模式皆在已知試題受到何種參數影響下進行模式建構，屬資料導向，故本研究在測驗模型使用上，為假定所有試題皆受試題鑑別參數與試題閾值參數影響，不存在其他參數影響下進行討論；然而，在「多向度社會支持量表」所得各向度之社會支持程度分布圖（如圖 6 與圖 7）及各填答者在試題的答題反應中，發現部分高度社會支持填答者會在部分試題上選擇低社會支持程度的選項，而部分低社會支持填答者則在部分試題上選擇高社會支持程度選項，上述狀態於 Smith（1996）於著作專欄中表明該現象為選項序列存在非預期回答（noisy progression），並說明此現象反映測驗本身具有良好的因素結構，但可能存在更高階的因素影響填答者的答題反應。儘管符號卡方檢定顯示其模型適配度良好，但仍需考量樣本數過大造成卡方檢定之型一錯誤率上升問題。

而在試題訊息量的使用上，為協助捕捉在特定填答者能力下對其能力具有最小估計偏誤的試題工具。其結果在分析的四份測驗中，可發現部分試題在能力中等時存在高的試題訊息量，其數值大於能力偏高下所有試題訊息量數值之總和，表明部分試題極適合提供能力中等的填答者施測。綜上所述，本研究採用試題訊息量協助捕捉適合特定填答者群體試題的工具，且採用算則為對填答者具有最小能力估計偏誤變異法，在結果呈現上亦可提供試題簡化的協助；然試題訊息量數值本身僅反映單一試題在特定能力下的能力估計偏誤程度，基於不同試題反應理論測量模型得到的數值涉及的模型參數不同，無法跨測驗進行數值比較，但在單一測驗中可藉由試題訊息量的相對大小選擇測驗中較佳施測試題。

壹、基於上述選題策略之試題效度討論

針對上述進行的題目簡化結果，其最後簡化的試題如表 25 呈現。簡化試題中，「多向度感知社會支持量表」與「正向因應資源」測量的因素概念近似，皆可分類成三個因素：自我資源、家庭資源與師友資源。

本階段將進行簡化試題同時效度的討論。首先，針對 2024 年同時納入題庫施測的測驗進行選擇，並找出對應相同測量概念的測驗或測驗向度進行相關係數的計算。如表 1 所示，除「多向度感知社會支持量表」與「正向因應資源」彼此可對應測量因素向度，預期兩者對應的測量因素間具有高度正相關外，因「寂寞感量表」(ULS-5) 為測量填答者之人際疏離感程度，本研究預期「寂寞感量表」會與上述兩量表之簡化結果存在負相關。另外，「負向情緒反應」簡化結果則預期與「憂鬱—焦慮—壓力量表」(DASS-9) 存在正相關；同時，「自我滿意度」簡化結果則分別和「生活滿意度量表」(SWLS) 存在正相關，和「世界衛生組織生活品質量表八題簡版」(EUROHIS-QOL 8-item index) 存在正相關。

後續針對上述提及之量表進行皮爾森相關係數的計算。「多向度感知社會支持量表」與「正向因應資源」在家庭資源向度具有高度正相關 ($r = .97, p < .001$)，師友資源向度具有高度正相關 ($r = .99, p < .001$)，ULS-5 的結果如表 23 所示。上述表明家庭資源向度、師友資源向度具有高度重疊性，本研究為簡

化身心理健康評量系統試題，考量「正向因應資源」試題組中多測量填答者之「自我資源」下，建議可僅採用「正向因應資源」，以達到簡化評量系統試題與對填答者的因應資源支持程度釐清。

另外在「負向情緒反應」結果上，存在與 DASS-9 的高度正相關，反應簡化後的試題可反應個體在憂鬱、焦慮和壓力上的情緒感受；「自我滿意度」中和 SWLS 及 EUROHIS-QOL 8-item index 皆具有高度正相關，反應簡化後的試題可表徵填答者在生活滿意程度，並與世界衛生組織生活品質的定義具有高度相關性。

其餘測量身心狀態的測驗上，心理韌性量表（The Connor-Davidson Resilience Scale [CD-RISC]; Connor, & Davidson, 2003）與多向度感知社會支持量表存在低相關，但與其他量表存在高度相關，反應正向因應資源在自我資源的納入與心理韌性欲測量的概念相近；安適幸福感量表（Peace of Mind [POM]; Lee et al., 2013）中，與負向情緒反應具有高度負相關，與自我滿意度存在高度正相關，與正向因應資源及多向度感知社會支持呈現中度正相關，反應負向情緒與自我滿意度對個體感知幸福感可預測性，亦即高負向情緒者，可能形成幸福感偏低情況，反之自我滿意度高者，可能存在幸福感偏高。

同時，表 24 提供簡化試題與原先量表的得分一致性，結果指出簡化試題總分與原先長版量表總分具有高度相關，表明經本研究進行試題簡化流程之簡化版本可提供與完整版量表相近的分數傾向。相較原先版本存在較多試題需進行填答，簡化版本儘管提供較少試題填答，但仍提供測驗實務者對測驗填答者在「正向因應資源」、「負向情緒反應」、「多向度感知社會支持量表」與「自我滿意度」方面的狀態釐清。

綜上所述，基於上述選題策略的試題簡化內容進行效度分析，針對分量表之分向度進行同時效度分析上，其結果皆與研究預期相符，反應既有簡化試題的策略可有效維持與原先欲測量的心理變項（如：家庭資源、負向情緒反應等）的相關性，並與其他身心量表之間存在中度或高度相關性，表明簡化的試題結果可對填答者的心理韌性與幸福感進行預測。然本身心理健康評量系統目標為捕捉存在高度自傷風險的填答者，故在最終效度評估上，可將既有捕捉存在

高度自傷風險的填答者工具與簡化試題結果進行比較，以評估簡化試題的系統對捕捉高風險填答者的能力。



表 23

試題簡化之同時效度報表：以皮爾森相關係數為效度檢驗工具

	正向因應資源	負向情緒反應	多向度感知 社會支持	自我滿意度
ULS-5	-.52	.54	-.54	-.49
DASS-9	-.44	.72 ^b	-.36 ^a	-.62
SWLS	.58	-.61	.49	.74 ^b
EUROHIS- QOL 8-item index	.61	-.64	.48	.77 ^b
CD-RISC	.51	-.62	.37 ^a	.64
POM	.56	-.72 ^b	.42	.70 ^b

註：ULS-5 為寂寞感量表；DASS-9 為憂鬱－焦慮－壓力量表；SWLS 為生活滿意度量表；EUROHIS-QOL 8-item index 為世界衛生組織生活品質量表八題簡版；CD-RISC 為心理韌性量表；POM 為安適幸福感量表。

^a 具有低度相關。

^b 具有高度相關。

表 24

完整量表與簡化量表間相關：以皮爾森相關係數為填答者得分一致性檢驗工具

	正向因應資源	負向情緒反應	多向度感知 社會支持	自我滿意度
正向因應資源	.96 ^b			
負向情緒反應	-.52	.95 ^b		
多向度感知 社會支持	.71 ^b	-.38 ^a	.91 ^b	
自我滿意度	.56	-.69	.42	.93 ^b

