

合作問題解決自我效能量表發展與 科學學習成就關係之探討

王淑卿^{1,*} 葉宗一² 王國華¹

¹國立彰化師範大學 科學教育研究所

²國立清華大學 數學系

摘要

本研究旨在發展國中合作問題解決自我效能(Collaborative Problem Solving Self-Efficacy, CPS-SE)量表作為合作問題解決(Collaborative Problem Solving, CPS)課堂評量工具，並應用以檢驗量表與學生科學學習成就的關係。研究對象是方便取樣自臺灣五個縣市的男女合班國中生，共1,045人。基於學生能力國際評量計畫(Programme for International Student Assessment, PISA) 2015 CPS評量架構而建構量表的構面及題項。量表分析過程包括五個階段：一、項目分析；二、探索性因素分析；三、信度分析；四、驗證性因素分析；和五、複核效度分析。研究發現三個值得注意的結果：首先，建構效度的過程中，萃取三個因素，包括共享理解、行動解決和團隊組織。其次，每個分量表的Cronbach's α 係數皆超過.95，透過結構方程模式(Structural Equation Modeling, SEM)分析競爭模型，選擇適配度最佳且最簡約的模型，顯示模型和資料的適配度良好，表示CPS-SE量表具備良好的信度和效度。第三，自然科學學習成就表現的高成就生與中成就生在共享理解、行動解決和團隊組織等三個分量表，皆顯著優於低成就生，高成就生與中成就生則無顯著差異。

關鍵詞：自我效能、合作問題解決評量、科學學習成就、量表編製、驗證性因素分析

壹、緒論

全球經濟型態由工業經濟轉變為知識經濟和資訊經濟的時代，大量研究發現生活中合作問題解決(Collaborative Problem Solving, CPS)能力的重要性。越來越多的非常規情境問題必須藉CPS能力解決，因此教育應培養學生具有CPS能力，以滿足未來成功工作

與生活的需求(Fiore et al., 2017)。隨著職場與生活對CPS的需求，教育界普遍理解合作和問題解決技能的重要性，然而結合此兩種技能成為CPS的相關研究，直至20世紀末與21世紀初才開始受到重視。關於CPS的評量則因為複雜且不易評測，直至澳洲墨爾本大學提出21世紀技能評量與教育(Assessment and Teaching of 21st-Century Skills, ATC21S)

*通訊作者：王淑卿，wsching123@gmail.com

(投稿日期：民國109年9月8日，修訂日期：民國109年11月27日，接受日期：民國109年11月27日)

計畫，開始評量結合問題解決與合作技能的CPS能力(Griffin & Care, 2015; Griffin, Care, & McGaw, 2012)。

過去20年來曾出現大量有關CPS能力需求的調查和研究報告(Fiore et al., 2017)，但不同學者對於CPS並沒有一致的定義或構面，對其構成要素也有不同觀點(Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995)。而且許多專有名詞也混合使用，例如：合作學習(cooperative learning)、協作學習(collaborative learning)、小組問題解決(small group problem solving)和CPS或稱協作問題解決等，甚至有時相同的專有名詞卻意指不同的技能(Oliveri, Lawless, & Molloy, 2017)。

CPS的定義與構面發展，最早由Stevens與Campion (1994)提出CPS是小組團隊合作(teamwork)所需的知識、技能和能力(Knowledge, Skill, and Ability, KSAs)。強調CPS的KSAs並非個人的人格特質，而是團隊特質。但主要以個人層面進行分析，此概念有助於日後CPS評量著重在個人層次的發展。並依據CPS定義提出團隊合作需要五種能力：衝突解決、合作解決困難、溝通、目標設定和績效管理、計畫和任務協調。此五種團隊合作能力不僅可轉移(transportable)，也可推廣應用於各種情境(Stevens & Champion)。Cannon-Bowers等(1995)將CPS所需的KSAs能力擴展為八種，包括適應力、共享理解、績效監控和回饋、領導、人際關係、協調、溝通，以及決策。KSAs能力依團隊組成可分為：特定團隊能力和通用團隊能力，前者是個人在特定團隊中所需的能力；後者則通用於任何團隊所需的能力。KSAs能力依任務性質可分為：特定任務能力和一般任務能力，前者是與特定任務有關的能力；後者則適用於各種任務有關的能力。

之後學者提出共享心智模型(shared mental model)作為團隊績效的核心，強調在共享心智的過程中，團隊組員可以分享彼此的觀點、知識、資源和解決問題的策略與方法，並注重團隊領導。主張CPS是動態的過程，應制定團隊合作的目標和策略，監控和協調，過程中應重視組員間的社交互動，以及每個人的角色與地位(Aguado, Rico, Sánchez-Manzanares, & Salas, 2014; Chiu, 2000; Cooke, Salas, Cannon-Bowers, & Stout, 2000)。Kyllonen (2012)定義CPS為要求一群學生共同合作解決問題的表現活動。Hesse, Care, Buder, Sassenberg與Griffin (2015)則定義CPS是由團體互動執行多個步驟的活動，以達成共同的目標。綜上，CPS的定義與構面隨著不同的發展，逐漸系統性地建立在其他研究的基礎上。

CPS逐漸被教育單位與國際評量組織所重視，因為真實生活大多是複雜多元的非常規問題，需要各領域高度整合與合作方能解決問題。因此，提高學生的CPS能力是21世紀教育、職場和生活中成功的關鍵技能(De Boeck & Scalise, 2019)。然而如何評量學生的CPS能力？兩大國際評量機構包括經濟合作暨發展組織(Organization for Economic Co-operation and Development [OECD])提出學生能力國際評量計畫(Programme for International Student Assessment, PISA)，以及澳洲墨爾本大學與跨國企業和專家學者共同推動的ATC21S，分別提出CPS評量架構(construct) (Griffin et al., 2012; Organization for Economic Co-operation and Development [OECD], 2013)。兩大國際CPS評量架構雖然有許多共同的特徵，但在微觀層面仍有些差異。

CPS牽涉認知(問題解決)和社會(合作技能)兩個面向，如何評量複雜的CPS能力是值

得探討的問題。慣用的教育評量方法有紙筆測驗、量表、問卷、實作及晤談等，而透過電腦進行線上評量為最近的趨勢。ATC21S的CPS評量是以電腦為介面評測人與人之間的互動過程，雖然較接近真實生活，但是如何識別對話內容並標準化評分是值得注意的課題(Fiore et al., 2017)。另一方面，PISA 2015的CPS評量則設計標準化的任務環境與評分機制，應用電腦化測驗，藉由學生與電腦代理人的對話，評測學生個人的CPS能力(OECD, 2013)。這種方式雖然較能直接測量CPS能力，但是，除了出題不易，過程繁瑣，必須在電腦上操作，不易在普通教室進行。因此學者質疑PISA所設計人為概念性選擇題的CPS問題，是否可代表學生真實的CPS能力(Fiore et al.; Harding, Griffin, Awwal, Alom, & Scoular, 2017)，這也是必須思考的問題。

眾人對於CPS能力的重要性皆有共識，然而對於CPS評量卻缺乏共識(Fiore et al., 2017; Hesse et al., 2015)。CPS雖然是個人的外顯能力，但也包含複雜而不容易直接觀察的認知和社會技能(social skills)等內在的心理歷程。因此，如何直接評量學生的CPS能力是很大的挑戰，到目前也僅有少數教育學者討論如何在課堂評量學生的CPS能力(Arieli-Attali, 2019; Care, Griffin, & Wilson, 2018; Griffin & Care, 2015; von Davier, Zhu, & Kyllonen, 2017)。初步解決方式或許可先利用效化的量表或問卷，間接調查學生CPS能力的自覺，此法雖然較主觀但比較簡便，只是必須假設學生會準確報告其感受(國立編譯館，2000)。另外，也可以從自我效能角度切入，自我效能是學習者對自己某一能力的感知，是判斷自我能力和顯示出來的自信程度(Bandura, 2006)。因此，透過自陳量表，讓學生報告其自我具備的CPS能力和自信程度，可間接理解學生CPS能力的表現。如此，有

助於學生省思CPS能力的學習和成長，並有助教師調整教學與輔導(國立編譯館)。

針對教師在課堂評量學生CPS能力的需求，本研究目的在發展和編製一份適用於中學課室的合作問題解決自我效能(Collaborative Problem Solving Self-Efficacy, CPS-SE)量表，作為CPS評量的工具，檢驗其信效度，並探究國中生科學學習成就與CPS自我效能的關係(relationships)。

貳、文獻探討

一、CPS評量的意涵

(一)ATC21S CPS評量的定義與架構

2009年由澳洲墨爾本大學推動的跨國評量與教育的ATC21S計畫，倡議CPS是21世紀公民必需的核心技能，是成功職業所需的最重要技能之一。定義CPS為要求一群學生共同合作解決問題的表現活動(Kyllonen, 2012)。同時提出CPS評量架構，包含社會技能和認知技能(skills in the cognitive)兩個向度(Fiore et al., 2017; Hesse et al., 2015)。

首先，社會技能(圖1)包括三種子能力(capability)，1.參與(participation)：包含行動、互動和完成任務三個元素，如從事任務與他人互動並回應他人，堅持完成任務或部分任務的能力。2.採取觀點(perspective-taking)：包含適應性反應及聽眾意識(audience awareness)和相互建模等元素，前者是整合隊員貢獻的能力；後者是能根據他人需求調整行為或使他人理解自己的能力。3.社會規範(social regulation)：包含(1)後設記憶(metamemory)，能認識自己的優勢和劣勢；(2)交互記憶(transactive memory)，能認識隊員的優勢與弱點；(3)談判協商，可幫助隊員協調觀點解決問題或達成妥協；(4)責任倡

