

分析科學習作評估示例

鄭美紅

中國 香港

香港教育學院科學系

電郵：maycheng@ied.edu.hk

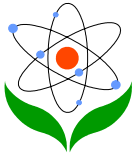
收稿日期：二零零二年十一月八日(十二月廿三日再修定)

內容

- [摘要](#)
 - [科學習作評估性質的轉變](#)
 - [另類科學評估模式](#)
 - [可信有效的評估方法之條件](#)
 - [評估課業的開展模式](#)
 - [評分準則的運用及發展](#)
 - [方法](#)
 - [結果及討論](#)
 - [綜合分析習作示例](#)
 - [參考文獻](#)
-

摘要

香港在近期的教育改革把評估重點從總結評估轉移到進展評估(Curriculum Development Council, 2001)。學習評估除用以檢視學生的學習表現外，亦反映了教師的教學情況，並能培養學生的自學能力及溝通技巧。面對這種評估模式的轉移及新需求，教師須考慮使用同儕互評的形式、選用另類的學生習作、設計及運用合適的評分指引。為加強教師採用新評估方法的信心，本文提供了不同國家的科學習作示例，並作一分析。所選例子是從一項名為「寰宇學校計劃」的國際性研究收集得來。計劃旨在讓教師主動討論及交流習作樣本，從而促進教師的專業發展。

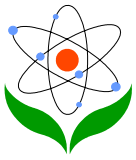


科學習作評估性質的轉變

本文以香港最近推行的教育改革為背景。根據政府的教育改革指引，學生的學習經驗可分為八個主要學習領域，科學教育屬其中一個；九種共通能力亦列為教育重點，包括協作能力、溝通能力、創造力、批判性思考能力、運用資訊科技能力、運算能力、解決問題能力、自我管理能力和研習技能(Curriculum Development Council, 2002)。學與教的文化亦轉向強調學會學習、全人發展、綜合學習、多元化的學與教資源、走出課室的學習、以及為學生提供機會，發掘自己的學習性向和潛能。因此，教學與評估必須重新配合這些教育重點。科學教育學習領域的主要關注項目強調評估的重要性，藉以提升學生的科學推理能力，加強科學過程技能，鼓勵學生主動設計和參與實驗，探討科學概念及將之應用於日常生活(Curriculum Development Council, 2002)。故此，教師應留意標準化測驗的種種限制，例如未能量度學生的特質、把知識應用於現實生活的技能(Torff, 1997)、學習科學科的興趣(Baron, 1991)、也未能鼓勵學生深入及全面理解知識(Kulm, 1991, Haetrel, 1991; Raizen and Kaser, 1989)。沒有一種評估形式能評核學生每一個學習層面的表現，就如單一的教學策略並不能涵蓋所有學習情況 (Doran, Chan and Tamir, 1998)。要包羅更多不同評估形式，量度更複雜的技巧，則須利用各式各樣的評估課業。不過，這並不代表背誦資料這種較低層次的學習形式對學習過程毫無作用，而是學生熟讀內容的程度，不應視為單一的評估範疇。

香港的教育制度較偏重常模參照的評估模式。學生要在總結評估(測驗及考試)取得理想成績，透過公開考試成績作為編班及批准入學的依據。常模參照評估模式，是指以其他接受評估學生所取得的等級作為評估比較(Marsh and Morris, 1991)。如教師採用常模參照評估而非標準參照評估，學生之間則較常出現互相比較的情形。久而久之，成績較好的學生會感到學習缺乏挑戰，而成績較遜色的學生則會感到灰心(Black, 1998)。面對教育改革，教師應重新判斷評估學生習作的根本價值。因此，Baird (1997)提出的評估要點實在值得教師詳細考慮。首先，評估應以改善學生的學習為目標。學生可根據學習的回饋改善學習。第二，評估與教學應緊密連繫起來。教師可透過評估學生習作，了解教學步驟、活動、課本及其他以學習成果為本的教材是否有效。教師應多根據這些資料來更新教學計劃。當教師了解學生的已有知識和尚未掌握的知識後，便可在教學策略上作出相應改動，支援學生的學習。Haetrel (1991) 亦強調了教學與評估的密切關係，指出評估課業應融入日常的課堂活動中。另外，評估亦可用以了解學生是否已全面掌握知識內容。最後，評估亦可讓教師為班內學生訂定實際的評核準則或標準。

