

高中生情境式 STEM 跨領域問題 解決能力評量工具之發展 ——以南方澳跨港大橋情境為例

林坤誼 謝雨蓁 許瑛珺 陳柏熹 吳心楷 楊凱琳

現階段推動STEM教育的主要問題之一是缺乏適切的評量工具，以致於難以評估學生在STEM課程中的學習表現。本研究主要以南方澳跨港大橋情境為例，開發一個可用於評估高中生STEM跨領域問題解決能力的評量工具，其構面包含情境解構、類比推理、預測分析、量化思維、反思能力等5項能力。本研究的施測對象為844位高中生，採用Rasch二元計分模式進行考驗，以提出一份能夠有效評量STEM跨領域問題解決能力的評量工具。整體而言，高中生在STEM跨領域問題解決能力的整體表現有待改善，尤其是在預測分析和反思能力這兩項能力需要更進一步的學習和努力。依據前述研究結果，本研究進一步討論可能的研究限制，並針對未來相關研究所應重視的議題提出討論，並且針對現場教師在課堂中引用與實施提供相關建議，以供未來研究參考。

關鍵詞：跨領域問題解決能力、STEM 教育、評量工具、南方澳跨港大橋、高中生

收件：2023年3月30日；修改：2023年6月5日；接受：2023年6月13日

林坤誼，國立臺灣師範大學科技應用與人力資源發展學系特聘教授
謝雨蓁，國立臺灣師範大學科技應用與人力資源發展學系博士候選人，E-mail: umi.sic168@gmail.com
許瑛珺，國立臺灣師範大學科學教育研究所講座教授
陳柏熹，國立臺灣師範大學教育心理與輔導學系教授
吳心楷，國立臺灣師範大學科學教育研究所講座教授
楊凱琳，國立臺灣師範大學數學系教授

Development of a Contextualized STEM Interdisciplinary Problem-Solving Skill Tool for Senior High School Students: Utilizing the Nanfang'ao Cross-Harbor Bridge as an Exemplar

Kuen-Yi Lin Yu-Jen Sie Ying-Shao Hsu
Po-His Chen Hsin-Kai Wu Kai-Lin Yang

A challenge in advancing STEM education is the absence of appropriate assessment tools for evaluating the learning performance of students in STEM courses. This study aimed to develop an assessment tool for the STEM interdisciplinary problem-solving skills of high school students. The Nanfang'ao Bridge serves as the case study for this tool. The proposed research tool consists of five dimensions: contextualization, analogical reasoning, predicting, quantitative thinking, and reflective ability. The study surveyed 844 high school students and employed the Rasch dichotomous scoring model to develop an effective assessment tool for the STEM interdisciplinary problem-solving competencies of these students. The findings suggest that high school students need to improve their competencies, especially those in predictive analysis and reflective ability. This study also discusses potential limitations and highlights critical considerations for future investigations in this field. Furthermore, it provides recommendations for classroom teachers regarding the implementation and integration of the assessment tool, and the results serve as a reference for future research.

Keywords: interdisciplinary problem-solving skill, STEM education, assessment tool, Nanfang'ao bridge, senior-high-school student

Received: March 30, 2023; Revised: June 5, 2023; Accepted: June 13, 2023

Kuen-Yi Lin, Distinguished Professor, Department of Technology Application and Human Resource Development, National Taiwan Normal University.

Yu-Jen Sie, Ph. D. Candidate, Department of Technology Application and Human Resource Development, National Taiwan Normal University, E-mail: umi.sie168@gmail.com

Ying-Shao Hsu, Chair Professor, Graduate Institute of Science Education, National Taiwan Normal University.

Po-His Chen, Professor, Department of Educational Psychology and Counseling, National Taiwan Normal University.

Hsin-Kai Wu, Chair Professor, Graduate Institute of Science Education, National Taiwan Normal University.

Kai-Lin Yang, Professor, Department of Mathematics, National Taiwan Normal University.

壹、前言

科學——科技——工程——數學教育（Science, Technology, Engineering, and Mathematics Education, 以下簡稱 STEM 教育）的評量是推動 STEM 教育的關鍵課題，現階段推動 STEM 教育的主要問題之一是缺乏適切的評量工具，以致於難以評量學生在 STEM 課程中的學習表現（Gao et al., 2020; Kayan-Fadlelmula et al., 2022; Margot & Kettler, 2019）。然而，缺乏 STEM 評量工具的主要原因之一，在於推動 STEM 教育的目標不夠明確，且無法凸顯 STEM 教育的價值。Gao 等人（2020）認為認知領域的 STEM 教育學習目標可從 2 個面向切入，一個面向是以單一領域為主、其他領域為輔，協助學生學習 STEM 領域中的單一或跨領域的內容知識，例如，使用工程設計以培養學生數學或科學知識等（Becker & Park, 2011; Guzey et al., 2014; Sanders, 2009, 2012）。而另外一個面向便是培養學生應用跨領域知識的能力，例如，許多研究會使用複雜問題解決的情境，以培養學生應用跨領域知識的能力（Hurley, 2001; Moore et al., 2014）。因此，針對不同教育目標在發展 STEM 評量工具上會有不同，才能適當地檢視多元的 STEM 教學設計下學生認知能力的發展情形。

根據 STEM 教育相關研究的文獻回顧分析指出，過去的 STEM 的相關研究已經普遍認同 STEM 的跨學科特性，專注於提升學生跨學科的理解、以及應用 STEM 跨領域知識的能力等，以作為 STEM 教育目標。然而，過去的 STEM 教育相關研究卻大多僅針對單一學科的知識面向、單一學科的情意面向以及超學科的情意面向等方面進行評量。儘管過去研究雖強調跨領域的學習，但很少有研究採用適當的 STEM 評量工具，以評估學生對跨學科學習目標的達成情況，形成課程、教學、與評量的不一致的狀況，也成為 STEM 教育發展的一項隱憂。因此，未來在推動 STEM 教育的同時，如何解決現階段缺乏適切的評量工具以評量學生在跨領域的學習表現，應是推動 STEM 教育時面臨且必須克服的關鍵問題（Herro et al., 2017）。

爲了更全面地評估學生在 STEM 教育中的成效，STEM 教育的研究需要致力於開發出適合的評估工具，以更適切的評量學生在跨領域方面的表現。舉例來說，現階段已有相關研究從 21 世紀關鍵能力的角度切入，透過發展解決真實問題的評量工具，以評量學生應用跨領域知識的能力（Amalina & Vidákovich, 2022; Priemer et al., 2020）。而爲了發展不同類型的 STEM 問題解決評量工具，許多研究者從不同的角度切入，有些學者著重在開發一個通用的認識論框架，透過從事科學過程所應用的方法與技能來評量學生的 STEM 問題解決能力（Kind & Osborne, 2017）；此外，也有學者著重在從科學、科技、工程與數學的問題解決能力切入，並研提共通的 STEM 跨領域問題解決能力，以作爲發展評量工具的基礎（Lin et al., 2022）。無論從何種角度切入，跨領域 STEM 問題解決評量工具仍有待更進一步深入的探討與評估（Amalina & Vidákovich, 2022; Chusinkunawut et al., 2018）。

總上所述，本研究基於前述 STEM 教育缺乏適切評量工具的問題，將研究目的聚焦在開發一個可評測學生針對真實情境所發展之跨領域問題解決能力的評量工具。依據此一研究目的，本研究的研究問題：（一）以真實問題爲情境的跨領域 STEM 問題解決能力評量工具之內涵爲何？（二）此一跨領域 STEM 問題解決能力評量工具的信、效度爲何？（三）高中生在此一跨領域 STEM 問題解決能力評量的實際表現情形爲何？

貳、文獻探討

一、STEM 教育目標

隨著 STEM 教育愈來愈受到重視，許多研究機構和學者從不同的角度探討 STEM 教育的意義和價值，或是針對未來 STEM 教育的發展契機進行深度的討論（Tanenbaum, 2016）。例如，美國在推動 STEM 教育的主要目標包括培養 STEM 領域的專業人才、強化 STEM 領域人才

的能力、培養所有學生的 STEM 素養，以及讓更多女性或弱勢族群能夠投入 STEM 領域（National Research Council, 2011）。然而，不同領域的專家學者對於 STEM 教育的目標存在差異，這可能導致推動 STEM 教育時失去方向，進而在推動過程中遭遇許多困難（Holmlund et al., 2018; Thibaut et al., 2018）。