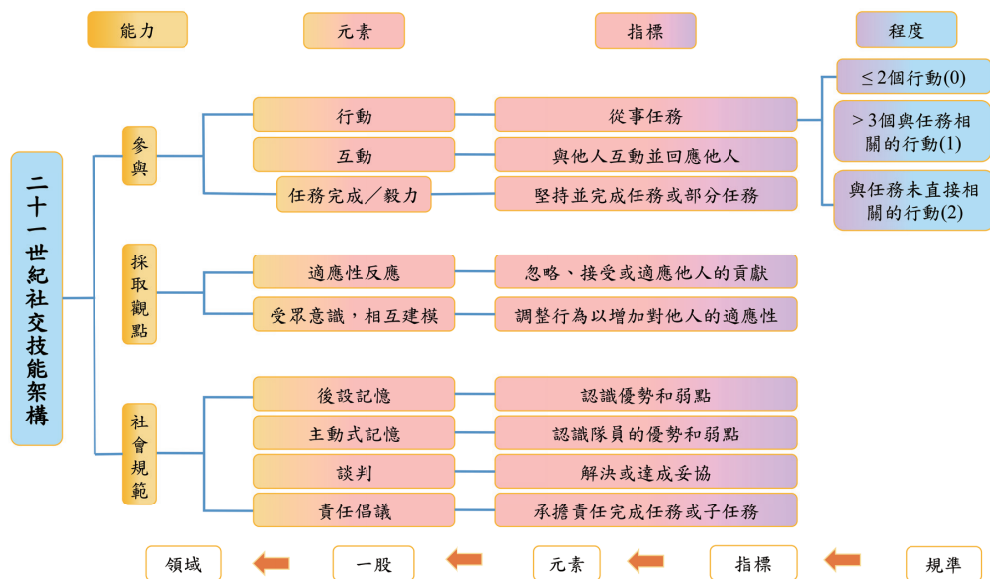


圖1：21世紀技能評量與教育(ATC21S)合作問題解決的社會技能評量架構

資料來源：Collaborative problem solving: Considerations for the National Assessment of Educational Progress (p. 17), by S. M. Fiore, A. Graesser, S. Greiff, P. Griffin, B. Gong, P. Kyllonen, et al., 2017, Alexandria, VA: National Center for Education Statistics.

議，個人能主動以不同方式承擔責任完成任務。每個元素透過可觀察的不同程度的行為指標描述，制定評分規準來評量其技能的程度(Fiore et al., 2017; Hesse et al., 2015)。

其次，認知技能包括兩種子技能，1.任務調節：包含問題分析、目標設定、資源管理、靈活性、數據收集、系統性等元素。靈活性是指對模糊不確定事物的容忍、關注和溝通、協商並理解。系統性指能澈底並有效解決問題。2.建構知識和學習：係指能建立關係和模型，能共享不同資訊及解釋模型間連結關係的能力。對突發事件的相關資訊連結與因果關係的溝通能力，能建立規則以解決問題的能力(Fiore et al., 2017; Hesse et al., 2015)。

ATC21S於2010年開始進行CPS評量，設計不同任務，以電腦為介面引導學生透過螢

幕與其他學生聊天溝通，來評量學生的CPS能力(Griffin & Care, 2015; Hesse et al., 2015)。

(二)PISA 2015 CPS評量的定義與架構

PISA 2015將PISA 2012「問題解決」擴大為“CPS”的評量。評量的目的是，理解15歲學生在教室內外面臨CPS情境所表現的CPS能力，以及他們對未來工作和就學的生活準備。強調CPS是種內在而複雜的機制，除了合作之外還有個人問題解決過程的發現認知內容，包括(A)探究理解問題、(B)表徵闡述問題、(C)計畫與執行、(D)監控與反思。應用問題解決策略、自我調節和後設認知過程來監控實現目標的完成進度。然而，更重要的是團隊小組成員參與CPS任務所需另外的(additional)認知和社會技能。PISA 2015定義CPS為個人能有效地與兩位或以上的團隊成員(或代理人)，共享理解和努力，集中彼此的

KSAs，嘗試解決問題的能力。並參酌其他研究，包括ATC21S的CPS評量架構和PISA 2015 CPS架構草案，2017年提出PISA 2015 CPS評量架構(OECD, 2013, 2017a)，如表1。

PISA 2015的CPS「合作」技能的三個核心能力說明如下：1.建立維持共享理解。能發現自己與團隊組員對問題的觀點及能力，並瞭解自己及組員在任務上的優勢和劣勢；能分享知識建立共享表徵並協商問題的意義；能與團隊組員溝通準備／正在執行的任務和行動；能與組員共同監控對問題的觀點與想法達成共識或修正共識。2.採取適當行動解決問題。能發現問題解決的合作互動類型及目標，以採適當的步驟解決問題；能辨識與描述任務，與組員溝通和合作進行解釋、辯論或協商問題與任務的類型和步驟；能合作制定解決問題的行動計畫；能根據問題與任務的限制和規則，採取不同的活動，並監控行動結果和評估問題解決的成效。3.建立維持團隊組織。能瞭解自己及組員的角色以解決問題；能描述組員的角色和團隊組織，並溝通協議團隊的參與規則；能遵守參與規則，並促使其他組員能執行其任務；能監控團隊組織，並提供回饋，促進溝通、

理解、服從組織領導及適應組織和角色，完成任務並維持團隊組織(OECD, 2017a)。

雖然之前PISA 2012問題解決評量有大量研究成果，但仍缺乏評量個人CPS能力的方法，也缺乏詳盡的評量標準應用於CPS評量(Griffin & Care, 2015)。而影響CPS能力的主要因素有許多，包括合作、問題解決、學生背景如其先備知識和人際關係，學生個人特徵如態度、情感和人格等，以及評量的情境如問題情境、團隊組成、任務特徵和試題媒介等(OECD, 2013, 2017a)。因此，PISA 2015 CPS評量決定僅針對學生基本的先備知識，不強調其專業知識。

綜合上述，ATC21S和PISA的CPS評量架構都包含兩種技能，合作技能(社會技能)和問題解決技能(認知技能)。皆由兩種技能雙維交叉組成，交叉點皆有特定的子技能，透過質性描述不同行為程度的評分標準來評量。兩者皆以個人的層面來評量學生的CPS能力，因為CPS的成效取決於每個成員的CPS能力，小組的成功是基於個人的成功。CPS能力涉及合作的社會技能和認知的問題解決技能，多是個人內部的心智活動。因此必須設計人為的CPS情境，利用個人與組員在互動過

表1：PISA 2015合作問題解決評量架構

歷程	(1)建立維持共享理解	(2)採取適當行動解決問題	(3)建立維持團隊組織
(A)探究理解問題	(A1)發現團隊組員的觀點及能力	(A2)發現問題解決的合作互動類型及目標	(A3)瞭解角色以解決問題
(B)表徵闡述問題	(B1)建立共享表徵並協商問題的意義	(B2)辨識與描述任務	(B3)描述角色和團隊組織(溝通協議／參與規則)
(C)計畫與執行	(C1)與團隊組員溝通要／正在執行的行動	(C2)制定計畫	(C3)遵循參與規則(促使其他組員執行其任務)
(D)監控與反思	(D1)監控與修正共識	(D2)監控行動結果並評估問題解決的成效	(D3)監控提供回饋及適應團隊組織和角色

註：PISA：學生能力國際評量計畫(Programme for International Student Assessment)。修改自“PISA 2015 collaborative problem solving framework,” by Organization for Economic Co-operation and Development, 2017 (<https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>).

程中的外在表現，來評測個人的內在心智活動，並推斷個人思考的認知過程。

分析ATC21S的CPS社會技能評量架構(圖1)與PISA 2015的CPS評量架構(表1)，ATC21S的「參與」能力可對應PISA 2015 CPS「合作」技能維度的三個核心能力：建立維持共享理解、採取適當行動解決問題和建立維持團隊組織。透過參與、分享、溝通、行動、監控與回饋，完成任務。「採取觀點」能力可對應PISA三個核心能力，藉由採取觀點，分享知識，與組員溝通協商問題和行動，以維持團隊組織。「社會規範」能力也可對應PISA的三個核心能力，能共同監控、理解、服從及維持團隊的共識、行動、組織及角色(Griffin & Care, 2015; OECD, 2017a)。

綜合上述，因此本研究選擇從PISA 2015 CPS評量架構「合作」維度的三個核心能力，發展出CPS-SE量表的三個構面：共享理解、行動解決和團隊組織，用以探究學生複雜而動態的CPS能力。其中「共享理解」構面：可理解小組成員的觀點與能力，溝通、處理衝突，即建立維持共享理解的能力。「行動解決」構面：可發現問題、制訂計畫、監控行動結果、管理任務進度，即採取適當行動解決問題的能力。「團隊組織」構面：可瞭解彼此的角色與地位、溝通協商參與的規則、遵循規則、完成任務、監控回饋和適應組織角色，即建立維持團隊組織的能力(OECD, 2017a)。藉由此三個構面來評測理解學生複雜而動態的CPS能力，對於教育是一種重要的評量方式。因為身處21世紀資訊化、科技化的變革時代，有許多複雜且非常規的問題，需要跨領域人才合作共同解決，CPS能力對於是否能成功完成任務至關重要。因此教育應透過CPS評量理解學生的CPS能力，培養並提高學生的CPS能力。因為

21世紀公民必需具備的核心技能之一CPS能力，是決定教育、職場和生活成功的關鍵技能，具有CPS能力有助於滿足未來成功工作與生活的需求(De Boeck & Scalise, 2019; Fiore et al., 2017; Griffin et al., 2012)。

