

臺灣四年級與八年級學生於TIMSS 2011試題中 建模能力之比較

林靜雯* 余阮清

國立東華大學 教育暨潛能開發學系暨科學教育研究所

摘要

有鑒於目前全面性探討學生建模實務評量之不足，本研究採用「建模能力分析指標」國際數學與科學教育成就趨勢調查 (Trends in International Mathematics and Science Study, TIMSS) 修正版，分析TIMSS 2011四、八年級科學試題以瞭解其中適合檢測建模能力之試題於「建模歷程」與「建模階層」之分布情形，而後再探討臺灣四、八年級學生於此測驗中建模能力之表現。結果發現：一、TIMSS 2011四、八年級試題皆集中於「模型選擇」與「模型分析與應用」兩階段；四年級試題於建模階層分布尚稱平均，八年級則集中於階層二及三。二、在「模型選擇」方面，臺灣四年級學生可掌握至階層二「多因素之選擇」，八年級學生則進展到階層三「簡單關係的選擇」；在「模型分析與應用」方面，四年級和八年級學生皆可掌握至階層三「簡單關係的應用」，但八年級學生可處理較多因素及較為抽象、微觀之變數。研究建議若運用TIMSS試題作為建模實務評量，尚需補充「模型建立」、「效化」及「調度」階段之試題，方能檢測學生各面向的建模能力。設計以建模為基礎之課程時，在「模型選擇」及「模型分析與應用」兩階段可參考本研究之結果發展學習進程。

關鍵詞：TIMSS 2011、建模能力分析指標TIMSS修正版、建模實務評量、學習進程

壹、前言

「模型」(model)以簡化的方式表徵系統或複雜的實體(Danusso, Testa, & Vicentini, 2010)，是科學探究的結果，而「建模」(modeling)則是科學探究的必要過程(Namdar & Shen, 2015)，兩者的關係密不可分，在科學研究及科學教育中皆扮演著重要的角色，

是近年來國際重視的課程元素(Schwarz et al., 2009)。例如：我國預計於2019年實施的十二年國民基本教育課程綱要自然科學領域課程綱要，其核心素養中的「探究能力」便包含「思考智能」與「問題解決」兩面向(國家教育研究院，2016)，兩面向各自有四個要素，「建模」不但為思考智能下一要素，邱美虹(2016)更認為「建模」是溝通「思考智能」與

*通訊作者：林靜雯，jingwenlin@gms.ndhu.edu.tw

(投稿日期：民國106年7月6日，修訂日期：民國106年10月15日，接受日期：民國106年10月16日)

「問題解決」的重要橋梁。透過循序漸進地模型建立、效化與應用，學生初始的心智模式將最終達致「科學模型」，進而解決所遭遇的問題與任務(Buckley & Boulter, 2000)。然而，許多教師缺乏「科學建模」的經驗，對於學生在建模實務(modeling practices)的能力亦缺乏理解(Harrison & Treagust, 2000; Justi & Gilbert, 2002)，因此，若能有設計良好的建模實務評量協助教師檢測及瞭解學生建模實務的能力及建模相關學習進程便成了當務之急。

但系統性建模能力評量回顧研究(Namdar & Shen, 2015; Nicolaou & Constantinou, 2014)卻指出建模實務評量十分缺乏，若有，亦集中於特定建模階段。此外，研究者從建模能力評量回顧性研究中亦發現，多數建模能力評量僅侷限於單一、特定科學主題之檢測，且不同研究對建模能力的階層評分亦缺乏一致性的架構可以適用於不同科學主題或反應於整體課程架構。相對的，張志康與邱美虹(2009)經文獻探討，區分了專家「心智建模」(例如：Justi & Gilbert, 2002)、生手心智建模(例如：Sins, Savelsbergh, & van Joolingen, 2005)與「科學建模」(例如：Halloun, 1996)後，奠基於Halloun「科學建模」過程設計「建模能力分析指標」(Modeling Ability Analytic Index, MAAI)。此指標含括建模歷程(modeling process)及建模階層(level of modeling)兩向度。前者意指建立模型時所經歷的過程，又區分成若干階段。後者則區分了建模歷程中所需不同程度的複雜認知能力，因而具有若干不同難度的階層。張志康與邱美虹主張MAAI涵蓋建模歷程「所有」階段，可運用此架構設計建模試題，或建立能力分析編碼以區辨學生於各建模階段所達到的能力層次。基於MAAI這

兩項重要特質，已有學者將之微調用以檢測各種國際大型測驗情境中之建模能力。例如：國際學生能力評量計劃(Programme for International Student Assessment, PISA) (邱美虹, 2015)、國際數學與科學教育成就趨勢調查(Trends in International Mathematics and Science Study, TIMSS) (林靜雯, 2015; Lin & Yu, 2015, 2017)，以建立適合評量學生建模能力之資料庫。本文即是以MAAI之TIMSS修正版(MAAI-T)為分析架構，分析TIMSS 2011。MAAI-T (林靜雯; Lin & Yu)乃整合TIMSS認知領域分類、國際基準點及其他學者之觀點，修改張志康與邱美虹所架構之MAAI而來。據此，本研究一方面藉此挑選其中適合檢測建模能力的試題建立資料庫(林靜雯, 2014)，使教師欲設計建模能力各階段試題時能有範例參酌，一方面藉以瞭解臺灣四、八年級學生於TIMSS 2011試題中展現之建模能力以回饋設計建模相關課程時，學習進程之發展。

TIMSS設計之初衷主要乃欲提供一歸納各國課程之架構，讓各國檢視並瞭解課程因素對學生成就之影響(Robitaille et al., 1993)。TIMSS已公開許多資訊供研究使用，這些研究大致分為三類，其一，運用TIMSS資料庫中之原始資料探討各式背景資訊對學生成就之影響或相關。例如：學生學習態度對學生成就因素之影響或相關(Liou & Liu, 2015)；其二，運用新的質性架構進行TIMSS試題之重新分類，再運用其他推論性統計進行試題分析。例如Liu與McKeough (2005)便挑選了27題與能量相關的TIMSS問題，並用自行設計的能量學習進程加以分類，最後再用項目反應理論的技術，評估其自行設計的能量學習進程之適當性；其三，運用新的質性架構重新分析TIMSS相關試題以描述性統計呈

現新架構之趨勢，例如：Zuzovsky與Tamir (1999)探索以色列四年級與八年級生在TIMSS 1999年各科試題中科學解釋的表現。他們針對科學解釋的理論架構，重新分類TIMSS試題，並計算四、八年級於五個科學解釋類別中之題數，並輔以國際平均分數作為解釋學生達到階層之參考。在此，「科學解釋」之理論架構並非TIMSS科學試題之架構，但藉著新架構，TIMSS國際性的資料得以為科學解釋提供新詮釋。本研究取向屬於第三類。如前所述，「建模」既為國際新近重視的課程元素，亦是十二年國教新課綱之重點，因此本研究挑選TIMSS以檢視建模相關試題在其中所扮演的角色。此外，TIMSS的科學試題架構，乃以「科學探究」歷程(形成問題與假設、設計調查、表徵資料、分析與解釋資料、下結論及發展解釋)作為測量學生學科知識的基礎(Martin, Mullis, Foy, & Stanco, 2012)，而建模是科學探究的必要過程。模型在歷屆TIMSS科學評量架構中皆占有一席之地，應用領域中，TIMSS架構更明書「使用模型」這項能力(Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009)，因此研究者將TIMSS與建模結合，具可行性及合理性。TIMSS評量學科包括數學和科學，主要施測對象為四年級與八年級學生。本研究分析對象為TIMSS 2011的科學，以下皆以此為探討範疇。

首先，研究者運用MAAI-T架構分析TIMSS 2011年公布之四、八年級之科學領域試題，先分類建模階段，再根據題目內容之因素與關係分類建模階層。接著，分析TIMSS四、八年級試題在不同學科之建模歷程各階段及建模階層的分布情形。最後，分析並對照臺灣四、八年級學生在TIMSS不同學科於建模歷程各階段所能掌握之最高階

層，以探索這兩個學習階段在建模能力上適當的學習進程。據此，本研究具體的研究問題為：

- 一、運用MAAI-T分析TIMSS 2011四年級與八年級試題在不同學科之建模歷程各階段及建模階層的分布情形為何？
- 二、臺灣四年級與八年級學生於TIMSS 2011試題中所展現建模能力之比較為何？

貳、文獻探討

本研究採用建模能力分析指標TIMSS修正版(MAAI-T) (林靜雯，2015；Lin & Yu, 2015, 2017)為分析架構，以下先介紹MAAI-T之架構，接著運用TIMSS科學試題簡介MAAI-T中建模歷程各階段、階層與內容，最後介紹建模能力學習進程之相關研究。

