

多面向道德情意量表之發展與信效度 考驗*

李琪明

國立臺灣師範大學
公民教育與活動領導學系

本研究目的在發展「多面向道德情意量表」(multidimensional affective morality scale, 簡稱 MAM scale), 並考驗其信效度。研究者汲取國內外道德學界既有經驗與成果, 針對國中至研究所學生為對象, 並奠基臺灣的在地脈絡, 以作為情意層面道德範疇概況或成效展現的具體指標, 希冀對於國內外道德教育理論與實踐有所助益。本研究方法與歷程包括先行探究理念架構以擬定量表初稿、繼而針對初稿評定其內容效度、再經修改後進行量表預試與正式施測。研究者區分四個階段進行統計分析, 並採用 IBM SPSS Statistics 23 與 Amos 25 Graphics 為分析軟體: 一是將預試樣本 ($N = 585$) 進行項目分析, 計算其內部一致性信度; 二是將預試樣本區分為二群, 以樣本_1 ($N = 283$) 進行探索性因素分析; 三是以樣本_2 ($N = 302$) 進行驗證性因素分析; 四是以正式施測樣本 ($N = 708$) 進行複核效化檢驗。本研究發展完成的 MAM 量表包括詳版與簡版, 詳版通過內容效度與內部一致性考驗, 包括四則道德故事與四個構面, 計有 44 個題項; 簡版則是奠基詳版, 並具有良好的模式適配度(二階單因子驗證性分析模式), 以及理想的組合信度、建構效度與複核效化檢驗, 其內容包括三則道德故事與三個構面, 計有 12 個題項。詳版可顯示道德情意量表的脈絡性與豐富性, 簡版則可展現結構性與精簡性。藉此量表的發展, 本文綜整結論以彰顯 MAM 量表的特點: 一是本量表凸顯情意層面道德範疇, 豐富且具深度的多面向意涵; 二是本量表與文獻所探討的量表, 有其相似度均凸顯道德情意內涵, 亦有其故事所關注道德議題的獨創性與在地關聯; 三是本研究嘗試以多種類的信效度, 對於道德情意量表進行堅實的考驗。最後, 本研究基於道德教育的角度, 提供在量表運用與後續研究的建議, 以期增進各教育階段學生的多面向道德情意發展。

關鍵詞: 多面向道德情意量表、信效度、探索性因素分析、複核效化、驗證性因素分析

* 本篇論文通訊方式: 李琪明, 通訊方式: tli023@ntnu.edu.tw。
本文係科技部 105 年度【臺灣學生道德情意發展之質性與量化探究】(105-2410-H-003-041-SS3) 三年期專題補助計畫的部分成果。感謝在研究歷程與投稿中接受諮詢的專家學者, 以及協助的研究助理和工讀生。另感謝匿名審查人的評論與建議, 使得本文更加豐富且具意義。

道德向為教育理論與實踐中屬抽象較難教學，且不易觀其成效的範疇；而其有關情意層面的道德評測，更是國內外學界極少觸及的主題。Lee 與 Taylor (2013) 曾針對國際間具道德研究指標的《道德教育期刊》(Journal of Moral Education)，自其 1971 年創刊至 2011 年共 40 年間所刊登的 945 篇論文，以內容分析法剖析其主題及其相關因素的發展趨勢，該研究發現有關情意層面的道德教育論文僅占 199 篇 (21.6%)，另外該刊歷年在占有 43.2% ($N = 408$) 的實徵研究中，最常被運用的評測工具亦偏於道德判斷與推理等認知取向，例如「界定議題測驗」(defining issues test, 簡稱 DIT) 等測驗。因此，情意層面的道德評測，可謂亟待開拓的研究領域，尤其處在當代紛繁複雜以及價值衝突的無國界社會氛圍，其重要性實不容忽視。

論及情意層面的道德量表發展，仍需師法數十年來國際間有關道德認知層面評測的豐富成果。自從美國學者 L. Kohlberg 於 1958 年提出其博士論文，並建構道德認知發展理論以來，道德兩難困境 (moral dilemma) 的故事陳述，似乎成為道德測驗的必備元素 (Colby & Kohlberg, 1987)。誠如 Kohlberg 所編製的「道德判斷訪談」(moral judgment interview, 簡稱 MJI) 評測工具，其中最著名的為「漢茲兩難」(Heinz dilemma) 的虛擬故事，該評測工具推出之後的三十年間，Colby 與 Kohlberg (1987) 仍一直在反思其理論基礎與檢測其研究的有效性，且出版《道德判斷測量：理論基礎與研究效度》(The measurement of moral judgment: Theoretical foundations and research validation) 一書，由此可見道德認知評測有其可行性但也有其限制。前述 DIT 稍晚於 MJI 測驗，於 1970 年代由美國明尼蘇達大學 J. Rest 為首 (包括 M. Bebeau, D. Narvaez, S. Thoma 等學者)，並於 1999 年號稱「新柯伯格取向」(Neo-Kohlbergian approach) 的研究群，延續 Kohlberg 道德認知發展主軸，但對其理論有所修正與超越 (Rest et al., 1999)。其評測工具發展迄今有 DIT-1 與 DIT-2 兩個版本，亦是採道德兩難故事的題型，以及排序 (ranking) 和評定 (rating) 方式填答，因其用紙筆且可大量施測，使用上較 MJI 便利且推廣較易，其關注的是成熟道德判斷層次 (亦即原則期) 的分數高低，且認為道德發展是「基模」而非「階段」(Rest et al., 1999)。另有 Yacker 與 Weinberg (1990) 在「關懷與正義道德取向：一個評量工具」(Care and justice moral orientation: A scale for its assessment) 一文，亦發展了 12 個道德兩難故事及各自的 4 個選項，由填答者將之排序。至於在臺灣學界關於道德認知層面的量表，多半是以翻譯國外工具，例如沈六 (1986)《道德發展與行為之研究》採用 Kohlberg 的 MJI 量表，單文經 (1982)《道德教育初探》運用 DIT-1 量表等，近年來則有林慧慈等人 (2017) 針對道德思考，以及 Lee 與 Thoma (2018) 針對道德思辯，分別編製屬於臺灣在地化的道德故事量表。因此，本研究乃師法國內外有關道德認知數十年來的發展基礎，但將關注情意層面，且著重量表內容得以彰顯與臺灣的關聯性。

再者，基於道德範疇的豐富樣態，新柯伯格學派以多面向的「四成分模式」(four-components model, 簡稱 FCM)，拓展了道德判斷／推理的單一面向 (Rest et al., 1999)。FCM 強調知行統整歷程中有四個交互為用的成分，包括成分一道德敏銳度 (moral sensitivity)、成分二道德判斷 (moral judgment)、成分三道德動機 (moral motivation)、成分四道德品格 (moral character) (Rest et al., 1999)。Chambers (2011) 基於前述乃針對 FCM 特點，在牙醫專業研究領域中，發展 40 題選擇題型的「道德技能檢測」(moral skills inventory, 簡稱 MSI) 工具。此外，Casali (2011) 亦在「發展倫理決定的多面向量表」(Developing a multidimensional scale for ethical decision making) 一文，為商管領域決策的研究，發展多面向的「管理道德概況量表」(managerial ethical profile, 簡稱 MEP scale)，該量表包括八個分量表例如效益主義與義務論等，並且進而以驗證性分析證明量表的效度，與前述量表多半採取排序或選擇題，以致無法考驗其建構效度已有極大不同。因此，本研究所關注的道德情意量表，亦希冀參考道德認知量表的多面向構念，以及具有堅實信效度的考驗。

基於前述，本研究目的為發展具有信效度的多面向道德情意量表，其重點有二：一是將汲取國內外研究成果，並以道德量表慣用的故事敘說方式，展現其理念基礎的豐富性，以及與臺灣脈絡的關聯性；二是採用多面向成分，期能凸顯道德情意構念的多樣性。希冀藉此量表的發展，可作為日後進一步檢視情意層面道德發展概況，或是成效展現的具體指標，期能對於國內外道德教育理論與實踐有所助益。

文獻探討

本研究發展的「多面向道德情意」量表 (multidimensional affective morality scale, 簡稱 MAM scale), 乃參考李琪明 (2017) 歸納國內外文獻諸多相似概念, 著重在中小學課程模式建構, 所提出的道德情意發展能力多元面向, 包括道德情感 (moral emotion)、道德同理 (moral empathy)、道德認同 (moral identity)、道德關懷 (moral caring) 四者。本文為凸顯道德情意層面的完整性與豐富性, 乃以該四者為潛在變項的構面分類, 且著重在量表發展。因而, 本研究的文獻探討將分別探究其意涵且加以綜整, 另列舉實徵研究中相關量表, 以利加以轉換與運用。

一、道德情感構面之意涵與量表探究

道德情感可說是道德哲學、道德心理學、人類學以及女性主義等泰半會提及的概念, 其大抵可謂情感與道德兩個範疇的交集, 也就是情感並非均與道德有關, 道德也非僅有情感元素。道德情感常被提及的有恥感、罪惡感、驕傲、羞愧、感恩等概念, 這些情感都與認知與行為有密切關聯 (Tangney et al., 2007)。然而, Nussbaum (2008) 提及即使是喜、怒、哀、樂等情感, 也在道德範疇佔有一席之地, 其可以透過各種方式阻斷判斷, 或使人變得不穩定和過度需要。這在 Eisenberg (2000) 對於道德情感的實徵研究中早已發現, 其指出道德情感中常被提及的羞愧、罪惡感與恥感等, 與道德行為有著不同程度的關聯, 而快樂、生氣與悲傷等情緒與氣質, 亦會影響一個人的道德言行。其次, 道德情感並非僅限於私人領域亦可擴展至公共領域, 誠如 Nussbaum (2013) 在《政治的情感: 愛對於正義為何重要》(Political emotions: Why love matters for justice) 一書, 強調愛對於正義的重要。再者, 道德情感的意涵可謂與「社會情感學習」(social-emotional learning, 簡稱 SEL) 所強調的自我覺知 (self-awareness) (了解自己的情感、個人目標與價值), 以及社會覺知 (social awareness) (了解社會規範與文化脈絡以產生同情與悲憫) (Durlak et al., 2015), 也有類似之處。