註：對角線為完整量表與簡化量表間相關，下三角為各測驗簡化版本間相關。

^a 具有低度相關。

^b 具有高度相關。

表 25

基於本研究試題簡化策略下的評量系統試題列表

簡化 編號	試題來源與編號	試題來源 所屬向度	試題描述
1	正向因應資源 13	F1 家庭資源	我滿意家人對我的情緒（喜、怒、哀、樂）表示關心和愛護的方式
2	正向因應資源 14	F1 家庭資源	我和家人可以共度愉快的時光
3	正向因應資源 17	F3 師友資源	我滿意老師、同學或朋友關懷我情緒的方式
4	正向因應資源 11	F2 自我資源	一般而言，我喜歡自己，滿意自己
5	正向因應資源 12	F1 家庭資源	當我有麻煩和煩惱時，我可以從家庭得到幫助
6	正向因應資源 15	F1 家庭資源	當我嘗試新事物時，我會感到家人的接受與支持
7	正向因應資源 16	F3 師友資源	當我遇到麻煩，能從老師、同學或朋友獲得協助
8	正向因應資源 18	F3 師友資源	我能和老師、同學或朋友共度愉快的時光
9	正向因應資源 4	F2 自我資源	我能夠從容地應付我的課業
10	正向因應資源 6	F2 自我資源	我是個心情愉快的人
11	負向情緒反應 17 ^a	F2 無望感	我覺得自己比不過別人
12	負向情緒反應 5 ^a	F1 因應情緒行為	我會有一段時間心情不好、容易苦惱、動怒、生氣
13	負向情緒反應 4 ^a	F1 因應情緒行為	我會有一段時間心情低落、憂鬱、鬱悶
14	負向情緒反應 20	F2 無望感	我覺得自己是一文不值的人
15	負向情緒反應 18	F2 無望感	我不滿意自己的學業或工作表現
16	負向情緒反應 3 ^a	F1 因應情緒行為	我會沒有理由地感到害怕、驚慌、不安
17	負向情緒反應 15 ^a	F1 因應情緒行為	我睡眠有困難，如：難以入睡、易醒來或早醒
18	負向情緒反應 22 ^a	F3 自傷想法行為	我有尋死（自殺）的想法
19	負向情緒反應 25	F3 自傷想法行為	我有尋死的計畫
20	負向情緒反應 26	F3 自傷想法行為	我有做過傷害自己的事情



表 25 (續)

基於本研究試題簡化策略下的評量系統試題列表

簡化 編號	試題來源與編號	試題來源 所屬向度	試題描述
21	多向度感知社會支持 4	F2 家庭資源	我能從家人得到所需的情緒上的協助與支持
22	多向度感知社會支持 3	F2 家庭資源	我的家人真的會試圖幫助我
23	多向度感知社會支持 8	F2 家庭資源	我可以告訴家人我的煩惱
24	多向度感知社會支持 11	F2 家庭資源	我的家人願意協助我做決定
25	多向度感知社會支持 7	F3 朋友資源	當萬事不順遂時，我可以依賴我的朋友
26	多向度感知社會支持 6	F3 朋友資源	我的朋友真的會試圖幫助我
27	多向度感知社會支持 12	F3 朋友資源	我可以告訴我的朋友我的煩惱
28	自我滿意度 4	F2 日常生活滿意	我對於自己的生活與人生目標感到
29	自我滿意度 3	F2 日常生活滿意	我對於自己的生涯規劃感到
30	自我滿意度 10	F1 身心狀態滿意	我對於自己的身心健康感到
31	自我滿意度 2	F2 日常生活滿意	我對於自己的學業或工作表現感到
32	自我滿意度 13	F1 身心狀態滿意	我對於自己的休閒活動安排感到
33	自我滿意度 15	F1 身心狀態滿意	我對於自己的睡眠狀況感到

註：粗體部分為本研究經效度檢驗後建議移除試題。

^a 為心情溫度計試題。



第三節 研究限制與未來研究方向



本研究著重於試題選擇策略，然而在測驗編製或評量系統建置流程中，觀看最終效度仍有其必要性。評量系統建置目標為捕捉存在高度自傷風險的填答者，可將既有捕捉存在高度自傷風險的填答者工具與簡化試題結果進行比較，以評估本研究基於兩測驗理論與臨床實務應用提供試題選擇的建議與建置目標的一致性。

除此之外，本研究在試題分數轉換上仍需進行進一步的討論。「正向因應資源」、「負向情緒反應」與「自我滿意度」為五點李克特氏量尺，「多向度社會支持量表」則為七點李克特氏量尺，兩者不論在古典測驗理論或試題反應理論上的分數轉換仍需進行研析。分數等化與轉換之統計策略上，Kolen 與 Brennan (2014) 在其著作討論不同的測驗間的分數等化與轉換方式。基於古典測驗理論的分數轉換上，除了採用線性方式進行轉換外，亦可採用百分等級方式進行轉換。線性方式為假定兩測驗的分數間為線性關係，其具有一等式關係，百分等級方式則透過對兩測驗總分進行排序，依次對應其分數關係。然上述僅涉及兩測驗，多項測驗上的分數轉換仍需要有更多研究或文獻支持。在試題反應理論上，Kolen 與 Brennan (2014) 在其著作提到試題反應理論測量模型為非線性測量模型，其分數轉換上可採用上述百分等級方式進行轉換。

相較上述分數等化與轉換之統計因應策略，部分學者會使用定錨題 (Anchor item) 等非統計策略進行分數等化或轉換。相較上述統計因應策略在多項測驗分數等化或轉換上的繁瑣程序，因定錨題為一在多個測驗皆存在相同測量概念的試題，後續可先基於跨測驗間的定錨題進行試題分數等化或轉換，再各自進到測驗內分析其餘試題和定錨題間的關係。當未來需納入與現今測量概念相近的測驗時，可藉由兩測驗間的定錨題進行分數等化或轉換。

本研究僅針對建置計畫中的四份測驗進行討論，然在建置計畫中仍有多組測驗，其題目選擇的方法上仍可有其他的簡化策略、標準或方案。Gregory (2015) 及 Kaplan 與 Saccuzzo (2018) 在其著作中表明題目選擇方法策略上具有多樣性：可基於試題難度、試題鑑別度進行題目選擇或簡化，標準使用上亦可根據不同測驗需求進行設定下，本研究採取的方法並非唯一可使用的簡化策

略或標準，未來身心健康評量系統建置過程中亦可針對不同身心變項的測驗進行討論，惟本研究建議除考量古典測驗理論與試題反應理論之測驗理論標準與統計標準外，仍需考量臨床上的實務使用需求。對此，針對未來建置後保留的題目可藉由精神臨床相關實務專家進行題目檢核，以評估建置系統試題是否仍需要加入其餘可協助區辨高身心風險狀態填答者。

同時，本研究建議未來身心健康評量系統分析量表上，除可考量題目數量偏多的測驗進行分析外，亦可搭配建置系統目標：「提供一完整而有效的身心健康評量系統，並在學生不同的學習時間點能了解學生的狀況以提供必要的協助」進行分析。部分量表題目數量少於本研究納入分析的量表題數十題，但實務上仍可評估對填答者身心情緒感受或外在社會資源支持的測驗為建議可分析與納入評量系統，如：2024 年編號 9 的自尊量表（Rosenberg Self-Esteem Scale [RSE]; Rosenberg, 1965），除具有自我滿意向度外，亦具有自我負面評價向度，其可與 2024 年編號 10 的害怕負面評價量表（Fear of Negative Evaluation Scale [FNE]; Watson & Friend, 1969）一同分析，以評估個體負面評價對填答者身心情緒感受的相關；另 2024 年編號 11 簡版完美主義量表（Frost Multidimensional Perfectionism Scale [FMPS]; Frost et al., 1990）具有目標導向完美主義向度，其可與 2024 年編號 17 主客觀學業成就量表一同納入分析，除評估填答者在學業目標與目標導向完美主義間的相關外，在主客觀學業成就量表中的填答亦可輔助校方輔導與教務單位適時提供填答者學習關懷與預警。

綜上所述，本研究著重在試題簡化與選擇，在研究限制與未來研究方向上可針對簡化試題的試題效度、試題間分數轉換進行討論外，亦可持續基於不同的統計策略或方法進行試題簡化方式的優化，以期完善國立臺灣大學身心健康評量系統建置，並提供第一線輔導人員一客觀評量填答者之身心狀態的工具。