若從認知面向來看，STEM 教育的基本學習目標之一，是協助學生學習 STEM 領域中單一或更多領域的內容知識。例如，STEM 教育的常見方法是使用工程設計培養學生數學或科學知識（Becker & Park, 2011; Guzey et al., 2014; Sanders, 2009, 2012），這個方法以 STEM 的單一領域為主，其他領域為輔（Gao et al., 2020）。因此，許多學者認為培養學生的 STEM 素養應是推動 STEM 教育的重要學習目標。然而，不同學者對於培養 STEM 素養有不同的看法，例如 Tang 與 Williams（2019）便指出 STEM 素養並非僅是將原本的科學素養、科技素養、工程素養以及數學素養結合起來，而是應該清楚地定義出透過 STEM 教育所能夠培養出的真正關鍵技能與實務能力（skills and practices）。有鑑於此，從認知角度來反思 STEM 教育的學習目標，除了關切 STEM 領域的知識學習之外，更值得關注如何培養出學生真正的關鍵技能與實務能力。

若從情意面向來看，美國在推動 STEM 教育的初期，其國內 STEM 領域的高階研究人才多為外移人口，而並非源自於國內的人才。為提高國人從事高階 STEM 領域的工作，美國透過 STEM 教育的推廣，旨於培養更多對 STEM 領域有興趣者才投入相關的工作中，培育出更多優秀的 STEM 創新人才（Raju & Clayson, 2010）。因此，許多學者紛紛針對培養學生對 STEM 領域的正向態度以及促進 STEM 職涯興趣進行探究，有些研究著重在發展 STEM 職涯興趣量表，另一些著重在探討參與社團、營隊等活動對 STEM 興趣的助益，還有一些則聚焦在探討不同 STEM 實作活動對 STEM 態度的助益（English, 2017; Ho et al., 2020; Lin et al., 2020; Lin et al., 2021; Wahono et al., 2020）。Martín-Páez 等人（2019）在其後設分析的研究中也指出，過去的相關研究在情意方面證

實以下幾個重點，包含 STEM 教育能夠鼓勵學習者獲得 STEM 學位、提升他們對 STEM 學科的認同感、增加他們對 STEM 學科的興趣與好奇心，以及培養學習者對 STEM 學科的正向態度。因此，在情意方面的 STEM 教育著重吸引更多對 STEM 領域有興趣的人才投入相關工作。

若從能力面向來看，STEM 教育十分強調學習者在不同情境（如能源、環境等）下進行跨領域問題解決，強調 STEM 教育應著重於培育學習者在真實情境下發展出跨領域問題解決能力。因此，許多 STEM 教育的研究學者指出，STEM 教育應著重在培養學生的 21 世紀關鍵能力（Lee et al., 2019; Lynch et al., 2018），而 Lin 等人（2022）分析不同學者對於透過 STEM 教育培養 21 世紀關鍵能力方面時，也歸納出 STEM 教育在此一方面的目標至少應包含以下幾項：（一）終身學習能力；（二）領導能力；（三）創造力；（四）適應力；（五）創業力；（六）解決問題能力；（七）批判性思考能力；（八）團隊合作問題解決能力。此外，許多教育家指出以往的教學將知識與能力分離，因此如何評量學生整合知識與能力的學習成果，便成為未來教育改革的重要課題（National Research Council, 2014）。

聯合國教育、科學及文化組織曾將能力定義為個人或群體為了實現利益的發展技能（Marope et al., 2017），如此一來，可以得知能力為一項動態的概念，是為了在複雜的環境中能夠合理運用知識、技能、情意和技術進而順利執行任務或改善生活。其中，許多 STEM 教育使用複雜的真實情境，讓學生進行問題解決，即是透過複雜的非結構問題，以培養與引導學生跨領域問題解決的能力（Hurley, 2001; Moore et al., 2014）。因此，未來需要有更多研究投入探討如何培養學生跨領域的問題解決能力，以呼應 STEM 教育目標所強調培育能夠在複雜真實情境下進行問題解決的創新人才。總上所述，無論從認知、情意或者能力的面向來檢視 STEM 教育目標，這些都是未來發展 STEM 評量工具所應該關切的重點。尤其是在 STEM 跨領域問題解決能力方面，更值得思考如何發展適切的評量工具，以剖析學習者在此 5 項關鍵轉化能力的表現情形。

二、STEM 評量工具之發展現況

Nadelson 等人（2013）的研究指出，有超過 40% 的老師認為缺少 STEM 課程的評量工具，因此若要評量學生在 STEM 課程中的學習表現便十分困難，也無法掌握學生在 STEM 學習歷程中對於不同學科概念的理解程度（Asghar et al., 2012; Dare et al., 2014）。Harwell 等人（2015）也認為現有的 STEM 教育認知評量側重於單一學科的知識，且很少關注知識的應用，因此如何衡量學習者在 STEM 課程中對內容知識的理解、STEM 學科之間的聯繫關係、以及 STEM 學科知識的實踐等，都是未來重要的課題。而許多老師認為缺乏有品質的 STEM 評量工具、授課時間以及 STEM 學科知識，是推動 STEM 教育的重大挑戰和障礙（Margot & Kettler, 2019）。然而，STEM 評量的設計與發展包含許多重要的工作，宜針對評量工具的設計與發展等環節進行詳細的規劃，例如，針對評量目的、形式等方面進行探討，並轉化為可推廣、可採用的形式，以供 STEM 教育的現場人員使用，方能更有助於未來 STEM 教育的推動。

首先，針對評量目的來看（如知識、技能、實務以及情意等），STEM 評量工具必須能夠依據評量目的設計評量內容，藉此確實檢視學生在跨領域方面的學習表現，以及精確地檢視推動 STEM 教育的成果（Gao et al., 2020; Kayan-Fadlelmula et al., 2022）。舉例來說，近年來有許多學者投入 STEM 職涯興趣量表的發展（Cantrell & Ewing-Taylor, 2009; Holmes et al., 2017; Kier et al., 2014; Sadler et al., 2012; Tuijl & Molen, 2015; Tyler-Wood et al., 2010），顯示如何協助學生試探 STEM 職涯興趣是一個十分值得探討的重要課題。Tyler-Wood 等人（2010）便曾發展 STEM 語義量表（STEM Semantics Scale）和 STEM 職業興趣問卷（Career Interest Questionnaire）2 項工具，其中 STEM 語義量表採李克特 7 點量表，透過 5 組對立的形容詞，如「迷人（fascinating）—平凡（mundane）」等，讓受試者選擇最符合其對科學、科技、工程、數學

及 STEM 職業看法；而 STEM 職業興趣問卷則採李克特 5 點量表，內容著重在調查受試者對科學職業、科技職業、工程職業、數學職業的興趣。Tyler-Wood 等人（2010）的工具看起來雖然完整，但是 Kier 等人（2014）指出 Tyler-Wood 等人所發展的 STEM 語義量表和 STEM 職業興趣問卷缺乏理論基礎。因此，Kier 等人（2014）結合社會認知職涯理論（social cognitive career theory），發展 STEM 職業興趣問卷（STEM-CIS），該問卷著重調查受試者對科學、科技、工程和數學職業的自我效能（self-efficacy）、預期結果（outcome expectations）、目標（goals）、興趣（interests）、情境支持（contextual supports）和個人意向（personal disposition）等 6 個面向的內涵。儘管已有相關的 STEM 職涯興趣量表工具，但如何評估學生對於未來新興的 STEM 職業興趣，仍需要進一步努力。

此外，若從評量形式的角度來看，傳統評量包含選擇題、簡答題、申論題等，但隨著資訊科技的發展，對於 STEM 跨領域學習的評估，便需要透過更多元的方式，以評估學生的多元學習成果，包括學科知識、科學方法、工程設計、技術應用、創造力和解決問題能力等方面。Wu 等人（2015）在探討評量項目的呈現方式時發現，採動態或靜態方式對於學生作答情形有一定程度的影響，呈現方式通常有 2 種，一種是動態呈現，較頻繁被使用的是動畫或互動元素，另一種是靜態呈現，使用固定的圖像或文字。研究發現較年輕的學生可能更容易理解和處理靜態的題目呈現方式，從靜態的圖像或文字中提取信息，而無需應對動態顯示所帶來的額外認知負擔，相較之下，年長的學生可能具備更多的認知資源，熟悉並能夠處理複雜的動態顯示，並從中提取有用的資訊，成功解決動態顯示的評估任務。因此，在設計評估任務時，應考慮到學生的年級和其認知程度的不同選擇不同的評量形式。因此，STEM 學習的評量應該是全面的、多元的、動態的和反思性的，同時，教師和教育工作者需要建立一個具包容性的評量系統，以實現學生在不同環境下綜合應用知識、解決問題和創造方案的能力（Fang & Hsu, 2019）。