至於在道德情感的量表方面, 較為相關者為前述 SEL 並多為自陳量表, 例如「中學年輕人社會情感學習評量工具」(social-emotional learning assessment measures for middle school youth) (Haggerty et al., 2011) 與「社會情感學習的評量工具資源」(resources for measuring social and emotional learning) (Hagen, 2014) 等。此外, Lietz 等人 (2011) 發表「同理評量指標: 多元面向同理模式的驗證性分析」(The empathy assessment index [EAI]: A confirmatory factor analysis of a multidimensional model of empathy) 一文, 該研究重點雖在同理, 但其五個因素中情意回應 (affective response, AR)、情感控制 (emotion regulation, ER)、對自我與他人的覺知 (self-other awareness, SOA) 三者都與道德情感有關, 例如其中第 14 題「當我看到朋友哭泣時, 我也會感到難過。」(情意回應)。再者, 早在 Davis (1980) 「人際反應指數」(interpersonal reactivity index, 簡稱 IRI) 共 28 題量表, 例如第 18 題「當我看到某人被不公平地對待時, 我有時並不會對他們感到同情。」(反向題), 其題型與題意對於道德情感構面, 都甚具參考價值。

二、道德同理構面之意涵與量表探究

對於道德同理闡釋完整且清晰的論著, 首推 Hoffman (2000) 《同理心與道德發展: 對於關懷與正義的啟示》(Empathy and moral development: Implications for caring and justice) 一書。Hoffman 強調道德同理兼具認知與情感兩個面向, 其中情感性同理的適度激發 (避免過度喚起), 可以產生利社會作用, 並可進而連結正義與關懷兩個道德原則的實踐 (Hoffman, 2000, p. 3)。此外, Segal 等人 (2012) 在「社會同理指標之發展: 一個試探性因素分析」(Developing the social empathy index: an exploratory factor analysis) (簡稱 SEI 量表) 一文, 提出社會同理一詞, 是指能瞭解人們的多元

背景，洞悉制度不公與剝削的脈絡，以激發正向的社會改變且促進社會幸福，此與前述 Hoffman 將同理以正義與關懷的角度關注，亦與本研究的道德同理意涵相通。至於道德同理的實徵性研究，Lietz 等人（2011）建構的 EAI 同理模式五因素中，「觀點取替」（*perspective taking, PT*）以及「同理的態度」（*empathic attitudes, EA*）二者，其題目可具參考價值，例如第 2 題「我能在別人的立場想像是甚麼樣的感覺」。另 Segal 等人（2012）在其社會同理一文，經由探索性因素分析歸結兩因子共 18 題，因素 1 為系統障礙的脈絡性理解，因素 2 為宏觀意識到自我他人的與觀點取替，題目則類似「我相信出生貧困的人，在獲得經濟福祉方面，會比非出生於貧困的人，存在更多障礙。」亦可供本研究參考。

三、道德認同構面之意涵與量表探究

有關道德認同的理念可謂以 Blasi（1984）在其《道德認同：其在道德功能中的角色》（*Moral identity: Its role in moral functioning*）一書為著，其指出道德的判斷與行動之所以常無法連結，關鍵因素為缺乏自我對於道德的認同，但該書主要由社會認知的角度出發，較少提及情意面向。Reimer（2012）在《青少年百科》（*Encyclopedia of adolescent*）中對於道德認同的探究有了新詮釋，他提及道德認同是統整自我意識的一種承諾，且該自我意識是激發利社會行動的重要因素，以往道德認同相關研究較侷限於理性思維與認知發展理論，但近年諸多學者將道德自我與認同加入了人格特質、目標、情緒與自我經驗等面向，以展現其情意層面。然而，道德認同並非侷限自我的個人範疇，亦同時須擴展至社會，也就是由個人要成為一個怎樣的人（自我認同），進而關心要怎樣的理想社會（群體認同），方可達到道德意識進化中所謂「美好生活的理念」（*idea of the good life*）（Habermas, 1979, p. 89）。

針對道德認同的實徵性研究，Aquino 與 Reed II（2002）在其「道德認同之自我重要性」（*The self-importance of moral identity*）論文中，指出其分別以青少年、大學生與成人的六個調查研究，檢視其道德認同之自我重要性，及其與道德認知、道德行為之間的關係，進而依據道德認同的心理計量特質，建構出具有道德認同量表（*moral identity measure*，簡稱 *MI measure*）。該 *MI* 量表（Reed et al., 2007, p. 191）題型為先列出九個德行（例如關懷、助人等）為題幹，然後有 10 個待答題目，例如勾選「我很希望自己擁有以上這些特質」的同意度，以表達其道德認同。此外，有關道德的集體認同相關量表，可參考「各層面認同問卷」第四版（*aspect of identity questionnaire-IV*，簡稱 *AIQ-IV*）（Cheek & Briggs, 2013），共有 45 題分屬個人、關係、社會與群體的認同，例如第 29 題「我對於身為自己國家的公民感到自豪。」可為參考。

四、道德關懷構面之意涵與量表探究

本研究的道德關懷構面非指實際行動階段，而是指行動激發動機且充分知能的準備。其意涵誠如 Noddings（2005）編著《教育公民具有全球關注之心》（*Educating citizens for global awareness*）提及，關懷（*care for*）並非限於親近的人而應擴及全人類與生態，因此對於人類的貧富差距以及戰爭和平等議題，需要瞭解的是其間是否違反了經濟與社會的公平正義與基本人權，並需理解族群與宗教的差異性。Held（2006, p. 157）在《關懷倫理：個人的、政治的與全球的》（*The ethics of care: Personal, political, and global*）一書，也同樣指出關懷倫理在公共領域中提供了不同的視野與互補作用，而且擺脫了道德情意在傳統理論中有害於正義達成的負面形象。

關於評測道德關懷的熱誠與動力的量表，可參考 Doolittle 與 Faul（2013）所編製的「公共參與量表」（*civic engagement scale*，簡稱 *CES*），計有態度構面 8 題與行動構面 6 題，前者對本研究較有參考價值，例如第 3 題「我相信我有責任要幫助窮困與飢餓者。」、第 4 題「我願意為社區服務。」然而，Bandura（1991）提出道德疏離（*moral disengagement*）概念，或可成為道德關懷的另類思考，其意謂著情意元素、自我控制能力，以及外在環境變項等，往往是道德認知與行動脫

節的關鍵因素。因此，道德疏離或可視為道德關懷的反向思考，亦即要使道德關懷付諸實踐需要避免道德疏離。Bandura 等人（1996, p. 374）進而發展 32 題的道德疏離（MD）量表，其中包括一些似是而非以及推卸責任的「合理化」藉口，例如「未經他人允許騎走腳踏車只是借用而已。」、「取笑他人又不會真正地傷害到他。」均可作為道德關懷構面的反向題。

五、小結

依據前述文獻探討的理念剖析，本研究「多面向道德情意」是指聚焦情意層面的道德範疇，並融合其多元豐富且有內在統整關係的面向，包括道德情感、道德同理、道德認同、道德關懷等元素，該些元素或是分立也可能有所重疊，其與認知及行動層面亦是相互影響，並且同時關注個人與公共領域的展現。其具體意涵為：對於社會議題或道德現象，具有情感的敏銳度與道德感；能站在多方立場及其社會結構脈絡，加以換位思考、想像與體會，且引發適當情感回應；對於自我道德價值或行動擁有自信與榮耀感，對於典範他人或理想社群有著學習與期待；並且對於道德實踐具有承諾與動力，建立良好關懷關係，以發揮仁慈心與正義感。因此，本研究乃奠基文獻探討的理論與參考國外相關既有量表（國內自創道德情意相關量表幾乎闕如），進而擬訂本量表理念架構與題項初稿，量表總分的得分高/低，將表示前述多面向道德情意具體意涵展現度的高/低，另將視統計驗證結果以區別分量表及其意涵。至於道德情意的發展，是否如同道德認知發展理論一般有其序階或是基模，以及施測對象所呈現的道德情意概況，則非本文所研究範圍。