參考文獻



- 林青慧 (2003):《臺灣簡短 36 (SF-36) 健康量表工具信效度及常模之建立》
(碩士論文, 中國醫藥學院), 臺灣博碩士論文知識加值系統,
<https://hdl.handle.net/11296/kk45qt>
- 高莉 (2025):《完美主義對臺大學生心理健康之影響:探討自我慈憫與社會支持的中介角色》(未出版碩士論文), 國立臺灣大學。
- 張雅筑 (2023):《研究生情緒困擾的成因及其對身心健康的影響:探討情緒調節自我效能之調節機制》(未出版之碩士論文), 國立臺灣大學。
- 陳伊柔 (2024):《探討社會心理因素對臺大學生心理健康之影響:情緒調節自我效能感的預測效果》(碩士論文, 國立臺灣大學), 華藝線上圖書館,
<https://doi.org/10.6342/NTU202401054>
- 陳禹昕 (2022):《自我認同與環境支持對經濟弱勢大學生校園生活適應之影響》(碩士論文, 國立臺灣大學), 華藝線上圖書館,
<https://doi.org/10.6342/NTU202201567>
- 陳慶餘, 吳英璋 (1987):〈以生物心理社會模式探討臺大新生之身心症狀〉。
《中華心理衛生學刊》, 3 (1), 89-105。
- 臺灣衛生福利部 (2021a):《世界心理健康日衛生福利統計通報》。臺灣衛生福利部統計處。<https://www.mohw.gov.tw/dl-72387-864eb8f0-64ef-42db-8a12-63169b8b97fc.html>
- 臺灣衛生福利部 (2021b):《108 年死因統計年報》。臺灣衛生福利部統計處。
<https://www.mohw.gov.tw/dl-69453-7815f16e-47b9-49cb-bec5-17497ccac32d.html>
- 賴宜琳 (2024):《正向因應資源與負向情緒反應對網路成癮之影響:以臺大大學部新生為例》(碩士論文, 國立臺灣大學), 華藝線上圖書館。
<https://doi.org/10.6342/NTU202401712>

- Beesdo-Baum, K., Klotsche, J., Knappe, S., Craske, M. G., Lebeau, R. T., Hoyer, J., Strobel, A., Pieper, L., & Wittchen, H. U. (2012). Psychometric properties of the dimensional anxiety scales for DSM-V in an unselected sample of German treatment seeking patients. *Depression and anxiety*, 29(12), 1014-1024.
<https://doi.org/10.1002/da.21994>
- Caprara, G. V., & Gerbino, M. (2001). *Regulatory Emotional Self-Efficacy Scale (RESE)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t23438-000>
- Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, 48(6), 1-29. <https://doi.org/10.18637/jss.v048.i06>
- Chen, S.-H., Weng, L.-J., Su, Y.-J., Wu, H.-M., & Yang, P.-F. (2003). *Chinese Internet Addiction Scale--Revised (CIAS, CIAS-R)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t44491-000>
- Clark, D. M. (1988). A cognitive model of panic attacks. In S. Rachman & J. D. Maser (Eds.), *Panic: Psychological perspectives* (pp. 71-89). Lawrence Erlbaum Associates, Inc.
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319. <https://doi.org/10.1037/1040-3590.7.3.309>
- Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press. <https://doi.org/10.4324/9781315827506>
- Connor, K. M., & Davidson, J. R. T. (2003). Development of a new resilience scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, 18(2), 76-82. <https://doi.org/10.1002/da.10113>
- Conroy, D. E., Willow, J. P., & Metzler, J. N. (2002). Multidimensional Fear of Failure Measurement: The Performance Failure Appraisal Inventory. *Journal of Applied Sport Psychology*, 14(2), 76-90.
<https://doi.org/10.1080/10413200252907752>

- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78(1), 98-104. <https://doi.org/10.1037/0021-9010.78.1.98>
- Costanza, R., Fisher, B., Ali, S., Beer, C., Bond, L., Boumans, R., Danigelis, N. L., Dickinson, J., Elliott, C., Farley, J., Gayer, D. E., Glenn, L. M., Hudspeth, T., Mahoney, D., McCahill, L., McIntosh, B., Reed, B., Rizvi, S. A. T., Rizzo, D. M., Simpatico, T., & Robert, S. R. (2007). Quality of life: An approach integrating opportunities, human needs, and subjective well-being. *Ecological Economics*, 61(2:3), 267-276. <https://doi.org/10.1016/j.ecolecon.2006.02.023>.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49(1), 71-75. https://doi.org/10.1207/s15327752jpa4901_13
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates Publishers.
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy*, 9(2), 171-180. <https://doi.org/10.1111/j.1752-0606.1983.tb01497.x>
- Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). *Frost Multidimensional Perfectionism Scale (FMPS)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t05500-000>
- Fletcher, T. D. (2023). *psychometric: Applied Psychometric Theory* (Version 2.4) [R package]. <https://CRAN.R-project.org/package=psychometric>
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37(6), 504-528. [https://doi.org/10.1016/S0092-6566\(03\)00046-1](https://doi.org/10.1016/S0092-6566(03)00046-1)
- Gray-Little, B., Williams, V. S. L., & Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 23(5), 443-451. <https://doi.org/10.1177/0146167297235001>

- Gregory, R. J. (2015). *Psychological Testing: History, Principles, and Applications* (7th ed.). Pearson.
- Harvey, R. J., & Murry, W. D. (1994). Scoring the Myers-Briggs Type Indicator: Empirical comparison of preference score versus latent-trait methods. *Journal of Personality Assessment*, 62(1), 116-129. https://doi.org/10.1207/s15327752jpa6201_11
- Horn, J. L. (1965). A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika*, 30(2), 179-185. <https://doi.org/10.1007/BF02289447>
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Iribarra, D. T., & Freund, R. (2014). *Wright Map: IRT item-person map with ConQuest integration*. (Version 1.0) [R package]. <https://github.com/david-ti/wrightmap>
- Ju, G.-F., & Duann, Y.-S. (2002). The Development of the College Students' Adjustment Check List and the Analysis the Students' Problems. *Journal of Research on Measurement and Statistics*, (10), 1-38. <https://doi.org/10.6773/JRMS.200212.0001>
- Kaplan, R. M., & Saccuzzo, D. P. (2018). *Psychological testing: Principles, applications, and issues* (9th ed.). Wadsworth/Thomson Learning.
- Kelderman, H. (1996). Multidimensional Rasch models for partial-credit scoring. *Applied Psychological Measurement*, 20(2), 155-168. <https://doi.org/10.1177/014662169602000205>
- Kolen, M. J., & Brennan, R. L. (2014). *Test Equating, Scaling, and Linking: Methods and Practices*. (3rd ed.). New York: Springer.
- Lai, Y.-C., & Wu, P.-H. (2018). The Development of the Epistemological Beliefs Scale-Mathematics for Taiwan Junior High School Students. *Psychological Testing*, 65(3), 315-339.

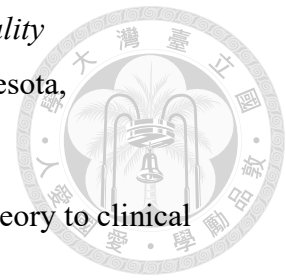
- Lee, E. H., Moon, S. H., Cho, M. S., Park, E. S., Kim, S. Y., Han, J. S., & Cheio, J. H. (2019). The 21-Item and 12-Item Versions of the Depression Anxiety Stress Scales: Psychometric Evaluation in a Korean Population. *Asian Nursing Research*, 13(1), 30-37. <https://doi.org/10.1016/j.anr.2018.11.006>
- Lee, M. B., Liao, S. C., Lee, Y. J., Wu, C. H., Tseng, M. C., Gau, S. F., & Rau, C. L. (2003). Development and verification of validity and reliability of a short screening instrument to identify psychiatric morbidity. *Journal of the Formosan Medical Association = Taiwan yi zhi*, 102(10), 687-694.
- Lee, Y.-C., Lin, Y.-C., Huang, C.-L., & Fredrickson, B. L. (2013). The construct and measurement of peace of mind. *Journal of Happiness Studies: An Interdisciplinary Forum on Subjective Well-Being*, 14(2), 571-590. <https://doi.org/10.1007/s10902-012-9343-5>
- Leng, C.-H., & Yao, G. (2024). Developing an abbreviated WHOQOL-BREF for patients with major chronic diseases with the assistance of Item Response Theory (IRT). *Taiwan Journal of Public Health*, 43(1), 32-45. [https://doi.org/10.6288/TJPH.202402_43\(1\).112130](https://doi.org/10.6288/TJPH.202402_43(1).112130)
- Li, C., Lin, Y., Tosun, B., Wang, P., Guo, H. Y., Ling, C. R., Qi, R., Luo, Q. Y., Wang, Y., Huang, F., Wang, J., Ma, S. H., Xu, D. F., Wu, S. Z., & Zhang, L. (2025). Psychometric evaluation of the Chinese version of the BENEFITS-CCCSAT based on CTT and IRT: a cross-sectional design translation and validation study. *Frontiers in Public Health*, 13, 1532709. <https://doi.org/10.3389/fpubh.2025.1532709>
- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335-343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Lue, B.-H., Chen, H.-J., & Kao, M.-Y. (2006). Perceived Stress in Medical Students and Its Relationship to Mental Health. *Journal of Medical Education*, 10(1), 25-33. [https://doi.org/10.6145/jme.200603_10\(1\).0004](https://doi.org/10.6145/jme.200603_10(1).0004)