二、CPS評量工具及相關研究

國內外有關評量CPS這種21世紀核心技能的研究還在起步階段，除了ATC21S和PISA 2105有發展大型且標準化的CPS評量外，僅少數研究是有關CPS能力的課堂評量(Care et al., 2018; von Davier et al., 2017)。如前所述，ATC21S的CPS評量架構中的每個元素，透過可觀察的行為指標描述，制定評分規準來評測該行為的能力程度。以電腦為評量介面，設計不同任務引導學生透過螢幕與其他學生聊天溝通，來評量學生的CPS能力(Fiore et al., 2017; Griffin & Care, 2015; Hesse et al., 2015)。而PISA 2105的CPS評量，基於CPS能力的複雜動態過程很不容易評測，所以決定將每個學生置於人為的CPS情境中，以電腦程式設計不同任務，團隊成員受到完全控制，透過學生與電腦代理人的互動，評量CPS架構中的各種子技能(OECD, 2013, 2017a)。

以上兩種評量的核心皆在於「合作」歷程，勝於解決問題的歷程。最重要的是，兩種評量的CPS定義都將「評量」視為其優先事項，皆強調任務與標準化的評量。評量的理念皆以概念構成題項，然後測試學生對題項的反應(Fiore et al., 2017; OECD, 2017a)。然而，兩者的最主要差異在於，ATC21S使用綜合方法將社會技能和認知技能融合一起，產生不同等級的評量粒度(granularity)和準確性(accuracy)。ATC21S評量學生之間的互動，優點是較接近真實情境，缺點是其評量情境的標準化程度較低。因為兩個學生合作時的

行為難以控制，評量時應如何識別兩人的對話與互動？如何利用此對話來標準化計分？反觀，PISA評量學生與電腦代理人之間的互動，優點是利用人為設計情境使任務更標準化。首先將CPS任務過程中的每個轉折標準化評分，每個單元有5 ~ 30個獨立測量點，每個題項有三個不同的選擇，代表不同分數，即不同的CPS能力。缺點是電腦代理人的行為必須事先編寫程式，因而限制學生的選擇，無法對題項做出適當的真實反應。尤其有人質疑這種人為的概念性CPS問題，是否可以評量出真實的CPS能力(Griffin & Care, 2015)。雖然此兩個國際評量各有不同的CPS評量方式，但各有其重要的意涵。

另外，也有研究是以自陳量表來進行CPS評量。例如：哈佛醫學院曾發展一套「合作解決問題的一致性和影響力量表」(Collaborative Problem Solving Adherence & Impact Measures, CPS-AIMs)。以Likert七點尺度量表自陳方式評估個人與CPS相關的信念和感受(Wang & Pollastri, 2019)。CPS-AIMs共有三種形式，其中CPS-AIM-E是提供學校教育者使用，可線上填寫或手寫，分為4個分量表，共32題，包含「遵守CPS哲學」、「感知正向積極的影響」、「感知倦怠」和「CPS感知能力」。平均得分越高，表示教育者感知自己的CPS能力越高。文中建議未來可繼續針對「CPS感知能力」進行研究，並使用更完整的題項來評測自我感知的CPS能力(Wang & Pollastri)。

總之，目前CPS評量方式仍處於研究發展階段。主要有強調任務與標準化的CPS評量和自我評估的自陳量表等兩類評量方法。值得關注的是，標準化CPS評量的評分模型是取決於之前評量的什麼內容？評量的方式是如何？尤其設計和開發評量時是否符合預

期的評分模型(Fiore et al., 2017)？這些都是標準化評量的重要挑戰。此外，自我評估CPS能力的自陳量表是否有良好的信效度？以及適用對象的CPS能力與相關因素的預測性是否良好？因此，儘管有不同方式的CPS評量，只要能協助理解學生的CPS能力，藉以分析學生CPS能力的優勢與弱點，提供適當的教學或輔導介入的依據，教育者皆可依其不同需求選擇適合的CPS評量方式。

三、自我效能評量與個人能力的相關研究

直接評量個人的CPS能力須面臨很大的挑戰，如果採用自我效能間接推斷個人的CPS能力，是否能克服挑戰？自我效能源自於Bandura (1986)提出的社會認知理論(social cognitive theory)，主張自我效能是個人、環境與行為三者交互作用的結果，是個人對自己是否有能力完成某特定工作的自我判斷與評估。

自我效能是對自我能力、認知及行為表現的主觀評價，相信自己能克服環境與困難的信念，是行為動機的來源，往往會決定行動，可以預測個人的能力、認知及行為表現(Bandura, 1986)。例如：Judge, Erez, Bono與Thoresen (2002)研究發現，自我效能和自尊是相同的概念，因為成功經驗會增強自尊感和自我效能感，反之也會因為失敗經驗而降低。高自我效能的人，有較高的自尊感，因而較能自我肯定，對自我能力也較有自信，面對困難任務反而更努力，有較高的機會解決問題達成目標。然而，低自我效能的人自尊感較低，容易感到壓力、無力感與沮喪，面對困難時，會懷疑自己的能力，感到挫折導致努力降低並放棄(Judge et al.)。另外，Schunk (1995)綜合1982 ~ 1992年間自我效能

的相關研究結果，指出藉由自我效能可以預測學生的學習動機與學習成就。呼應Bandura, Caprara, Barbaranelli, Gerbino與Pastorelli (2003)的研究發現，11 ~ 14歲的青少年若能透過自我效能以自律學習(self-regulation learning)或自律而達成學習目標，會直接影響其學業成就。因此，自我效能和自律是影響學生學習和成就的關鍵過程(Bandura, 2001)。

自我效能是學習者在特定學習或行為表現的自我感知能力(Bandura, 1997)，而自律則是自我產生系統性的思想、感覺和行為而影響個人的知識和技能的學習(Zimmerman, 2000)。高自我效能者會透過自律，影響學習動機、學習興趣而影響學習，而提高學業成就表現(Sorić & Palekčić, 2009)。透過自我效能和自律學習的建模過程，有助於提升學習動機、認知策略與後設認知策略，因而提高學業成就(Schunk & Zimmerman, 2007)。Ramdass與Zimmerman (2011)也發現，自律學習是自我效能的外在應用，高自我效能者由於自律學習而提升其學業成就，因為自律學習會影響學習行為，而決定學習目標與學業成就。

總而言之，雖然自我效能僅能反映個人主觀對自我能力的信念，與個人的能力無直接關聯，但會影響之後個人的記憶、判斷、思考、態度、選擇、歸因及情緒等行為過程與方式，也會影響自尊和自律、學習興趣和學習動機，進而影響學習行為和學習成就。因此，縱使影響個人行為的表現、能力有各種因素，自我效能是直接影響個人行為動機的主要因素(Bandura, 1986; Miwa, 2005; Zimmerman, 2000)。換言之，自我效能與能力、認知及行為表現雖然沒有直接的因果推論，但是有明顯的相關，而且對於個人的能力及學習成就也具有預測性。

有些研究以Bandura (1986)社會認知理論的核心概念(自我效能)來編製量表，以預測特定行為或能力。例如：Hammack與Ivey (2017)利用教師工程自我效能量表和教師工程教學自我效能量表，測量542位幼稚園 ~ 小學五年級的科學教師，並指出教師的低工程自我效能和低工程教學自我效能與他們所具有工程的學科教學知能(Pedagogical Content Knowledge, PCK)有關。另外，Robertson Evia, Peterman, Cloyd與Besley (2018)編製並效化「科學家對公眾參與科學的自我效能量表」，提供科學家對於進行民眾參與科學活動的自我效能與自我反思的工具，評估科學培訓、民眾參與科學及科學傳播的成效。林碧芳與邱皓政(2008)編製「創意教學自我效能感量表」，研究指出創意教學自我效能感是預測教師創意教學行為的具體指標。彭淑玲(2019)發現，當學生的學業自我效能程度提升時，會促使其學習更趨向精熟目標，導致更努力學習而降低學習無聊感。

綜合本章文獻探討，CPS不僅是外顯的能力表現，也是複雜不容易直接觀察的內在心理歷程，很難直接準確測量。因此，利用自陳量表，並從自我效能的角度切入，間接調查學生CPS能力的表現，應是可行的做法。自我效能是學習者對自己某一能力的感知，是一種判斷自我能力和顯現出來的自信程度(Bandura, 2006)。雖然，不是直接評量學生的CPS能力，但是自我效能是相信自己有能力克服困難完成任務，會影響自尊與自律學習，決定行為的動機或興趣，往往可預測個人的行為或能力表現。因此，研究者試圖發展一份適用於中學生的CPS-SE量表，作為課堂的CPS評量工具，幫助學生評測其自我感知的CPS能力，並評估其適用性。

參、研究方法

一、研究對象

本研究採方便取樣自男女混合常態編班的七～九年級學生，分成預試樣本與正式樣本共1,045人。取樣對象是來自「CPS探究教學教師學習社群」教師所任教的班級學生，取樣學生會接受其自然教師於一學年間使用CPS探究教學策略進行二個自編跨領域統整課程融入自然課程的CPS探究學習。此社群源自於10多年前研究者與幾位跨縣市、跨校的國中自然領域教師所組成的「奈米探究教學教師社群」，後來發展成跨領域CPS探究教學的教師專業學習社群。透過一次預試、兩次正式施測進行量表的發展與效化，並再次收集資料檢核此量表與自然科學學習之間的關係。取樣對象與分析方法統整如表2，以預試樣本190人發展量表初稿，進行項目分析(item analysis)、探索性因素分析(Exploratory