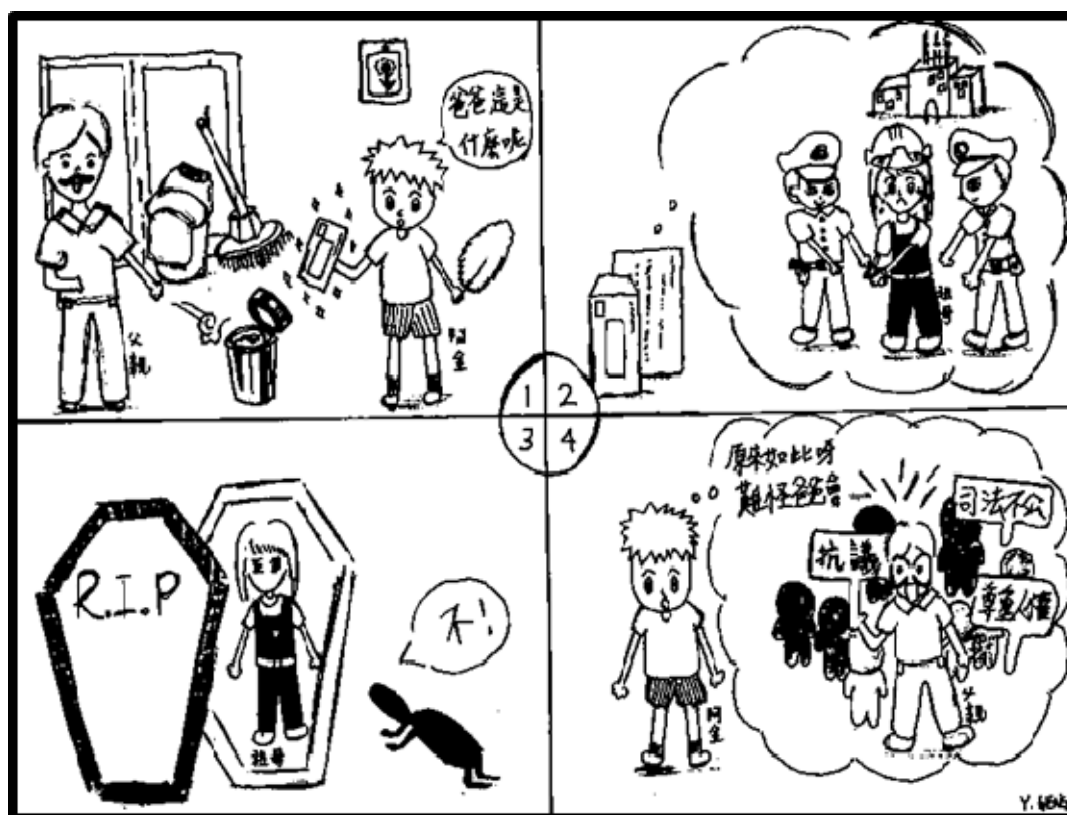
研究方法與歷程

本研究發展多面向道德情意量表，自 2017 年 8 月起至 2019 年 7 月歷時 2 年，其方法與歷程包括奠基理念架構以擬定量表初稿、針對初稿邀約專家學者以評定其內容效度、經修改後進行量表預試，以及統計後進而正式施測。由於在量表的信效度方面，近年來國內有關量表發展，已有運用驗證性分析（confirmatory factor analysis, 簡稱 CFA）進行測驗模型考驗，例如甘孟龍等人（2016）、程姿瑩等人（2016）等論文。因此，本研究信效度考驗歷經多方統計測試與反思後，本量表的發展將採兩種類別的信效度考驗：一為將採用內容效度與內部一致性信度為檢驗標準，形成多面向道德情意量表（詳版）；二為採用建構效度與組合信度為信效度考驗，以驗證性因素分析加以統計，形成較為精簡型的多面向道德情意量表（簡版）。以下將各個歷程分別詳述。

一、理念架構與擬定初稿

本研究歷經文獻與時事分析，且經召開多次專家諮詢與助理會議討論後，乃以四則與臺灣當前社會有深度關聯，且分別在公與私領域均顯重要的四個道德故事（A：分享照片的風波、B：遲來的真相、C：吃不到的巧克力、D：不平靜的野餐）加以組合。其分別為關於網路言論自由的界線、轉型正義、全球化下的童工人權與貧窮、以及國際移工衍生的族群等議題，以彰顯道德範疇的多樣性與複雜性，及其內容的豐富性與脈絡性。每則故事都基於相同概念，初步各先擬定 27 個題項，並以 Likert 五點量表自陳方式勾選同意度（包括非常不同意、不同意、普通、同意、非常同意）。此外，本量表將每則故事內容繪製成四格漫畫，以增加量表對於填答者的易讀性與親近性。茲以 B 故事為例，呈現其漫畫與敘事（見圖 1），並將各題項與構面的內涵重點加以對照（見表 1）。

故事 B：遲來的真相



阿金在整理舊家時看見一封信，問了父親才知道是祖母遺留下來的。父親這才提起，祖母身在威權專制不依法審判的時代，有一天下午工廠突然出現保安大隊（負責維護當時公共安全的機關），他們奉上級命令前來追捕，並宣稱有人檢舉祖母參加地下組織（秘密、非法組織），所以要將祖母當場逮捕。保安大隊表示如果祖母願意供出其他參與者就可以減輕其刑，但無論怎麼嚴刑逼供，她仍是全程保持靜默，經過沒多久以後，苦等消息的家人只收到前來領取遺體的通知……。阿金終於明白為何常見父親看到祖母照片時掉淚，還有聽聞侵犯人權時總是很激動，並且積極參與人權促進活動的原因。

圖 1 量表故事 B 之圖示與內容

表 1 量表理念架構與題項對照表—以 B 故事為例 (ACD 故事均雷同)

構面	內涵重點	量表題項初擬
道德情感	1. 對於個案中人物處境，引發自我的情緒／感受	1. 看到故事中阿金父親在哭的時候，我通常會覺得這沒有什麼關係。(反向題)。
	2. 個案中人物若是親友，引發自我的情緒／感受	2. 看到朋友像阿金父親失去親人時，我通常會產生負面情緒。
	3. 克制自己因個案引發過度聯想，而導致情緒感染	3. 當我看到類似阿金父親的故事，很容易被他的情緒感染，開始去想自己也很負面的事。(反向題)。
	4. 針對個案人物的處境，將自我情感連結道德評價	4. 看到故事中阿金的祖母被嚴刑逼供，我覺得這很殘忍。
	5. 針對類似通案，將自我情感連結道德評價	5. 當我看到政府濫用公權力使人民權益受損時，我會覺得這令人氣憤。
道德同理	6. 想像個案主角他／她的心情	6. 我能夠想像故事中阿金得知真相後產生的負面心情。
	7. 體會個案中人物若是自己親友，他／她們的心情	7. 這類被迫失去親人的事不會發生在我親友身上，因此我無法體會他們的心情。(反向題)
	8. 對於個案中各個角色，加以换位思考其心情	8. 我能夠想像故事中保安大隊人員必須處理這類情形的複雜心情。
	9. 想像個案中不同立場者，他／她們的心情和想法	9. 我認為在人權議題當中，某方是對的，我就不會浪費時間去關心其他人的心情和想法。(反向題)
	10. 體會通案中類似角色的心情	10. 我能夠體會威權專制時代下遭受壓迫的無奈心情。
	11. 對於類似通案的可能原因，加以多元想像	11. 我通常可以想像有人在威權時代下遭受壓迫的原因。
	12. 對於不同群體所處的環境脈絡，予以同情地理解	12. 我通常可以理解政治受害者面臨的人權壓迫。
道德認同	13. 個案雖有其多元複雜性，但認同道德判斷的重要性	13. 在威權時代下大家的處境有所不同，所以我很認同不必判斷誰對誰錯的說法。(反向題)
	14. 聚焦個案中某些作為，且產生道德期待與認同	14. 有人在威權時代下犧牲生命避免牽連他人，會讓我感動產生認同感。
	15. 聚焦個案中某些人物，且產生道德感動與認同	15. 如果我是故事中的保安大隊，我也會帶走阿金的祖母。(反向題)。
	16. 對於類似通案激發道德承諾	16. 我所認同的社會是：即使為了社會安定和秩序，政府仍應保障人民生命安全。
	17. 不接受將違反道德的作為，成為追求某些價值的藉口	17. 政府為了維持政局穩定，可以採取高壓統治，我很認同這種做法。(反向題)
道德關懷	18. 有心了解個案的背後成因	18. 我願意付出心力，多加了解轉型正義案例的來龍去脈。
	19. 避免既往不究的消極作為	19. 我覺得現在的社會和諧比較重要，再回頭去追究當時政府的作為是沒有意義的。(反向題)
	20. 避免似是而非的比較以卸責	20. 我覺得保安大隊抓走阿金祖母，相較於戰爭的大量死傷，這類的政治迫害是可以接受的。(反向題)
	21. 避免希望受害者隱忍／遺忘的消極作為	21. 我覺得我們應該忘掉過去的歷史傷痛，以前的事就讓它過去。(反向題)
	22. 避免盲從多數人或與己無關的心態	22. 我覺得威權政府的壓迫只是時代的產物，與現在的我們沒有關係。(反向題)
	23. 願意積極維護親友權益	23. 當我得知自己好友的長輩曾受政治迫害，我願意付出時間心力為他們爭取權益。
	24. 願意積極彰顯公益／公義	24. 當我遇到有人被剝奪人權時，在可能與他人意見分歧的情況下，我仍願意挺身而出。
	25. 願意積極向親友倡議宣導正向價值	25. 雖然可能因為政治傾向不同而使親友不諒解，我仍願意和周遭的人分享政治迫害對於人權侵害的嚴重性。
	26. 願意承擔責任為公義發聲	26. 在必須承擔責任的情況下，我仍願意為政治受難者的人權議題發聲。
	27. 激發自己與他人道德行動的承諾	27. 即使可能遭受威脅，我仍願意與更多人一起關心人權議題。

註：標註粗體的 11 項，為量表（簡表）最終 12 題項所彰顯的理念重點。

二、內容效度評定

本研究擬定量表題項初稿之後，為求其內容周延且符合研究對象的理解與經驗，乃邀請 15 位至少具碩士級學歷的專家學者，分別 5 位任職國中、5 位任職高中、5 位任職大專院校，且分布於臺灣多縣市。本研究邀請專家學者以其教育專業與實務經驗，對於量表初稿加以評定題項的適當性（包括「非常適當」、「適當」、「不適當」、「非常不適當」），以及提供該些道德故事內容與題項，是否適合學生閱讀與理解的具體文字修改建議。經計算內容效度指數（Content Validity Index, CVI）（評定「非常適當」與「適當」的人數，除以填答人數的百分比），統計結果為故事 A 整體 CVI = 0.95，各題均落在 0.80~1.00，故事 B 整體 CVI = 0.97，各題均落在 0.80~1.00，故事 C 整體 CVI = 0.98，各題均落在 0.87~1.00，故事 D 整體 CVI = 0.98，各題均落在 0.93~1.00。雖然四則故事的所有題目都合乎 CVI 的 0.80 以上，但是參考專家學者的修改建議後，乃將題數加以刪減與潤飾內容，以形成預試問卷四個故事（A、B、C、D），各有 12 個題項（分別標示 1~12）合計 48 題。