一、建模能力分析指標TIMSS修正版的架構

近年來，科學教育政策主張科學建模為核心實務之一，在建模取向教學中，學生學習著像科學家一樣建構、修改及使用模型以增進科學知識與探究技巧(Namdar & Shen, 2015)。張志康與邱美虹(2009)所建立之MAAI，便是以「科學建模實作」為首要考量，並因應不同研究情境之修正，應用於科學建模各實務層面之分析架構。如：國際大型測驗(邱美虹，2015)、實務評量設計(張志康、邱美虹；Lin, Cheng, Yu, & Wu, 2016)、教科書分析與設計(劉俊庚、邱美虹，2010；Jong, Chiu, & Chung, 2015)、學生科展活動及成品(鍾建坪、邱美虹，2014)。本研究即是採用MAAI之TIMSS修正版(MAAI-T)為分析架構(表1)，此架構乃統整TIMSS之認知領域及次領域之分類、國際基準點及其他領域

表1：MAAI-T之架構

建模歷程	國際基準點、 建模階層 TIMSS 認知領域	階層一 (單一因素)		階層二 (多重因素)		階層三 (簡單關係)	階層四 (延伸關係)
		四	初級	中級	高級	優級	無對應國際基準點
		八	初級	中級		高級	優級
模型選擇	回憶／辨識、定義、描述、範例展示、演示科學工具的知識(理解)	指出單一物件或特徵，並據此選擇模型。		指出多重物件或特徵，並據此選擇模型。		指出簡單因果關係，並據此選擇模型。	指出多個因果關係後，進一步延伸抽象的關係，並據此選擇模型。
模型建立	發現解法、解釋資訊(應用)；下結論、通則化(推理)	NA		NA		建立具簡單因果關係或結構的模型。	建立多個因果關係後，進一步延伸抽象的關係或結構的模型。
模型效化	關聯、比較／對照／分類(應用)；設計、評鑑(推理)	使用單一物件或特徵去評鑑、測試或比較已建立模型的內在一致性。		使用多重物件或特徵去評鑑、測試或比較已建立模型的內在一致性。		使用簡單因果關係去評鑑、測試或比較已建立模型的內在一致性。	使用多個因果關係後，進一步延伸抽象的關係去評鑑、測試或比較已建立模型的內在一致性。
模型分析與應用	分析、判斷(推理)；使用模型、解釋(應用)	分析已效化模型中單一物件或特徵之適用性以解決相似情境的問題。		分析已效化模型中多重物件或特徵之適用性以解決相似情境的問題。		分析已效化模型中的簡單因果關係適用性以解決相似情境的問題。	分析已效化模型中多個因果關係進一步延伸抽象關係的適用性，以解決相似情境的問題。
模型調度	假設／預測、統整／綜合(推理)	使用已效化模型中的單一物件或特徵以解決新情境的問題。		使用已效化模型中的多重物件或特徵以解決新情境的問題。		使用已效化模型中的簡單因果關係以解決新情境的問題。	使用已效化模型中的多個因果關係，做進一步延伸抽象的關係以解決新情境的問題。

註：僅是單一或多個物件、特徵無法建立模型，故以NA註記。

學者對於建模歷程及階層之最新研究，重新詮釋、修正MAAI的建模歷程及建模階層兩向度而得。在建模歷程向度，因TIMSS試題鮮少題組，不具循環建模觀點，故刪除「重建」階段，另因模型分析與模型應用之認知歷程相近，常並行發生(Halloun, 2006; Justi & Gilbert, 2002; Oh & Oh, 2011)，因此將「模型分析與應用」合併為一。最後形成之各階段為模型：(一)選擇、(二)建立、(三)效化、(四)分析與應用，及(五)調度；在建模階層向度，刪除張志康、邱美虹研究中「答非所問」的階層零及「科學理論」的階層五，因TIMSS

出題不可能有答非所問的反射經驗以及單一題目完成科學家探究「科學理論」之試題，並重新詮釋階層一至四使之適用於TIMSS建模能力框架之比較。以下，由「建模歷程」各階段及「建模階層」兩向度簡述MAAI-T之架構。

(一)建模歷程

1. 模型選擇

張志康與邱美虹(2009)認為模型選擇為建模歷程的第一步，學生於此階段能清楚「描述」並決定模型被建構之目的，並根據個人經驗、知識、態度，自發性地「回

憶」，從熟悉的模型來源中「辨識」新情境之若干特徵以界定目標模型之範圍(Halloun, 1996, 2006)，最終達成問題解決。據此，MAAI-T定義「模型選擇」為指出若干物件或特徵，並據此選擇模型。與TIMSS認知領域與次領域定義(Mullis et al., 2009)對照，其可與TIMSS「理解」認知領域中的：(1)回憶／辨識、(2)定義、(3)描述、(4)範例展示、(5)演示科學工具的知識對應。

2. 模型建立

學生於前一階段已覺知到目標模型範圍及既有知識有限性，張志康與邱美虹(2009)指出模型建立需確認或重製先前所選擇之模型成分與結構，藉由教師引導其「發展(數學)模型」，並激發討論、「解釋最接近目標模型相關資訊」，亦即理論、概念組成、「通則化」或特定化結構以進行問題解決(Halloun, 1996, 2006)，進而透過從「實驗現象的結果」建構可解釋的模型(Oh & Oh, 2011)。MAAI-T定義「模型建立」為建立具因果關係或結構的模型，並可與TIMSS「應用」認知領域的發現解法、解釋(圖表)資訊，以及「推理」認知領域中的下結論、通則化(Mullis et al., 2009)對應。

3. 模型效化

完成建立個人初步模型後，張志康與邱美虹(2009)認為學生將運用不同方針檢驗模型內部一致性結構，精進探究方法以探索解決在個人模型與科學模型中之主要不可共量(incommensurabilities)之處(Halloun, 1996, 2006)，進入引證階段之可執行模型、調查「設計」兩個步驟，透過交互「比較、對照」代表相同現象或問題的模型，「評鑑」其優點與限制以選擇最適合模型(Oh & Oh, 2011)，充份經歷了回憶與辨識特定屬性、行為、概念之「相關知識」。MAAI-T定義「模

型效化」為使用若干物件、特徵或關係去評鑑、測試或比較已建立模型的內在一致性，並可與TIMSS「應用」認知領域的關聯、比較／對照／分類，以及「推理」認知領域中的分析、判斷(Mullis et al., 2009)對應。

4. 模型分析與應用

此階段與張志康與邱美虹(2009)定義有些微不同，學生不僅需「分析」已效化模型之建構以「判斷」是否能滿足其模型目的(Halloun, 1996)，還需「著手使用」模型，盡力運用有效資料、所學科學原理引導出欠缺之規則並進行理性推論，以「解釋」科學模型的定義(Halloun, 2006)。MAAI-T定義「模型分析與應用」為分析已效化模型中若干物件、特徵或關係之適用性以解決相似情境的問題，並可與TIMSS「應用」認知領域的使用模型、解釋，以及「推理」認知領域中的設計、評鑑(Mullis et al., 2009)對應。

5. 模型調度

此階段模型已被分析及效化，符合模型建立目的，並可應用於問題解決，張志康與邱美虹(2009)認為學生試圖發展轉換技巧，回溯至最初目的，「預測」其是否能遷移至其他適當目的中(Halloun, 1996)，並透過「統整」描述、解釋、預測及控制結構或行為等方式，獲得新的實證、理性見解之科學模型意義，希望能帶領此模型至新的理論、典範之視野(Halloun, 2006)。MAAI-T定義「模型調度」為使用已效化模型中的因素、因果或互動關係以解決新情境的問題，並可與TIMSS「推理」認知領域中的假設／預測、統整／綜合(Mullis et al., 2009)對應。

(二) 建模階層

MAAI-T之建模階層乃奠基於張志康與邱美虹(2009)之研究架構(Biggs & Collis,

1982; Grosslight, Unger, Jay, & Smith, 1991; Hempel, 1958), 並更新「觀察學習結果架構」(Structure of the Observed Learning Outcome, SOLO)分類法之最新文獻(Biggs & Tang, 2007), 再輔以參照各領域學者之理論架構(Crowe, Dirks, & Wenderoth, 2008; Krathwohl, 2002; Rodgers, 2002)、TIMSS使用量尺定錨(scale anchoring)技術訂定四、八年級學生所達科學成就之四個國際基準點(Martin et al., 2012), 重新調整建模能力階層所得。茲說明如下:

1. 階層一：單一因素

此階層為學習之最基層，張志康與邱美虹(2009)認為學生若能對某一特定問題反應出「一種」與理論相關的因素即為階層一。與TIMSS之國際基準點對應，四年級學生於「初、中級」擁有各科之初階知識並應用於實務情境，八年級學生之「初級」則除具有各科之基本概念外，尚可進一步解釋簡單圖表、示意圖之訊息(Martin et al., 2012)。據此，MAAI-T定義階層一為TIMSS各科之初階知識，學生在面對試題時，能反應出「單一」種與模型相關的物件或特徵。

值得一提的是，四年級的初級與中級具有重疊性，此乃因為不同學科的特質不同造成難度(國際答題正確率)不一，舉例而言，生命科學的變因通常較為具象，而地球科學的變因則可能觀察尺度較大。同理，中級與高級，高級與優級之間亦有重疊，以下便不再贅述原因。