除了前述從評量目的、形式等不同角度進行 STEM 評量工具的現況之探討外，Priemer 等人（2020）針對 STEM 教育提出一個評量問題解決能力的架構，作為檢視學生在問題解決能力之重要參考依據。此外，Kind 與 Osborne（2017）為了評量學生的問題解決能力，開發了一個 STEM 領域問題解決過程的理論架構，藉此從跨學科的觀點來思考如何培養學生的問題解決能力，並從問題解決的過程更細膩地檢視學生在過程中的實際表現。Amalina 與 Vidákovich（2022）也同樣指出許多研究者著重於發展不同的 STEM 問題解決評量工具，但仍然有能夠努力的空間，例如，如何從單一學科的角度轉換為跨學科的角度，以開發更適切的 STEM 跨領域問題解決評量工具。因此，透過前述 STEM 評量工具的相關研究現況來看，如何著重在發展 STEM 跨領域問題解決能力的評量工具，以確實掌握學生的跨領域問題解決能力，便應該是現階段 STEM 教育評量值得探究的重點。

三、STEM 跨領域問題解決能力之評量

依據前述的分析與探討，雖然近年來有關跨領域的 STEM 學習評量已經有重要的進展，但仍有很長的路要走，尤其是如何聚焦在提出 STEM 跨領域問題解決能力之評量。Gao 等人（2020，頁 12）針對 STEM 學習評量的建議中也指出：「STEM 學習過程和實務十分多元與複雜，尤其是不同學科在過程和實務中所可能呈現的不同特徵」。在前述有關跨領域問題解決能力評量工具之探討中，也明確指出如何從跨學科的觀點出發，思考評估學生在跨領域問題解決能力過程中表現的方法，是未來發展 STEM 評量工具所必須突破及最重要的課題（Amalina & Vidákovich, 2022; Kind & Osborne, 2017）。

針對此一方面的挑戰，多位學者也提出幾項重要能力，以落實問題解決能力的訓練，可以藉由 Lin 等人（2022）從科學、科技、工程與數學等不同的角度剖析不同領域的問題解決歷程，並提出 STEM 跨領域問題解決能力所應具備的 5 項關鍵轉化能力，歸納如下：（一）情境解構：

在面對日常生活的問題解決情境時，如何解構情境以找出問題解決的關鍵要素，是非常重要的核心能力之一。能夠妥善地解構情境，才能夠更適切地界定問題與分析問題（Kirschner et al., 2010）；（二）預測分析：在解決問題的過程中，經常會需要提出不同解決問題的方案，如何妥善地應用證據或者資料評估這些不同的方案，以預測分析不同方案可能遭遇的困難與預期效益（Olewnik et al., 2020）；（三）類比推理：在面對日常生活中的不同問題時，我們通常需要仰賴先備知識和經驗以構思如何解決問題。從過去經驗中，找到可類比的思維或是模型，作為解決問題時詮釋資料或是推理出結論的參考依據。因此，學習者若能夠具備類比推理的能力，那麼就能夠適切地將先備知識和經驗連結至其所面臨的問題情境中（Songer et al., 2009）；（四）量化思維：在面對日常生活的問題情境時，學習者經常因為缺乏量化思維，而導致其無法適切地權衡所提出的解決方案，繼而無法提出具體可行的解決方案或成品。因此，學習者如何妥善處理資料、觀察資料，應用數學運算模式和運算思維等進行量化的推理與論證，便是 STEM 跨領域問題解決過程中重要的轉化能力之一（Quinnell et al., 2013）；（五）反思能力：在日常生活的問題解決情境中，通常沒有標準答案，因此學習者必須努力尋找最佳化的問題解決方案。換句話說，當學習者所提出的問題解決方案經過測試後，便需要運用反思能力進行檢討和改進，以期精進自身的問題解決能力並達到最佳化解決方案的目標（Shanta & Wells, 2020）。

Lin 等人（2022）指出的 STEM 跨領域問題解決之轉化能力，能夠從跨學科的觀點評估學生在跨領域問題解決過程中的情境解構、預測分析、類比推理、量化思維以及反思能力之表現。因此，本研究主要以南方澳跨港大橋的情境為例，依據前述 5 項 STEM 跨領域問題解決能力的定義，設計情境式的 STEM 跨領域問題解決評量試題，以評估學習者在 STEM 跨領域問題解決能力之表現。

參、研究方法

一、研究對象

爲了分析高中生在 STEM 跨領域問題解決能力方面的表現情形，本研究於臺灣北、中、南、東區之普通高中學生中進行抽樣，最後共有 13 所普通高中，共 42 班參加，總共抽樣 1,435 名學生，回收樣本 1,253 名（回收率 87%）。經刪除資料遺漏者、重複作答者、以及連續作答者後，最終有效樣本爲 844 名（有效回收率 67%），其中男性 419 人、女性 425 人；1 年級 302 人、2 年級 372 人、3 年級 170 人（詳見表 1）。

二、研究工具

（一）情境式 STEM 跨領域問題解決能力評量試題設計理念

本研究主要根據 Lin 等人（2022）所指出的 STEM 跨領域問題解決之 5 項轉化能力，開發情境式 STEM 跨領域問題解決能力評量試題。依據此 5 項轉化能力，本研究設計情境式的線上評量題組，以協助教師評估高中學生在此 5 項轉化能力的表現。以南方澳跨港大橋爲例，本研究設計之情境式線上評量題組提供相關的情境圖片與說明（如圖 1），以利學生了解實際情況。

表 1 性別與年級交叉列表

性別 \ 年級	1 年級	2 年級	3 年級	總計
男	114	220	85	419
女	188	152	85	425
總計	302	372	170	844



重建宜蘭南方澳跨港大橋方案

橋梁除了交通運輸外，亦兼具美觀藝術的功能，橋梁工程一方面可以縮短兩地交通時間，建立四通八達的運輸網路；另一方面具美觀性的橋梁也能成為地方上的代表指標。1998年興建完成的南方澳大橋，是國內獨一單鋼拱橋，更是亞洲第一座雙叉式單拱橋。橋長140公尺，橋寬15公尺，跨漁港航道，以解決大型漁船無法通行之困境，使得南方澳漁港得以長遠發展，並作為豆腐岩海岬及東側新建碼頭對外連絡之捷徑。橋柱上的魚蝦蟹造型，畫龍點睛道出南方澳的海洋意象，觀景平台可讓遊客欣賞漁船進出之景象，是南方澳觀光漁業發展的重要里程碑。

觀看試卷

觀看成績

圖 1 重建南方澳跨港大橋情境

資料來源：取自情境式 STEM 問題解決轉化能力評量系統
 (<https://stem.tahrd.ntnu.edu.tw/zh/>)。

針對此一情境，本研究依據 Lin 等人（2022）的 5 項轉化能力設計相關試題，藉此評估學生在各項轉化能力方面的表現。舉例來說，在情境解構方面，系統針對同一事件提供不同的新聞報導資訊（含新聞影片與新聞報導文字），呈現不同角色人物對該事件的剖析觀點，例如新北市土木技師公會理事長、負責南方澳大橋設計案的公司、以及基隆港務分公司主任秘書等。透過類似的多元設計，可確實掌握學生在 STEM 跨領域問題解決能力的轉化能力之表現。此外，當學生完成相關測驗之後，系統會提供學生在 5 項轉化能力的測驗分數，並同時透過雷達圖呈現實際測驗結果，以利學生掌握自己可能不足與有待改善之處。以圖 2 為例，該位學生在量化思考的表現便有待進一步的強化。

（二）試題內容架構

本測驗以宜蘭南方澳跨港大橋重建案作為試題情境，建立 STEM 問題解決能力架構，包含（情境解構、類比推理、預測分析、量化思維與



圖 2 STEM 跨領域問題解決能力測驗結果

資料來源：取自情境式 STEM 問題解決轉化能力評量系統
(<https://stem.tahrd.ntnu.edu.tw/zh/>)。

反思能力 5 項能力)。全部題目包含 14 題選擇題，計分方式為答對給 1 分，答錯 0 分。

本題組情境設定主軸為南方澳大橋吊索斷裂事故為起始情境，此一橋梁形式在世界各地非常普遍，學生先要從情境敘述、新聞影片、媒體報導等資訊，推論橋梁最有可能斷裂的原因，而後再以大橋重建作為主要任務，藉由評量系統所提供的參考資料，在限制的條件下擬定重建大橋方案以解決問題。為了能夠妥善的測量學生在 STEM 跨領域問題解決能力方面的表現，本研究主要依據 Lin 等人 (2022) 所提出的 5 項重要轉化能力進行試題設計，分別說明如下。