三、預試與正式施測

（一）研究對象

本研究母群體為臺灣國中至大專院校等教育階段學生，取樣施測分成預試樣本與正試樣本，均採分層立意取樣方式，且力求分布區域的均衡。預試施測學校包括北區四所國中、四所高中、二所大學，中區二所國中、二所高中、一所大學，南區二所國中、二所高中、一所大學，以及東區一所大學。預試樣本發出 606 份問卷全數回收，廢卷 21 份（廢卷率 3.5%，廢卷類型包括整題組未填答 5 份，以及填答特定數值 16 份），故有效樣本為 585 份，包括教育階段（國中 240、高中 240、大學 33、研究所 93）、區域（北區 276、中區 150、南區 150、東區 30）、性別（男 273、女 332、1 人未填）。正式施測學校包括北區二所國中、二所高中、二所大學，中區三所國中、二所高中、二所大學，南區一所國中、二所高中、二所大學，以及東區二所國中、二所高中、二所大學。正式樣本發出 737 份問卷且全數回收，有效樣本 708 份，包括教育階段（國中 235、高中 242、大學 114、研究所 117）、區域（北區 176、中區 171、南區 181、東區 180）、性別（男 292、女 415、1 人未填）。

（二）分析軟體與統計方法

本研究採用 IBM SPSS Statistics 23 與 Amos 25 Graphics 為分析軟體，針對量表預試與正式施測進行統計分析，計有四個階段：第一階段是將預試樣本以 SPSS 計算其內部一致性信度，並基於預試前的內容效度檢驗，形成量表的詳表。接續的第二至四階段是將預試樣本以 SPSS 隨機分割為兩個樣本，預試樣本_1（pre_G1）有 283 份與預試樣本_2（pre_G2）有 302 份，以探求簡表的建構效度與組合信度。其中，第二階段原本是預計將預試樣本_1 直接進行驗證性因素分析，卻產生大量題項需刪減，才能符合模型檢驗的情形，為避免以驗證性分析誤作為探索性分析的不妥，故而將預試樣本_1 以 SPSS 先進行探索性因素分析（exploratory factor analysis, EFA），但仍秉持原理念的道德情意構面加以適切組合，與其他未加預設因子數或理念基礎者有所不同。第三階段接續前階段結果，將預試樣本_2 以 Amos 進行驗證性因素分析。第四階段為接續前述階段結果，將正式樣本以 Amos 多群組分析考驗模式進行複核效化（cross-validation）檢驗，確定該量表模式的複製與穩定性。

結果

一、第一階段（詳版量表）：以預試樣本進行信效度考驗

本研究奠基於文獻理念基礎與專家學者的內容效度，進而將預試樣本（ $N = 585$ ）進行項目分析後，計算其內部一致性信度。原本四則故事 48 題刪除 4 題後剩下 44 題，其總量表 Cronbach's $\alpha = .922$ 。其次，四個構面分別形成的分量表：道德情感分量表 10 題（刪 2 題 A2、C2），其 Cronbach's $\alpha = .741$ ；道德同理分量表 11 題（刪 1 題 A6），其 Cronbach's $\alpha = .821$ ；道德認同分量表 11 題（刪 1 題 A7），其 Cronbach's $\alpha = .731$ ；道德關懷分量表 12 題（未刪題），其 Cronbach's $\alpha = .831$ 。因此，整體而言本研究發展的多面向道德情意量表（詳版），有四則道德故事計有 44 個題項，其不僅符合理念基礎與內容效度，而且內部一致性信度總量表為 .922，四個分量表介於 .731~.831，均達要求標準 0.7 以上。

二、第二階段（簡版量表）：以預試樣本_1 進行探索性因素分析

為發展簡版量表，本研究將預試樣本區分為兩群體，以預試樣本_1（ $N = 283$ ）進行探索性因素分析，先行以項目分析的高低分組相關性 CR 值，以及所有題項與總分的相關性進行初步篩選，原本 48 題剩下 46 題。接著以因素分析的主軸因子法（principal axis factoring, PAF）直交轉軸最大變異量（varimax）多次進行測試，其中包括三因子、四因子與五因子的測試，不過因為故事 A 的題項在各個測試中都幾乎僅剩極少題數，因此最後乃決定刪除故事 A 的 12 題，持續以三個故事 B、C、D 共 36 題進行三因子測試。經 PAF 測試，其 Bartlett's 球形檢定卡方值為 1913.057（ $p < .001$ ），KMO 值為 .907，表示相關矩陣中存有共同因素，具有良好因素分析適切性，另則保留題項因素分析負荷量至少 0.45 者，並達到整體之累積解釋變異量達 56.024%。結果顯示萃取三個因子計 17 題項，F1 有 8 題其內部一致性信度 Cronbach's $\alpha = .876$ ，F2 有 5 題內部一致性信度為 .793，F3 有 4 題內部一致性信度為 .724，總量表內部一致性信度為 .897，均達要求標準 0.7 以上（見表 2）。由於本研究乃奠基道德情意理念，經與原構念的四個面向對照後，可維持道德情感、道德同理與道德關懷三個構面，但道德認同則需拆解分屬 F1 與 F3 兩個因子，因而將 F1 命名為「道德承諾與關懷」、F2 為「道德同理」、F3「道德情感與期待」。

表 2 預試樣本_1 探索性因素分析負荷矩陣及信度摘要表

題項	Factor			Cronbach's Alpha (N of Items)	
	1	2	3		
D11	.763	.268	.048	0.876 (8)	
C12	.739	.228	.127		
D12	.714	.174	.246		
D9	.694	.118	.061		
C11	.560	.249	.220		
C9	.542	.153	.327		
C10	.489	.249	.324		
B12	.454	.214	.372		
C6	.270	.689	.109		0.793 (5)
C3	.229	.647	.141		
C7	.335	.591	.285		
C1	.212	.522	.251		
B7	.071	.498	.280		
B5	.129	.130	.681	0.724 (4)	
B2	.126	.144	.608		
B4	.225	.270	.542		
B9	.149	.236	.500		
整體解釋量 56.024%				0.897 (17)	

註：主軸因子最大變異量分析 (extraction method: principal axis factoring, rotation method: varimax with kaiser normalization.)；旋轉在 5 次疊代中收斂 (rotation converged in 5 iterations.)。

三、第三階段 (簡版量表)：以預試樣本_2 進行驗證性因素分析

接續預試樣本_1 的探索性因素分析結果，本階段進一步以樣本_2 ($N = 302$) 進行驗證性因素分析，求取模式的適配度，以及組合信度與建構效度。以 Amos 運算評鑑測量模型時，主要有數個步驟，分別為常態性檢定、檢驗違犯估計、檢驗模型適配度、檢驗證收效度、檢驗區別效度，且需符合理想標準 (陳寬裕、王正華，2016, pp. 177-178, 203-204)。本研究依據前述步驟多次反覆測試以及修改指數 (modification index, MI) 顯示的建議，最終刪除 5 題而剩下 12 個題項的「多面向道德情意量表」二階單因子模式。茲將本研究量表自原四個構念的 48 題 (四則故事)，至預試樣本_1 的三個因子 17 題 (三則故事)，再至預試樣本_2 的三個因子 12 題而確認此模型，其變化一覽表見表 3。

表 3 簡版量表題項變化一覽表

原構念	原題項	Pre_G1	Pre_G2	原構念	原題項	Pre_G1	Pre_G2
ME1	A1	x	x	MI	A7	x	x
	A2	x	x		A8	x	x
	A5	x	x		A9	x	x
	B1	x	x		B3	x	x
	B2	F3	F3		B8	x	x
	B5	F3	F3		B9	F3	F3
	C1	F2	x		C4	x	x
	C2	x	x		C8	x	x
	C5	x	x		C9	F1	F1
	D1	x	x		D4	x	x
	D2	x	x		D8	x	x
D6	x	x	D9	F1	F1		
A3	x	x	A10	x	x		
A4	x	x	A11	x	x		
ME2	A6	x	x	A12	x	x	
	B4	F3	F3	B10	x	x	
	B6	x	x	B11	x	x	
	B7	F2	x	B12	F1	x	
	C3	F2	F2	C10	F1	x	
	C6	F2	F2	C11	F1	F1	
	C7	F2	F2	C12	F1	F1	
	D3	x	x	D10	x	x	
	D5	x	x	D11	F1	x	
	D7	x	x	D12	F1	F1	

註：原構念 ME1 道德情感、ME2 道德同理、MI 道德認同、MC 道德關懷；新構念 F3 道德情感與期待、F2 道德同理、F1 道德承諾與關懷；灰色底粗體字表示新構面題項；原構念有 48 題，Pre_G1 有 17 題，Pre_G2 有 12 題。

根據最終預試樣本₂的 12 個題項驗證結果，並對照陳寬裕與王正華（2016）一書所論及的標準，其結果分析如次：

（一）12 個觀察變項均符合常態性檢定，因為偏態的絕對值小於 3，峰度的絕對值小於 8，並且 Mardia 係數為 85.829，小於題數*（題數+2）（12*14 = 168）。

（二）本模式沒有違反估計，因為檢驗 12 題都沒有負的誤差變異數、沒有標準化迴歸加權係數超過或大於 0.95 而接近 1、也沒有太大的標準誤，因為 C.R. (t 值) 的 p 值都達顯著（小於 0.001）（見表 4）。