- Luthar, S. S., Kumar, N. L., & Zillmer, N. (2020). High-achieving schools connote risks for adolescents: Problems documented, processes implicated, and directions for interventions. *The American Psychologist*, 75(7), 983-995. <https://doi.org/10.1037/amp0000556>
- McBrien, A., Wild, M., & Bachorowski, J. A. (2020). Social-Emotional Expertise (SEE) Scale: Development and Initial Validation. *Assessment*, 27(8), 1718-1730. <https://doi.org/10.1177/1073191118794866>
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum Associates Publishers.
- McDonald, R. P., & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods*, 7(1), 64-82. <https://doi.org/10.1037/1082-989X.7.1.64>
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, 14, 59-71. <https://doi.org/10.1177/014662169001400106>
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16(2), 159-176. <https://doi.org/10.1177/014662169201600206>
- Muraki, E. (1997). A Generalized Partial Credit Model. In: van der Linden, W. J., & Hambleton, R. K. (Eds.), *Handbook of Modern Item Response Theory*. Springer. https://doi.org/10.1007/978-1-4757-2691-6_9
- Murphy, K. R., & Davidshofer, C. O. (2004). *Psychological Testing: Principles and Applications* (6th ed.). Prentice Hall.
- National Taiwan University. (2023). *Number of male and female students, sorted by college*. National Taiwan University. <https://annualstatistics.cc.ntu.edu.tw/data/acct2023/acct2023e/acct1/27.pdf>
- Orlando, M., & Thissen, D. (2003). Further investigation of the performance of $S-X^2$: An item fit index for use with dichotomous item response theory models. *Applied Psychological Measurement*, 27(4), 289-298. <https://doi.org/10.1177/0146621603027004004>

- Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., & Lozano, G. (2012). The Depression Anxiety Stress Scales-21 (DASS-21): further examination of dimensions, scale reliability, and correlates. *Journal of Clinical Psychology*, 68(12), 1322-1338. <https://doi.org/10.1002/jclp.21908>
- Piedmont, R. L. (2014). Inter-item Correlations. In: Michalos, A. C. (Eds.), *Encyclopedia of Quality of Life and Well-Being Research*. Springer. https://doi.org/10.1007/978-94-007-0753-5_1493
- Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., Thissen, D., Revicki, D. A., Weiss, D. J., Hambleton, R. K., Liu, H., Gershon, R., Reise, S. P., Lai, J. S., Cella, D., & PROMIS Cooperative Group. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical Care*, 45(5 Suppl 1), S22-S31. <https://doi.org/10.1097/01.mlr.0000250483.85507.04>
- Reise, S. P., & Waller, N. G. (1990). Fitting the two-parameter model to personality data. *Applied Psychological Measurement*, 14(1), 45-58. <https://doi.org/10.1177/014662169001400105>
- Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology*, 5, 27-48. <https://doi.org/10.1146/annurev.clinpsy.032408.153553>
- Revelle, W. (2025). *psych: Procedures for Psychological, Psychometric, and Personality Research*. (Version 2.5.3) [R package]. <https://CRAN.R-project.org/package=psych>.
- Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton University Press.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36. <https://doi.org/10.18637/jss.v048.i02>
- Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, Validity, and Factor Structure. *Journal of Personality Assessment*, 66(1), 20-40. https://doi.org/10.1207/s15327752jpa6601_2

- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34(4, Pt. 2), 100.
- Santor, D. A., Ramsay, J. O., & Zuroff, D. C. (1994). Nonparametric item analyses of the Beck Depression Inventory: Evaluating gender item bias and response option weights. *Psychological Assessment*, 6(3), 255-270. <https://doi.org/10.1037/1040-3590.6.3.255>
- Schmidt, S., Mühlan, H., & Power, M. (2006). The EUROHIS-QOL 8-item index: Psychometric results of a cross-cultural field study. *European Journal of Public Health*, 16(4), 420-428. <https://doi.org/10.1093/eurpub/cki155>
- Schulenberg, S. E., Schnitzer, L. W., & Buchanan, E. M. (2011). The Purpose in Life Test-Short Form: Development and Psychometric Support. *Journal of Happiness Studies*, 12, 861-876. <https://doi.org/10.1007/s10902-010-9231-9>
- Segall, D. O. (1996). Multidimensional adaptive testing. *Psychometrika*, 61(2), 331-354. <https://doi.org/10.1007/BF02294343>
- Segrin, C., Hanzal, A., Donnerstein, C., Taylor, M., & Domschke, T. J. (2007). Social skills, psychological well-being, and the mediating role of perceived stress. *Anxiety, Stress, and Coping*, 20(3), 321-329. <https://doi.org/10.1080/10615800701282252>
- Smith, R. M. (1996). Polytomous mean-square fit statistics. *Rasch Measurement Transactions*, 10(3), 516-517.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of Internal Medicine*, 166(10), 1092-1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Syed, H. R., Zachrisson, H. D., Dalgard, O. S., Dalen, I., & Ahlberg, N. (2008). *Hopkins Symptoms Check List-10 (HSCL-10)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t12173-000>
- Tabachnick, B. G. & Fidell, L. S. (2019). *Using multivariate statistics*. Pearson Education.

- Tellegen, A. (1982). *Brief Manual for the Multidimensional Personality Questionnaire*. Unpublished Manuscripts, University of Minnesota, Minneapolis, 1031-1010.
- Thomas, M. L. (2019). Advances in applications of item response theory to clinical assessment. *Psychological Assessment*, 31(12), 1442-1455.
<https://doi.org/10.1037/pas0000597>
- Walter, O. B., Becker, J., Bjorner, J. B., Fliege, H., Klapp, B. F., & Rose, M. (2007). Development and evaluation of a computer adaptive test for ‘Anxiety’ (Anxiety-CAT). *Qual Life Res*, 16 (Suppl 1), 143-155.
<https://doi.org/10.1007/s11136-007-9191-7>
- Wang, C., & Chang, H.-H. (2011). Item Selection in Multidimensional Computerized Adaptive Testing—Gaining Information from Different Angles. *Psychometrika*, 76, 363-384. <https://doi.org/10.1007/s11336-011-9215-7>
- Watson, D., & Friend, R. (1969). *Fear of Negative Evaluation Scale (FNE)* [Database record]. APA PsycTests. <https://doi.org/10.1037/t00794-000>
- World Health Organization. (1998, February 12-13). *Wellbeing measures in primary health care/the DepCare Project: report on a WHO meeting: Stockholm, Sweden, 1998*. Wellbeing measures in primary health care/the DepCare Project: report on a WHO meeting: Stockholm, Sweden.
- Wright, B., Linacre, J. & Gustafson, J. (1994). Reasonable Mean-Square Fit Values. *Rasch Measurement Transformation*, 8, 370.
- Zimet, G. (2016). *Multidimensional Scale of Perceived Social Support (MSPSS) - Scale Items and Scoring Information*.
https://www.researchgate.net/publication/311534896_Multidimensional_Scale_of_Perceived_Social_Support_MSPSS_-_Scale_Items_and_Scoring_Information
- Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment*, 52(1), 30-41. https://doi.org/10.1207/s15327752jpa5201_2