1. 情境解構

此能力主要用於了解與解決複雜的問題，情境解構的組成可以包括能夠明確地描述問題或情境，以便更好理解問題，以及能夠確定解決問題的目標，此目標可以是具體的、也可以是可衡量的，同時是符合情境需求的。因此本題組在情境解構能力共擬定 3 題，請學生個別依據不同專家所述，歸納出造成南方澳大橋坍塌的可能因素為何，並從中選出最關鍵影響之因素，如圖 3 所示。

一座好的橋梁，除了美觀性和需求性考量以外，更重要的是建造出安全、穩固的橋梁。南方澳大橋於2019年10月1日傳出不幸倒塌的消息，而斷裂的主要原因眾說紛紜。以下提供三段新聞內容，請參考新聞中的敘述，並回答下列問題。

敘述一	新北市土木技師公會理事長直言「監視器會說話」，從監視器畫面可以看到橋是從中間的橋面處先斷落，再把兩側橋墩拉垮，推測可能是鋼索銲接處風化生鏽，再加上颱風側風力高，油罐車又剛好經過，三項「巧合」加總起來，造成南方澳跨港大橋造成斷橋的原因。
敘述二	負責南方澳大橋設計案的公司表示，初步觀察橋台結構正常，但還是要分析橋面與鋼索連結處為何會斷裂，不過考量現場狀況混亂，以救災為優先，斷裂點已沉入海裡，有待相關單位拉撈上岸才能進一步判斷。
敘述三	基隆港務分公司主任秘書表示，橋梁坍塌的原因還要和工程人員討論才能有答案。然而，有部分地方人士質疑可能是去年為修復被颱風吹毀的海堤，於重修時載運大量消波塊去經南方澳大橋壓壞的。

請依據提供之影片與三則擷取自新聞的文字敘述，針對此一事件的情境背景進行分析後，回答以下幾個問題。

圖 3 情境解構試題

資料來源：取自情境式 STEM 問題解決轉化能力評量系統
(<https://stem.tahrd.ntnu.edu.tw/zh/>)。

2. 類比推理

此能力是為解析問題與其相關的資訊，通常是基於將一個已知情境或問題其具有明確的特徵或屬性，進一步與另一個相似但未知的情境或問題進行比較與分析出共同特徵或屬性，以便更好的理解問題和設計解決方案。因此本題組在類比推理能力共擬定 2 題，請學生依據所提供的資料進行分析，進一步比較 3 座橋梁形式的差異，並判斷坍塌的南方澳大橋與何種橋梁屬於同種形式，如圖 4 所示。

3. 預測分析

可以應用於評估不同解決問題方案的可能性和效益，將所蒐集與提取後的重要特徵創建出若干種預測方案，以便解決問題。因此本題組在預測分析能力共擬定共 2 題，請學生評估何種橋梁形式較適用於重建宜蘭南方澳大橋，並考量南方澳漁港之特色進而設計重建大橋，同時也能夠判斷其優、缺點為何。

橋梁	地理位置	環境條件	需求	橋樑型式	橋長	橋高
關渡大橋	新北市	淡水河流域靠近出海口	位出海口，有行船和排洪需要	拱橋	539公尺	12公尺
中沙大橋	彰化縣 雲林縣	橫跨濁水溪	大跨度且橋面須配合道路彎曲	梁橋	2345公尺	24公尺
鵬灣大橋	屏東縣	橫跨大鵬灣潟湖中唯一的出海口	考量潟湖內遊艇港及帆船中心的發展	開啟式斜張橋	195公尺	17公尺

請就以上資料進行分析，比較三座橋梁形式的差異，並選出下列敘述何者為**錯誤**？（單選題）

圖 4 類比推理試題

資料來源：取自情境式 STEM 問題解決轉化能力評量系統
 (<https://stem.tahrd.ntnu.edu.tw/zh/>)。

4. 量化思維

此轉化能力是指利用量化分析和數據驅動的方法，以便更好地理解問題與做出決策。因此本題組在量化思考能力共擬定 5 題，請學生分析橋梁設計的結構模擬數據，並選出何者較適合作為重建的橋梁選擇，以避免橋梁再次發生坍塌事件，如圖 5 所示。

5. 反思能力

是基於對問題的理解和蒐集的資料，設計一個解決方案。這個解決方案應該能夠實現目標並解決問題。評估所設計的解決方案，並確定其是否符合情境的需求。如果需要調整解決方案並重新評估。則能夠針對不同的解決問題方案進行評量。這個過程通常包括將預測結果應用於不同方案，並比較各方案可能遭遇的困難與預期效益。因此本題組在量化思考能力共擬定 2 題，請學生選出解決方案中可能的優點和缺點，以確認是否能解決重建大橋的任務。

宜蘭縣政府目前規劃重建的經費是430,000 (千元) , 須連接相隔150公尺的兩地, 同時讓漁船可以自由進出港口, 漁船高度約在18公尺以下, 且橋梁要可乘載2000噸運輸量載重, 又因橋梁為重要運輸建設, 本次重建之橋梁安全係數 (極限受力/許用受力) 需大於或等於2.5。

表3、橋梁模擬圖與數據

編號	類型	造價 (千元)	橋梁數據
甲	拱橋	430,000	橋長：180 橋寬：15 最大跨度：180 距海平面：25 建築材料：鋼筋混凝土 橋墩：鋼筋混凝土 結構可支撐：5000噸 橋面約重：2500噸 極限受力：9,000噸

圖 5 量化思維試題

資料來源：取自情境式 STEM 問題解決轉化能力評量系統
 (<https://stem.tahrd.ntnu.edu.tw/zh/>)。

(三) STEM 跨領域問題解決轉化能力評量編製

本研究開發情境式 STEM 跨領域問題解決能力評量試題的開發流程主要說明如下，如圖 6 所示。

1. 第一階段：試題內容研發

STEM 跨領域問題解決轉化能力評量試題初稿由本研究團隊研發，核心團隊成員包含多位來自 STEM 領域的教授，以及中學科學、科技與數學領域的教師。試題經由核心團隊成員共同研討發展，並經過多次命題研習會議發展完成。在完成試題初稿後，為確認試題是否符合高中程度作答，邀請 6 位高中學生參與紙本預試與訪談，並邀請參與者針對測驗時間、答題表現、5 項能力個別表現狀況等提供看法與建議。紙本預試後，針對學生反應題目冗長、耗費時間多、題目過難等問題，與出題教師討論、修正。

2. 第二階段：線上系統研發

為了能夠善用科技媒體以使學生更融入情境式的 STEM 跨領域問題解決情境，本研究團隊依據第一階段所設計的 STEM 跨領域問題解

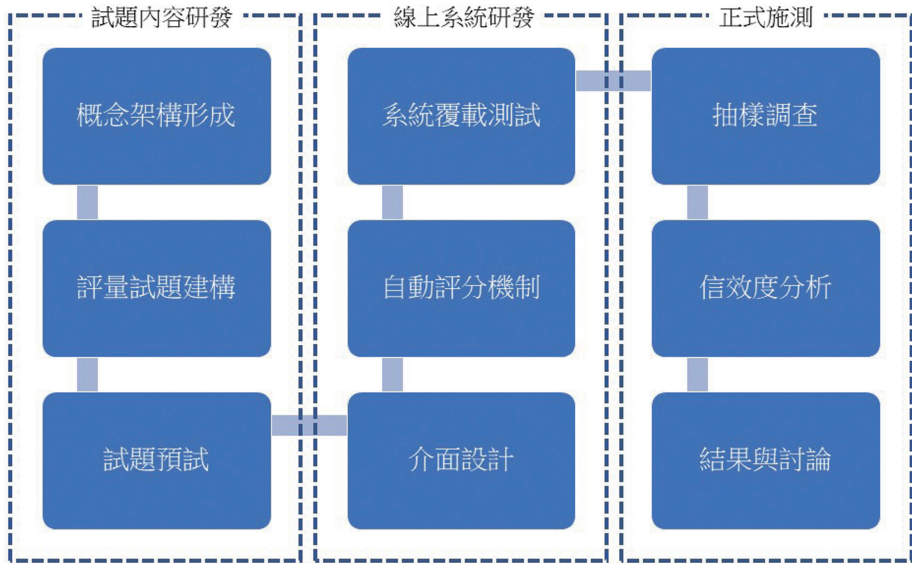


圖 6 情境式 STEM 跨領域問題解決能力評量試題研究流程

決能力評量試題，研發線上評量系統。為了解題組情境與敘述表達適當與否，邀請 5 名高中學生針對系統流程、功能操作等問題進行預試，並根據預試結果將系統介面大幅修正更新，以及加入自動評分機制，以助提升評量效率。此外，為檢驗評量題目是否有專業知識上的錯誤，以及是否符合命題向度，邀請 3 位中學現場實務教師進行題目檢驗與系統測試，修正線上試題呈現方式，增加互動性與趣味性界面，營造試題情境，再次確立題目信效度。