2. 階層二：多重因素

學生已逐步開始瞭解，但並非達到「全然理解」(Krathwohl, 2002)，張志康與邱美虹(2009)認為學生若能對某一特定問題反應出「兩種以上」與理論相關的因素即為階層

二，與TIMSS之國際基準點對應，四年級學生於「中、高級」基準點能夠應用知識解釋日常生活中自然現象，八年級學生於「中級」基準點則可進一步拓展運用至多樣化的脈絡中，並透過簡短的描述去溝通自身理解(Martin et al., 2012)。據此，MAAI-T定義階層二為TIMSS各科在不同科學實用處境中的基本知識，學生在面對試題時，能反應出「多重」與模型相關的物件或特徵。

3. 階層三：關係層次

進入經驗分析階段，開始生成問題的可能解釋(Rodgers, 2002)，張志康與邱美虹(2009)認為學生若能對某一特定問題反應出「因素間的關係、交互作用或影響」即為階層三，與TIMSS之國際基準點對應，四年級學生於「高、優級」能展現對於科學探究知識與過程之理解，八年級於「高級」基準點學生除具備解讀科學過程、系統與原理之能力，還能進行科學研究(Martin et al., 2012)。據此，MAAI-T定義階層三為TIMSS各科之過程與探究知識，學生在面對試題時，能反應出模型相關物件或特徵之「簡單因果關係」。

4. 階層四：延伸關係

透過結合、綜析不同訊息來源以創造「新的概念」(Crowe et al., 2008)，張志康與邱美虹(2009)認為學生若面對某一特定問題，能將多個因素關係間的概念作進一步延伸抽象的反應，即為階層四，與TIMSS之國際基準點對應，八年級學生於「優級」基準點能溝通、理解各學科中之抽象複雜知識，並整合多元資訊進行解決問題、推導結論，可提供書面解釋及科學知識的溝通；但四年級科學成就尚無對應之「國際基準點」。據此，MAAI-T定義階層四為TIMSS各科之複雜與抽象概念，學生在面對試題時，能反應出

模型相關物件或特徵之「複雜因果或互動關係」。

張志康與邱美虹(2009)站在科學教育實務觀點，主張MAAI可協助教師理解「科學模型」、「建模能力」，在設計建模試題過程，探索並發現學生可能存在之素樸模型以檢測學生建模能力。在建模階段方面，教科書的設計及教師教學雖會利用模型，但卻極少讓學生參與模型「建立或修正」的設計(Van Driel & Verloop, 1999, 2002)；此外，學生較不擅長於科學模型之「效化」及「調度」，可歸因於傳統教學歷程，教師較不重視效化與調度兩步驟(Halloun, 1996)，在接受科學知識後，可能產生「混合模型」或「記憶答案」以應付考試(Greca & Moreira, 2000)，無法擁有科學家透過科學社群溝通的機會，以達成調整、更新想法並建立科學模型(Nersessian, 1995)。Davis等(2008)的研究亦指出科學建模活動中，調度、遷移為一較容易被忽視的環節。故本研究分析TIMSS 2011四、八年級試題建模能力各面向之結果，研究者預測亦將反應「建立」、「效化」及「調度」等面向的缺乏。

二、運用TIMSS科學試題簡介MAAI-T中建模歷程各階段、階層與內容

TIMSS設計之初衷乃想要提供一個歸納各國課程而獲得之精簡架構，讓各國用這個架構來檢驗自己國家的課程，藉以進一步探討課程因素對學生成就的影響情形、比較各參與地區或國家的教育成效，以及評估他們能否掌握投入社會所需的知識與技能(Robitaille et al., 1993)。臺灣共參加1999、2003、2007與2011年四次，並於2015年五、六月進行最近一次施測。由於TIMSS 2011測

驗主要評估在各內容領域下，學生於「科學探究」歷程中所使用之知識與技巧，其所設立之四個國際基準點，可提供有意義的描述去體現學生科學成就所達能力(Martin et al., 2012)。各國近年來逐漸重視學生建模能力，TIMSS多項「認知向度」及四個「國際基準點(難易程度)」，都能適當與MAAI-T中之「建模歷程」及「建模階層」相互呼應，故本文以MAAI-T分析TIMSS之題目，藉此架起TIMSS試題與建模能力之橋梁，希望能挑選出適當的建模試題作為範例，以協助第一線教師在現行課程、課程試題與設計建模相關試題中順利轉化。

以下，因TIMSS 2011八年級釋出試題數較多且涵蓋內容較豐富，故此，本研究優先整理八年級「國際基準點範例試題」及其「國際答題正確率」(Martin et al., 2012)說明建模歷程各階段、階層之試題形式與內容，讀者於解構、分析試題時可輔以參照表1，以對於各階段、階層定義有更清晰之理解。而若部分建模階段(如：建立、調度)缺乏相應國際基準點的範例，才改以其他非基準點說明。

(一)模型選擇——從現有的概念或模型中，選出適當的物件或基本模型

編號S042073試題(圖1)，TIMSS歸類為「理解」認知領域下之化學學科(化學的分類與分子組成)選擇題(International Association for the Evaluation of Educational Achievement

哪一個是二氧化碳的化學式？

- (A) CO
- (B) CO₂
- (C) C
- (D) O₂

圖1：模型選擇範例

[IEA], 2013c; Martin et al., 2012)，該題主要測驗學生是否能運用「理解」領域中較基礎之認知能力，亦即「回憶」所學知識與經驗去「辨識」碳原子(C)、氧原子(O)所形成「二氧化碳」分子之鍵結關係，方能進一步從題目所提供的選項中，選擇正確之二氧化碳之分子式(CO_2)，亦即由「兩個氧原子」與「一個碳原子」所連接而成，正確答案為B。

此題為TIMSS 2011釋出「初級基準點」範例試題，國際正確率為85%，於此基準點上，學生於「化學」領域具備化學式的基礎知識。分析其建模能力階層，其內容僅包含一個與理論相關之因素(碳與氧的连接方式)，因此屬「階層一」。

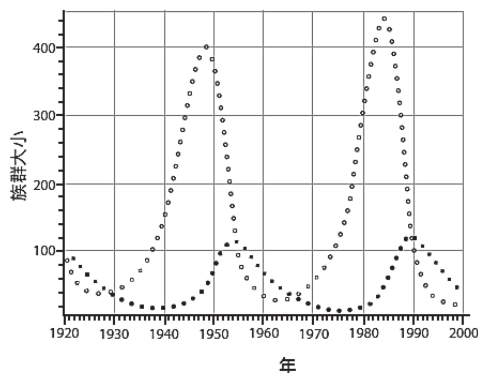
(二)模型建立——為所挑選的物件、組成、基本模型建立關係或結構

編號S042051B試題(圖2)，TIMSS歸類為「應用」認知領域下之生物學科(生態系統)問

答題(IEA, 2013c; Martin et al., 2012)，該題主要希望學生瞭解不同物種之間依存的食性關係(食物鏈、食物網)。但實際答題時，在題A需要的是圖表閱讀能力，學生尚未真正建立模型，但在題B，盼學生能運用「應用」領域中較進階之認知技巧，也就是挑選適當物件進行「解釋資訊」，並進一步「發現解法」以達成建立模型的「通則化」關係。此題包含「兔子的數量」與「狐狸的數量」兩個變數，其正確描述為：兔子的數量增加，狐狸有充足的食物，便會隨之增加，但增加到某一種程度，兔子便會遭到較大量的獵捕，因此數量銳減，少了食物的狐狸，數量也難能擴張。

此題非TIMSS介紹之國際基準點範例，但因為是TIMSS 2011八年級中，模型建立唯一試題，因此在此引介。其國際正確率為27%，於此基準點上，學生在「生物」領域能解釋生態系之食物鏈並認知到競爭與捕食

在某個人跡罕至的地區，住了一群兔子和狐狸，狐狸沒有任何天敵。科學家長期計算兔子和狐狸的數量，並將結果繪製如下圖。



A. 哪一年兔子的數量最多？

B. 請描述兔子和狐狸族群大小改變情形的相關性。

圖2：模型建立範例

關係(Martin et al., 2012)。分析其建模能力階層，其內容反應因素間之影響並形成簡單關係，屬「階層三」。

(三)模型效化——使用模型已建立的關係或結構去判斷、測試或比較一個模型的內在一致性

編號S032542試題(圖3)，TIMSS歸類為「應用」認知領域下之生物學科(特徵、分類與有機體的生命歷程)選擇題(IEA, 2013c;