（三）檢驗模式的適配度模型（見圖 2）及其絕對、增量和精簡適配指標的相關指數幾乎都達標準（見表 5）。

（四）檢驗建構效度中的收斂效度結果，12 個觀察變數因素負荷量介於 0.51~0.85 都超過 0.5 的標準、個別信度介於 0.26~0.67 都超過 0.2 的標準，三個因子的組合信度由 F1 至 F3 依序為 .851、.813 與 .753 都達到超過 0.6 的標準、平均變異數抽取量(AVE)分別為 .536、.592 與 .439，以 0.5 的標準而言僅有一個稍弱但可接受，另 AVE 的平方根分別為 .732、.770 與 .662，都超過 0.5 的標準，整體量表的組合信度為 .843、平均變異數抽取量為 .641、AVE 的平方根為 .802，所以收斂效度均達標準，表示本量表內在結構的聚斂效果佳。

（五）檢驗建構效度中的區別效度結果，三個構面前述的 AVE 的平方根，均大於各個構面間的相關係數，表示本量表具有區別效度（見表 6）。綜而言之，多面向道德情意量表（簡版）運用驗證性因素分析結果，理論模式足以適切解釋實際的觀察資料，且具有良好的組合信度與建構效度。

表 4 預試樣本_2 觀察變項之 Regression Weights (標準化迴歸係數)

構面與題項	總量表與構面	Estimate	S.E.	C.R. (p-value)
道德承諾與關懷	<--- 多面向_道德情意	1.463	.289	5.069***
道德同理	<--- 多面向_道德情意	1.419	.277	5.121***
道德情感與期待	<--- 多面向_道德情意	.874	.130	6.702***
C9	<--- 道德承諾與關懷	1.582	.228	6.947***
C11	<--- 道德承諾與關懷	1.419	.206	6.894***
C12	<--- 道德承諾與關懷	1.963	.279	7.036***
D9	<--- 道德承諾與關懷	1.516	.235	6.449***
D12	<--- 道德承諾與關懷	1.736	.250	6.944***
C3	<--- 道德同理	1.772	.246	7.216***
C6	<--- 道德同理	1.932	.285	6.771***
C7	<--- 道德同理	1.760	.248	7.110***
B2	<--- 道德情感與期待	1.622	.176	9.194***
B4	<--- 道德情感與期待	2.029	.184	11.045***
B5	<--- 道德情感與期待	1.973	.193	10.200***
B9	<--- 道德情感與期待	1.234	.161	7.642***

表 5 預試樣本_2 測量模型適配度指標檢核表

統計檢定量	標準值	Pre_G2
χ^2	越小越好 ($p \geq \alpha$ 值)	142.661 ($p < .001$)
絕對適配指標		
χ^2/df	1~5 之間	2.797
GFI	大於 0.9	.927
AGFI	大於 0.9	<u>.888</u>
RMR	小於 0.08	.041
SRMR	小於 0.08	.055
RMSEA	小於 0.08	.077
增量適配指標		
NFI	大於 0.9	.906
NNFI (TLI)	大於 0.9	.918
CFI	大於 0.9	.973
RFI	大於 0.9	<u>.878</u>
IFI	大於 0.9	.938
精簡適配指標		
PNFI	大於 0.5	.700
PGFI	大於 0.5	.606
CN (HOELTER0.01)	大於 200	<u>164</u>

註：□表示未達標準，底線表示接近標準，其餘均達標準。

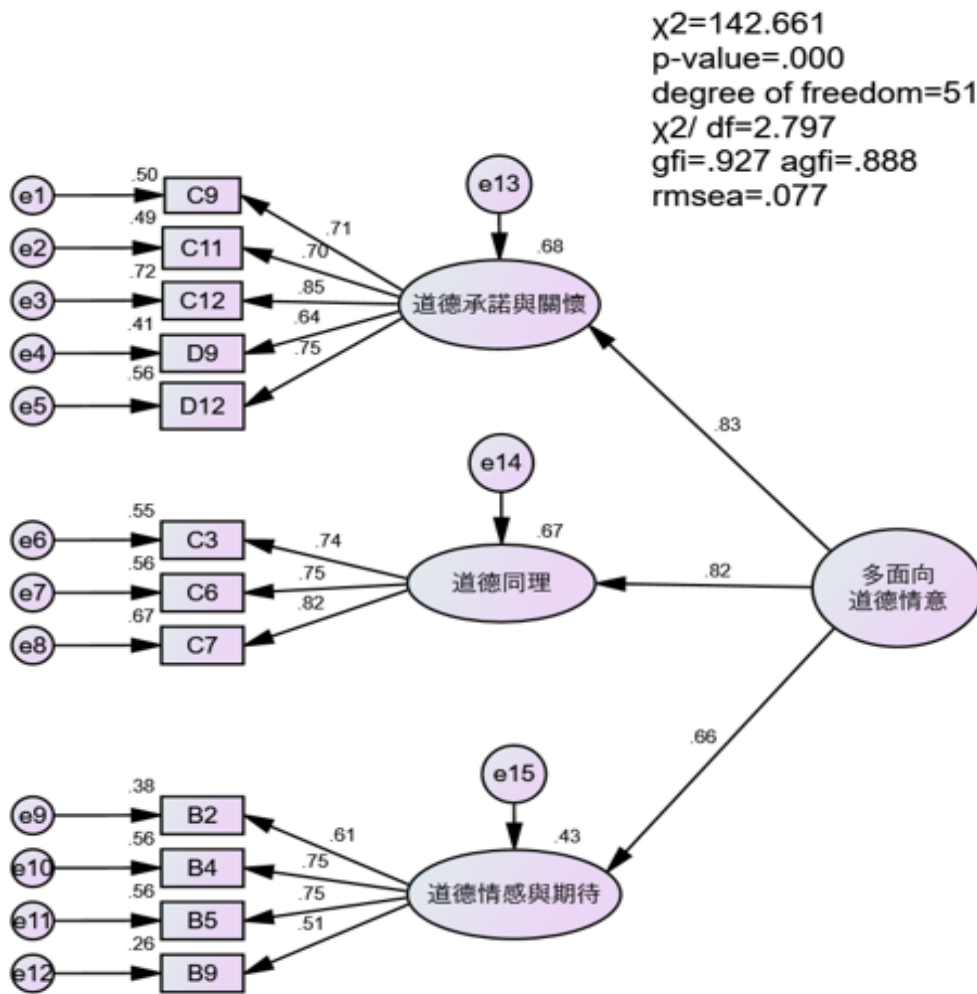


圖 2 預試樣本_2「多面向道德情意量表」二階單因子模式

表 6 預試樣本_2 量表區別效度檢定表

構面	題數	道德承諾與關懷	道德同理	道德情感與期待
道德承諾與關懷	5	0.732		
道德同理	3	.561**	0.770	
道德情感與期待	4	.451**	.415**	0.662

註：各細項代表量表三個構面的相關係數，且均達 0.01 顯著水準；對角線粗體灰底表示 AVE 平方根，顯示該值均大於非對角線之值。

四、第四階段（簡版量表）：以正式施測樣本進行複核效化檢驗

本研究接續前階段的探索性與驗證性因素分析後，乃形成正式施測的多面向道德情意量表（簡版），包括三則故事（分別為故事一：遲來的真相、故事二：吃不到的巧克力、故事三：不平靜的野餐）的組合，計有三個構面（道德情感與期待、道德同理、道德承諾與關懷），共有 12 題項（重新編碼為 X1~X12），並分別對照原理念中的重點（見表 7）。

表 7 正式施測樣本新舊故事編碼、理念與簡版量表題目對照表

構面名稱	對照原理念 (另見表 1)	原故事	原題項	題目	新題項	新故事
F3 道德情感與期待	4. 針對個案人物的處境，將自我情感連結道德評價	故事 B	B2	看到故事中阿金的祖母被逮捕後最終死亡，我覺得這樣很殘忍。	X1	故事一
	7. 體會個案中人物若是自己親友，他／她們的心情		B4	如果這個故事中失去親人的事發生在我自己親友身上，我能體會他們的心情。	X2	
	5. 針對類似通案，將自我情感連結道德評價		B5	當我在現實生活中看到政府濫用公權力使人民權益受損時，我會覺得這個行為令人氣憤。	X3	
	17. 不接受將違反道德的作為，成為追求某些價值的藉口		B9	我期待的是一個即使為了安定和秩序，人民基本權利仍受到保障的社會。	X4	
F2 道德同理	6. 想像個案主角他／她的心情	故事 C	C3	我能夠想像故事中傑克被關在工廠過勞受虐而產生的絕望心情。	X5	故事二
	10. 體會通案中類似角色的心情		C6	我能夠體會在現實生活中，工作時遭受虐待的痛苦心情。	X6	
	12. 對於不同群體所處的環境脈絡，予以同情地理解		C7	無論是否為社會底層，我可以理解他／她們所遭受到的生存困境。	X7	
F1 道德承諾與關懷	16. 對於類似通案激發道德承諾（1）貧窮弱勢	故事 C	C9	聽聞有人積極關心貧窮國家中受苦的童工時，我會因而感動。	X8	故事二
	23. 願意積極維護親友權益		C11	假使我的親友因疾病陷於貧困，我願意主動付出時間心力提供協助。	X9	
	26. 願意承擔責任為公義發聲	C12	即便必須承擔責任，我仍願意與志同道合的人一起為貧窮的人推動某些制度改變。	X10		
	16. 對於類似通案激發道德承諾（2）國際移工族群議題	故事 D	D9	我認為自己是一個願意關心移工權益的人。	X11	
	27. 激發自己與他人道德行動的承諾		D12	即使和主流文化會有所衝突，我仍願意與他人合作一起促進族群平權。	X12	