而為了測試前述所開發的線上系統，邀請臺北市與新北市 2 間會考平均成績位於中段程度的高中生進行系統測試，每校各抽樣 2 個班級，共 100 位同學進行測驗。但由於學校連線不穩與系統負載等問題，影響回收資料數量，依據樣本篩選標準篩選後，最終有效樣本為 43 份。而後進行系統升級，再一次進行負載測試後，可容納 160 人同時進行測驗，得最終版本再次進行正式測驗。

3. 第三階段：正式施測

本研究於臺灣北、中、南、東區之普通高中學生中進行抽樣，最後共有 13 所普通高中，共 42 班參加施測，總共抽樣 1,435 名學生，回收樣本 1,253 名（回收率 87%）。經刪除資料遺漏者、重複作答者、以及連續作答者後，最終有效樣本為 844 名（有效回收率 67%）。並進行信效度分析，針對結果進行討論與建議。

三、資料分析與詮釋

本研究從 2010 年 10 月至 2021 年 1 月進行，為期 4 個月，共回收有效樣本 844 份。試題皆為選擇題，採 Rasch 二元計分模式計算得分。本研究以 ConQuest 軟體進行試題分析，採用貝式估計法中的期望後驗法（expected a posteriori, EAP）進行能力估計。在效度分析方面，採用 infit 和 outfit 均方誤適配指標（mean square error, MNSQ）作為適配度檢驗指標。信度方面，採用試題反應理論（item response theory, IRT）的測驗訊息量（test information）、受試者分離信度（separation reliability），以及古典測驗理論的內部一致性係數，以檢視題目難度分布的適當性及測驗結果的精準度（Wright & Master, 1982; Wright et al., 1994）。

肆、研究結果

本研究以南方澳跨港大橋的情境為範例，檢驗此一情境式 STEM 跨領域問題解決能力評量試題的信度與效度，並剖析高中生在此一情境式 STEM 跨領域問題解決能力的實際表現情形，以作為未來開發 STEM 跨領域問題解決能力評量工具之參考。

一、情境式 STEM 問題解決能力評量工具之內涵

本研究主要以 STEM 跨領域問題解決轉化能力為核心，並依據以下 4 個重點進行規劃，以藉此發展情境式 STEM 問題解決能力評量工具。

(一) 情境設定

情境設定在評量過程中扮演著重要角色，要求學生將問題設置在真實情境之中，例如考慮當前社會問題、企業需求、科學研究等情境，讓學生能夠更好地理解問題的背景和涵義 (Roehrig et al., 2021)。透過情境設定便為多學科的整合奠定基礎，學生能夠在真實問題背景下，更好地將先前的知識、以及新的經驗和技能應用於解決問題。本研究以南方澳跨港大橋為主要情境，透過對這一事件的分析和研究，可以讓學生更深入地了解橋梁斷裂事件的背景和相關知識，從而更好地進行問題分析和解決方案的設計，以理解 STEM 知識和技能的應用。

(二) 問題選擇與分析

問題選擇是評估工具的首要考量，是選擇真實問題作為評量的基礎，由於真實問題具有現實性和體現 STEM 應用的特點，學生在問題解決過程中需要運用 STEM 知識、技能和方法，才能夠有效地解決問題；問題分析要求學生對問題進行識別、分析和蒐集相關資訊等工作。學生需要具備 STEM 科學思維和分析能力，才能夠有效地進行問題分析，找到解決問題的方法。在本研究中，選擇「請歸納出造成南方澳大橋坍塌的可能因素」作為一系列問題主軸，由於這個問題涉及到多個學科領域，是一個典型的跨領域 STEM 問題，學生需要具備跨學科領域的知識和技能，才能夠有效地進行問題分析和解決，從多個角度對問題進行分析，找到可能的因素並加以歸納，透過上述評量問題，學生能夠在跨領域的 STEM 學科背景下，培養解決問題的能力，並學習如何運用 STEM 知識和技能解決現實問題，促進學生的 STEM 跨學科學習。

（三）解決方案設計

解決方案設計是考慮學生在解決方案設計方面的能力。要求學生嘗試多種可能的方案進行數據分析、技術應用、優化方案，並考慮社會、環境、經濟等多種因素，設計出最佳的解決方案，確保解決方案能夠有效地實施並解決問題（Wyatt-Smith & Cumming, 2010）。本研究以「請考量各式橋梁形式較適用於重建宜蘭南方澳大橋，並判斷其優缺點」等問題作為評量工具的解決方案之設計主軸，透過這一系列的問題，學生學習了解在評量所提到的橋梁的相關概念，包括橋梁的設計、施工、維護等以發展最佳解決方案。

（四）跨領域 STEM 問題解決

跨領域 STEM 問題解決是 STEM 評量的重要標誌之一，也鼓勵學生從多個角度和層面來思考問題，培養學生的綜合思維和解決問題的能力，因為它要求學生結合跨學科的 STEM 知識和技能，並在解決問題的過程中應用這些知識和技能（Tripp et al., 2020; Zollman, 2012）。本研究以情境解構、類比推理、預測分析、量化思維、反思能力等 5 項能力作為 STEM 問題解決轉化能力評量工具之關鍵能力，旨在評估學生在跨領域 STEM 問題解決能力。這些能力代表了學生能夠將跨學科的 STEM 知識和技能應用到實際問題中，並轉化成解決問題的能力，透過這些能力的評估，可以更全面、客觀地了解學生的 STEM 解決問題能力，進而指導和改進 STEM 教育的教學和評估。

二、STEM 跨領域問題解決能力評量試題信度分析

本測驗受試者分離信度為 0.995，分離信度用以判斷試題在分隔受試者能力或位置上是否正確（Wright & Master, 1982），此係數愈大（即愈接近 1），表示題目可以測到能力分布較廣的受試者。而受試者能力估計信度（EAP/PV reliability）為 0.518，傳統古典信度（coefficient Alpha）為 0.52，表示此評量是大致可信，但可能評量到多元能力而非

單一能力，需要進一步做刪題或修改。

在試題反應理論部分，當題目難度越符合參與者的能力值時，代表測驗有較高的精準度，如圖 7 是有關本測驗的訊息量分配，橫軸代表學生能力，縱軸表示測驗訊息量，從結果可知整體測驗對於中等程度的考生能力所提供的訊息較高。

三、STEM 跨領域問題解決能力評量試題效度分析

(一) 試題配適度

本研究以二元計分 Rasch 模型分析學生的作答，各試題 MNSQ 落在 1 ± 0.3 之間，且 t 檢定值介於 ± 2.0 之間，符合 Wright 等人（1994）所提出的門檻值，表示模式與資料良好擬合。整體而言，運算經過第 26 次遞迴估計達到收斂。各試題鑑別度在 0.2 至 0.5 之間，顯示試題的鑑別度優良。難度參數估計值 -1.59 至 1.50 之間，然而 T1、T8 的 MNSQ 雖落在 1 ± 0.3 之間，但 t 檢定值超過 ± 2 的範圍區間，意指其 t 值絕對值大於 2 以上，顯示該試題達顯著不適配，需要修改或刪除。經評估