Factor Analysis, EFA)與信度分析(reliability analysis) (邱皓政, 2019)。正式樣本共三筆，分別進行驗證性因素分析(Confirmatory Factor Analysis, CFA)和信度分析、複核效度分析與量表的應用。

研究取樣依據邱皓政(2019)量表編制的建議，樣本數是題目數的5～10倍或超過150人，則可進行EFA。本量表共25題，預試樣本190人，符合理論標準。李茂能(2019)建議結構方程模式(Structural Equation Modeling, SEM)樣本數介於250～500人，本量表的SEM CFA與信度分析、複核效度分析分別為276、305人，量表與自然科學學習成就的相關應用以274人進行檢測，符合理論標準。

二、CPS-SE量表的編製與效化

CPS-SE量表發展初稿參考ATC21S和PISA的CPS評量架構，發現有很多相同之處。最後，主要依據PISA 2015 CPS評量架構(表1)，

表2：CPS-SE量表發展的取樣對象與分析方法統整

項目	預試樣本	正式樣本		
		一	二	三
取樣縣市	臺中市(1縣市)	臺中市、彰化縣和高雄市(3縣市)	臺中市、彰化縣、南投縣、臺南市和高雄市(5縣市)	臺中市、彰化縣和高雄市(3縣市)
取樣學校(所)	4	6	6	5
取樣人數(人)	190	276	305	274
有效樣本率(%)	91	95	97	97
統計分析	項目分析、探索性因素分析、信度分析	驗證性因素分析、信度分析	複核效度分析	One-way ANOVA
統計軟體	SPSS	Amos	Amos	SPSS
目的	分析預試題項與確認正式題項、萃取因素個數與命名、檢驗總量表與分量表的Cronbach's α 信度分析	檢驗建構效度之收敛效度與區別效度、信度	檢驗複核效度	應用CPS-SE量表

註：CPS-SE：合作問題解決自我效能(Collaborative Problem Solving Self-Efficacy)；one-way ANOVA：單因子變異數分析(one-way Analysis of Variance)。

分析文本中「合作」維度的三個核心能力中每項子技能，探討每項子技能的能力表現描述(OECD, 2017a)與中學生實際CPS能力的相關表現，作為題項文字編制的依據。過程中邀請二位科學教育教授、五位有經驗的國中科學教師和國中生，多次討論與協助題項文字的審題和修訂，進行表面效度和內容效度的檢核。並於四所國中進行試驗研究(pilot study)，分析預試結果，最終將CPS-SE量表分為三個分量表：共享理解、行動解決、團隊組織，分別是8題、8題、9題，共25題。並以自我效能的角度，由學生在這些有關CPS能力的題項自評其具備程度。學生要自評所具備的CPS能力程度，需要具備自我判斷其能力表現的自信心(Bandura, 2006)。據此，「共享理解」構面旨在讓學生自評建立維持共享理解能力的自信和表現程度，如例題：我能瞭解團隊組員的想法及能力。「行動解決」旨在讓學生自評採取適當行動解決問題能力之自信和表現，如例題：我能與團隊成員溝通所要執行的行動。「團隊組織」旨在讓學生自評建立維持團隊組織合作能力的自信和表現，如例題：我能幫助組員依約定完成他們的任務。依據吳毓瑩(1996)建議在態度意見測量採四點量表，可減少被方法變異量所干擾，因此採Likert四點量表記分，選項有「非常具備」、「普通具備」、「稍未具備」、「尚未具備」，分別給予4、3、2、1分。採用SPSS 22.0統計軟體進行項目分析、EFA和信度分析。

三、CPS-SE量表的驗證與應用

藉由Amos 22.0 SEM進行CFA，檢驗建構效度和複核效度(李茂能, 2019)。首先依據EFA所獲得的分量表因素結構，以第一筆正式樣本進行CFA，建立不同假設的競爭模型，檢驗不同模型的適配度。分析比較不同競爭模型的適配程度，以確認最終模型的CFA信

效度(李茂能)。再以模型內在結構適配度(fit of internal structure model)評斷量表題項的內部一致性，由題項信度、平均變異數抽取量和組合信度(Composite Reliability, CR)等指標判斷建構效度(Hair, Black, Babin, & Anderson, 2010)。以第二筆正式樣本進行複核效度分析(cross-validation analysis)，檢驗不同樣本的複核效化，驗證量表的理論模型是否有效度延展性(Byrne, 2010)。最後以第三筆正式樣本探討不同科學學習成就學生在CPS-SE量表差異之分析，探究國中生科學學習成就與CPS自我效能的關係。

肆、研究結果

一、量表的信效度分析

(一)項目分析

首先以190位國中生進行量表的项目分析，採四種評判指標(邱皓政, 2019)，分析結果如下。第一，極端組檢驗，分別以總量表總分最高的前27%和最低的後27%學生為高分組和低分組，進行每個題項的獨立樣本 t 檢驗，比較高低分組平均數之差異。25題極端組檢驗 t 值介於13.30 ~ 29.95，差異顯著性 p 值皆 $< .01$ ，顯示25題皆有良好的鑑別度。第二，題項與總分相關係數，25題與總量表的Pearson相關係數介於.74 ~ .88 ($> .30$)，所有 p 值均 $< .01$ ，顯示所有題項皆測量同一種概念，具有同質性。第三，校正後題項與總分的相關係數，25題都介於.72 ~ .87 ($> .30$)， $p < .01$ ，各題皆有同質性。第四，刪題後的Cronbach's α 係數，25題皆為.98 ($> .70$)，具有良好信度。綜上，25題皆無不良指標，不必刪題，全部保留。

(二)EFA

經Kaiser-Meyer-Olkin (KMO)檢驗和

Bartlett球形檢驗結果， $KMO = .96$ ($> .90$ ，極佳)，代表變數間有共同因素。Bartlett球形檢驗卡方值為5,743.02，達顯著水準($p < .01$)。綜上適合進行EFA，結果如表3。因素萃取採主成分分析法、直接斜交轉軸法、選取特徵值 > 1 的因素進行轉軸，萃取出三個因素。未轉軸的三個因素特徵值分別為16.80、1.52和1.05，解釋變異量為67.20、6.06和4.20%，累

積解釋變異量達77.45%。25題的轉軸後因素負荷量介於.57 ~ .85之間($> .50$)，通過因素負荷量高於.50則不必刪題的標準，表示個別題項與因素間有高相關性且影響力佳(邱皓政，2019)。

(三)信度分析

將EFA所通過的25題進行Cronbach's α 信

表3：合作問題解決能力量表之探索性因素分析結果($N = 190$)

題號	轉軸後因素負荷量			未轉軸的特徵值	未轉軸的解釋變異量	未轉軸的累積解釋變異量
	因素一 團隊組織	因素二 行動解決	因素三 共享理解			
17	.64	.45	.33	16.80	67.20	67.20
18	.66	.42	.37			
19	.70	.44	.34			
20	.71	.41	.39			
21	.74	.38	.33			
22	.77	.33	.37			
23	.79	.29	.33			
24	.85	.24	.29			
25	.81	.24	.29			
1	.24	.70	.34	1.52	6.06	73.25
2	.28	.74	.26			
3	.31	.72	.32			
4	.21	.78	.37			
5	.41	.74	.32			
6	.37	.75	.35			
7	.37	.70	.30			
8	.37	.74	.26			
9	.36	.48	.57	1.05	4.20	77.45
10	.31	.45	.59			
11	.32	.43	.73			
12	.34	.39	.74			
13	.43	.41	.66			
14	.42	.34	.74			
15	.42	.27	.75			
16	.36	.33	.73			

註：表中分析採用主成分分析法配合直接斜交轉軸法。

度分析，確定量表題項的內部一致性。結果發現，CPS-SE總量表 α 值為.98，三個分量表 α 值分別為.95、.96、.97，皆高於 $>.70$ 的標準，表示總量表與各分量表皆具有良好的信度(邱皓政，2019)。

綜合以上CPS-SE量表的EFA與信度分析結果，可知量表題項的關連性良好，信度佳。根據EFA所得之解釋變異量顯示，確認適合將CPS能力的構面拆成三個因素，並命名為「共享理解」、「行動解決」和「團隊組織」，然後繼續進行CFA。

二、CPS-SE量表的驗證

(一)CFA

以正式樣本第一筆276位國中生，根據EFA獲得的25題三因素進行CFA，檢驗建構效度。先以Amos軟體SEM繪製三個分量表的假設模型，進行CFA結果如圖2。CFA應用SEM檢驗的主要目的是，探究無法直接觀察或測量的抽象構面與其組成結構的關係，以潛在變項(因素)的形式，利用觀察變項(題項)的模型路徑分析來估計，透過SEM嚴謹的統計檢驗證明抽象構面的存在(李茂能，2019；邱皓政，2019)。CFA三種適配度評估結果如下。

第一種模型基本適配度(preliminary fit criteria)評估，結果所有題項的因素負荷量介於.67~.86之間($>.50$)，積差相關係數皆小於1，符合指標(Bagozzi & Yi, 1988)。

第二種模型整體適配度(overall model fit)評估，採絕對適配度指標(absolute fit measure index)、增值適配度檢驗(incremental fit measures)及精簡適配度指標檢驗(parsimonious fit measures)。絕對適配度指標的平均平方誤差平方根(Root-Mean-Square Error of Approximation, RMSEA)指標，是比較假設模型與最佳模型的差異程度，RMSEA數值為.00時為完美適配，