下表中的動物被區分成兩大類。

第一類	第二類
兔子	青蛙
長頸鹿	蜘蛛
大象	獅子

請問下列何者是這分類的依據？

- (A) 呼吸時所使用的器官
- (B) 食物的來源
- (C) 生殖的方式
- (D) 運動的模式

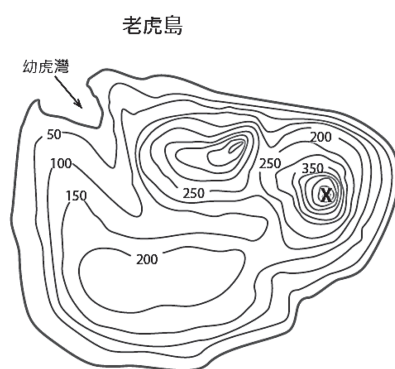
圖3：模型效化範例

Martin et al., 2012)，該題主要在測驗學生是否能採用「應用」領域中較進階之認知技巧，經由「比較、對照」圖3中動物之分類模式以建立模型，選項則提供檢驗與效化模型的機會，學生嘗試從呼吸器官、食物來源、生殖方式以及運動的模式「評鑑」原先自己設定模型與提供選項之間的「關聯」。

此題非TIMSS介紹之國際基準點範例，但因模型效化僅有兩題，故在此引介。其國際正確率為49%，於此基準點上，學生在「生物」領域能解釋有關於生物的特徵、分類和生命過程(Martin et al., 2012)。對照建模能力階層，因試題內容僅反應出「多個生物」之「多個生物特徵」等一個以上的因素，但並未涉及因素間之交互作用、關係，仍屬「階層二」。

(四)模型分析與應用——分析已效化模型之適用性及解決相同情境的問題

編號S032651A試題(圖4)，TIMSS歸類為「應用」認知領域下之地球科學學科(地球的



上圖是老虎島的地形圖。圖上相同高度的點連成的線稱為等高線，其高度單位為公尺。

A. X是什麼樣的地形？_____

B. 想一想河流的發源位置以及河水是如何流動的。請在圖上畫出一條在X點和幼虎灣之間河流的路徑，並以箭頭表示此條河水的流向。

圖4：模型分析與應用範例

構造及其物理特徵)問答題(IEA, 2013c; Martin et al., 2012)，該題主要在測驗學生是否能展現「應用」領域中較進階之認知技巧，亦即「使用圖中所呈現不同疏密程度與轉折線條之等高線模型」，並藉由圖中的數字去對應出等高線通過之處所代表的絕對高度，以判讀地形的種類。首先，學生必須先「分析」在等高線圖中，所呈現線條之起伏與高度，方能應用於題目中，去「判斷」特定位置(X)所代表的地形。答案是「高山、山丘、火山、最高點或其他類似的答案」。

此題為TIMSS釋出「高級基準點」範例試題，國際正確率為38%，於此基準點上，學生在「地球科學」能解釋等高線圖並辨認

山頂的地形表徵，其試題內容反應因素間簡單關係(地形、等高線圖)，屬「階層三」。

(五)模型調度——使用已效化的模型解決新情境的問題，是種遠遷移

編號S032665B試題(圖5)，TIMSS歸類為「推理」認知領域下之生物學科(生態系統)問答題(IEA, 2013c; Martin et al., 2012)。該題主要希望學生瞭解目前人口成長之衍生問題，並探討、認知到人類活動對環境所造成衝擊，需做好自然保育以維持生態系穩定，解題過程能運用「推理」領域較高階認知技巧，亦即「統整」題目中所有相關訊息並提出「假設」以「預測」結果，其「土地使用」問題，需從「乙國人口增加」推論出

全世界有 60 億人口共享地球的自然資源。下表為兩個假想國家（甲國和乙國）的一些資料。

	甲國	乙國
人口數（百萬人）	200	500
年出生率（每一千人中的出生人口數）	10	40
年死亡率（每一千人中的死亡人口數）	10	10
面積（平方公里）	2 000 000	2 000 000
穀物收成量（佔全世界的百分比）	40%	20%
石油使用量（佔全世界的百分比）	20%	5%

A. 請根據表中提供的資料，預測這兩個國家的人口在未來十年的變化。
（在每一橫列勾選一個方格）

	人口會增加	人口會減少	人口會保持一樣
甲國	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
乙國	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B. 請預測未來十年中，這兩個國家的人口對以下兩項將有何影響。

土地使用：

污染：

圖5：模型調度範例

「住宅用地增加」、「農業用地減少」；其「汙染」問題，需從「乙國人口增加」推論出「石油量使用、垃圾製造量增加」，因此，此題將人口增加的模型遷移至探討「人口增加與土地使用」、「人口增加與汙染」之間的關係複雜的新情境，故為「模型調度」。

此題非TIMSS介紹之國際基準點範例，但因為是TIMSS 2011八年級中，模型調度唯一試題，因此在此引介。其國際正確率為21%，約落於優級基準點。於此基準點上，學生在「生物」能理解生態的複雜性，並理解到人類人口的增加對於環境的衝擊(Martin et al., 2012)。對照建模能力階層，此題將人口增加的模型遷移至探討其與土地使用、汙染之間的新複雜情境，同時進一步延伸多個因素間關係，屬「階層四」。

三、建模能力學習進程

所謂的學習進程意指具目的性的教、學預期目標。這些目標橫跨多個發展階段、年齡或年級，通常用於課程標準中，藉以闡明學生在特定的學習內容或能力，在特定階段應達到怎樣的程度(The Glossary of Education Reform, 2013)。評量學生位於學習進程的哪個定點，對於提升學習進程之研究品質，以及發展適當的學習進程而言，是極為關鍵的一步(Wilson, 2009)。Gilbert與Justi (2016)主張真實的科學教育應由一系列關注特定科學知識與技能之學習進程組成，透過引導學生去發展、運用這些能力。而「建模之學習進程」便是此種教育中，活化所有學生能力的重要組成。雖然如此，但有關模型與建模能力之學習進程研究並不多見，Newberry Online Limited (2018)發展國民教育課程時，以「進步路徑」(progress pathway)取代依

據Bloom, Engelhart, Furst, Hill與Krathwohl (1956)教育目標分類發展之「階層山」(levels mountain)。其進步路徑明確保留模型與建模的概念，並提出在國小階段，在基本建模能力學習上，須能分類及聚物件件，並預期他們應能描述與解釋事件發生的原因，若能回憶事實並正確使用科學字彙便超越該階段預期達到的目標；而在國中階段學習，在基本建模能力上，則為運用數學計算結果，並預期該階段學生應能以適當圖表解釋資料，若學生能運用多個核心概念以解釋不熟悉情境便超越了該階段預期達到的目標。與MAAI-T對照，Newberry Online Limited的進步路徑，在國小階段約略可對應於單一因素(階層一)的建模，而在國中階段則可以處理多因素(階層二)的模型建立。

Schwarz等(2009)執行設計建模促進科學學習(Modeling Designs for Learning Science, MoDeLS)計畫，亦發展複雜性遞增之科學建模學習進程，其中在建模實務向度上，其團隊嘗試讓國小高年級及國中學生嘗試理解重要科學實務之涵義，其目標為利用「模型作為預測與解釋的生產工具」，又分成四個階層。階層一，學生能建構和使用模型以展示單一現象的部分情形。此時，學生尚不會將模型視為產生新知識的工具，但應能將模型視為顯示其觀察現象的展示工具。階層二，學生能建構和使用模型以解釋現象成因，並提出證據支持解釋。此階段，學生將模型視為溝通其理解的工具，而較不會將其視為支持其思考的工具。階層三，學生建立和使用多個模型以解釋和預測一群相關現象的多個面向。此階段，學生將模型作為支持他們思考新現象的工具，並會基於不同模型的優點或缺點的分析解釋和預測不同模型的結果。階層四，學生在某一特定範疇中自發地建

構和使用模型以幫助他們思考。此階層，學生應能考量世界根據各種模型會如何運作，學生建構和使用模型以解決既存現象的新問題。

此研究讓國小高年級與國中學生進行為期六週不同單元之科學實作，再評估學生教學前後建模能力的表現。研究結果指出在經過相關建模實作教學後，國小及國中學生雖皆未能掌握至最高階層，但在投入建模實務中仍學習到關鍵要素。國小學生由階層一進步至階層二，能建構更多不可見解釋機制與歷程之模型，國中生則進一步延伸至階層三，可利用模型解釋現象之多個層面，對模型與現象一致有更複雜的理解，較不會執著於模型和現象的表面相似性。整體而言，國中小學生更自己建構模型去表達現象的機制，並發展與現象多個層面一致的模型，也能理解到須透過比較與修改去建立共識模型。與MAAI-T相較，Schwarz等(2009)的建模學習進程混雜了建模階段與建模階層，統稱階層，容易使人將建模歷程與複雜度混為一談，亦不易於大型國際評測中使用。與MAAI-T對應的建模階段，Schwarz等主要著重於模型的建立與應用階段，其階層三之後加入模型的效化與評鑑，階層四又加入了對模型限制與使用範疇的察覺及於新情境的使用，約等同於模型調度。至於與MAAI-T對應的建模階層，其階層一與二都較為側重在單一因素或現象，階層三之後，逐步探討多因素及其之間的關係。