附錄 1 原有量表施測試題



表 1-1

正向因應資源原有施測試題

題 號	題目	1 從 未 如 此	2 很 少 如 此	3 有 時 如 此	4 經 常 如 此	5 總 是 如 此
1	家人對我的生活方式感到滿意					
2	家人對我的學習與學業目標感到滿意					
3	我的生活目標同時符合我的與父母的需求					
4	我能夠從容地應付我的課業					
5	我知道我為何要讀書					
6	我是個心情愉快的人					
7	在別人面前，我不怕說出自己的缺點或過去的失敗					
8	早上起床時，我覺得精神良好					
9	我對於自己休閒活動的安排感到滿意					
10	碰到困難時，我會冷靜下來去尋找解決困難的方法					
11	一般而言，我喜歡自己，滿意自己					
12	當我有麻煩和煩惱時，我可以從家庭得到幫助					
13	我滿意家人對我的情緒(喜、怒、哀、樂)表示關心和愛護的方式					
14	我和家人可以共度愉快的時光					
15	當我嘗試新事物時，我會感到家人的接受與支持					
16	當我遇到麻煩，能從老師、同學或朋友獲得到協助					
17	我滿意老師、同學或朋友關懷我情緒的方式					
18	我能和老師、同學或朋友共度愉快的時光					

表 1-2

負向情緒反應原有施測試題



題 號	題目	1 從 未 如 此	2 很 少 如 此	3 有 時 如 此	4 經 常 如 此	5 總 是 如 此
1	我會因為不安與生氣，造成我的身體不舒服 （如：拉肚子、頭痛、皮膚起疹子、手腳麻木等）					
2	我會覺得自己很神經質，很容易緊張，對事情都想得太多，考慮太多					
3	我會沒有理由地感到害怕、驚慌、不安					
4	我會有一段時間心情低落、憂鬱、鬱悶					
5	我會有一段時間心情不好、容易苦惱、動怒、生氣					
6	我會有一段時間（至少持續一週）情緒高昂，容易興奮					
7	日常生活中，我情緒會不可預測地忽高忽低，讓我感到困擾					
8	我對一般人不感到害怕的環境（如：高處、人群、封閉的地方或出家門），我卻感到極度害怕，以至於想要離開					
9	我曾因生活上的變化（如親人離別、與人吵架等）導致情緒不穩與痛苦，進而影響工作與人際關係等					
10	我曾因不必要的重複動作（如：洗手、關門、關瓦斯）或因停不下來的想法（如：一直以為手是髒的，或一直以為可能會傷害他人，但我又明知那是不可能發生的事情）而心理感到不安與痛苦					
11	我相信一些事情（如：有人跟蹤我，或有人要害我等等）但是別人認為那是不存在的，或是不可能發生的					
12	我會聽到一些別人聽不到的聲音，看到別人看不到的事物					
13	我有「將要生病」的感覺					
14	我覺得心情不好，沒辦法應付周圍發生的事					
15	我睡眠有困難，如：難以入睡、易醒來或早醒					
16	我覺得處在壓力中					
17	我覺得自己比不過別人					

表 1-2 (續)

負向情緒反應原有施測試題

題 號	題目	1 從 未 如 此	2 很 少 如 此	3 有 時 如 此	4 經 常 如 此	5 總 是 如 此
18	我不滿意自己的學業或工作表現					
19	我覺得自己無法做決定					
20	我覺得自己是一文不值的人					
21	我覺得人生很無望，對事物皆失去興趣					
22	我有尋死（自殺）的想法					
23	我有想傷害別人的念頭或衝動					
24	我對未來不抱甚麼期待					
25	我有尋死的計畫					
26	我有做過傷害自己的事情					

表 1-3

多向度感知社會支持量表原有施測試題



題 號	題 目	1 非 常 強 烈 不 同 意	2 強 烈 不 同 意	3 有 些 不 同 意	4 沒 意 見	5 有 些 同 意	6 強 烈 同 意	7 非 常 強 烈 同 意
1	每當我有需要時，有個特別的人會在我身邊幫助我。							
2	有個特別的人能讓我分享快樂與悲傷							
3	我的家人真的會試圖幫助我							
4	我能從家人得到所需的情緒上的協助與支持							
5	有位特別的人是我慰藉的主要來源							
6	我的朋友真的會試圖幫助我							
7	當萬事不順遂時，我可以依賴我的朋友							
8	我可以告訴家人我的煩惱							
9	我有朋友可以分享我的快樂與悲傷							
10	我生命中有個特別的人會關心我的感受							
11	我的家人願意協助我做決定							
12	我可以告訴我的朋友我的煩惱							

表 1-4

自我滿意度原有施測試題



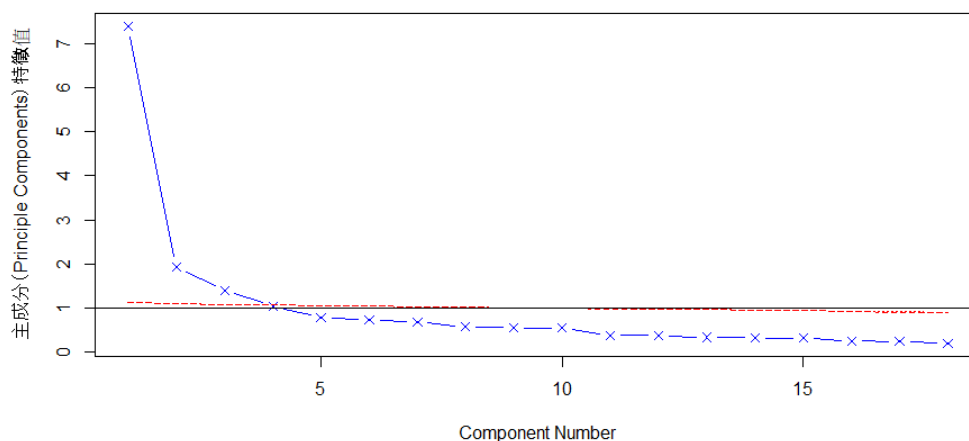
題 號	題目	1 非常 不滿意	2 不 滿意	3 普通	4 滿意	5 非常 滿意
1	我對於自己所就讀的科系感到					
2	我對於自己的學業或工作表現感到					
3	我對於自己的生涯規劃感到					
4	我對於自己的生活與人生目標感到					
5	我對於自己的時間管理能力感到					
6	我對於自己的人際關係感到					
7	我對於自己的情感關係感到					
8	我對於自己的情緒穩定度感到					
9	我對於自己的個人特質感到					
10	我對於自己的身心健康感到					
11	我對於自己的自我掌控能力感到					
12	我對於自己的家庭狀況感到					
13	我對於自己的休閒活動安排感到					
14	我對於自己的運動習慣感到					
15	我對於自己的睡眠狀況感到					
16a	我傾向認同的性別： ☐ 男性、 ☐ 女性、 ☐ 其他					
16b	我對於環境對我自己性別認同的支 持感到					