$<.05$ 為良好適配， $.05 \sim .08$ 為不錯適配， $.08 \sim .10$ 為中度適配， $>.10$ 為不良適配(Bentler & Yuan, 1999)。增值適配度指標是比較研究模型的配適度與假設模型的改善程度，其中標準適配度指標(Normed Fit Index, NFI)採Ullman (2001)建議 $NFI >.80$ 的標準，精簡適配度指標是判別模型的精簡度。三種競爭模型檢驗結果(表4)，一階直交模型的RMSEA = .09、配適度指標(Goodness of Fit Index, GFI) = .80、期望複核效度指標(Expected Cross-Validation Index, ECVI) = 3.33、NFI = .84、比較適配度指標(Comparative Fit Index, CFI) = .88、增值適配度指標(Incremental Fit Index, IFI) = .88、簡約標準適配度指標(Parsimony Normed Fit Index, PNFI) = .77、簡約比較適配度指標(Parsimonious Comparative Fit Index, PCFI) = .81、赤池信息量準則(Akaike's Information Criterion, AIC) = 917.79，皆達適配標準。但是CFI = .88低於.90的適配標準，和IFI = .88低於.90的適配標準，因此CFI和IFI未通過適配標準。一階斜交模型的RMSEA = .07、GFI = .85、ECVI = 2.57、NFI = .88、CFI = .93、IFI = .93、PNFI = .80、PCFI = .84、AIC = 705.33，皆達適配標準。

一階直交模型與一階斜交模型經模型整體適配度檢驗比較，結果顯示：一階直交模型未通過整體適配度檢驗，因此推翻一階直交模型。然而，一階斜交模型通過檢驗且符合理論，代表一階三個因素之間有相關性。同時，二階三因素模型整體適配度檢結果：RMSEA = .07、GFI = .85、ECVI = 2.57、NFI = .88、CFI = .93、IFI = .93、PNFI = .80、PCFI = .84、AIC = 705.33，皆達標準。顯示二階三因素模型符合理論，且有二階的共同因素(李茂能，2019)。雖然整體適配度評估結果(圖2)，一階斜交模型和二階三因素模型皆達適配標準且指標數值也相同，但是，參

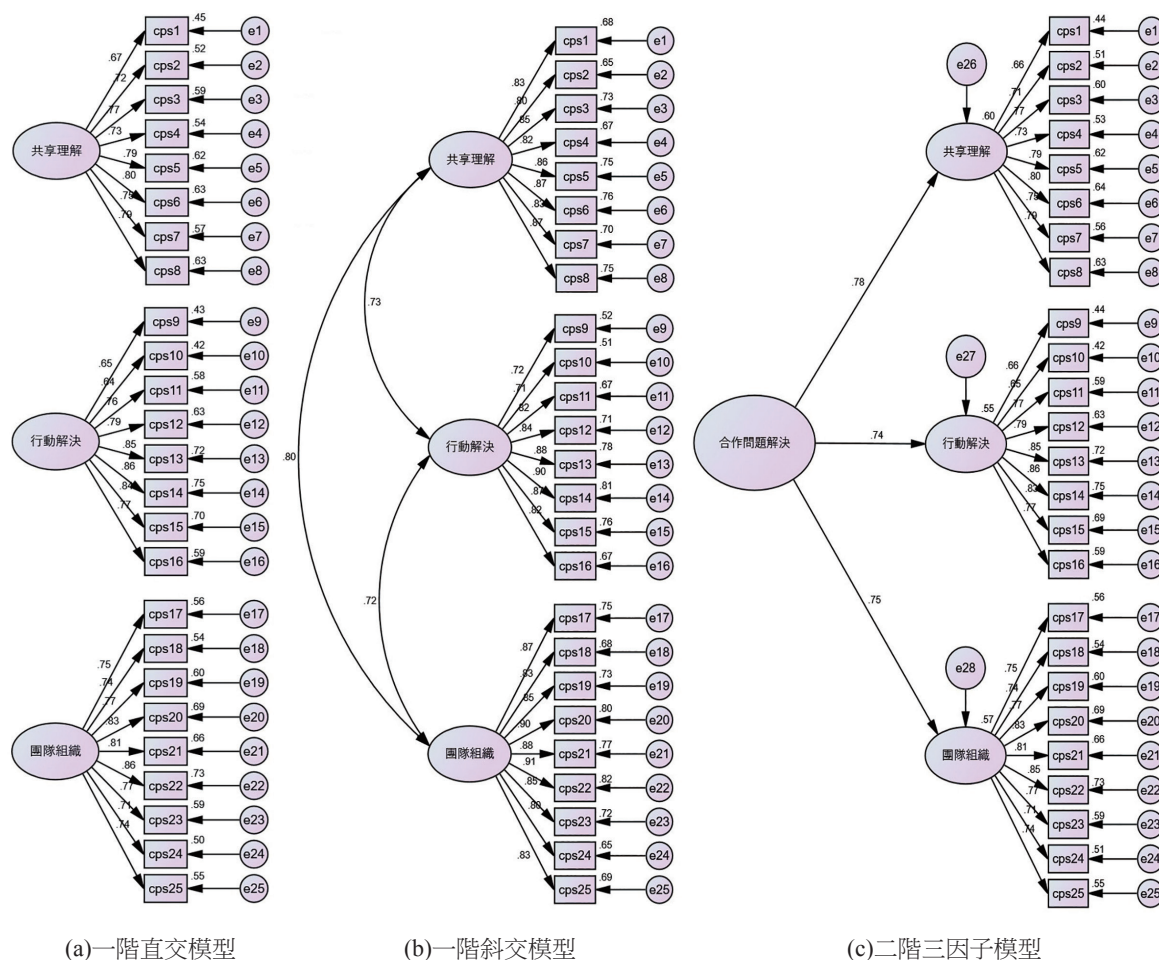


圖2：合作問題解決自我效能(Collaborative Problem Solving Self-Efficacy, CPS-SE)量表
三個分量表的三種競爭模型驗證性因素分析結果($N = 276$)

酌OECD (2017a) PISA 2015的CPS架構，本研究選擇二階三因素模型作為CPS-SE量表的最終模型，支持在「共享理解」、「行動解決」、「團隊組織」的三個因素之上，有二階的CPS構面存在。

第三種模型內在結構適配度評估，目的是模型內在品質評估。將二階三因素模型進行分析結果(表5)：25題的因素負荷量介於.65 ~ .86之間，> .50的標準。所有題項Cronbach's α 值介於.44 ~ .75間，除了題1、9、10的 α 值略< .50，其餘皆> .50的可信標準。三個分量

表的CR，又稱建構信度(construct reliability)為.90、.91、.92，> .70的標準，顯示三個構面具有內部一致性，即建構信度。三個分量表的平均變異數抽取量(Average Variance Extracted, AVE)分別為.54、.57、.57，> .50的標準，代表潛在變項(因素)可解釋觀察變項(題項)的解釋力總和達標(邱皓政，2019)。表示相同構面的題項間有高度相關，具有建構效度的收斂效度(convergent validity)。區別效度(discriminant validity)檢驗結果如表6的構面相關係數矩陣，「共享理解」AVE平方根 =

表4：三種競爭模型整體適配度評估之比較($N = 276$)

適配度指標	適配標準	一階直交模型	一階斜交模型	二階三因素模型
絕對適配度指標				
RMSEA	< .10	.09	.07	.07
GFI	> .80	.80	.85	.85
ECVI	越小越好	3.33	2.57	2.57
增值適配度指標				
NFI	> .80	.84	.88	.88
CFI	> .90	.88	.93	.93
IFI	> .90	.88	.93	.93
精簡適配度指標				
PNFI	> .50	.77	.80	.80
PCFI	> .50	.81	.84	.84
AIC	越小越好	917.79	705.33	705.33

註：RMSEA：平均平方誤差平方根(Root-Mean-Square Error of Approximation)；GFI：配適度指標(Goodness of Fit Index)；ECVI：期望複核效度指標(Expected Cross-Validation Index)；NFI：標準適配度指標(Normed Fit Index)；CFI：較適配度指標(Comparative Fit Index)；IFI：增值適配度指標(Incremental Fit Index)；PNFI：簡約標準適配度指標(Parsimony Normed Fit Index)；PCFI：簡約比較適配度指標(Parsimonious Comparative Fit Index)；AIC：赤池信息量準則(Akaike's Information Criterion)。

.74，大於與「行動解決」和「團隊組織」構面間的相關係數.58和.59。「行動解決」AVE平方根 = .75，大於與「團隊組織」構面間的相關係數.56。結果代表不同構面間的題項的相關性低，具有區別性，即具有建構效度的區別效度(Hair et al., 2010)。整合上述結果，支持CPS-SE量表的二階三因素模型具有良好的建構信效度。

綜合以上EFA、信度分析與CFA的檢驗結果，CPS-SE量表的25道題項和三個因素「共享理解」、「行動解決」和「團隊組織」所組成的二階三因素模型，具有良好的建構信效度(李茂能, 2019)。

(二)複核效度分析

再以不同的305人為樣本，進行複核效度三種分析，結果如下。第一種相同型態模型檢驗，支持二階三因素模型。第二種測量模型檢驗，獲得三個因素，且三因素

之間有相同的共變數，其因素負荷量分別為.75、.72、.77 (> .50)，符合指標。第三種模型適配度檢驗，經絕對適配度、增值適配度和精簡適配度三種指標複核檢驗的結果：RMSEA = .07、GFI = .84、ECVI = 2.61、NFI = .88、CFI = .92、IFI = .92、PNFI = .79、PCFI = .84、AIC = 792.46，皆達適配度指標標準(Hair et al., 2010)。綜合上述，複核效度分析結果與CFA相同，都獲得三個因素，每個因素與每道題項的因素負荷量皆達標準；都獲得二階三因素模型，經模型適配度檢驗皆達適配度指標。顯示二階三因素模型的CPS-SE量表，具有複核效度，效度延展性佳。