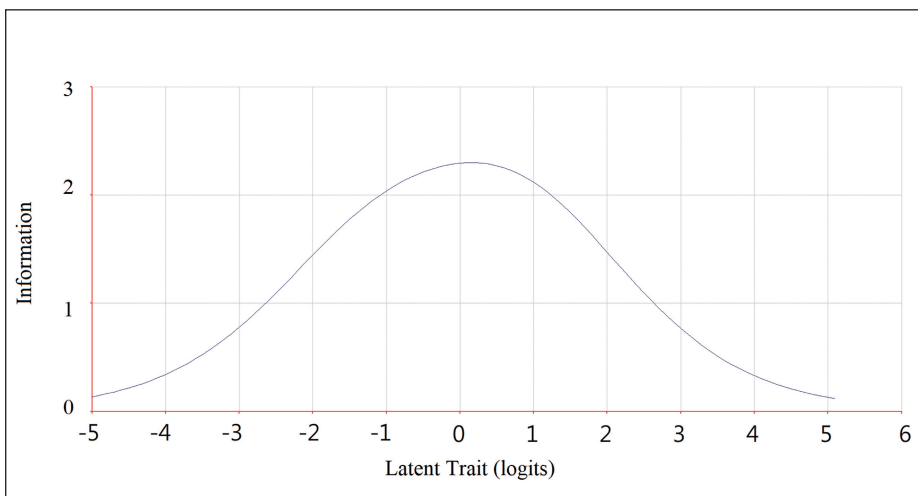


圖 7 測驗訊息量

後，刪除 T1、T8 後各項數據皆符合門檻值，因此，本評量符合單向度 STEM 問題解決轉化之單一構念解釋（如表 2）。

另從圖 8 可知，受測者能力與試題難度參數均分布在合理範圍（介於 ± 3 之間），難度在 -3 至 2 個 logits。虛線左邊的 X 表示學生能力分布，愈往上代表學生的能力愈高；虛線右邊的數字為評量題目，愈上面的題目難度愈高。整體而言，受測者參數分布範圍與試題難度範圍大致吻合，最低難度的題目在 T12，難度為 -1.59，僅有少數在極端低能力的學生可能無法答對該題。未來若要針對極低能力的受試者進行測驗，需要再增加極簡單的試題。

表 2 試題參數估計值

	難度	誤差	Outfit MNSQ	Infit MNSQ
T2	-1.43	0.08	1.02	1.01
T3	0.68	0.08	1.10	1.06
T4	0.44	0.08	1.07	1.04
T5	0.30	0.07	1.05	1.04
T6	1.29	0.09	1.04	1.02
T7	1.14	0.09	0.91	0.96
T9	-0.85	0.07	0.97	0.98
T10	-1.22	0.08	1.04	1.04
T11	-1.08	0.07	0.98	0.99
T12	-1.59	0.08	0.89	0.94
T13	0.82	0.08	0.93	0.94
T14	1.50	0.10	0.93	0.99

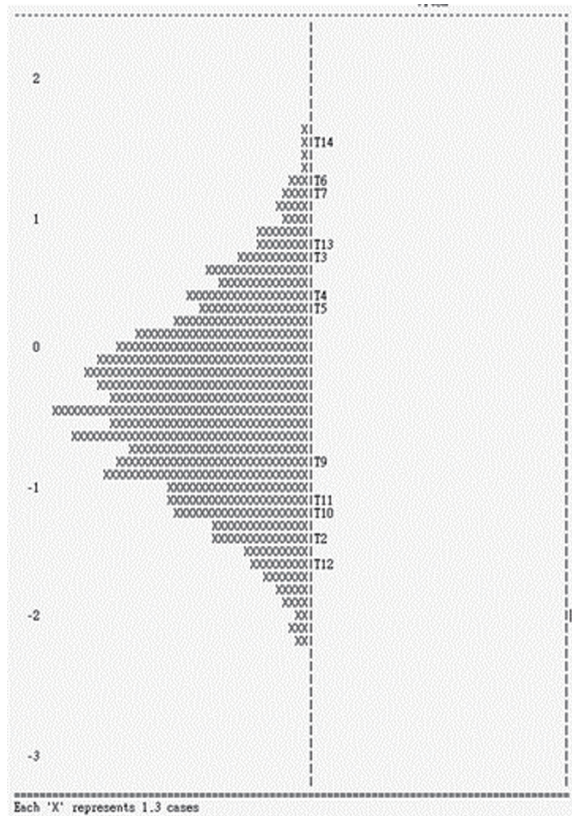


圖 8 個人試題難度參數圖

(二) 性別差異試題功能分析

若試題具有 DIF 現象 (differential item functioning)，即表示同一試題針對兩群受試者展現不同的功能，因此本研究進一步檢驗 STEM 問題解決轉化能力評量是否存在 DIF 現象。針對男女兩群體能力參數估計值分析的卡方值為 10.85，自由度為 11， p 值為 0.456，大於 .05，顯示本研究男女學生間不存在能力差異。由表 3 可知，在能力參數估算值與估計標準誤的比值計算後，其絕對值皆未大於 3，符合 Wu 等人 (1998) 的建議，表示本測驗的題目不會因為男、女性別的不同而產生不公平的現象，其對不同性別的評量功能皆相同。

表 3 性別差異分析

	性別	能力參數估算值	估計標準誤	比值
T2	男性	0.01	0.08	0.14
	女性	-0.01		
T3	男性	0.16	0.08	2.05
	女性	-0.16		
T4	男性	0.03	0.08	0.39
	女性	-0.03		
T5	男性	0.02	0.07	0.27
	女性	-0.02		
T6	男性	-0.05	0.09	-0.60
	女性	0.05		
T7	男性	-0.16	0.09	-1.91
	女性	0.16		
T9	男性	0.01	0.07	0.14
	女性	-0.01		
T10	男性	-0.01	0.08	-0.11
	女性	0.01		
T11	男性	-0.07	0.07	-0.97
	女性	0.07		
T12	男性	0.09	0.08	1.18
	女性	-0.09		
T13	男性	0.001	0.08	-0.01
	女性	0.001		
T14	男性	-0.03	0.09	-0.27
	女性	0.03		

四、高中生在 STEM 跨領域問題解決能力的表現

依據前述的信度與效度分析，本研究刪除 T1 與 T8 試題。保留後的正式試題包含 2 題情境解構、2 題類比推理、2 題預測分析、4 題量化思考及 2 題反思能力，總計 12 題。針對本研究所回收的 844 份有效樣本進行分析後，發現高中生的 STEM 跨領域問題解決能力表現如表 4 所示。依據此一分析結果，發現學生在面對 STEM 跨領域的情境時，其在 5 項轉化能力中，預測分析 ($M=0.38, SD=0.58$) 和反思能力 ($M=0.40, SD=0.59$) 的表現有待改善，顯示高中生可能不易運用此 2 項能力來解決 STEM 跨領域的情境問題。此外，從 5 項轉化能力的平均整體表現來看，受試學生在 12 個試題中，平均答對題項為 5.11 題，代表在面對情境式的 STEM 跨領域問題時，高中生可能無法適切地展現不同的轉化能力。

表 4 高中生 STEM 跨領域問題解決能力描述性統計

	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
情境解構	0.99	0.65	0.00	2.00
類比推理	0.67	0.67	0.00	2.00
預測分析	0.38	0.58	0.00	2.00
量化思維	2.67	1.18	0.00	4.00
反思能力	0.40	0.59	0.00	2.00
STEM 跨領域問題 解決能力整體表現	5.11	2.12	0.00	11.00

註：Min、Max 欄位是指高中生在評量系統中的填答表現；除了量化思考之滿分為 4 分外，其餘構面都是 2 分，整體滿分是 12 分。

伍、結論與建議

一、結論

(一) 跨領域 STEM 問題解決能力評量工具的內涵可以 5 項轉化能力為核心

本研究以跨領域 STEM 問題解決能力的情境解構、類比推理、預測分析、量化思維和反思能力這 5 項能力作為核心，據此發展情境式跨領域 STEM 問題解決能力的評量工具。根據前述針對 STEM 跨領域問題解決能力評量試題的研發，以及信度與效度的考驗結果，我們發現要設計出一套評量工具，以檢視學生在 STEM 跨領域問題解決過程中的實際表現，確實如 Kind 與 Osborne (2017) 等人所提出的觀點一樣非常困難。本研究雖然引用 Lin 等人 (2022) 所提出的 STEM 跨領域問題解決能力之轉化能力作為評量工具的開發基礎，以解決 Amalina 與 Vidákovich (2022) 所強調需要從跨學科的角度進行開發的問題，但未來的研究需要持續努力，以更有效地透過不同的評量形式來更適切地評估這 5 項轉化能力。然而，本研究因採用自動化的評量，無法應用多元的評量形式，所以侷限於以選擇題或是非題的形式命題。因此，在未來開發 STEM 跨領域問題解決評量工具時，克服本研究的限制，以開發出具備多元評量形式的評量工具，並透過該工具更精細地檢視學生在 STEM 跨領域問題解決過程中 5 項轉化能力之表現，是十分值得探討的重要。

(二) 本研究的跨領域 STEM 問題解決能力的評量工具具有信、效度

本研究開發 STEM 問題解決能力評量工具，並進行了信度和效度分析。結果顯示，試題的分離信度和受試者能力估計信度相對較高，顯示這些試題能夠測量到各種能力水平的受試者。在試題反應理論方面，試題的配適度良好，且試題的鑑別度和難度也處於適當的範圍內。整體而言，評量符合 STEM 問題解決能力的單一構念解釋。此外，還檢驗了試題在不同性別間的功能差異，結果顯示男女學生的能力沒有顯著差異，