註：故事一請見「圖 1 量表故事 B 圖示與內容」示例，故事二與三見附錄。

為進一步驗證簡版量表的穩定性並是否可複製模式以適合不同的樣本，乃以預試樣本₂ 做為測定樣本（ $N = 302$ ），而正式施測樣本（ $N = 708$ ）做為效化樣本，進行複核效化驗證。本研究先以測定樣本的模式套入正式施測樣本，經統計所得相關係數（見圖 3 與表 8）乃獲實徵資料支持，顯示其適配性。接續，以 Amos 多群組分析的巢套模式比較法，依據 Byrne (2010, pp. 257-268) 主張若要計算複核效度可採用五種模型，亦即模型 1 因素負荷量（measurement weights）、模型 2 結構係數（structural weights）、模型 3 結構共變異數（structural covariances）、模型 4 結構殘差（structural residuals）、模型 5 測量殘差（measurement residuals），並且這 5 種模式是分別奠基在前一模式的數值加以比較其是否同值。本研究的統計結果（見表 9），五個模型的 ΔCFI 均小於 0.01、

ΔTLI 也都小於 0.02，僅有模型 4 的 $\Delta\chi^2$ 的 p 值小於 0.05，整體而言，五種模型的複核效化考驗，皆可獲得實徵資料以及模式複製的支持。

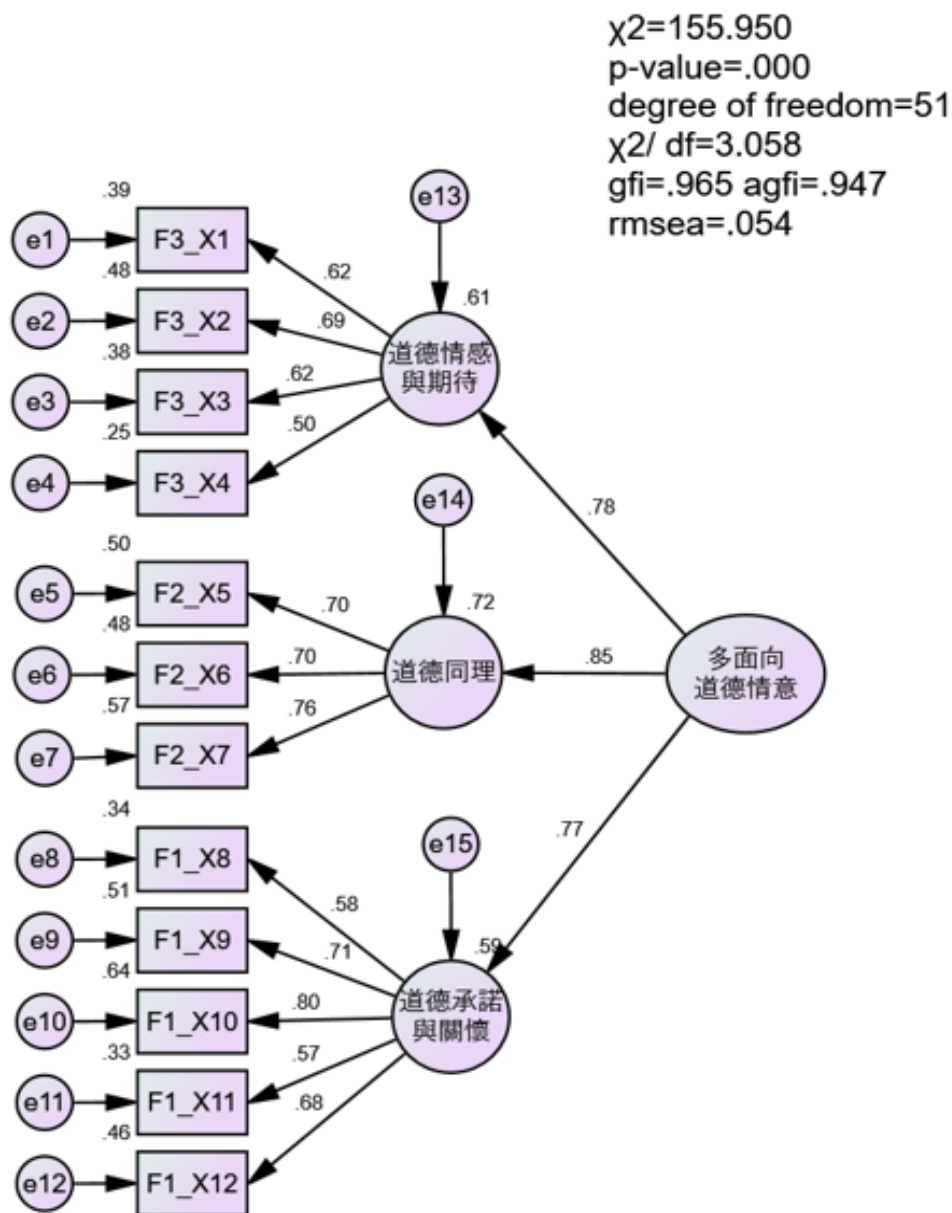


圖 3 正式施測樣本「多面向道德情意量表」二階單因素模式

表 8 正式施測樣本之測量模型適配度指標檢核表

	統計檢定量	標準值	Formal_G
絕對適配指標	χ^2	越小越好 ($p \geq \alpha$ 值)	155.950 ($p < .001$)
	χ^2/df	1~5 之間	3.058
	GFI	大於 0.9	.965
	AGFI	大於 0.9	.947
	RMR	小於 0.08	.027
	SRMR	小於 0.08	.041
	RMSEA	小於 0.08	.054
增量適配指標	NFI	大於 0.9	.941
	NNFI (TLI)	大於 0.9	.947
	CFI	大於 0.9	.959
	RFI	大於 0.9	.923
	IFI	大於 0.9	.959
精簡適配指標	PNFI	大於 0.5	.727
	PGFI	大於 0.5	.631
	CN (HOELTER0.01)	大於 200	351

註：□表示未達標準，其餘未標示者表示均達標。

表 9 複核效化五種模型檢驗表

模式	χ^2 (df)	$\Delta\chi^2$ (Δdf)	P 值	CFI	ΔCFI	TLI	ΔTLI
基線 (Unconstrained)	298.712 (102)	—	—	.951	—	.937	—
模型 1 因素負荷量 (Measurement weights)	312.633 (111)	13.921 (9)	0.125	.950	-0.001	.940	0.003
模型 2 結構係數 (Structural weights)	314.533 (113)	1.899 (2)	.387	.950	0.000	.941	0.001
模型 3 結構共變異數 (Structural covariances)	316.978 (114)	2.445 (1)	.118	.950	0.000	.942	0.001
模型 4 結構殘差 (Structural residuals)	330.071 (117)	13.093 (3)	□.004	.947	-0.003	.940	-0.002
模型 5 測量殘差 (Measurement residuals)	342.936 (129)	12.865 (12)	.379	.947	0.000	.946	0.006

註：準則為 $\Delta\chi^2$ 的顯著性大於 0.05，且 ΔCFI 小於 0.01、 ΔTLI 小於 0.02；□為未達標準。

結論與建議

一、結論

(一) 多面向道德情意量表詳版與簡版的信效度考驗

本研究發展多面向道德情意量表，先基於理念架構以擬定量表初稿，而後邀約 15 位專家學者針對初稿評定其內容效度，經修改後列有四則道德故事、四個構面，以及 48 個題項，接續進行量表預試與正式施測。歷經四個階段對於研究結果的剖析，本研究分別發展詳版與簡版量表，並均通過信效度考驗。

首先就詳版量表而言，乃以預試樣本進行信效度考驗，維持原設計的四則道德故事（A、B、C、D），與四個構面（道德情感、道德同理、道德認同、道德關懷）組成的四個分量表，共計有 44 個題項。效度為奠基文獻理念基礎與專家學者的內容效度，信度採內部一致性信度，總量表為 .922，四個分量表介於 .731~.831。

其次就簡版量表而言，將預試樣本進行統計時分割為兩個樣本，預試樣本₁（pre_G1, $N = 283$ ）以 SPSS 進行探索性因素分析，結果顯示計三則故事（刪除原設計故事 A 後，原 B、C、D，改為故事一、二、三），且萃取三個因子（「道德承諾與關懷」、「道德同理」、「道德情感與期待」），合計 17 個題項最為適切，其總量表內部一致性信度為 .897。接續將預試樣本₁ 結果，套用預試樣本₂（pre_G2, $N = 302$ ），並以 Amos 進行驗證性因素分析，經過常態性檢定、違犯估計檢驗、模型適配度（二階單因子模式 12 題項， $\chi^2 = 142.661$ 、 $p < .001$ 、 $\chi^2/df = 2.797$ 、GFI = .927、AGFI = .888、SRMR = .055、RMSEA = .077），符合標準；其三因子的組合信度分別為 .851、.813 與 .753（都達到超過 0.6 的標準）、整體量表的組合信度為 .843、平均變異數抽取量為 .641、AVE 的平方根為 .802，表示簡版量表內在結構的聚斂效果佳；檢驗建構效度中的區別效度結果，三個構面前述的 AVE 的平方根，均大於各個構面間的相關係數，表示本量表具有區別效度。再者，接續前二階段結果，進行正式施測以及正式樣本（ $N = 708$ ）統計，並以 Amos 多群組分析考驗模式進行複核效化檢驗，經統計所得相關係數獲實徵資料支持，顯示其模型適配度良好（二階單因子模式 12 題項， $\chi^2 = 155.950$ 、 $p < .001$ 、 $\chi^2/df = 3.058$ 、GFI = .965、AGFI = .947、SRMR = .041、RMSEA = .054），並且以多群組分析的巢套模式比較，五種模型的複核效化考驗，可獲得實徵資料以及模式複製的支持。綜合而言，本研究發展的簡版多面向道德情意量表，除了具有內容效度與內部一致性考驗歷程，亦有良好的模式適配度，以及通過組合信度、建構效度與複核效化檢驗。