(三)CPS自我效能與科學學習成就的關係

以三縣市五所國中共10個班級274位國中生，進行量表應用的分析，旨在理解不同科學學習成就學生在CPS自我效能的差異。依教師提供的各班級學生前一學期在學校的

表5：二階三因素模型內部結構適配度分析結果($N = 276$)

題號	因素負荷量	題項信度	組合信度	平均變異數抽取量
1	.66	.44		
2	.71	.51		
3	.77	.60		
4	.73	.53		
5	.79	.62		
6	.80	.64		
7	.75	.56		
8	.79	.63		
9	.66	.43		
10	.65	.42		
11	.77	.58		
12	.79	.63		
13	.85	.72		
14	.86	.75		
15	.83	.70		
16	.77	.59		
17	.75	.56		
18	.74	.54		
19	.77	.60		
20	.83	.69		
21	.81	.66		
22	.85	.73		
23	.77	.59		
24	.71	.51		
25	.74	.55		
總量表			.95	.44
共享理解分量表	.78	.60	.90	.54
行動解決分量表	.74	.55	.91	.57
團隊組織分量表	.75	.57	.92	.57

表6：二階三因素模型區別效度分析結果($N = 276$)

構面	共享理解	行動解決	團隊組織
共享理解	.74 ^a		
行動解決	.58	.75 ^a	
團隊組織	.59	.56	.75 ^a

註：^a數值代表各構面平均變異數抽取量(Average Variance Extracted, AVE)的平方根，其餘為各構面標準化的相關係數。

自然科學領域的學期成績，各班學生區分為三組成就生。全班總分最高和最低的27%學生為高成就生和低成就生，其餘為中成就生。三組學生的總量表平均分數以單因子變異數分析(one-way Analysis of Variance, one-way ANOVA)分析結果(表7)： $F = 15.63$ 、 $p < .01$ ，達顯著水準，表示不同科學學習成就學生在CPS-SE量表的CPS自我效能有顯著差異。再將三組學生在三個分量表的平均分數以ANOVA分析的結果：共享理解向度 $F = 14.61$ 、 $p < .01$ ，行動解決向度 $F = 14.52$ 、 $p < .01$ ，團隊組織向度 $F = 14.45$ 、 $p < .01$ ，皆達顯著水準。表示不同科學學習成就學生在共享理解、行動解決和團隊組織三個向度皆有顯著差異。進行Scheffé事後檢驗，結果發現：在CPS總量表的平均分數高成就生顯著高於低成就生，中成就生顯著高於低成就生，高成就生與中成就生之間則無顯著差異。接著比較三組學生在三個向度的差異情形，結果顯示：共享理解、行動解決和團隊組織三個向度都是高成就生顯著高於低成就生，中成就生顯著高於低成就生，高成就生與中成就生之間則無顯著差異。

綜上可知，低成就生在CPS自我效能和共享理解、行動解決和團隊組織的三個構面，皆顯著低於中成就生和高成就生，高成就生與中成就生之間則無顯著差異。換言

之，科學學習成就會影響國中生CPS的自我效能。

伍、討論與結論

一、編製的CPS-SE量表信效度良好

本研究的CPS-SE量表共25題，包含共享理解、行動解決和團隊組織三個因素。經量表編製與效化，發現三個因素的累積解釋變異量達77.45%，證實CPS-SE量表有三個因素，每個因素與每個題項對應的因素負荷量高，且與總量表有高度相關，顯示CPS-SE量表具有穩定良好的因素結構。經量表驗證，評估三種假設的競爭模型，發現二階三因素模型的適配度符合各項指標，兼具外在與內在結構適配度。且呼應PISA 2015 CPS評量架構的三種核心能力：建立維持共享理解、採取適當行動解決問題和建立維持團隊組織(OECD, 2017a)，因此選擇二階三因素模型為最終模型，並證實所建構的理論模型和實際收集的資料適配良好，具有良好的建構效度。CPS-SE量表和三個分量表的Cronbach's α 皆高於.95，證實具有良好的信度。經複核效度分析也皆通過指標標準，證實本研究提出的CPS-SE量表具有良好的複核效度。

比較CPS-SE量表與國際ATC21S和PISA 2015的CPS評量，本量表缺乏CPS情境任務

表7：不同科學學習成就的國中生在CPS-SE量表平均分數的ANOVA分析($N = 274$)

構面	低成就生 ($n = 83$)		中成就生 ($n = 115$)		高成就生 ($n = 76$)		F值	p值	事後比較
	M	SD	M	SD	M	SD			
共享理解	2.39	0.87	2.79	0.67	3.00	0.63	14.61	< .01	高 > 低、中 > 低
行動解決	2.39	0.91	2.80	0.62	2.98	0.63	14.52	< .01	高 > 低、中 > 低
團隊組織	2.37	0.90	2.78	0.65	2.97	0.63	14.45	< .01	高 > 低、中 > 低
總量表	2.38	0.86	2.79	0.62	2.98	0.61	15.63	< .01	高 > 低、中 > 低

註：CPS-SE：合作問題解決自我效能(Collaborative Problem Solving Self-Efficacy)；ANOVA：變異數分析(Analysis of Variance)；高：高成就生；中：中成就生；低：低成就生。

與標準化試題。此兩種評量皆將CPS評量架構的理念轉化成概念，設計試題測試學生的反應。透過行為指標描述，制定評分規準，直接評測學生個人的CPS能力程度。皆利用學生於電腦介面與其他學生或電腦代理人之互動過程的外在表現，來推斷個人內在的合作、溝通、思考和認知能力等心智活動。然而這種人為設計的外在概念性CPS標準化評量，真的可以代表學生複雜而動態的CPS能力嗎(Fiore et al., 2017)？而且龐大繁瑣的命題設計，一般教師的專業能力是否能負擔？這些質疑，目前為止尚缺乏足夠的研究證據支持，有待更進一步探討。

本研究的CPS-SE量表與哈佛醫學院的CPS-AIM-E量表有許多相似之處，皆以線上填寫或手寫的簡易量表，主觀地自我評估個人的CPS信念和感受，皆不是直接評測CPS能力，而是由個人內在的CPS感知或自我效能，評估自己能克服困難完成任務能力的自信程度，間接去推斷外在的CPS能力。因為自我效能會影響個人的學習興趣、動機、自律、自尊和行為，可預測個人外在的行為或能力表現，所以自我效能與能力有明顯的相關性(Bandura, 2006; Miwa, 2005; Zimmerman, 2000)。然而，CPS-AIM-E量表採Likert七點制，由教育者填寫，並未調查學生的CPS感知。其CPS感知能力只是一個分量表，有4個題項，沒有提出具體的理論基礎，信度較不佳。本研究CPS-SE量表採Likert四點制，共25題，有良好的建構信效度與複核效度。

無論是藉由電腦標準化的CPS評量，以外在的行為表現來推估學生內在的CPS心智活動，或是藉由自陳量表，以個人內在的CPS自我效能信念或感知來推斷其外在CPS能力的評量，都無法真正評測出個人複雜動態的CPS能力，也無法設計出詳盡真實的

CPS標準化評量(Griffin & Care, 2015; OECD, 2017a)。因此只要善用評量方式符合學習目標和教師的專業能力，克服教育現場的限制，選擇適合的評量方法，均能幫助教師或學生理解其CPS的自我效能或能力。

二、CPS-SE量表能有效應用於評估不同科學學習成就學生的CPS自我效能差異

研究結果分析顯示，自然科學學習成就對國中生CPS自我效能的共享理解、行動解決和團隊組織等能力，具有良好的預測力，能評估不同科學學習成就學生CPS自我效能的差異。換言之，自然科學學習成就會影響國中生CPS的自我效能，中、高成就生的CPS自我效能較高；反之，低成就生的CPS自我效能較低。結果呼應Bandura (2006)和Zimmerman (2000)主張高自我效能的青少年會尋求問題解決的方法以完成目標，因而影響其學業成就，以及知識和技能的學習與表現。而低自我效能者，缺乏對自我能力與表現的信心，容易放棄而無法完成目標。支持Sorić與Palekčić (2009)和Ramdass與Zimmerman (2011)認為自我效能會影響學業成就，自我效能與學業成就、知識和技能的表現有關聯性。趙珮晴與余民寧(2012)研究結果相似，自我效能和學業成就有正相關。中、高成就生在CPS的自我效能優於低成就生的原因可能是，中、高成就生具有較多的成功經驗，可增強其自我效能和自尊感，較能自我肯定、自律學習，對自我能力也較有自信。即使面對困難任務，也能更努力去解決問題達成目標，因此有較高的CPS自我效能；反之，低成就生因缺乏成功經驗，容易降低自我效能和自尊感，缺乏自我肯定與自信，如果又缺乏自律學習，面對困難任務容易產生挫折感而放棄，所以CPS自我效能也較低。影響CPS能

力的因素除學業成就之外，包括性別、幸福感、生活滿意度、成就動機指數、受霸凌指數和自我導向的完美主義等(OECD, 2017b)。所以有關中、高成就生的CPS自我效能並無顯著差異的原因，需更多研究支持。