本研究則以MAAI-T分析TIMSS 2011年中試題，分析四、八年級試題於建模階段、階層之分布情形，並比較臺灣四、八年級學生在不同科目(生命科學／生物、地球科學、物質科學／物理與化學)之建模能力，希冀藉由大型國際評測中更大樣本群的表現以提供

教育工作者在設計建模評量與學習進程時參考。

參、研究方法

一、資料來源

TIMSS 2011四年級、八年級部分含子題共釋出78及97題，其中，臺灣四年級(S031084)及八年級(S042311、S032184)中分別有1、2題未進行計分(IEA, 2013a)，以及八年級有3題(S042304、S042051A、S042238C)僅測得讀圖能力，故不列入分析。最後，四年級與八年級各分析77、92題，共計169題。

二、建模能力分析指標TIMSS修正版(MAAI-T)之效度及信度

建模能力分析指標TIMSS修正版(MAAI-T)包括建模歷程各階段及建模能力階層，如文獻探討一節之分析整理，建模歷程最後分為：「模型選擇」、「模型建立」、「模型效化」、「模型分析與應用」、「模型調度」五階段；建模能力階層則分為四個階層，分別為「階層一：單一因素」、「階層二：多重因素」、「階層三：關係層次」及「階層四：延伸關係」。

在效度部分，MAAI-T分成兩個向度，其一為「建模階段」，主要透過文獻探討一節中之探析與定義對照建立文獻基礎，而後，此分類架構請兩位科教相關且具建模研究背景之教授審閱建模階段依據文獻分類之適當性、與TIMSS認知領域架構分類呼應之合宜性，直至兩專家全然同意，因此具內容效度。其二為「建模階層」，除具備文獻基礎與內容效度外，另以效標關聯效度(criterion-related validity)檢視MAAI-T中建模階層之意義性與有效性。效標關聯效度進行

方式主要根據TIMSS國際基準點之定義，以及其釋出之基準點範例相應之國際正確率，推估每一等級國際答題正確率之區間，例如：四年級之生命科學，初級之國際正確率為60%以上，中級為40～59%，高級為39%以下，因此每個TIMSS題目，可以推估出其國際基準點之落點，亦可由MAAI-T推估其所屬之建模階層。而後，再將各題經MAAI-T分析之階層的結果與國際基準點之結果進行效標關聯效度，結果為結果顯示MAAI-T建模階層與四、八年級試題之整體Spearman相關分別為0.7 ($p < .001$)、0.6 ($p < .001$)，皆具備良好之效標關聯效度。

在信度部分，本研究先訓練兩名熟悉建模相關研究，且具碩士學位之研究者(其中之一為本研究第二作者)，先各取三分之一TIMSS四、八年級試題作為練習，提出編碼過程中較有爭議、模糊題項，與本研究第一作者討論並抵定編碼後，其建模歷程與階層之編碼正確度皆達90%，再由第一、第二作者全部試題的編碼後，進行評分者間信度(inter-rater reliability)的計算，分別將「一致題數」除以「總題數」之百分比，其信度值為90.59%。

三、資料分析

(一)TIMSS四年級與八年級各學科在建模歷程各階段與建模階層之分布情形

本研究以描述性統計呈現TIMSS 2011四、八年級試題之分布情形。主要包含兩部分：1.「學科」、「建模歷程」、「建模階層」之試題分布；2.「學科」於「建模歷程各階段」與「建模階層」之試題分布。其中，四年級科學領域包含生命科學、地球科學、物質科學三學科，八年級則包含生物、地球科學、物理、化學四學科，本研究將八

年級「物理」與「化學」合併為「物質科學」進行統計與比較。

(二)臺灣四年級與八年級學生於TIMSS 2011試題中所展現建模能力之表現

在分析TIMSS 2011釋出四年級、八年級與建模能力相關題項中(共169題)，本研究主要參照TIMSS 2011公布四、八年級各題項之「正確答題率」(IEA, 2013b)，並設定「50%」作為掌握該題項建模能力之判定基準，分別計算臺灣四、八年級學生在TIMSS 2011各學科在建模歷程各階段、建模階層之「整體通過率」(通過題數／總題數之比率)，以探討臺灣四、八年級學生在各建模階段與階層之表現。「整體通過率」若大於50%，則視為可以掌握某建模階段某建模階層。整體通過率之所以選擇以「50%」為標準，乃因TIMSS雖運用多種判準(例如：65%、55%、50%)來定錨各類題型(例如：選擇題、及屬於建構反應試題的填充題、問答題)是否屬於某一國際基準點，但皆運用50%這個判準判斷國際基準點的適當性，亦即難易程度是否合宜，因此與本研究是否掌握某建模階段的某階層之能力目的相當。亦即，若某一試題之答題正確率高於50%，則於某一基準點被評為出題適當，低於50%，則被視為出題太難(Mullis, 2012; Mullis, Erberber, & Preuschoff, 2008)。Lin (2017)進行概念演化之跨年級研究，亦是使用50%作為某一概念是否發展之指標。

肆、研究結果

一、TIMSS 2011四年級與八年級試題在不同學科之建模歷程各階段及建模階層的分布情形

TIMSS 2011四年級(77題)與八年級(92題)

試題於各學科、建模歷程各階段與階層之分布情形如表2所示。以下括弧中之數值，將先報導題數，再報導其占總題數之比例。

在四年級部分($N = 77$)，綜觀「學科」、「建模階段」、「建模階層」三大向度之分布情形，首先，「學科」由高至低依序為生命科學(33, 42.86%)、物質科學(28, 36.36%)、地球科學(16, 20.78%)。接著，「建模階段」由高至低依序為模型分析與應用(38, 49.35%)、模型選擇(37, 48.05%)、模型建立(1, 1.30%)及模型效化(1, 1.30%)，而在模型調度則未包含任何試題。最後，「建模階層」由高至低為階層二(27, 35.06%)、階層一(26, 33.77%)、階層三(24, 31.17%)。

依「學科」於「建模歷程各階段」與「建模階層」之分布情形觀之，首先，「生命科學」($N = 33$)包括模型選擇、建立、分

析與應用三階段，其於該學科所占之比例，以模型分析與應用(18, 54.55%)及模型選擇(14, 42.42%)為主，模型建立僅1題(3.03%)。至於建模階層，則各階層分布平均，由高至低依序為階層二(13, 39.39%)、階層一與階層三(10, 30.33%)。唯，模型選擇之試題集中於階層二，但模型分析與應用之試題則分別集中於階層一和三。第二，「地球科學」($N = 16$)則包含模型選擇、效化、分析與應用三階段，其於該學科所占之比例，以模型選擇(11, 68.75%)為主，模型分析與應用次之(4, 25.00%)，模型效化僅1題(6.25%)，這也是四年級有關效化唯一的一題。至於建模階層，各階層分布依序為階層一(7, 43.75%)、階層二(5, 31.25%)與三(4, 25%)。唯模型選擇之試題集中於階層一，但模型分析與應用之試題則集中於階層三。第三，「物質科學」($N = 28$)僅包含模型選擇(12, 42.86%)、分析與應用

表2：TIMSS 2011四年級與八年級題項在建模歷程各階段之描述性統計表

學科 建模 歷程	建模 階層	生命科學／生物		地球科學		物質科學／ 物理、化學		總計	
		四年級	八年級	四年級	八年級	四年級	八年級	四年級	八年級
選擇	一	3	7	7	4	7	4	37 (48.05%)	45 (48.91%)
	二	9	4	3	3	4	11		
	三	2	2	1	1	1	9		
建立	三	1	1	—	—	—	—	1 (1.30%)	1 (1.09%)
效化	二	—	1	1	—	—	—	1	2
	三	—	1	—	—	—	—	1 (1.30%)	2 (2.17%)
分析與 應用	一	7	1	—	—	2	1	38 (49.35%)	42 (45.65%)
	二	4	6	1	2	5	3		
	三	7	5	3	5	9	14		
	四	—	1	—	2	—	2		
調度	四	—	2	—	—	—	—	—	2 (2.17%)
總計		33 (42.86%)	31 (33.70%)	16 (20.78%)	17 (18.48%)	28 (36.36%)	44 (47.82%)	77 (100.00%)	92 (100.00%)

註：「—」表示該建模階段、階層未包含試題。

(16, 57.14%)，而其於建模階層一至三的分布上皆約近三成，尚稱平均，唯模型選擇以階層一為主，但模型分析與應用則以階層三為主。大體來說，四年級整體試題僅包含階層一至三，雖平均分配，但各科於模型選擇仍以較低階層為主，於模型分析與應用則相對以高階層為主。

至於八年級部分($N = 92$)，綜觀「學科」、「建模階段」、「建模階層」三大向度之分布情形，首先，「學科」由高至低依序為物質科學(44, 47.82%)、生物(31, 33.7%)、地球科學(17, 18.48%)。接著，「建模階段」由高至低依序為模型選擇(45, 48.91%)、模型分析與應用(42, 45.65%)、模型效化與調度(2, 2.17%)、模型建立(1, 1.09%)。最後，「建模階層」由高至低為階層三(38, 41.30%)、階層二(30, 32.61%)、階層一(17, 18.48%)、階層四(7, 7.61%)。