對照前述動機與文獻探討中列舉的道德範疇評測工具，本研究發展的多面向道德情意量表可彰顯若干特色：一是本研究發展的詳版與簡版量表，相較單一構念的評測工具，如道德情感的 SEL 相關量表（Haggerty et al., 2011; Hagen, 2014）、道德認同的 MI 量表（Reed et al., 2007）與 AIQ-IV 量表（Cheek & Briggs, 2013），以及道德關懷的 CES 量表（Doolittle & Faul, 2013）與 MD 量表（Bandura et al., 1996）等，兩種版本量表較可凸顯情意層面道德範疇的多面向，以及豐富且具深度的意涵。二是本研究兩種版本量表兼具統計的信效度檢驗，尤其簡版突破大多既有道德認知發展相關量表，無法以結構方程模型的測量模型，加以檢測其信效度的限制。即便如道德心理學與教育學界數十年來運用廣泛的 DIT-2，以其六個效度規準觀之，包括：不同教育階段的區辨、縱貫性探究、與認知能力測驗的相關、道德教育介入的敏銳度、與道德行為及專業判斷的相關性，以及對於政治選擇與態度的預測力等（Thoma & Dong, 2014），其均為效標關聯效度而非建構效度，因而本研究可謂是新的成功嘗試。

(二) 多面向道德情意量表詳版與簡版的特點與意涵

本研究發展的詳版多面向道德情意量表，其意涵基本上符合文獻理論構念初衷。就其內涵而言包括四則故事，為故事 A「分享照片的風波」、故事 B「遲來的真相」、故事 C「吃不到的巧克力」、故事 D「不平靜的野餐」的組合，另有「道德情感」、「道德同理」、「道德認同」與「道德關懷」四個構面，合計 44 個題項，可謂有其脈絡性與豐富性。至於簡版多面向道德情意量表，就其內涵而言包括三則故事，分別為故事一「遲來的真相」、故事二「吃不到的巧克力」、故事三「不平靜

的野餐」的組合，另有「道德情感與期待」、「道德同理」、「道德承諾與關懷」三個構面，計有 12 個題項，可謂具有結構性與精簡性。

將詳版與簡版量表加以對比可發現，故事 A「分享照片的風波」所關注的網路言論自由的界線之道德議題，在驗證性分析統計中無法納入簡版量表，但可保留在詳版量表。其次，由於各則故事題項的擬定是重複環繞在共通的概念架構（見表 1），故簡版量表顯示出各個構面會有較為適切的故事脈絡得以彰顯，並且因為以建構效度加以計算，故題項往往會十分精簡。再者，簡版量表原本四個構面中道德認同的意涵，經探索性因素分析後，分散至道德情感以及道德關懷二個構面，無法單獨成為一個構面，因此根據文獻理念再加界定其意涵如次：構面一「道德情感與期待」的意涵，是指對於道德故事中的個案人物處境、推衍至自己親友、以及在真實生活中的類似通案，可引發蘊含道德評價的適切情感展現，以及進而對於某些道德價值的期待。構面二「道德同理」的意涵，是指能夠想像道德故事中個案人物的心情，對於真實生活中的通案類似角色也能加以體會，而且對於不同群體所處的環境脈絡，得以同情理解與換位思考。構面三「道德承諾與關懷」的意涵，是指對於真實生活中類似道德故事的通案，能激發自我感動或與他人合作的道德承諾，願意積極維護權益，且承擔責任為公義發聲。

此外，本研究發展的多面向道德情意量表，對照前述文獻探討有數點研究發現：一是簡版的道德情感與期待構面以及詳版的道德情感與認同構面，較為接近 IRI (Davis, 1980)，針對人際之間互動所產生的道德評價與反應，以及 AIQ-IV (Cheek & Briggs, 2013) 的道德集體認同與期待，但與 SEL 所關切的問題較無法連結。二是詳版與簡版共通的道德同理構面，能充分彰顯 Lietz 等人 (2011) EAI 的觀點取替以及同理的態度，且與 Hoffman (2000) 連結正義與關懷原則的道德同理精神相通，另則凸顯了 Segal 等人 (2012) 社會同理概念中強調的系統障礙的脈絡性理解。三是道德承諾與關懷構面與詳版的道德認同與道德關懷構面，與 AIQ-IV (Cheek & Briggs, 2013) 關於社會與群體的認同與感動，以及 CES (Doolittle & Faul, 2013) 關於公共參與的量表內容十分接近，不過稍嫌遺憾的是簡版量表對於道德疏離 (Bandura, 1991) 此類的反向題較無法顯現。

二、建議

(一) 多面向道德情意量表的運用

本研究旨在發展「多面向道德情意量表」，希冀對於國內外道德教育理論與實踐有所助益。因此，對於本研究具信效度的 MAM 量表（包括詳版與簡版），可於多方面加以運用：一是在教育評測方面，可作為情意層面道德範疇概況的具體指標，藉以了解受試者道德情意的成熟度及其發展，此施測範圍可由一個班級、擴展至整個國家層面；或是作為某些課程與教學方案介入實施前後的成效展現。二是在課程與教學方面，尤其針對詳版，教育者可藉由這些故事所關注的道德議題及其題項，作為師生相互討論的素材與內涵；亦可對於言論自由與網路霸凌、轉型正義、全球化與國際移工等道德議題加以延伸探究。三是在道德教育理論方面，可針對量表的各個構面，持續在國內外文獻進行探究，以建構更為豐富、完整與具深度的多面向道德情意意涵與教育。

(二) 多面向道德情意量表的後續研究

本研究的 MAM 量表，在後續研究上仍有諸多得以拓展與延續之處可資參考：首先就詳版量表而言，可針對正式施測所展現的結果加以探究與詮釋；進而針對其人口背景變項加以檢測其差異性，或可運用多元迴歸分析檢視相關變項的解釋量。其次就簡版量表而言：一是可針對不同性別與教育階段等人口背景變項，進行測量恆等性 (measurement invariance) 的檢驗，進一步確立本量表在不同樣本間的穩定度；甚而日後可考量針對不同族群與國家進行施測，並藉由測量恆等性統計，以測試本量表簡版的多元文化適用性，抑或是仿照本量表發展方式，編製與建構其他的道德議題故事；二是可參考 DIT 等量表探究其效標關聯效度的做法，例如 MAM 量表與年齡/教育階段增長，以及本量表與民主及人權意識是否有其正向關聯；三是後續可採用 SEM 的結構模式，對於 MAM 量表與其他量表的關係進行探究，尤其是與認知或行為層面的道德量表加以連結與比較等。總之，本研究對於情意層面的道德量表發展及其教育，是一個起點而非終點。

參考文獻

- 甘孟龍、黃秀霜、劉峯銘 (2016)：「國中生憤怒反應量表」之發展與信效度考驗。《測驗學刊》，63(1)，1-30。[Gan, M.-L., Huang, H.-S., & Liu, F.-M. (2016). The development of anger response scale for junior high school student. *Psychological Testing*, 63(1), 1-30.]
- 沈六 (1986)：《道德發展與行為之研究》。水牛。[Sheen, L. (1986). *A research on moral development and behavior*. Buffalo Plaza.]
- 李琪明 (2017)：我國中小學道德情意發展之課程模式建構。《課程與教學季刊》，20(3)，1-26。[Lee, C.-M. (2017). Constructing a curriculum model of moral-affective development for Taiwan's primary and secondary education. *Curriculum and Instruction Quarterly*, 20(3), 1-26.]
- 林慧慈、林文瑛、王震武 (2017)。從局部統合到整體統合的發展—以道德思考為例。《教育心理學報》，48(3)，329-349。http://doi.org/10.6251/bep.20160128 [Lin, H.-T., Lin, W.-Y., & Wang, J.W. (2017). From local to global coherence: The development of moral thinking. *Bulletin of Educational Psychology*, 48(3), 329-349. http://doi.org/10.6251/bep.20160128]
- 陳寬裕、王正華 (2016)：《結構方程模型分析實務：AMOS 的運用》。五南。[Chen, K.-Y., & Wang, Z.-H. (2016). *Structural Equation Modeling analysis practice: Application of AMOS*. Wunan.]
- 單文經 (1982)：《道德教育初探》。復文。[Shan, W.-J. (1982). *A pilot study of moral education*. Fuhwen.]
- 程姿螢、張詠婷、宋曜廷 (2016)：大專學生工作價值組合量表之編製及信效度研究。《教育心理學報》，48(1)，91-112。http://doi.org/10.6251/BEP.20151020。[Cheng, T.-Y., Chang, Y.-T., & Sung, Y.-T. (2016). The development of the work value assembly scale for college students. *Bulletin of Educational Psychology*, 48(1), 91-112. http://doi.org/10.6251/BEP.20151020]
- Aquino, K., & Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality & Social Psychology*, 83(6), 1423-1440. http://doi.org/10.1037//0022-3514.83.6.1423
- Bandura, A. (1991). Social cognitive theory of moral thought and action. In W. M. Kurtines, & J. L. Gewirtz (Eds.), *Handbook of moral behavior and development*, (Vol. 1, pp.45-103). Erlbaum.
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in exercise of moral agency. *Journal of Personality & Social Psychology*, 71(2), 364-374.
- Blasi, A. (1984). Moral identity: Its role in moral functioning. In W. M. Kurtines, & J. L. Gewirtz (Eds.), *Morality, moral behavior, and moral development* (pp. 128-139). Wiley.
- Byrne, B. M. (2010). *Structural equation modelling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge.