附錄 2 探索性因素分析因素個數判別圖：平行分析法



圖 2-1

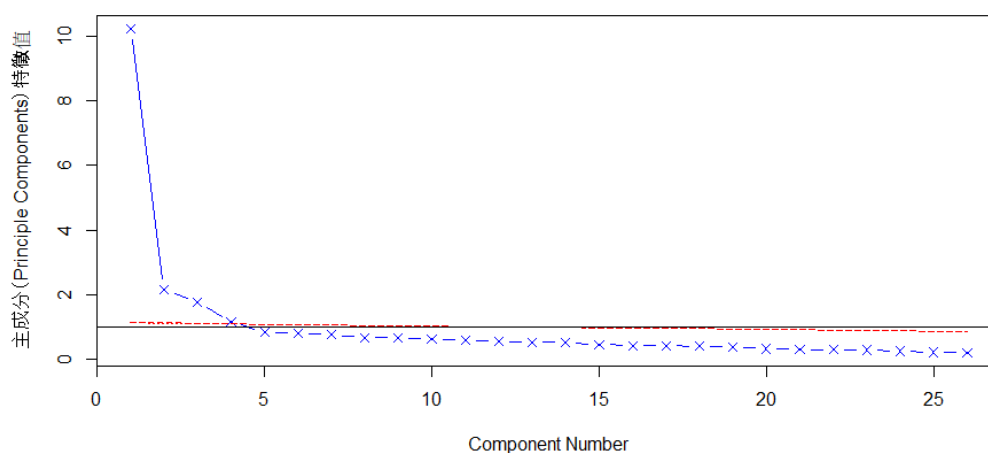
正向因應資源因素個數判別圖



註：藍色線為實際資料之主成分特徵值曲線，紅色線為模擬資料之主成分特徵值曲線。

圖 2-2

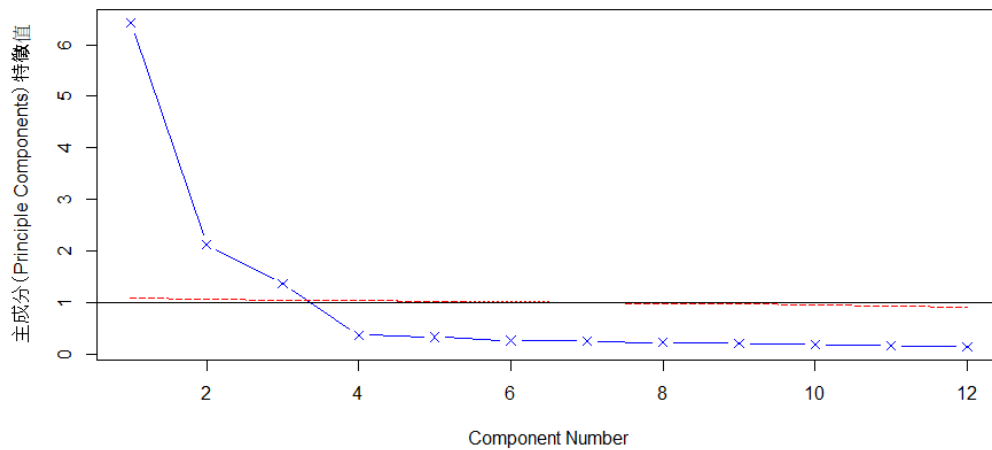
負向情緒反應因素個數判別圖



註：藍色線為實際資料之主成分特徵值曲線，紅色線為模擬資料之主成分特徵值曲線。

圖 2-3

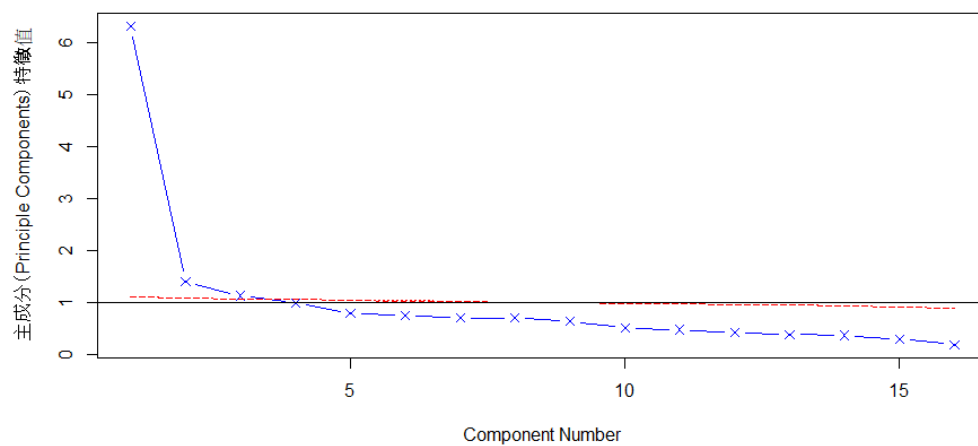
多向度感知社會支持量表因素個數判別圖



註：藍色線為實際資料之主成分特徵值曲線，紅色線為模擬資料之主成分特徵值曲線。

圖 2-4

自我滿意度量表因素個數判別圖



註：藍色線為實際資料之主成分特徵值曲線，紅色線為模擬資料之主成分特徵值曲線。

附錄 3 本研究使用之程式操作語法



```
##descriptive
```

```
write.csv(describe(data[, c(col.1:col.2)]), file = 'test_desc.csv')
```

```
write.csv(item.exam(data[, c(col.1:col.2)]), file = 'test_item_exam.csv')
```

```
##exploratory factor analysis
```

```
init_mat <- matrix(correlation matrix, nrow = # of factors , ncol = # of factors, byrow  
= TRUE)
```

```
KMO(data)
```

```
fa.parallel(data, fm = 'ml', fa = 'both', SMC = TRUE)
```

```
explo_Q <- efa(data, nfactors, sample.cov = init_mat, rotation = 'promax')
```

```
explo_Q_summary <- summary(explo_Q, cutoff = 0)
```

```
explo_Q_loading <- explo_Q_summary[["model.list"]][["# of factors  
"]][["efa"]][["lambda"]][[1]]
```

```
write.csv(explo_Q_loading, file = 'QEFA.csv')
```

```
##confirmatory factor analysis
```

```
c_model <- mirt.model('
```

```
  F1 = items with higher loading on F1
```

```
  F2 = items with higher loading on F2
```

```
  .....
```

```
  Fn = items with higher loading on Fn
```

```
  COV = F1*F2, ....., F(n-1)*F(n)
```

```
')
```

```
conf_Q <- data
```

```
conf_modQ <- mirt(conf_Q, c_model)
```

```
coef(conf_modQ) ## MDISC & MDIFF
```

```
conf_modQ_itemfit <- itemfit(conf_modQ, fit_stats = "S_X2")
```



```
##dimensionality
mod1 <- mirt(data[, c(col.1:col.2)], 1)
mod2 <- mirt(data[, c(col.1:col.2)], 2)
mod3 <- mirt(data[, c(col.1:col.2)], 3)
anova(mod1, mod2, mod3)

##specify IRT model
mod3_gpcm<- mirt(data[, c(col.1:col.2)], 3, itemtype = "gpcm")
mod3_graded <- mirt(data[, c(col.1:col.2)], 3, itemtype = "graded")
anova(mod3_gpcm, mod3_graded)

##item fit/ parameters
summary(mod, rotate = 'promax')
coef(mod, rotate = 'promax')
itemfit(mod, fit_stats = "infit")

##multidimensional IRT parameters
write.csv(MDISC(mod), file = 'test_IRT_MDISC.csv')
write.csv(MDIFF(mod), file = 'test_IRT_MDIFF.csv')

##information function with D-optimality with 3 factors
init_mat <- matrix(data = correlation matrix, nrow = 3, ncol = 3, byrow = TRUE)
x <- seq(-4, 4, 0.1)
y <- seq(-4, 4, 0.1)
z <- seq(-4, 4, 0.1)
Theta <- as.matrix(expand.grid(x, y, z))
LTheta <- subset(Theta, Theta[, 1]< -0.5 & Theta[, 2]< -0.5 & Theta[, 3]< -0.5)
HTheta <- subset(Theta, Theta[, 1]> 0.5 & Theta[, 2]> 0.5 & Theta[, 3]> 0.5)
MTheta <- subset(Theta, Theta[, 1]<= 0.5 & Theta[, 2]<= 0.5 & Theta[, 3]<= 0.5 &
```

```

Theta[, 1]>= -0.5 & Theta[, 2]>= -0.5 & Theta[, 3]>= -0.5)
wtM <- dnorm(seq(-.5, .5, 0.1))
wtL <- dnorm(seq(-4, -.6, 0.1))
wtH <- dnorm(seq(.6, 4, 0.1))
wtMM <- as.matrix(expand.grid(wtM, wtM, wtM))
wtLM <- as.matrix(expand.grid(wtL, wtL, wtL))
wtHM <- as.matrix(expand.grid(wtH, wtH, wtH))

## lower ability level
info.L <- c()
tempinfo.L <- c()
for(i in 1:total item number){
  for(j in 1: nrow(LTheta)){
    ab <- matrix(LTheta[j, ], nrow = 1, ncol = 3)
    test <- iteminfo(extract.item(mod, i), ab, degrees = c(45,45,45), multidim_matrix =
TRUE)
    tempinfo.L[j] <- det(solve(init_mat)+test) *
      as.numeric(wtLM[j, 1])* as.numeric(wtLM[j, 2])* as.numeric(wtLM[j, 3])
  }
  info.L[i] <- sum(tempinfo.L)
}

##moderate ability level
info.M <- c()
tempinfo.M <- c()
for(i in 1: total item number){
  for(j in 1: nrow(MTheta)){
    ab <- matrix(MTheta[j, ], nrow = 1, ncol = 3)