依「學科」於「建模歷程各階段」與「建模階層」之分布情形觀之。首先，「生物」($N = 31$)包含模型選擇、建立、效化、分析與應用、調度完整之五階段，其於該學科所占之比例以模型分析與應用(13, 41.94%)及模型選擇(13, 41.94%)為主，模型建立僅1題(3.23%)，模型效化及調度僅2題(6.45%)。至於建模階層，則集中於階層二，由高至低依序為階層二(11, 35.48%)、階層三(9, 29.03%)、階層一(8, 25.81%)、階層四(3, 9.68%)。唯模型選擇之試題集中於階層一及二，但模型分析與應用之試題則集中於階層二和三。第二，「地球科學」(17)僅包含模型選擇(8, 47.06%)、分析與應用(9, 52.94%)，而其於建模階層一至三呈遞增分配，由高至低依序為階層三(6, 35.29%)、階層二(5, 29.41%)、階層一(4, 23.53%)、階層四(2, 11.76%)，唯模型選擇以階層一及二為

主，但模型分析與應用則以階層三為主。第三，「物理與化學」(44)僅包含模型選擇(24, 54.55%)、分析與應用(20, 45.45%)，而其於建模階層一至三亦呈遞增分配，由高至低依序為階層三(23, 52.27%)、階層二(14, 31.82%)、階層一(5, 11.36%)、階層四(2, 4.55%)，唯模型選擇以階層二為主，但模型分析與應用則以階層三為主。大體而言，八年級完整包含階層一至四，其試題不僅於階層一至三呈遞增分配，各科試題亦於模型選擇以較低階層為主、於模型應用與分析則以較高階層為主，亦顯示相較於四年級試題，八年級試題更著重測量較高階層之建模能力。

綜析而言，在「學科」向度，四、八年級分別以生命科學、物理與化學試題涵蓋比例最高，以地球科學最低；在「建模階段」向度，TIMSS 2011四年級、八年級分布於各建模階段之試題極不平均，所有學科皆集中於模型選擇、模型分析與應用兩階段；而在「建模階層」向度，四年級僅分布於階層一至三，呈大致平均，八年級則分布於階層一至四，相對較不平均，集中於階層二及三；在「學科」於「建模歷程各階段」所測建模階層，四年級部分，生命科學在模型選擇以測量「多個因素」為主，而地球科學、物質科學僅測量「單一因素」；生命科學在模型分析與應用，同時測量「單一因素」及「因素間簡單關係」，地球科學、物質科學則測量「因素間簡單關係」。八年級部分，生物及地球科學在模型選擇測量「單一因素」，而在物理與化學則延伸測量「多個因素」；生物於模型應用與分析注重「多個因素」，而地球科學、物理與化學則延伸測量「因素間簡單關係」。另言之，四年級試題以「生命科學」在模型選擇難度較高，在模型應用與分析則以「地球科學」、「物質科學」難

度較高；八年級試題以「物理與化學」在模型選擇難度較高，而在模型應用與分析則以「物理與化學」、「地球科學」兩科難度較高，此外，「生物」雖是唯一涵蓋完整建模能力各階段之科目，但其難度在模型選擇、應用與分析兩階段皆為最低。

二、臺灣四、八年級學生於TIMSS 2011試題中所展現建模能力之比較

本研究關注建模能力之學習進程，以下將先分別探討並比較臺灣四、八年級學生於TIMSS 2011試題中所展現建模能力，最後再整合結果，於結論與建議處對兩學習階段之學習進程提出說明與建議。

表3以各科試題通過率呈現臺灣四、八年級學生在TIMSS 2011科學領域各學科之建模能力表現。所謂「通過率」係參考TIMSS 2011臺灣四、八年級學生之答題正確率，並以「50%」作為是否通過之基準，進行計算（通過題數／總題數）。以下先比較四、八年級學生在不同建模歷程、階層之整體通過率，再分析不同建模歷程、階層中之各科通過率。但因「模型建立」與「模型效化」二階段包含題數過低（僅一至兩題），恐未能以通過率評定學生之實際建模能力，故本研究僅以「模型選擇」、「模型分析與應用」兩階段為探討重點。

分析臺灣學生於「模型選擇、分析與應用」兩階段所能勝任之「建模階層」，在

表3：臺灣四年級與八年級學生在TIMSS 2011各科試題通過率描述性統計表

建模歷程	學科 建模階層	生命科學 (生物)		地球科學		物質科學 (物理／化學)		小計 ^a	
		四年級	八年級	四年級	八年級	四年級	八年級	四年級	八年級
模型選擇	一	100.00%	85.71%	71.43%	100.00%	100.00%	100.00%	88.24%	93.33%
	二	55.56%	50.00%	66.67%	100.00%	50.00%	81.82%	56.25%	77.78%
	三	50.00%	50.00%	0.00%	100.00%	0.00%	55.56%	25.00%	58.33%
模型建立	三	0.00%	0.00%	—	—	—	—	0.00%	0.00%
模型效化	二	—	100.00%	0.00%	—	—	—	0.00%	100.00%
	三	—	100.00%	—	—	—	—	—	100.00%
模型分析與應用	一	85.71%	0.00%	—	—	100.00%	100.00%	88.89%	50.00%
	二	100.00%	100.00%	100.00%	50.00%	80.00%	33.33%	90.00%	72.73%
	三	57.14%	60.00%	66.67%	80.00%	55.56%	50.00%	57.89%	58.33%
	四	—	0.00%	—	50.00%	—	0.00%	—	20.00%
模型調度	四	—	0.00%	—	—	—	—	—	0.00%
小計 ^b	一	90.00%	75.00%	71.43%	100.00%	100.00%	100.00%	88.46%	88.24%
	二	69.23%	81.82%	60.00%	80.00%	66.67%	71.43%	66.67%	76.67%
	三	50.00%	55.56%	50.00%	83.33%	50.00%	52.17%	50.00%	57.89%
	四	—	0.00%	—	50.00%	—	0.00%	—	14.33%
總計 ^c		69.70%	64.52%	62.50%	82.35%	71.43%	61.36%	68.83%	66.33%

註：^a代表「整體學科」在各建模階段、階層之通過率；^b代表「建模歷程、建模階層」在各學科之通過率；^c代表「整體建模歷程」在各學科之通過率；「—」表示該建模階段、階層未包含試題。

「模型選擇」階段，四年級學生於階層一(88.24%)與二(56.25%)之整體通過率皆高於50%，但階層三則僅有25%的通過率，這顯示此階段學生可掌握至階層二；而八年級學生於階層一(93.33%)、二(77.78%)與三(58.33%)之通過率皆高於50%，亦即八年級學生可掌握至階層三。接著，在「模型分析與應用」階段，四年級學生於階層一(88.89%)、二(90.00%)與三(57.89%)皆高於50%，亦即四年級學生可掌握至階層三；而在八年級學生於階層一(50.00%)、二(72.73%)與三(58.33%)之通過率亦高於50%，但在階層四僅有20.00%通過率，亦即八年級學生亦能掌握至階層三。

進一步分析臺灣四、八年級學生於「不同學科」之建模能力表現。首先，在「模型選擇」階段之各科通過率，四年級學生除生命科學(50.00%)可掌握至階層三，在地球科學(66.67%)及物質科學(50.00%)僅能掌握至階層二；八年級學生在生物(50.00%)、地球科學(100.00%)及物理與化學(55.56%)三科皆可掌握至階層三。第二，在「模型分析與應用」階段之各科通過率，四年級學生在生命科學(57.14%)、地球科學(66.67%)、物質科學(55.56%)皆可掌握至階層三；八年級學生僅能掌握生物階層二(100.00%)及三(60.00%)，未能掌握階層一(0.00%)及四(0.00%)，地球科學(50.00%)可掌握至階層四，而物理與化學僅掌握至階層三(50.00%)，未能掌握階層四(0.00%)。

臺灣在不同年級之建模能力整體表現，在模型選擇階段，四年級學生可掌握至多個因素(階層二)，八年級學生則延伸至掌握因素間簡單關係(階層三)；在模型分析與應用階段，四年級與八年級學生則皆可掌握至因素