- Casali, G. L. (2011). Developing a multidimensional scale for ethical decision making. *Journal of Business Ethics, 104*, 485-497. <http://doi.org/10.1007/s10551-011-0923-9>
- Chambers, D.W. (2011). Developing a self-scoring comprehensive instrument to measure Rest's four-component model of moral behavior: The moral skills inventory. *Journal of Dental Education, 75*(1), 23-35.
- Cheek, J. M. & Briggs, S. R. (2013). Aspects of identity questionnaire (AIQ-IV). *Measurement Instrument Database for the Social Science*. <http://dx.doi.org/10.13072/midss.562>
- Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgment*. Cambridge University Press.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.
- Doolittle, A. & Faul, A. C. (2013). Civic engagement scale: A validation study. *SAGE Open, (July-September)*, 1-7. <https://doi.org/10.1177/2158244013495542>
- Durlak, J. A., Domitrovich, C. E., Weissberg, R. P. & Gullotta, T. P. (Eds.). (2015). *Handbook of social and emotional learning: Research and practice*. The Guilford Press.
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annual Review of Psychology, 51*(1), 665-697.
- Habermas, J. (1979). *Communication and the evolution of society*. (T. McCarthy, Trans.). Beacon Press. (Original work published 1976)
- Hagen, E. (2014). *Resources for measuring social and emotional learning*. Retrieved Dec. 20, 2015 from <https://conservancy.umn.edu/bitstream/handle/11299/195179/issue-brief-measurement-resource.pdf?sequence=1&isAllowed=y>
- Haggerty, K., Elgin, J., & Woolley, A. (2011). *Social-emotional learning assessment measures for middle school youth*. Retrieved Dec. 19, 2015, from <http://www.search-institute.org/sites/default/files/a/DAP-Raikis-Foundation-Review.pdf>
- Held, V. (2006). *The ethics of care: Personal, political, and global*. Oxford University.
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. Cambridge University.
- Lee, C.-M., & Taylor, M. J. (2013). Moral education trends over 40 years: A content analysis of the Journal of Moral Education (1971-2011). *Journal of Moral Education, 42*(4), 399-429. <https://doi.org/10.1080/03057240.2013.832666>
- Lee, C.-M., & Thoma, S. J. (2018). Developing and testing a scale of moral thinking and communication (MTC) functioning: a preliminary study and its implications for moral development and education. *Asia Pacific Journal of Education, 38*(1), 32-50. <https://doi.org/10.1080/02188791.2017.1386618>

- Lietz, C. A., Gerdes, K. E., Sun, F., Geiger, J. M., Wagaman, A. M., & Segal, E. A. (2011). The empathy assessment index (EAI): A confirmatory factor analysis of a multidimensional model of empathy. *Journal of the Society for Social Work Research, 2*(2), 104-124.
- Noddings, N. (2005). *Educating citizens for global awareness*. Teachers College.
- Nussbaum, M. (2008). *Morality and emotions*. Retrieved Dec. 21, 2015, from <http://virtueethicsinfocentre.blogspot.tw/2008/02/morality-and-emotions.html>
- Nussbaum, M. (2013). *Political emotions: Why love matters for justice*. Harvard University.
- Reed, A., Aquino, K., & Levy, E. (2007). Moral identity and judgments of charitable behaviors. *Journal of Marketing, 71*(1), 178-193. <https://doi.org/10.1509/jmkg.71.1.178>
- Reimer, K. (2012). Moral identity. In Levesque, R. J. R. (Ed.), *Encyclopedia of adolescence* (pp. 11-19). Springer. Retrieved Dec. 22, 2015, from https://link.springer.com/referenceworkentry/10.1007%2F978-3-319-32132-5_30-2
- Rest, J., Narvaez, D., Bebeau, M. J., & Thoma, S. J. (1999). *Postconventional moral thinking: A neo-Kohlbergian approach*. Lawrence Erlbaum Associates.
- Segal, E. A., Wagaman, M. A., & Gerdes, K. E. (2012). Developing the social empathy index: An exploratory factor analysis. *Advances in Social Work, 13*(3), 541-560.
- Tangney, J. P., Stuewig, J. & Mashek, D. J. (2007). *Moral emotions and moral behavior*. *Annual Review of Psychology, 58*, 345-372. <https://doi.org/10.1146/annurev.psych.56.091103.070145>
- Thoma, S. J., & Dong, Y. (2014). The defining issues test of moral judgment development. *Behavior Development Bulletin, 19*(3), 55-61. <http://dx.doi.org/10.1037/h0100590>
- Yacker, N., & Weinberg, S. L. (1990). Care and justice moral orientation: A scale for its assessment. *Journal of Personality Assessment, 55*(1-2), 18-27. <https://doi.org/10.1080/00223891.1990.9674043>

附錄：本量表四則故事

詳版故事 A：分享照片的風波

阿美在個人臉書發佈一張個人獨照，班上同學阿翹看到後覺得這張照片很漂亮，因此將此張照片轉載至「表特版」(Beauty，網友會張貼帥哥美女圖片的網路論壇，供人觀看、點評)，不料引來諸多網友的負面評論，甲說：「這樣也敢 PO 喔！」乙提到：「這照片修圖修太多了吧！」丙也說：「臉形跟身材比例不搭，這是整過吧！」。某天，阿美無意間在表特版發現自己的照片及網友們的言論，感到十分的自卑，開始不太想去學校上課。從小與阿美一起長大的好友阿芳發現她的異狀後，不斷追問原因，才赫然明白事情的真相。

詳版故事 B (簡版故事一)：遲來的真相 (見圖 1 量表故事 B 圖示與內容)

詳版故事 C (簡版故事二): 吃不到的巧克力

臺灣的無國界記者阿凱到一個盛產可可豆的貧窮國家，發現 10 歲的傑克從來沒上過學，為了負擔家計，被迫和父母分開、住在擁擠的工廠裡，每天都要在大太陽下採收可可豆長達十多個小時。面臨各種壓力的廠長因為心情不佳，時常毆打與虐待這些童工，甚至沒有正當理由扣減他們的薪水，傑克每天晚上都躲起來哭卻無法離開，因為只有採收可可豆才能勉強活下去。當傑克無意間吃到阿凱自掏腰包買來送童工的巧克力時，才知道這些採收的可可豆被賣到國外，做成昂貴美味的巧克力，可是他卻從來沒有吃過。記者阿凱心疼這群童工的處境，將傑克的故事寫成了報導，希望喚起大眾的重視。

詳版故事 D (簡版故事三): 不平靜的野餐

由東南亞國家至臺灣工作的勞工阿平，放假時和同鄉的朋友聚集在公園唱歌、跳舞，分享故鄉的美食。坐下來沒多久，就有警察因為接獲附近民眾連續檢舉而前來關切，並表示有民眾甲抗議：「外勞太吵了！」、民眾乙投訴：「他們坐在地上喝酒吃東西，會把公園搞得髒亂又噁心！」、民眾丙抱怨：「這裡聚集太多外勞了，我的小孩會被嚇到」等。警察希望他們的活動可以提早結束，阿平和朋友們只好開始收拾、準備離開。這時路人小榮對警察說：「公園是開放空間，聚會野餐的音量也沒有過分影響到其他人，為什麼要驅趕他們？況且在公園內不僅有外籍移工，也還有其他居民在野餐或從事休閒活動。」

收稿日期：2019 年 07 月 22 日

一稿修訂日期：2019 年 10 月 21 日

二稿修訂日期：2019 年 11 月 18 日

三稿修訂日期：2019 年 11 月 27 日

接受刊登日期：2019 年 11 月 27 日

Bulletin of Educational Psychology, 2020, 51(4), 5671-583

National Taiwan Normal University, Taipei, Taiwan, R.O.C.

Multidimensional Affective Morality Scale: Development and Reliability and Validity Evaluation

Angela Chi-Ming Lee

Department of Civic Education and Leadership

National Taiwan Normal University

This study developed the Multidimensional affective morality (MAM) scale by testing its reliabilities and validity. On the basis of both national and global academic experiences and research achievements, this study selected students from junior high school to graduate school in Taiwan as subjects to construct a concrete indicator for affective morality and enhance the theory and practice of moral education. This study first explored the rationale underlying the scale's draft. The researcher then examined the content validity and conducted the pretest and formal survey. Statistical analysis was performed utilizing the IBM SPSS Statistics (version 23) and AMOS Graphics (version 25). In Stage 1, an item analysis was performed and the internal reliability was computed for the pretest group sample ($N = 585$). Stage 2 comprised an exploratory factor analysis (EFA) for pretest_1 group ($N = 283$). Stage 3 comprised a confirmatory factor analysis (CFA) for pretest_2 group ($N = 302$). In Stage 4, the cross-validation for the formal survey_group was assessed ($N = 708$). The researcher developed two versions of the MAM scale. The long version, comprising 4 moral stories, 4 constructs, and 44 items, passed the test for content validity and internal consistency. The short version, comprising 3 moral stories, 3 constructs, and 12 items, demonstrated goodness of fit (a one-factor second-order CFA model) and good composite/component/construct reliability, construct validity, and cross-validity. The long version of the MAM scale demonstrated contextualization and richness, whereas the short one demonstrated structure and refinement. Thus, the MAM scale has the following characteristics: (1) The MAM scale has good richness and offered an in-depth understanding of the multiple dimensions of affective morality. (2) Similar to other related scales, the MAM scale demonstrated the concepts of affective morality. In addition, the scale encompassed localization and creativity pertaining to moral issues. (3) The MAM scale's reliability and validity could be tested using various evaluation methods. Finally, the researcher offered several suggestions for the scale's application and further research to improve the development of affective morality among students in different educational levels.

Keywords: confirmatory factor analysis (CFA), cross-validation, exploratory factor analysis (EFA), multidimensional affective morality (MAM) scale, reliability and validity