```



```

test <- iteminfo(extract.item(mod, i), ab, degrees = c(45,45,45), multidim_matrix =
TRUE)

tempinfo.M[j] <- det(solve(init_mat)+test) *
  as.numeric(wtMM[j, 1])* as.numeric(wtMM[j, 2])* as.numeric(wtMM[j, 3])
}
info.M[i] <- sum(tempinfo.M)
}

```

```
##higher ability level
```

```

info.H <- c()
tempinfo.H <- c()
for(i in 1: total item number){
  for(j in 1: nrow(HTheta)){
    ab <- matrix(HTheta[j, ], nrow = 1, ncol = 3)
    test <- iteminfo(extract.item(mod, i), ab, degrees = c(45,45,45), multidim_matrix =
TRUE)
    tempinfo.H[j] <- det(solve(init_mat)+test) *
      as.numeric(wtHM[j, 1])* as.numeric(wtHM[j, 2])* as.numeric(wtHM[j, 3])
  }
  info.H[i] <- sum(tempinfoQ2.H)
}

```

```
##output
```

```

info.frame <- data.frame(info.M,info.H, info.L)
write.csv(info.frame, 'info.csv')

```

附錄 4 驗證性因素分析標準化參數報表

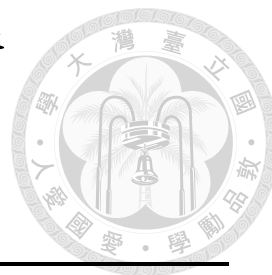


表 4-1-1

正向因應資源進行分析之試題表

題號	題目
Item 4	我能夠從容地應付我的課業
Item 5	我知道我為何要讀書
Item 6	我是個心情愉快的人
Item 7	在別人面前，我不怕說出自己的缺點或過去的失敗
Item 8	早上起床時，我覺得精神良好
Item 9	我對於自己休閒活動的安排感到滿意
Item 10	碰到困難時，我會冷靜下來去尋找解決困難的方法
Item 11	一般而言，我喜歡自己，滿意自己
Item 12	當我有麻煩和煩惱時，我可以從家庭得到幫助
Item 13	我滿意家人對我的情緒（喜、怒、哀、樂）表示關心和愛護的方式
Item 14	我和家人可以共度愉快的時光
Item 15	當我嘗試新事物時，我會感到家人的接受與支持
Item 16	當我遇到麻煩，能從老師、同學或朋友獲得到協助
Item 17	我滿意老師、同學或朋友關懷我情緒的方式
Item 18	我能和老師、同學或朋友共度愉快的時光

表 4-1-2

正向因應資源標準化之因素負荷量顯著性報表

預測因素	分析變項	估計因素負荷量	標準誤	t	顯著性
家庭資源					
	Item 12	.832	.006	140.800	<.001
	Item 13	.890	.005	196.500	<.001
	Item 14	.851	.005	156.400	<.001
	Item 15	.813	.006	127.800	<.001
自我資源					
	Item 4	.582	.012	48.906	<.001
	Item 5	.522	.013	40.391	<.001
	Item 6	.773	.008	96.408	<.001
	Item 8	.553	.012	44.548	<.001
	Item 9	.619	.011	55.214	<.001
	Item 10	.616	.011	54.780	<.001
	Item 11	.806	.007	110.600	<.001
師友資源					
	Item 16	.857	.006	146.200	<.001
	Item 17	.888	.005	167.300	<.001
	Item 18	.811	.007	119.300	<.001

表 4-1-3

正向因應資源外生變項變異數之標準化結果報表

分析變項	估計因素負荷量	標準誤	<i>t</i>	顯著性
因素				
家庭資源	1.000			
自我資源	1.000			
師友資源	1.000			
因素間比較				
家庭資源、自我資源	.534	.014	38.934	<.001
家庭資源、師友資源	.524	.013	38.827	<.001
自我資源、師友資源	.627	.012	51.226	<.001
分析變項所屬誤差				
<i>E</i> _{item 12}	.308	.010	31.364	<.001
<i>E</i> _{item 13}	.207	.008	25.668	<.001
<i>E</i> _{item 14}	.276	.009	29.769	<.001
<i>E</i> _{item 15}	.339	.010	32.766	<.001
<i>E</i> _{item 4}	.662	.014	47.805	<.001
<i>E</i> _{item 5}	.727	.014	53.905	<.001
<i>E</i> _{item 6}	.403	.012	32.567	<.001
<i>E</i> _{item 8}	.694	.014	50.627	<.001
<i>E</i> _{item 9}	.617	.014	44.509	<.001
<i>E</i> _{item 10}	.620	.014	44.713	<.001
<i>E</i> _{item 11}	.350	.012	29.791	<.001
<i>E</i> _{item 16}	.266	.010	26.470	<.001
<i>E</i> _{item 17}	.212	.009	22.459	<.001
<i>E</i> _{item 18}	.343	.011	31.127	<.001

表 4-2-1

負向情緒反應進行分析之試題表



題號	題目
Item 1	我會因為不安與生氣，造成我的身體不舒服（如：拉肚子、頭痛、皮膚起疹子、手腳麻木等）
Item 2	我會覺得自己很神經質，很容易緊張，對事情都想得太多，考慮太多
Item 3	我會沒有理由地感到害怕、驚慌、不安
Item 4	我會有一段時間心情低落、憂鬱、鬱悶
Item 5	我會有一段時間心情不好、容易苦惱、動怒、生氣
Item 6	我會有一段時間（至少持續一週）情緒高昂，容易興奮
Item 7	日常生活中，我情緒會不可預測地忽高忽低，讓我感到困擾
Item 9	我曾因生活上的變化（如親人離別、與人吵架等）導致情緒不穩與痛苦，進而影響工作與人際關係等
Item 10	我曾因不必要的重複動作（如：洗手、關門、關瓦斯）或因停不下來的想法（如：一直以為手是髒的，或一直以為可能會傷害他人，但我又明知那是不可能發生的事情）而心理感到不安與痛苦
Item 11	我相信一些事情（如：有人跟蹤我，或有人要害我等等）但是別人認為那是不存在的，或是不可能發生的
Item 12	我會聽到一些別人聽不到的聲音，看到別人看不到的事物
Item 15	我睡眠有困難，如：難以入睡、易醒來或早醒
Item 17	我覺得自己比不過別人
Item 18	我不滿意自己的學業或工作表現
Item 19	我覺得自己無法做決定
Item 20	我覺得自己是一文不值的人
Item 21	我覺得人生很無望，對事物皆失去興趣
Item 22	我有尋死（自殺）的想法
Item 23	我有想傷害別人的念頭或衝動
Item 24	我對未來不抱甚麼期待
Item 25	我有尋死的計畫
Item 26	我有做過傷害自己的事情

表 4-2-2

負向情緒反應標準化之因素負荷量顯著性報表

預測因素	分析變項	估計因素負荷量	標準誤	<i>t</i>	顯著性
因應情緒行為					
	Item 1	.599	.011	54.373	<.001
	Item 2	.716	.009	83.388	<.001
	Item 3	.778	.007	109.000	<.001
	Item 4	.824	.006	137.400	<.001
	Item 5	.828	.006	140.700	<.001
	Item 6	.414	.014	29.479	<.001
	Item 7	.749	.008	96.007	<.001
	Item 9	.619	.011	58.167	<.001
	Item 15	.555	.012	46.933	<.001
無望感					
	Item 17	.744	.008	91.656	<.001
	Item 18	.740	.008	90.124	<.001
	Item 19	.616	.011	56.933	<.001
	Item 20	.839	.006	143.200	<.001
	Item 21	.808	.007	122.400	<.001
	Item 24	.761	.008	98.529	<.001
自傷想法行為					
	Item 22	.842	.007	113.100	<.001
	Item 23	.527	.013	40.240	<.001
	Item 25	.783	.008	93.113	<.001
	Item 26	.668	.011	63.234	<.001
精神病狀					
	Item 10	.723	.011	65.877	<.001
	Item 11	.740	.011	68.774	<.001
	Item 12	.623	.012	49.980	<.001