間簡單關係(階層三)，但八年級學生能處理的簡單關係，實則蘊含微觀、抽象的概念本質，並包含較多因素。舉例而言：S032158為物理選擇題，試題內容為「把氣體加熱使其溫度增高後，氣體分子會怎樣？」，主要測量「溫度高低」對於「氣體分子」之影響(階層三)，正確答案為B——它們運動會變快，此題則為蘊含微觀、不可見之氣體分子概念。此結果和Newberry Online Limited (2018)設計雷同，其進步路徑指出國中階段更能處理多因素的模型關係。Schwarz等(2009)研究指出在建模實作教學後，國中學生能建構與使用模型預測一群相關現象之多個面向，其建模階層的複雜程度約略與MAAI-T的階層三相應，而國小學生能將模型視為溝通其理解的工具，建構與使用模型以解釋現象成因，其建模階層的複雜程度則大概與MAAI-T的階層二呼應。本研究之結果在建模階段上不盡相同，但在建模階層上，則亦約略可相互呼應。值得注意的是，在模型分析與應用階段，顯示四、八年級之整體答對趨勢階層一(88.89%, 50.00%)比階層二(90.00%, 72.73%)低，此結果看似矛盾，但同樣有機會在TIMSS不同國際基準點(難易程度)之答題正確率的反應中出現。Mullis (2012)指出，TIMSS為包含所有試題，僅能以特定百分比作為判準(建構反應試題之50%、選擇題之65%及50%)，例如：對於建構反應試題而言，通過此特定百分比便屬於該國際基準點，有時，並無法兼顧下一個層級試題的百分比表現。換句話說，兩種不同國際基準點試題皆符合50%之特定判準，但有可能高一層級的基準點試題之答題正確率為60%，而低一層級之答題正確率為55%，而造成高層級試題的答題正確率反而高於低層級試題的情形。

伍、結論與建議

一、TIMSS 2011四、八年級試題在「建模歷程」之分布比例懸殊，但四年級在「建模階層」分布上尚稱平均、八年級則集中於階層二及三

整體而言，TIMSS 2011四、八年級分布於「建模歷程」之試題比例極不均勻，所有學科皆集中於「模型選擇」與「模型分析與應用」兩階段。值得注意的是，僅八年級試題涵蓋模型調度階段，呼應了Schwarz等(2009)中建模學習進程，進入國中階段才出現對模型限制與使用範疇的察覺並運用於新情境(Schwarz等的建模學習進程之階層四)，這也建議了臺灣2018年新課綱雖在國小三年級便有建模課程的引入(國家教育研究院，2016)，但有關模型調度建議於國中以上課程出現。另一方面，四年級試題在MAAI-T「建模階層」之階層一至階層三，分布比例尚稱均勻；然而八年級試題則主要集中於階層二至三。在「不同學科」中，於相同建模階段，則反應出截然不同之階層面向、難度。四年級部分，生命科學整體難度較高，在「模型選擇」著重於測量「兩個以上因素」，在模型分析與應用則測量「單一因素」及「因素間簡單關係」；地球科學、物質科學在模型選擇難度較低，僅測量「單一因素」，而在模型分析與應用難度較高，延伸至測量「因素間簡單關係」。八年級部分，生物及地球科學在模型選擇測量「單一因素」，而在物理與化學則延伸測量「多個因素」；生物於模型應用與分析注重「多個因素」，而地球科學、物理與化學則延伸測量「因素間簡單關係」。上述結果顯示應用MAAI-T於TIMSS 2011四、八年級試題，亦與多數建模實務評量一樣，僅注重測量特定

建模實務類別或面向(Namdar & Shen, 2015; Nicolaou & Constantinou, 2014)。此外，過去文獻提及課程、教學及學生觀點較不重視模型「建立」(Van Driel & Verloop, 1999, 2002)、「效化」與「調度」(張志康、邱美虹，2009；Halloun, 1996)，而TIMSS之測驗便是各國現行課程之縮影，因此在測驗試題上亦反映了其在建模歷程比例分布之不均。研究結果除合乎本研究之預期結果，並同時建議教育工作者在教學過程，不應只重視片段知識的記憶與背誦，應基於「科學與建模實務」的首要考量，讓學生體察模型選擇、建立、效化、分析與應用及調度之循環建模歷程，透過提供小組討論的機會(Nersessian, 1995)、強調「科學理論」與「生活經驗」的關係(Greca & Moreira, 2000)以強化學生的效化、調度之建模能力；另一方面，若欲使用TIMSS試題檢測學生整體建模能力表現，尚需補充模型建立、效化及調度等方面的試題，這亦顯示了全面性建模實務試題設計之困難，未來應加強全面性探討學生建模能力實務測驗之設計與開發。但即使如此，有關模型選擇及模型分析與應用兩階段之試題，仍可作為教師設計這兩階段建模試題之參考。

二、臺灣四、八年級學生於「模型選擇」可分別掌握多個因素、因素間關聯，於「模型分析與應用」則皆可掌握至因素間關聯

臺灣四年級學生於「模型選擇」階段可掌握生命科學至階層三，地球科學、物質科學僅能掌握至階層二，而在「模型應用與分析」階段各科皆可掌握至階層三；臺灣八年級學生於「模型選擇」階段各科皆可掌握至階層三(生物、地球科學、物理與化學)，而在

「模型應用與分析」物理與化學、生物至階層三，地球科學可掌握至階層三。此結果呼應了Schwarz等(2009)研究指出國中學生所能掌握之階層較國小學生更進階，因不同教育階段所能理解模型與建模之程度有所差異，且期待亦不相同(Newberry Online Limited, 2018)。上述研究結果可供教科書編輯或教師設計建模導向教科書、教學或評量時，其所設計之建模能力，於模型選擇與模型分析與應用兩階段學習進程設計之參考。例如：根據學生於TIMSS四年級之表現，未來2019年課綱下，四年級建模課程的設計，便建議以階層三為學生需達成的基本學習目標。至於其餘建模階段(建立、效化及調度)之能力表現，則無法由TIMSS試題回應。

值得一提的是，本研究運用TIMSS 2011所釋出之四年級、八年級試題，瞭解其與建模歷程各階段及建模能力層級之對應情形後，再參照其「正確答題率」作為掌握建模能力之判定基準。是故原有TIMSS在抽樣或推論上的限制，亦為本研究之研究限制。於此限制上，本研究分析TIMSS 2011四、八年級試題結果顯示模型建立、效化、調度等相關歷程之試題缺乏，因此用此結果僅能解讀我國四、八年級學生於「模型選擇」、「模

型分析與應用」兩階段之建模能力，而未能全面檢視其餘各階段建模能力，此結果除與TIMSS原來架構在應用領域中，強調模型使用之結果(Mullis et al., 2009)相呼應外，亦為以TIMSS進行建模能力定位進而設計學習進程之限制。此外，因本研究係以臺灣四、八年級學生為分析對象，本結果僅能解釋、反映臺灣學生之建模能力，未能瞭解臺灣於其他國家之相對表現，建議未來研究可建基於本研究MAAI-T架構進行跨國比較。最後，此分析結果亦反映TIMSS 2011施測前之國際課程中雖有建模之元素但尚未廣泛強調建模之完整歷程。但目前，國際課程對建模重視的情形已有所變動，且TIMSS 2015的試題已經施測完畢，其題目雖尚未釋出，但評量架構對於模型之陳述內容略增(Mullis & Martin, 2013)，建議未來研究亦可奠基於MAAI-T之架構，持續分析並比較TIMSS試題對建模元素的重視情形，並充實建模實務評量之資料庫以協助科學教師相關專業知能之成長。

誌謝

本研究感謝科技部經費補助(計畫編號：NSC 102-2511-S-259-003-MY3, MOST-105-2628-S-259-001-MY3)，在此特致謝忱。

參考文獻

1. 林靜雯(2014)。以系統化方式建立模型與建模之線上教學與評量系統——探討科學課程、概念發展路徑與建模能力之研究——以網路化模型與建模系統協助教師科學建模教學與評量(第1年) (NSC 102-2511-S-259-003-MY3)。臺北市：行政院國家科學委員會。
2. 林靜雯(2015, 12月)。以TIMSS 2011試題及建模能力分析指標檢測學生物質分類建模能力之跨國比較。發表於中國化學會年會。花蓮縣：國立東華大學。
3. 邱美虹(2015)。以系統化方式進行模型與建模能力之線上教學與評量系統——探討科學課程、概念發展路徑與建模能力之研究(第2年) (NSC 102-2511-S-003-006-MY3)。臺北市：行政院國家科學委員會。