自然科學學業的中、高成就生在CPS-SE量表的共享理解、行動解決、團隊組織三個構面，皆高於低成就生，原因可能是當學生有較高自我效能同時也較有自信，自我評估有能力也願意與團隊成員建立維持共享理解、採取適當行動解決問題，以及建立維持團隊組織。即使遭遇困難，也願意面對並努力克服挑戰，解決問題達成任務；反之，自然科學學業的低成就生可能數學或有些學業表現也不佳，自尊感較低，較缺乏自信與自律學習，因而無法與成員建立維持共享理解。當遭遇困難任務時，容易懷疑自己的能力而產生挫折感，降低努力並放棄，缺乏自信評估有能力採取適當行動解決問題。於團隊中因為低科學學習成就而導致自我效能和自尊感較低，容易感到壓力與無力感，而缺乏自信、自尊與自律去建立維持團隊組織。因此，自然科學學習的低成就生在CPS-SE量表三構面皆顯著低於中、高成就生。此研究結果呼應PISA 2015 CPS的成果報告(*PISA 2015 Results: Collaborative Problem Solving*)：學生的CPS能力與PISA的三項核心學科(科學、閱讀和數學)的成就表現成正相關，學科表現較佳者在CPS的得分也較高。在每個參與國家／經濟體的三項核心學科表現與問卷調查，幾乎都是表現優勢的學生比表現劣勢的學生較重視關係，也較重視團隊合作，我國學生也是。當學生越重視關係則在CPS的表現就越好，也較重視團隊合作，較能維持對團隊的動力，較能根據彼此的角色行動，解決分歧和衝突，較能以有效的方法監控並解決問題

的進度(OECD, 2017b)。換言之，學生的CPS能力與自然科學的學習成就成正相關，科學學習成就優勢的學生比劣勢的學生在CPS的能力較佳，亦即在建立維持共享理解、採取適當行動解決問題、建立維持團隊組織等三個核心能力較佳。綜上所論，PISA 2015 CPS評量的結果支持本研究的發現：自然科學學習成就的低成就生在CPS-SE量表的表現低於中、高成就生，在共享理解、行動解決和團隊組織三構面也皆低於中、高成就生。

綜合上述，本研究編製的CPS-SE量表信效度良好，且能有效應用於評估不同自然科學學習成就學生的CPS自我效能差異，可作為中學CPS的課堂評量工具。本評量屬於一般領域(domain general)，且適用於自然科學領域學習。誠如Manitoba Education, Citizenship and Youth (2006)倡導的評量三大目的，透過本量表評量可促進學習(assessment for learning)，作為形成性評量工具，蒐集資料以瞭解學生的CPS能力的優缺點，診斷其可能遭遇的困難。本評量也是學習(assessment as learning)，提供學生自我評量、反思並監控自己的CPS能力並調整學習策略，培養其後設認知。然而CPS是複雜困難的技能，無法快速或自發地發展，教師可定期評量並監控學生的CPS後認知及能力表現，幫助學生理解並自信地達到CPS的期望內容。本量表也是學習的評量(assessment of learning)，評量學生在教學結束時是否達成CPS學習目標的總結性評量(Manitoba Education, Citizenship and Youth)。

雖然我國學生普遍缺乏CPS的實際經驗，然而十二年國民基本教育倡導核心素養、科學素養和科學探究能力(國家教育研究院，2018)，教師實施探究教學時若結合作學習，可培養學生自主行動以科學探究來解

決問題、能與人合作溝通互動和社會參與，也就是培養學生CPS能力的素養。因此，CPS-SE量表也可應用於學生的探究學習或合作學習的問題解決成效的評量。

三、建議

因為時間、人力與經費的限制，本研究為方便取樣，研究對象僅取樣自五個縣市的國中生，因此本量表的應用，須謹慎使用。根據研究結果，提出以下三點建議。

首先，有關教學的建議。教師可設計常規性且定義良好的問題與任務，引導並幫助低成就生解決問題完成任務，增加其精熟與成功經驗，並多給予鼓勵與讚美。同時提供機會，讓低成就生能與中、高成就生共同合作解決問題。使其觀摩他人的成功經驗，學習其自尊與自律，和正向積極主動面對困難且持續努力以克服困難，與人合作解決問題完成任務。此外，教師也可設計非常規且非定義良好的問題與任務，提供學生發揮批判思考與創意，引導學生於CPS歷程中，能將內在的CPS自我效能轉化成外在的CPS能力展現出來。也可結合跨領域／跨科探究課程，進行探究學習和合作學習，透過實做評量瞭解學生在真實CPS情境的表現及自我效能。學生可以異質性分組，能力較強的學生可以指導能力稍弱的學生，小組成員共同合作解決問題，合作完成學習單或專題報告等任務。教師藉由制訂評分規準分別評量團體和個人成績，以瞭解學生實際的CPS歷程與結果。

其次，關於資料蒐集與分析的建議。學

生填寫量表時，可透過放聲思考(think-aloud)促進反思。可增加質性訪談和個案研究等質性資料，理解學生在CPS的自我效能與能力間的直接或間接關聯性。深入理解低成就生與高成就生在CPS自我效能差異的原因，並參酌學生的CPS能力和其背景，規劃適性化的相關課程，以獲得最佳教學成效。CPS評量結果的統計分析可用創新建模的方法進行，以評量複雜而動態的CPS能力，並組成跨領域的CPS團隊，共同合作完成任務。

最後，未來研究方法的建議。若能參酌大型教育測驗，如PISA、國際數學與科學教育成就趨勢調查(Trends in International Mathematics and Science Study, TIMSS)和臺灣學生學習成就評量資料庫(Taiwan Assessment of Student Achievement, TASA)等，採二階段分層叢集抽樣設計(two-stage stratified cluster sample design)，第一階段以全國學生母群體學校分布進行分層隨機抽樣，第二階段再以校內學生進行叢集取樣。以階層線性模型(Hierarchical Linear Modeling, HLM)分析技術探討不同學校或地區之間CPS能力的差異，可增加結果的推論性和有效性。

誌謝

感謝國立臺灣師範大學生命科學系鄭湧涇教授與科學教育研究所林陳涌教授、國立彰化師範大學科學教育研究所郭重吉教授與簡頌沛教授，以及《科學教育學刊》的編輯團隊與審查教授，提供寶貴建議與協助，特致謝忱。

參考文獻

1. 李茂能(2019)。結構方程模式理論與實務：圖解AMOS取向。臺北市：五南。

- [Li, M.-N. (2019). *Principles and structural equation modeling with AMOS graphics*. Taipei, Taiwan: Wu-Nan.]
2. 吳毓瑩(1996)。量表奇偶點數的效度議題。調查研究，2，5-34。doi:10.7014/TCYC.199610.0005
[Wu, Y.-Y. (1996). Issues about the validity of scales with an odd or even number of points. *Survey Research*, 2, 5-34. doi:10.7014/TCYC.199610.0005]
 3. 邱皓政(2019)。量化研究與統計分析：SPSS與R資料分析範例解析(第六版)。臺北市：五南。
[Chiou, H.-J. (2019). *Quantitative research and statistical analysis: Analysis of SPSS and R data analysis examples* (6th ed.). Taipei, Taiwan: Wu-Nan.]
 4. 林碧芳、邱皓政(2008)。創意教學自我效能感量表之編製與相關研究。教育研究與發展期刊，4(1)，141-169。
[Lin, P.-F., & Chiou, H.-J. (2008). Construction and related study of the inventory of self-efficacy for creative teaching. *Journal of Educational Research and Development*, 4(1), 141-169.]
 5. 國立編譯館(2000)。教育大辭書(第三冊)。臺北市：文景。
[National Institute for Compilation and Translation. (2000). *Encyclopedic dictionary of education* (Vol. 3). Taipei, Taiwan: Win Join.]
 6. 國家教育研究院(2018年11月2日)。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。查詢日期：2020年5月20日，檢自https://www.k12ea.gov.tw/files/class_schema/%E8%AA%B2%E7%B6%B1/17-%E8%87%AA%E7%84%B6%E7%A7%91%E5%AD%B8/17-1/%E5%8D%81%E4%BA%8C%E5%B9%B4%E5%9C%8B%E6%B0%91%E5%9F%BA%E6%9C%AC%E6%95%99%E8%82%B2%E8%AA%B2%E7%A8%8B%E7%B6%B1%E8%A6%81%E5%9C%8B%E6%B0%91%E4%B8%AD%E5%B0%8F%E5%AD%B8%E6%9A%A8%E6%99%AE%E9%80%9A%E5%9E%8B%E9%AB%98%E7%B4%9A%E4%B8%AD%E7%AD%89%E5%AD%B8%E6%A0%A1%E2%94%80%E8%87%AA%E7%84%B6%E7%A7%91%E5%AD%B8%E9%A0%98%E5%9F%9F.pdf。
[National Academy for Educational Research. (2018, November 2). *Shiernian guomin jiben jiaoyu kecheng gangyao—Guomin zhongxiaoxue ji putongxing gaoji zhongdeng xuexiao: Ziran kexue lingyu*. Retrieved May 20, 2020, from https://www.k12ea.gov.tw/files/class_schema/%E8%AA%B2%E7%B6%B1/17-%E8%87%AA%E7%84%B6%E7%A7%91%E5%AD%B8/17-1/%E5%8D%81%E4%BA%8C%E5%B9%B4%E5%9C%8B%E6%B0%91%E5%9F%BA%E6%9C%AC%E6%95%99%E8%82%B2%E8%AA%B2%E7%A8%8B%E7%B6%B1%E8%A6%81%E5%9C%8B%E6%B0%91%E4%B8%AD%E5%B0%8F%E5%AD%B8%E6%9A%A8%E6%99%AE%E9%80%9A%E5%9E%8B%E9%AB%98%E7%B4%9A%E4%B8%AD%E7%AD%89%E5%AD%B8%E6%A0%A1%E2%94%80%E8%87%AA%E7%84%B6%E7%A7%91%E5%AD%B8%E9%A0%98%E5%9F%9F.pdf]