表 4-2-3

負向情緒反應外生變項變異數之標準化結果報表

分析變項	估計因素負荷量	標準誤	t	顯著性
因素				
因應情緒行為	1.000			
無望感	1.000			
自傷想法行為	1.000			
精神病狀	1.000			
因素間比較				
因應情緒行為、無望感	.721	.010	75.309	<.001
因應情緒行為、自傷想法行為	.519	.014	36.425	<.001
因應情緒行為、精神病狀	.538	.015	35.491	<.001
無望感、自傷想法行為	.596	.013	45.748	<.001
無望感、精神病狀	.431	.017	25.582	<.001
自傷想法行為、精神病狀	.509	.016	30.971	<.001
分析變項所屬誤差				
<i>E</i> _{item 1}	.641	.013	48.539	<.001
<i>E</i> _{item 2}	.488	.012	39.691	<.001
<i>E</i> _{item 3}	.395	.011	35.627	<.001
<i>E</i> _{item 4}	.321	.010	32.532	<.001
<i>E</i> _{item 5}	.314	.010	32.219	<.001
<i>E</i> _{item 6}	.829	.012	71.410	<.001
<i>E</i> _{item 7}	.438	.012	37.463	<.001
<i>E</i> _{item 9}	.617	.013	46.869	<.001
<i>E</i> _{item 15}	.692	.013	52.653	<.001
<i>E</i> _{item 17}	.447	.012	37.049	<.001
<i>E</i> _{item 18}	.453	.012	37.324	<.001
<i>E</i> _{item 19}	.620	.013	46.432	<.001
<i>E</i> _{item 20}	.297	.010	30.219	<.001
<i>E</i> _{item 21}	.348	.011	32.610	<.001
<i>E</i> _{item 24}	.422	.012	35.898	<.001
<i>E</i> _{item 22}	.292	.013	23.298	<.001
<i>E</i> _{item 23}	.722	.014	52.259	<.001
<i>E</i> _{item 25}	.387	.013	29.438	<.001
<i>E</i> _{item 26}	.553	.014	39.163	<.001
<i>E</i> _{item 10}	.477	.016	30.077	<.001
<i>E</i> _{item 11}	.453	.016	28.442	<.001
<i>E</i> _{item 12}	.611	.016	39.309	<.001

表 4-3-1

多向度感知社會支持量表進行分析之試題表

題號	題目
Item 1	每當我有需要時，有個特別的人會在我身邊幫助我
Item 2	有個特別的人能讓我分享快樂與悲傷
Item 3	我的家人真的會試圖幫助我
Item 4	我能從家人得到所需的情緒上的協助與支持
Item 5	有位特別的人是我慰藉的主要來源
Item 6	我的朋友真的會試圖幫助我
Item 7	當萬事不順遂時，我可以依賴我的朋友
Item 8	我可以告訴家人我的煩惱
Item 9	我有朋友可以分享我的快樂與悲傷
Item 10	我生命中有個特別的人會關心我的感受
Item 11	我的家人願意協助我做決定
Item 12	我可以告訴我的朋友我的煩惱

表 4-3-2

多向度感知社會支持量表標準化之因素負荷量顯著性報表

預測因素	分析變項	估計因素負荷量	標準誤	t	顯著性
重要他人	Item 1	.880	.004	202.600	<.001
	Item 2	.923	.003	280.300	<.001
	Item 5	.863	.005	181.200	<.001
	Item 10	.883	.004	206.700	<.001
家人	Item 3	.886	.004	204.300	<.001
	Item 4	.921	.004	257.700	<.001
	Item 8	.854	.005	166.800	<.001
	Item 11	.825	.006	140.400	<.001
朋友	Item 6	.863	.005	172.300	<.001
	Item 7	.879	.005	189.900	<.001
	Item 9	.866	.005	174.900	<.001
	Item 12	.856	.005	164.400	<.001

表 4-3-3

多向度感知社會支持量表外生變項變異數標準化結果報表

分析變項	估計因素負荷量	標準誤	<i>t</i>	顯著性
因素				
重要他人	1.000			
家人	1.000			
朋友	1.000			
因素間比較				
重要他人、家人	.440	.014	31.120	<.001
重要他人、朋友	.638	.011	59.128	<.001
家人、朋友	.431	.014	29.882	<.001
分析變項所屬誤差				
<i>E</i> _{item 1}	.226	.008	29.594	<.001
<i>E</i> _{item 2}	.149	.006	24.460	<.001
<i>E</i> _{item 5}	.255	.008	30.955	<.001
<i>E</i> _{item 10}	.221	.008	29.333	<.001
<i>E</i> _{item 3}	.214	.008	27.849	<.001
<i>E</i> _{item 4}	.154	.007	23.177	<.001
<i>E</i> _{item 8}	.270	.009	30.839	<.001
<i>E</i> _{item 11}	.320	.010	33.037	<.001
<i>E</i> _{item 6}	.255	.009	29.446	<.001
<i>E</i> _{item 7}	.227	.008	27.919	<.001
<i>E</i> _{item 9}	.250	.009	29.219	<.001
<i>E</i> _{item 12}	.268	.009	30.120	<.001

表 4-4-1

自我滿意度量表進行分析之試題表

題號	題目
Item 1	我對於自己所就讀的科系感到
Item 2	我對於自己的學業或工作表現感到
Item 3	我對於自己的生涯規劃感到
Item 4	我對於自己的生活與人生目標感到
Item 6	我對於自己的人際關係感到
Item 7	我對於自己的情感關係感到
Item 8	我對於自己的情緒穩定度感到
Item 9	我對於自己的個人特質感到
Item 10	我對於自己的身心健康感到
Item 11	我對於自己的自我掌控能力感到
Item 12	我對於自己的家庭狀況感到
Item 13	我對於自己的休閒活動安排感到
Item 14	我對於自己的運動習慣感到
Item 15	我對於自己的睡眠狀況感到

表 4-4-2

自我滿意度量表標準化之因素負荷量顯著性報表

預測因素	分析變項	估計因素負荷量	標準誤	t	顯著性
身心狀態滿意度					
	Item 1	.491	.013	37.460	<.001
	Item 2	.673	.010	68.741	<.001
	Item 3	.882	.005	161.300	<.001
	Item 4	.896	.005	170.600	<.001
日常生活滿意度					
	Item 6	.575	.012	48.820	<.001
	Item 7	.505	.013	38.901	<.001
	Item 8	.704	.009	76.526	<.001
	Item 9	.733	.009	85.616	<.001
	Item 10	.787	.007	107.000	<.001
	Item 11	.755	.008	93.571	<.001
	Item 12	.431	.014	30.551	<.001
	Item 13	.664	.010	66.079	<.001
	Item 14	.504	.013	38.756	<.001
	Item 15	.516	.013	40.332	<.001

表 4-4-3

自我滿意度量表外生變項變異數之標準化結果報表



分析變項	估計因素負荷量	標準誤	t	顯著性
因素				
身心狀態滿意度	1.000			
日常生活滿意度	1.000			
因素間比較				
身心狀態滿意度、 日常生活滿意度	.667	.011	60.371	<.001
分析變項所屬誤差				
<i>E</i> _{item 1}	.759	.013	58.862	<.001
<i>E</i> _{item 2}	.547	.013	41.556	<.001
<i>E</i> _{item 3}	.221	.010	22.908	<.001
<i>E</i> _{item 4}	.197	.009	20.938	<.001
<i>E</i> _{item 6}	.669	.014	49.323	<.001
<i>E</i> _{item 7}	.745	.013	56.763	<.001
<i>E</i> _{item 8}	.504	.013	38.859	<.001
<i>E</i> _{item 9}	.462	.013	36.793	<.001
<i>E</i> _{item 10}	.381	.012	32.966	<.001
<i>E</i> _{item 11}	.430	.012	35.239	<.001
<i>E</i> _{item 12}	.815	.012	67.097	<.001
<i>E</i> _{item 13}	.559	.013	41.818	<.001
<i>E</i> _{item 14}	.746	.013	56.902	<.001
<i>E</i> _{item 15}	.733	.013	55.449	<.001