試題對於不同性別的評量功能相同。總結而言，評量工具在信度和效度方面有一定程度的可靠性和有效性。然而，仍然需要進一步修改和改進評量工具，以提高對學生轉化能力的準確評估。

（三）高中生的跨領域 STEM 問題解決能力有待強化

根據研究結果顯示高中生在此情境式 STEM 跨領域問題解決能力的 5 項轉化能力整體表現略嫌不足，尤其是他們的預測分析和反思能力。因此，未來在推動 STEM 教育時，應該避免側重於單一學科知識，而要更關注學生知識應用的問題（Harwell et al., 2015），且應更聚焦在教導學生如何在 STEM 學習歷程中能夠確實應用不同學科概念（Nadelson et al., 2013）、以及應用跨領域的問題解決轉化能力（Lin et al., 2022），方能真正有效地培養學生的跨領域能力，並突顯 STEM 教育的真實價值。

二、研究建議

（一）對教學現場的建議

以橋梁結構的真實問題解決情境作為評量，根據國中已學過的機構與結構內容，加以延伸至高中學習目標的結構設計所需的基本概念，學生可以在本評量中所提供的情境敘述與參考資料，學習工程與科技、科學、數學等領域的重要內涵。因此，未來在 STEM 活動中，學生已具有結構設計的基本概念，應再教授 STEM 基本知識，透過引導學生在本評量中進行情境解構、類比推理、預測分析、量化思維、反思能力，藉以觀察學生在問題解決轉化能力之展現。

（二）對未來研究的建議

在情境的設定方面，本研究採用臺灣南方澳跨港大橋的情境為主。然而，近年來多元文化影響成為 STEM 教育研究的重要課題之一。長期以來，STEM 教育存在教育公平性的問題，造成弱勢族群的學生對於自己未來能夠成為 STEM 人才的認同感不足（Coxon et al., 2018）。因此，

STEM 教育應該著重讓每一位學生都能擁有公平地接受教育的機會，並增進他們對 STEM 的基本認知 (Jackson et al., 2021)。面對強調多元文化的課題，未來的 STEM 教育必須納入相關的學習內涵，所以在評量工具的設計上，若僅侷限在本土的學習情境，可能無法更適當地評估學生在面對多元文化的 STEM 跨領域問題解決情境之轉化能力表現。針對本研究的限制，建議未來在設計 STEM 跨領域問題解決能力評量工具時，應考慮納入多元文化情境，以更適切地掌握學生在多元文化情境下的轉化能力實際表現。

致謝

本文為教育部高等教育深耕計畫之特色領域研究中心補助專題研究計畫「中學 STEM 科際整合教學評量之跨國合作研究」之部分研究成果，由「國立臺灣師範大學學習科學跨國頂尖研究中心」支應，謹致謝忱。

參考文獻

- Amalina, I. K., & Vidákovich, T. (2022). Assessment in STEM problem-solving: A systematic review. *The International Journal of Assessment and Evaluation*, 29(2), 63-80. <https://doi.org/10.18848/2327-7920/CGP/v29i02/63-80>
- Asghar, A., Ellington, R., Rice, E., Johnson, F., & Prime, G. M. (2012). Supporting STEM education in secondary science contexts. *The Interdisciplinary Journal of Problem-Based Learning*, 6(2), 85-125. <https://doi.org/10.7771/1541-5015.1349>
- Becker, K., & Park, K. (2011). Effects of integrative approaches among science, technology, engineering, and mathematics (STEM) subjects on students' learning: A preliminary meta-analysis. *Journal of STEM Education: Innovation and Research*, 12(5/6), 23-37.
- Cantrell, P., & Ewing-Taylor, J. (2009). Exploring STEM career options through collaborative high school seminars. *Journal of Engineering Education*, 98(3), 295-303. <https://doi.org/10.1002/j.2168-9830.2009.tb01026.x>
- Chusinkunawat, K., Nugultham, K., Wannagatesiri, T., & Fakcharoenphol, W. (2018). Problem solving ability assessment based on design for secondary school students. *International Journal of Innovation in Science and Mathematics Education*, 26(3), 1-20.

- Coxon, S. V., Dohrman, R. L., & Nadler, D. R. (2018). Children using robotics for engineering, science, technology, and math (CREST-M): The development and evaluation of an engaging math curriculum. *Roeper Review*, 40(2), 86-96. <https://doi.org/10.1080/02783193.2018.1434711>
- Dare, E. A., Ellis, J. A., & Roehrig, G. H. (2014). Driven by beliefs: Understanding challenges physical science teachers face when integrating engineering and physics. *Journal of Pre-College Engineering Education Research*, 4(2), 47-61. <https://doi.org/10.7771/2157-9288.1098>
- English, L. D. (2017). Advancing elementary and middle school STEM education. *International Journal of Science and Mathematics Education*, 15(1), 5-24. <https://doi.org/10.1007/s10763-017-9802-x>
- Fang, S.-C., & Hsu, Y.-S. (2019). *Assessment challenges in STEM reforms and innovations*. In Y.-S. Hsu & Y.-F. Yeh (Eds.), *Asia-pacific STEM teaching practices: From theoretical frameworks to practices* (pp. 191-203). Springer. https://doi.org/10.1007/978-981-15-0768-7_12
- Gao, X., Li, P., Shen, J., & Sun, H. (2020). Reviewing assessment of student learning in interdisciplinary STEM education. *International Journal of STEM Education*, 7, Article 24. <https://doi.org/10.1186/s40594-020-00225-4>
- Guzey, S. S., Tank, K., Wang, H., Roehrig, G., & Moore, T. (2014). A high-quality professional development for teachers of grades 3-6 for implementing engineering into classrooms. *School Science and Mathematics*, 114(3), 139-149. <https://doi.org/10.1111/ssm.12061>
- Harwell, M., Moreno, M., Phillips, A., Guzey, S. S., Moore, T. J., & Roehrig, G. H. (2015). A study of STEM assessments in engineering, science, and mathematics for elementary and middle school students. *School Science and Mathematics*, 115(2), 66-74. <https://doi.org/10.1111/ssm.12105>
- Ho, M.-T., La, V.-P., Nguyen, M.-H., Pham, T.-H., Vuong, T.-T., Vuong, H.-M., Pham, H.-H., Hoang, A.-D., & Vuong, Q.-H. (2020). An analytical view on STEM education and outcomes: Examples of the social gap and gender disparity in Vietnam. *Children and Youth Services Review*, 119, 105650. <https://doi.org/10.1016/j.childyouth.2020.105650>
- Holmes, K., Gore, J., Smith, M., & Lloyd, A. (2017). An integrated analysis of school students' aspirations for STEM careers: Which student and school factors are most predictive? *International Journal of Science and Mathematics Education*, 16, 655-675. <https://doi.org/10.1007/s10763-016-9793-z>
- Holmlund, T. D., Lesseig, K., & Slavit, D. (2018). Making sense of "STEM education" in K-12 contexts. *International Journal of STEM Education*, 5, Article 32. <https://doi.org/10.1186/s40594-018-0127-2>
- Hurley, M. (2001). Reviewing integrated science and mathematics: The search for evidence

