

## 壹、緒論

科學論證表達、科學實驗設計、儀器操作技術，以及應用科學理論解決問題係為重要的科學能力。為了檢視學生在科學能力的學習發展狀況，開放式申論測驗題型既可深度反映學生的知識層次與應用，亦可評量其闡釋表達能力。基此，許多研究者開始聚焦在建構反應評量暨評分規準的研發與應用 (Toranj & Ansari, 2012)。試題類型大致可分為兩類：一、選擇反應型 (selecting response type)：受試者需從所提供答案中選出較適當的選項，如是非題、選擇題或配合題；二、建構反應型：受試者必須根據題幹敘述作答，如填充題、簡答題或申論題。其中，簡答題和申論題係要求受試者依據題幹說明，組織、建構暨闡釋想法，反映帶有文字敘述的答案，不僅能夠測量敘述與程序知識，還能評測記憶、組織與想法啟發等高階認知能力 (Valenti, Neri, & Cucchiarelli, 2003)。Valenti、Cucchiarelli 與 Panti (2002) 指出，現行多數評量工具均植基於明確標準答案的題型，如選擇題、複選題、簡答題、配合題等，而有些複雜的學習成就實難以用單一標準答案的問題來測量。Gronlund (1985) 即認為某些諸如回憶、組織、整合想法、文字表達，以及資料詮釋 (interpretation) 與應用等複雜成就層次的能力，即 Bloom (1956) 六種思考模式之綜合 (synthesis)、評鑑 (evaluation) 等高層次能力，並無法透過客觀性的題型測量，必須仰賴建構反應題型方能有效檢測。

現行有關科學評量，諸如 Temiz、Taşar 與 Tan (2006) 所發展之科學過程技能測驗 (science process skills test, SPS)，以及 Kuo、Wu、Jen 與 Hsu (2015) 所發展之多媒體科學探究能力評量 (multimedia-based assessment of scientific inquiry abilities, MASIA) 皆為具有信度與效度的評量工具，可提供良好的範例參考給從事科學過程技能或科學探究能力等課程研發或工具發展的研究人員。然由於本研究所欲檢核之科學能力範疇，還包括科學基本概念的理解能力，並非僅止於科學過程技能或科學探究能力這個向度。基此，本研究研擬另外發展「科學能力的建構反應評量」(constructed-response assessment of scientific abilities, CRASA)，並驗證評量之理論構念模式是否能為實徵資料所驗證支持，藉以檢視構念效度，期能提供教師瞭解學生科學學習狀況，繼而評估應用分析、闡釋表達，以及問題解決等高階認知能力。然則，Bennett 與 Ward (1993) 認為建構反應評量相較於選擇反應評量雖檢具很多優勢，但也存在評量客觀暨耗時費力的雙重疑慮問題，亦即知覺主觀的申論評分歷程往往導致評分回饋的變異，致使評分的客觀性受到質疑 (林小慧、曾玉村，2017)。因此，本研究第二個目的即在於發展建構反應評分量規準 (rubric)，藉以克服人為評量之主觀、疲勞暨偏誤的影響，建立一套評鑑科學能力的標準，裨益客觀公正地評測學生的學習表現。綜上所述，本研究目的係在發展 CRASA 暨其評分規準，並透過 CRASA，藉以檢視學生科學能力的學習發展狀況，包括「科學知識的記憶與瞭解」、「科學程序的應用與分析」、「科學邏輯的論證與表達」，以及「問題解決的評估與創造」。

## 貳、文獻探討

### 一、建構反應評量之基礎研究

#### (一) 建構反應評量的內涵

建構反應試題中的簡答和申論題，需要受試者依據題幹內容，在回答中建構與表達自身想法，可視為是寫作的一種表現。Witkin (2000) 指出，科學寫作偏重在理性的闡釋與論述，係以理論基礎為後盾，並檢具諸如問題陳述、文獻探討、方法、結果與討論等文章結構的形式，強調推理、解析和邏輯性。故而學生在產出建構反應的過程中，需植基於理論與邏輯，進行推演和論證，並在客觀的事實範疇內，展開理性思維與合理解析。申論係為評量學習結果有效的工具，其應用到受試者本身的回憶、組織與整合想法等認知能力，也是資料解釋及應用之自我書寫表達能力的展現 (Valenti et al., 2003)。因此 Carter、Ogle 與 Royer (1993) 提出教師可透過科學建構反映成品，檢視學生科學概念的理解程度暨思考歷程。換言之，教師可藉由科學申論寫作評量來檢視學生理解科學概念的程度 (Stepanek & Jarrett, 1997)。Miller 與 Calfee (2004) 指出大型寫作評量能檢測兩種截然不同卻呈高相關的認知能力，其一為閱讀理解，其二則是將理解轉換成寫作的能力。其中，學生如何將所獲得的想法與認知轉譯成寫作文字，係植基於有助於進行理解及撰寫之思考建構的基模理論 (schema theory)。Miller 與 Calfee 進一步將寫作評量分為兩類，包括文本取向寫作評量 (text-based assessments) 與單一寫作評量 (stand-alone assessments)。

#### 1. 文本取向寫作評量

文本取向寫作評量係植基於一篇閱讀文本，學生必須依據該文本完成寫作任務。Miller 與 Calfee (2004) 認為，若在未提供背景知識 (background knowledge) 的情況下，即期待學生植基個人經驗進行評量或鼓勵創作是不合理的任務。換言之，文本具有提供背景知識的功能，可作為學生進行思考、理解及組織的中介，促使從其不同的認知觀點來執行任務。

因此，舉凡學術性寫作評量均納入文本，藉以提供學生從中習得修辭結構暨語意關聯，促使連結先前經驗 (prior experiences)。換言之，透過文本聚焦闡述的寫作提示 (writing prompts)，不僅能夠活化學生的先備知識，還可激發得以融貫 (coherence) 及有效寫作思考的隱含模式。

#### 2. 單一寫作評量

Miller 與 Calfee (2004) 主張單一寫作評量僅包含寫作任務，係植基個體的先備知識或經驗進行寫作，並認為個體只有在理解狀態下，才能進行修辭、概念及語意觀點之動態心智活動 (dynamic mental entity)，致使能夠組織寫作任務。因此，開放性試題應提供作答指引，以