4. 邱美虹(2016)。科學模型與建模：科學素養中的模型認知與建模能力。臺灣化學教育，11。查詢日期：2017年1月16日，檢自<http://chemed.chemistry.org.tw/?p=14186>
5. 張志康、邱美虹(2009)。建模能力分析指標的發展與應用——以電化學為例。科學教育學刊，17(4)，319-342。
6. 國家教育研究院(2016)。十二年國民基本教育課程綱要：國民中小學暨普通型高級中等學校自然科學領域(草案)。查詢日期：2017年11月21日，檢自https://www.naer.edu.tw/ezfiles/0/1000/attach/90/pta_10118_2261414_00571.pdf
7. 劉俊庚、邱美虹(2010)。從建模觀點分析高中化學教科書中原子理論之建模歷程及其意涵。科學教育研究與發展季刊，59，23-53。
8. 鐘建坪、邱美虹(2014)。運用SOLO分類法探討科展活動之建模的類型——以八年級物理科展為例。教育實踐與研究，27(1)，31-64。
9. Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.
10. Biggs, J. B., & Tang, C. (2007). *Teaching for quality learning at university* (3rd ed.). Maidenhead, UK: Open University Press.
11. Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The cognitive domain*. New York: David McKay.
12. Buckley, B. C., & Boulter, C. J. (2000). Investigating the role of representations and expressed models in building mental models. In J. K. Gilbert & C. J. Boulter (Eds.), *Developing models in science education* (pp. 119-135). Dordrecht, The Netherlands: Kluwer.
13. Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE-Life Sciences Education*, 7(4), 368-381.
14. Danusso, L., Testa, I., & Vicentini, M. (2010). Improving prospective teachers' knowledge about scientific models and modelling: Design and evaluation of a teacher education intervention. *International Journal of Science Education*, 32(7), 871-905.
15. Davis, E. A., Kenyon, L., Hug, B., Nelson, M., Beyer, C., Schwarz, C., et al. (2008, January). *MoDeLS: Designing supports for teachers using scientific modeling*. Paper presented at the Association for Science Teacher Education. St. Louis, MO.
16. Gilbert, J. K., & Justi, R. (2016). Learning progression during modelling-based teaching. In J. K. Gilbert & R. Justi (Eds.), *Modelling-based teaching in science education* (pp. 193-222). Cham, Switzerland: Springer.
17. The Glossary of Education Reform. (2013). *Learning progression*. Retrieved November 21, 2017, from <http://edglossary.org/learning-progression/>
18. Greca, I. M., & Moreira, M. A. (2000). Mental models, conceptual models, and modelling. *International Journal of Science Education*, 22(1), 1-11.

19. Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28(9), 799-822.
20. Halloun, I. A. (1996). Schematic modeling for meaningful learning of physics. *Journal of Research in Science Teaching*, 33(9), 1019-1041.
21. Halloun, I. A. (2006). *Modeling theory in science education*. Dordrecht, The Netherlands: Springer.
22. Harrison, A. G., & Treagust, D. F. (2000). A typology of school science models. *International Journal of Science Education*, 22(9), 1011-1026.
23. Hempel, C. G. (1958). Fundamentals of concept formation in empirical science. In C. G. Hempel (Ed.), *International encyclopedia of unified science: Foundations of the unity of science* (Vol. 2, pp. 88-93). Chicago, IL: University of Chicago Press.
24. International Association for the Evaluation of Educational Achievement. (2013a). *TIMSS 2011 item information tables*. Retrieved May 6, 2017, from <http://timssandpirls.bc.edu/timss2011/international-database.html>
25. International Association for the Evaluation of Educational Achievement. (2013b). *TIMSS 2011 released items with percent correct statistics—Fourth grade/Eighth grade*. Retrieved May 6, 2017, from <http://timssandpirls.bc.edu/timss2011/international-database.html>
26. International Association for the Evaluation of Educational Achievement. (2013c). *TIMSS 2011 user guide for the international database*. Retrieved May 6, 2017, from https://timssandpirls.bc.edu/timss2011/downloads/T11_UserGuide.pdf
27. Jong, J.-P., Chiu, M.-H., & Chung, S.-L. (2015). The use of modeling-based text to improve students' modeling competencies. *Science Education*, 99(5), 986-1018.
28. Justi, R. S., & Gilbert, J. K. (2002). Modeling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education*, 24(4), 369-387.
29. Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212-218.
30. Lin, J.-W. (2017). A cross-grade study validating the evolutionary pathway of student mental models in electric circuits. *Eurasia Journal of Mathematics, Science and Technology Education*, 13(7), 3099-3137.
31. Lin, J.-W., Cheng, M.-F., Yu, R.-C., & Wu, W. (2016, April). *Comparing elementary and junior High school students' conceptual understanding and analogical modeling competence of electricity*. Paper presented at 2016 NARST Annual International Conference. Baltimore, MD.

32. Lin, J.-W., & Yu, R.-C. (2015, July). *Using TIMSS items to examine students' science modeling ability: A cross-countries comparison in the topic of classification of matter*. Paper presented at the 6th NICE Conference of Network for Inter-Asian Chemistry Educators. Tokyo, Japan.
33. Lin, J.-W., & Yu, R.-C. (2017, April). *Using revised modeling ability analytic index to compare cross-countries students' modeling ability in TIMSS items*. Paper presented at 2017 NARST Annual International Conference. San Antonio, TX.
34. Liou, P.-Y., & Liu, E. Z.-F. (2015). An analysis of the relationships between Taiwanese eighth and fourth graders' motivational beliefs and science achievement in TIMSS 2011. *Asia Pacific Education Review*, 16(3), 433-445.
35. Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy: Analysis of selected items from the TIMSS database. *Journal of Research in Science Teaching*, 42(5), 493-517.
36. Martin, M. O., Mullis, I. V. S., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 international results in science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
37. Mullis, I. V. S. (2012). Using scale anchoring to interpret the TIMSS and PIRLS 2011 achievement scales. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
38. Mullis, I. V. S., Erberber, E., & Preuschoff, C. (2008). The TIMSS 2007 international benchmarks of student achievement in mathematics and science. In M. O. Martin, I. V. S. Mullis, & F. Pierre (Eds.), *TIMSS 2007 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
39. Mullis, I. V. S., & Martin, M. O. (Eds.). (2013). *TIMSS 2015 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
40. Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
41. Namdar, B., & Shen, J. (2015). Modeling-oriented assessment in K-12 science education: A synthesis of research from 1980 to 2013 and new directions. *International Journal of Science Education*, 37(7), 993-1023.
42. Nersessian, N. J. (1995). Should physicists preach what they practice? *Science & Education*, 4(3), 203-226.
43. Newberry Online Limited. (2018). *Science pathways*. Retrieved January 16, 2017, from <http://www.sciencepathways.co.uk>
44. Nicolaou, C. T., & Constantinou, C. P. (2014). Assessment of the modeling competence: A systematic review and synthesis of empirical research. *Educational Research Review*, 13, 52-73.

45. Oh, P. S., & Oh, S. J. (2011). What teachers of science need to know about models: An overview. *International Journal of Science Education*, 33(8), 1109-1130.
46. Robitaille, D. F., Schmidt, W. H., Raizen, S. A., McKnight, C. C., Britton, E., & Nicol, C. (1993). *Curriculum frameworks for mathematics and science*. Vancouver, Canada: Pacific Educational Press.
47. Rodgers, C. (2002). Defining reflection: Another look at John Dewey and reflective thinking. *The Teachers College Record*, 104(4), 842-866.
48. Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Achér, A., Fortus, D., et al. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46(6), 632-654.
49. Sins, P. H. M., Savelsbergh, E. R., & van Joolingen, W. R. (2005). The difficult process of scientific modeling: An analysis of novices' reasoning during computer-based modeling. *International Journal of Science Education*, 27(14), 1695-1721.
50. Van Driel, J. H., & Verloop, N. (1999). Teachers' knowledge of models and modelling in science. *International Journal of Science Education*, 21(11), 1141-1153.
51. Van Driel, J. H., & Verloop, N. (2002). Experienced teachers' knowledge of teaching and learning of models and modelling in science education. *International Journal of Science Education*, 24(12), 1255-1272.
52. Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching*, 46(6), 716-730.
53. Zuzovsky, R., & Tamir, P. (1999). Growth patterns in students' ability to supply scientific explanations: Findings from the third international mathematics and science study in Israel. *International Journal of Science Education*, 21(10), 1101-1121.

Comparing Taiwanese Grade 4 and Grade 8 Students' Modeling Ability in TIMSS 2011 Items

Jing-Wen Lin* and Ruan-Ching Yu

Department of Education and Human Potentials Development, National Dong Hwa University

Abstract

Due to the scantiness of the whole perspectives of the assessments of modeling practices, this study adopted the Revised Modeling Ability Analytic Index (MAAI-T), to analyze the distribution of stages of modeling process and levels of modeling ability in the appropriate 4th and 8th TIMSS items, then exploring Taiwanese fourth and eighth grade students' performance of modeling ability in this assessment. The results indicate: 1. The fourth and eighth grade items in TIMSS 2011 mainly focus on the stages of "model selection" and "model analysis and application." The items equally distributed in each level of modeling ability in fourth grade, but mainly focus on level 2 and 3 in eighth grade. 2. In the stage of model selection, the fourth grade students in Taiwan could handle multiple factors (level 2); the eighth grade students could expand to operate simple relationships (level 3). In the stage of model analysis and application, both fourth and eighth grade students could handle simple relationships (level 3), but eighth grade students could deal with more multiple factors, abstract and microscope variables. This study suggests that using TIMSS items as modeling practice assessments still requires supplementing the items to model construction, validation and deployment stages in order to explore students' performances in whole modeling process. However, designing the modeling-based curriculum, educators can establish learning progression of modeling ability by referencing to the stages of model selection and model analysis and application in the study.

Key words: TIMSS 2011, Revised Modeling Ability Analytic Index (MAAI-T), Modeling Practice Assessment, Learning Progression

* Corresponding author: Jing-Wen Lin, jingwenlin@gms.ndhu.edu.tw