總括而言，香港教育改革提出改變評估方法，其實與最近的教育研究趨向一致。評估、教學和學習三者互有關連。要評估不同技能的發展及知識理解的程度，必須利用多元化的評估模式。教師亦應注意使用評分準則或表現標準的重要性。因此，評估課業的示例亦應能反映上述重點。



另類科學評估模式

爲了評估學生是否達到不同學習目標，教師應採用多元化的評估方法。在決定使用哪些評估形式的問題上，Jones (1998)提出三個主要考慮因素：首先，教師須細心留意所用形式是否配合學生的能力及年齡，並確保有充足的資源進行有關課業。第二，所選課業應切合所要量度的學習成果。第三，評估計劃及形式應互相配合，讓學生有機會展示其技能及學習成果。因此，評估計劃不應使用單一的評估形式（如多項選擇題）。就算採用多種評估形式，亦不應只利用一種評估技巧（如書面回應）。評估計劃應揉合各種形式，同時牽涉學生不同的技能。Meng (1990) 進一步探討有效的評估技巧與學習成果之間的關係。實務技能測試、觀察、討論、晤談及專題研習，都可更有效地量度科學能力及解決問題能力的發展進度。Doran, Chan and Tamir (1998)強調運用科學習作的「另類」評估模式。另類評估是指有別於傳統所採用的形式，通常要求學生製作、示範或表演，傾向把焦點放在學生身上，以學生爲本，而且著重真確性。這些課業形式通常爲學生提供機會，就某些問題找出多種解決方法，而非從既定的表格中選擇「對」或「錯」的答案。此外，另類評估亦可讓教師了解學生的智能發展進度(Enger and Yager, 2001)。

Doran, Chan and Tamir (1998)提出的另類評估形式包括：

口頭報告及辯論：口頭報告及辯論提供另類的評估形式，對於說話能力較書寫技巧強的學生尤爲適合。口頭報告可以個人、二人或多人小組進行。利用這兩種形式，學生有機會就某一特定的科學課題作研究，並向教師及同學展示研究結果。

組織圖是以圖畫展示認知結構及思考過程的圖表，例如：概念圖及溫氏圖(Venn diagrams)。

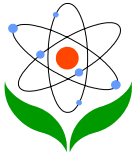
習作選輯集合了不同形式的學生習作，可包括一般家課、專題報告、實驗報告、測驗或其他習作樣本，以顯示學生的技能發展，及特定範疇內的知識及技能概念轉變。

晤談及會議：在晤談形式的評估過程中，教師會晤學生時，應以學生的概念轉變、所要糾正的錯誤觀念、及提供資料爲目的。在會議期間，可邀請學生參與自我評估或同儕互評。

技能清單是加強實驗室技巧及揉合評估與教學的理想方法。教師、學生及同儕之間都可於進行教學活動時利用該清單。

自我評估、二人評估及同儕互評：在小組學習的過程中，教師應請組內每位學生對解決難題作出貢獻，並樂於接受組員的意見，向組員學習。如學生有機會共同工作，上述行爲亦可推動學生完成評估課業。

由此可見，有效的評估示例應能如實反映學生的表現，並配合既定的學習目標。



可信有效的評估方法之條件

無論是哪種評估模式，所有評估方法都應可信而有效。所謂可信，是指所取得的評估結果必須準確、穩定，即以統一的方法評核所有參與評估的學生。Chicago Board of Education (2000)則指出有效的評估，是就某一特定目的而進行，所量度的範疇與教師原定要量度的範疇一致。評估效度的幾個重要層面，包括涵蓋內容、公平及是否能擴大其應用範圍。評估應與學習成果及教師的教學連成一線，覆蓋所有學習成果的重要層面。開展評估方法(Dougherty, 1997; Doran, Chan and Tamir, 1998)及使用評分準則都能有效達到這些評估目的。