- and definitions from new perspectives. *School Science and Mathematics*, 101(5), 259-268. <https://doi.org/10.1111/j.1949-8594.2001.tb18028.x>
- Herro, D., Quigley, C., Andrews, J., & Delacruz, G. (2017). Co-measure: Developing an assessment for student collaboration in steam activities. *International Journal of STEM Education*, 4, Article 26. <https://doi.org/10.1186/s40594-017-0094-z>
- Jackson, C., Mohr-Schroeder, M. J., Bush, S. B., Maiorca, C., Roberts, T., Yost, C., & Fowler, A. (2021). Equity-oriented conceptual framework for K-12 STEM literacy. *International Journal of STEM Education*, 8, Article 26. <https://doi.org/10.1186/s40594-021-00294-z>
- Kayan-Fadlemlula, F., Sellami, A., Abdelkader, N., & Umer, S. (2022). A systematic review of STEM education research in the GCC countries: Trends, gaps and barriers. *International Journal of STEM Education*, 9, Article 2. <https://doi.org/10.1186/s40594-021-00319-7>
- Kier, M. W., Blanchard, M. R., Osborne, J. W., & Albert, J. L. (2014). The development of the STEM career interest survey (STEM-CIS). *Research in Science Education*, 44(3), 461-481. <https://doi.org/10.1007/s11165-013-9389-3>
- Kind, P., & Osborne, J. (2017). Styles of scientific reasoning: A cultural rationale for science education? *Science Education*, 101(1), 8-13. <https://doi.org/10.1002/sce.21251>
- Kirschner, P. A., Sweller, J., & Clark, R. E. (2010). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41(2), 75-86. https://doi.org/10.1207/s15326985ep4102_1
- Lee, M.-H., Chai, C. S., & Hong, H.-Y. (2019). STEM education in Asia Pacific: Challenges and development. *Asia-Pacific Education Research*, 28, 1-4. <https://doi.org/10.1007/s40299-018-0424-z>
- Lin, K.-Y., Hsiao, H.-S., Williams, P. J., & Chen, Y.-H. (2020). Effects of 6E-oriented STEM practical activities in cultivating middle school students' attitudes toward technology and technological inquiry ability. *Research in Science and Technological Education*, 38(1), 1-18. <https://doi.org/10.1080/02635143.2018.1561432>
- Lin, K.-Y., Lu, S.-C., Hsiao, H.-H., Kao, C.-P., & Williams, P. J. (2021). Developing student imagination and career interest through a STEM project using 3D printing with repetitive modeling. *Interactive Learning Environments*, 31(5), 2884-2898. <https://doi.org/10.1080/10494820.2021.1913607>
- Lin, K.-Y., Yeh, Y.-F., Hsu, Y.-S., Wu, J.-Y., Yang, K.-L., & Wu, H.-K. (2022). STEM education goals in the twenty-first century: Teachers' perceptions and experiences. *International Journal of Technology and Design Education*, 33, 479-496. <https://doi.org/10.1007/s10798-022-09737-2>

- Lynch, S. J., Burton, E. P., Behrend, T., House, A., Ford, M., Spillane, N., Matray, S., Han, E., & Means, B. (2018). Understanding inclusive STEM high schools as opportunity structures for underrepresented students: Critical components. *Journal of Research in Science Teaching*, 55(5), 712-748. <https://doi.org/10.1002/tea.21437>
- Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education*, 6, Article 2. <https://doi.org/10.1186/s40594-018-0151-2>
- Marope, M., Grifn, P., & Gallagher, C. (2017). *Future competences and the future of curriculum: A global reference for curricula transformation*. UNESCO-IBE. https://fundaciobofill.cat/uploads/old-files/BOOK_COMPLETE_future_competences_and_the_future_of_curriculum.pdf
- Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education*, 103(4), 799-822. <https://doi.org/10.1002/sce.21522>
- Moore, T., Stohlmann, M., Wang, H., Tank, K., Glancy, A., & Roehrig, G. (2014). Implementation and integration of engineering in K-12 STEM education. In S. Purzer, J. Strobel, & M. Cardella (Eds.), *Engineering in pre-college settings: Synthesizing research, policy, and practices* (pp. 35-60). Purdue University Press.
- Nadelson, L. S., Callahan, J., Pyke, P., Hay, A., Dance, M., & Pfister, J. (2013). Teacher STEM perception and preparation: Inquiry-based STEM professional development for elementary teachers. *The Journal of Educational Research*, 106(2), 157-168. <https://doi.org/10.1080/00220671.2012.667014>
- National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. The National Academies. <https://doi.org/10.17226/13158>
- National Research Council. (2014). *Developing assessments for the next generation science standards*. The National Academies. <https://doi.org/10.17226/18409>
- Olewnik, A., Yerrick, R., Simmons, A., Lee, Y., & Stuhlmiller, B. (2020). Defining open-ended problem solving through problem typology framework. *International Journal of Engineering Pedagogy*, 10(1), 7-30. <https://doi.org/10.3991/ijep.v10i1.11033>
- Priemer, B., Eilerts, K., Filler, A., Pinkwart, N., Rösken-Winter, B., Tiemann, R., & Belzen, A. U. Z. (2020). A framework to foster problem-solving in STEM and computing education. *Research in Science & Technological Education*, 38(1), 105-130. <https://doi.org/10.1080/02635143.2019.1600490>
- Quinnell, R., Thompson, R., & LeBard, R. (2013). It's not maths; it's science: exploring thinking dispositions, learning thresholds and mindfulness in science learning. *International Journal of Mathematical Education in Science and Technology*, 44, 808-816.

- <https://doi.org/10.1080/0020739X.2013.800598>
- Raju, P. K., & Clayson, A. (2010). The future of STEM education: An analysis of two national reports. *Journal of STEM Education*, 11(5 & 6), 25-28.
- Roehrig, G. H., Dare, E. A., Ring-Whalen, E., & Wieselmann, J. R. (2021). Understanding coherence and integration in integrated STEM curriculum. *International Journal of STEM Education*, 8, Article 2. <https://doi.org/10.1186/s40594-020-00259-8>
- Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education*, 96(3), 411-427. <https://doi.org/10.1002/sce.21007>
- Sanders, M. E. (2009). STEM, STEM education, STEMmania. *The Technology Teacher*, 68, 20-26. <https://www.teachmeteamwork.com/files/sanders.istem.ed.ttt.istem.ed.def.pdf>
- Sanders, M. E. (2012). Integrative STEM education as “best practice.” In H. Middleton (Ed.), *Explorations of best practice in technology, design, & engineering education* (Vol. 2, pp. 103-117). Griffith Institute for Educational Research.
- Shanta, S., & Wells, J. (2020). T/E design based learning: Assessing student critical thinking and problem solving abilities. *International Journal of Technology and Design Education*, 32, 267-285. <https://doi.org/10.1007/s10798-020-09608-8>
- Songer, N., Kelcey, B., & Gotwals, A. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *Journal of Research in Science Teaching*, 46(6), 610-631. <https://doi.org/10.1002/TEA.20313>
- Tanenbaum, C. (2016, September 14). *STEM 2026: A vision for innovation in STEM education*. Office of Innovation and Improvement, US Department of Education. https://oese.ed.gov/files/2016/09/AIR-STEM2026_Report_2016.pdf
- Tang, K.-S., & Williams, P. J. (2019). Context and implications document for: STEM literacy or literacies? Examining the empirical basis of these constructs. *Review of Education*, 7(3), 698-700. <https://doi.org/10.1002/rev3.3161>
- Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., Boevende Pauw, J., Dehaene, W., Deprez, J., De Cock, M., Hellinckx, L., Knipprath, H., Langie, G., Struyven, K., Van de Velde, D., Van Petegem, P., & Depaepae, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education*, 3(1), Article 2. <https://doi.org/10.20897/ejsteme/85525>
- Tuijl, C., & Molen, J. H. (2015). Study choice and career development in STEM fields: An overview and integration of the research. *International Journal of Technology and Design Education*, 26(2), 159-183. <https://doi.org/10.1007/s10798-015-9308-1>

- Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *Journal of Technology and Teacher Education*, 18(2), 345-368.
- Tripp, B., Voronoff, S. A., & Shortlidge, E. E. (2020). Crossing boundaries: Steps toward measuring undergraduates' interdisciplinary science understanding. *CBE Life Sciences Education*, 19(1). <https://doi.org/10.1187/cbe.19-09-0168>
- Wahono, B., Lin, P.-L., & Chang, C.-Y. (2020). Evidence of STEM enactment effectiveness in Asian student learning outcomes. *International Journal of STEM Education*, 7, Article 36. <https://doi.org/10.1186/s40594-020-00236-1>
- Wright, B. D., & Master, G. N. (1982). *Rating scale analysis*. Mesa.
- Wright, B. D., Linacre, J. M., Gustafson, J.-E., & Martin-Löf, P. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8(3), 370.
- Wu, H.-K., Kuo, C. Y., Jen, T.-H., & Hsu, Y. S. (2015). What makes an item more difficult? Effects of modality and type of visual information in a computer-based assessment of scientific inquiry abilities. *Computers & Education*, 85, 35-48. <https://doi.org/10.1016/j.compedu.2015.01.007>
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1998). *ACER ConQuest: Generalized item response modeling software*. ACER.
- Wyatt-Smith, C. M., & Cumming, J. J. (2010). Curriculum literacies: Expanding domains of assessment. *Assessment in Education: Principles, Policy & Practice*, 10(1), 47-59. <https://doi.org/10.1080/09695940301690>
- Zollman, A. (2012). Learning for STEM literacy: STEM literacy for learning. *School Science and Mathematics*, 112(1), 12-19. <https://doi.org/10.1111/j.1949-8594.2012.00101.x>