7. 彭淑玲(2019)。知覺教師回饋，個人成就目標，學業自我效能與無聊之關係：中介效果與條件間接化效果分析。*教育心理學報*，**51**(1)，83-108。doi:10.6251/BEP.201909_51(1).0004
[Peng, S.-L. (2019). Perceived teacher feedbacks, personal achievement goals, and academic self-efficacy on boredom: The mediation effect and conditional indirect effect. *Bulletin of Educational Psychology*, *51*(1), 83-108. doi:10.6251/BEP.201909_51(1).0004]
8. 趙珮晴、余民寧(2012)。自律學習策略與自我效能、學習興趣、學業成就的相關研究。*教育研究集刊*，**58**(3)，1-32。doi:10.3966/10288708201209001
[Chao, P.-C., & Yu, M.-N. (2012). An empirical study of the relationships among self-regulated learning strategies, learning interests, and academic achievements. *Bulletin of Educational Research*, *58*(3), 1-32. doi:10.3966/10288708201209001]
9. Aguado, D., Rico, R., Sánchez-Manzanares, M., & Salas, E. (2014). Teamwork Competency Test (TWCT): A step forward on measuring teamwork competencies. *Group Dynamics: Theory, Research, and Practice*, *18*(2), 101-121. doi:10.1037/a0036098
10. Arieli-Attali, M. (2019). Book review: Innovative assessment of collaboration (Methodology of educational measurement and assessment) by von Davier, A. A., Zhu, M., & Kyllonen, P. C. (Eds.). *Applied Psychological Measurement*, *43*(1), 89-91. doi:10.1177/0146621618768292
11. Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*, 74-94. doi:10.1007/BF02723327
12. Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
13. Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman.
14. Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1-26. doi:10.1146/annurev.psych.52.1.1
15. Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 307-337). Greenwich, CT: Information Age.
16. Bandura, A., Caprara, G. V., Barbaranelli, C., Gerbino, M., & Pastorelli, C. (2003). Role of affective self-regulatory efficacy in diverse spheres of psychosocial functioning. *Child Development*, *74*(3), 769-782. doi:10.1111/1467-8624.00567
17. Bentler, P. M., & Yuan, K.-H. (1999). Structural equation modeling with small samples: Test statistics. *Multivariate Behavioral Research*, *34*(2), 181-197. doi:10.1207/S15327906Mb340203
18. Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York, NY: Routledge. doi:10.4324/9780203805534
19. Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining competencies and establishing team training requirements. In R. A. Guzzo & E. Salas (Eds.), *Team*

- effectiveness and decision making in organizations* (pp. 333-380). San Francisco, CA: Jossey-Bass.
20. Care, E., Griffin, P., & Wilson, M. (Eds.). (2018). *Assessment and teaching of 21st century skills: Research and applications*. Cham, Switzerland: Springer.
 21. Chiu, M. M. (2000). Group problem-solving processes: Social interactions and individual actions. *Journal for the Theory of Social Behaviour*, 30(1), 26-49. doi:10.1111/1468-5914.00118
 22. Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. J. (2000). Measuring team knowledge. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 42(1), 151-173. doi:10.1518/001872000779656561
 23. De Boeck, P., & Scalise, K. (2019). Collaborative problem solving: Processing actions, time, and performance. *Frontiers in Psychology*, 10. Retrieved May 20, 2020, from <https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01280/full>
 24. Fiore, S. M., Graesser, A., Greiff, S., Griffin, P., Gong, B., Kyllonen, P., et al. (2017). *Collaborative problem solving: Considerations for the National Assessment of Educational Progress*. Alexandria, VA: National Center for Education Statistics.
 25. Griffin, P., & Care, E. (2015). ATC21S method. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach* (pp. 3-33). Dordrecht, The Netherlands: Springer.
 26. Griffin, P., Care, E., & McGaw, B. (2012). The changing role of education and schools. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 1-16). Dordrecht, The Netherlands: Springer. doi:10.1007/978-94-007-2324-5
 27. Hair, J., Black, W., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
 28. Hammack, R., & Ivey, T. (2017). Examining elementary teachers' engineering self-efficacy and engineering teacher efficacy. *School Science and Mathematics*, 117(1-2), 52-62. doi:10.1111/ssm.12205
 29. Harding, S.-M. E., Griffin, P. E., Awwal, N., Alom, B. M. M., & Scoular, C. (2017). Measuring collaborative problem solving using mathematics-based tasks. *AERA Open*, 3(3), 1-19. doi:10.1177/2332858417728046
 30. Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach* (pp. 37-56). Dordrecht, The Netherlands: Springer. doi:10.1007/978-94-017-9395-7_2
 31. Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, 83(3), 693-710. doi:10.1037/0022-3514.83.3.693

32. Kyllonen, P. C. (2012). *Measurement of 21st century skills within the common core state standards*. Paper presented at the Invitational Research Symposium on Technology Enhanced Assessments. Washington, DC.
33. Manitoba Education, Citizenship and Youth. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Winnipeg, Canada: Author.
34. Miwa, M. (2005). Bandura's social cognition. In K. E. Fisher, S. Erdelez, & L. E. F. McKechie (Eds.), *Theories of information behavior* (pp. 54-57). Medford, NJ: Information Today.
35. Oliveri, M. E., Lawless, R., & Molloy, H. (2017). A literature review on collaborative problem solving for college and workforce readiness. *ETS Research Report Series*, 2017(1), 1-27. doi:10.1002/ets2.12133
36. Organization for Economic Co-operation and Development. (2013). *PISA 2015 draft collaborative problem solving framework*. Retrieved May 20, 2019, from https://www.oecd.org/callsfortenders/Annex%20ID_PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf
37. Organization for Economic Co-operation and Development. (2017a). *PISA 2015 collaborative problem solving framework*. Retrieved May 20, 2019, from <https://www.oecd.org/pisa/pisa-products/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>
38. Organization for Economic Co-operation and Development. (2017b). *PISA 2015 results: Collaborative problem solving* (Vol. 5). Paris, France: Author. doi:10.1787/19963777
39. Ramdass, D., & Zimmerman, B. J. (2011). Developing self-regulation skills: The important role of homework. *Journal of Advanced Academics*, 22(2), 194-218. doi:10.1177/1932202X1102200202
40. Robertson Evia, J., Peterman, K., Cloyd, E., & Besley, J. (2018). Validating a scale that measures scientists' self-efficacy for public engagement with science. *International Journal of Science Education, Part B*, 8(1), 40-52. doi:10.1080/21548455.2017.1377852
41. Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), *Self-efficacy, adaptation, and adjustment: Theory, research, and application* (pp. 281-303). New York, NY: Springer. doi:10.1007/978-1-4419-6868-5
42. Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly*, 23(1), 7-25. doi:10.1080/10573560600837578
43. Sorić, I., & Palekčić, M. (2009). The role of students' interests in self-regulated learning: The relationship between students' interests, learning strategies and causal attributions. *European Journal of Psychology of Education*, 24(4), 545-565. Retrieved May 20, 2020, from <https://link.springer.com/article/10.1007/BF03178767>

44. Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. *Journal of Management*, 20(2), 503-530. doi:10.1177/014920639402000210
45. Ullman, J. B. (2001). Structural equation modeling. In B. G. Tabachnick & L. S. Fidell (Eds.), *Using multivariate statistics* (4th ed., pp. 653-771). Needham Heights, MA: Allyn and Bacon.
46. von Davier, A. A., Zhu, M., & Kyllonen, P. C. (Eds.). (2017). *Innovative assessment of collaboration*. Cham, Switzerland: Springer.
47. Wang, L., & Pollastri, A. R. (2019). *User's guide to the collaborative problem solving adherence & impact measures (CPS-AIMs)*. Retrieved May 20, 2020, from <http://thinkkids.org/wp-content/uploads/2015/04/Users-Guide-to-Collaborative-Problem-Solving-Aims-Impact-Measures-CPS-AIMs.pdf>
48. Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82-91. doi:10.1006/ceps.1999.1016

Development of the Collaborative Problem Solving Self-Efficacy Scale With the Relationship on Science Learning Achievement

Shu-Ching Wang^{1,*}, Tsung-I Yeh² and Kuo-Hua Wang¹

¹Graduate Institute of Science Education, National Changhua University of Education

²Department of Mathematics, National Tsing Hua University

Abstract

The purpose of this study was to develop the Collaborative Problem Solving Self-Efficacy (CPS-SE) scale as a Collaborative Problem Solving (CPS) classroom assessment tool and to use it to examine the relationship between students' CPS-SE and their science learning achievement. Convenience sampling was used to select 1,045 students from coeducational secondary schools in five Taiwan counties. The dimensions and items of the scale were constructed based on the 2015 CPS assessment framework of the Programme for International Student Assessment (PISA). The scale analysis process comprised of five stages: (1) item analysis, (2) exploratory factor analysis, (3) reliability analysis, (4) confirmatory factor analysis, and (5) cross-validation. Three noteworthy results were found. First, there were three factors extracted through factor analysis during the establishment of construct validity: sharing understanding, taking action on problems, and team organization. Second, the Cronbach's α for each subscale exceeded .95. Applied Structural Equation Model (SEM) the best fit and most parsimonious model was chosen based on model comparisons. Statistical analysis of the model showed goodness-of-fit to the data, which indicated that CPS-SE scale had high consistency and reliability. Third, students who were high and medium academic achievers on science learning scored significantly higher than those who were low academic achievers on three subscales: sharing understanding, taking action on problems, and team organization. There was no significant difference between high-achieving and medium-achieving students on the three subscales.

Key words: Self-Efficacy, Collaborative Problem Solving (CPS) Assessment, Science Learning Achievement, Scale Construction, Confirmatory Factor Analysis

* Corresponding author: Shu-Ching Wang, wsching123@gmail.com