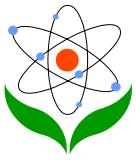
評估課業的開展模式

不斷開展評估課業可確保其信度及效度。Doran, Chan and Tamir (1998) 為科學教育工作者簡述了開展評估課業的八個非線性步驟，包括：列明目的、選取合適的課業形式、寫下或改動課業、闡明課業進行的步驟、訂立評分準則、測試課業、分析結果、修訂課業。上述步驟注重持續的發展過程，包括學生在測試課業提供的意見，以及教學與評估之間的關係。Dougherty (1997)在訂立發展科學教學的進展評估時，提出包含參與、發掘、解釋、進一步闡釋、評估這五個階段的模式，強調教師在協助學生建立科學概念時，可利用這五個階段進行進展評估。雖然單從學生習作上未能清楚看到這個發展過程，但這些模式都著重教學與評估的關係，並應在有效的評估示例中反映出來。

評分準則的運用及發展

一套統一的準則，應清楚界定及列明各項評估習作表現的準則定義(Enger and Yager, 2001)。Doran et. al. (1998)認為，評分指引應列明對學生在評估課業表現的評分或評級標準，而學生在這類課業的參與比重，遠勝選擇題形式的習作。有效的評分準則不但能清楚界定習作進行過程的標準，亦應列明對學科知識的要求。在開展進展評估模式時，Sadler (1998) 要求教師先深入了解所用評估課業的應有準則及標準。教師對學生表現的期望，應包括對全班學生的一般期望，以及對活動表現的特定期望，然後利用評估準則及標準，使評估課業與預期結果相配合(Erickson, 2001)。這種做法驗證了教學與評估的密切關係，以及評估模式的非線性發展過程。

一套評分準則可以是個有效的溝通工具，用以清楚訂明評核的動機，並就學生的表現作出回饋(Enger and Yager, 2001)。當教師、學生及家長共同使用時，評分指引就能把學校所注重的層面具體地表達出來(Chicago Board of Education, 2000)。它不但讓教師及其同事弄清優異表現的定義，將此定義傳達給學生，同時亦能就主觀評分的課業，提供理性的評級標準參考。評分指引應清楚闡明各項評核準則，好讓兩位教師按之評核同一份學生習作時，都會得到同一分數，以體現評估課業的可信性。這種統一的數據，對跨班評估及發展學校的整體改善計劃相當有幫助。



教師在設計及訂立評分準則時，須考慮幾方面的問題。首先，教師應決定要訂立一般或是特定的評分準則。教師可先設計一個整體評分架構，然後就特定的習作加以修訂。例如，用以評估一份實驗報告的評分準則，也可以用來評核其他報告寫作的習作。相反，個別習作的特定評分指引亦可加以調整，用來評核其他同一性質的習作。另一個要考慮的因素，是如何釐訂評估等級的分野。Enger & Yager (2001)認為三個等級已足以用來評估學生的表現，而通常使用的等級記號，則是 5、3 和 1。如學生習作被評為「不可接受」的類型，教師便應給予有建設性的意見，盡可能讓學生重交習作。表現優異的學生亦應得到教師的回應或意見，以確保學生在日後的習作有同樣理想的表現。Erickson (2001)指出教師往往把評估的重點放在學生未能達到的標準上，而不是為達到目標的學生給予鼓勵。其實，評分準則應以推動學生的學習動機、協助他們主動學習為目標。因此，每一個表現等級所訂立的標準均應以正面的字眼來描述，讓學生更積極面對自己的表現，把自己看成是不斷進步，而非總是達不到要求的學習者。教師評估習作時，應考慮學生邁進每一個標準期間，在每個發展階段所展示的學習能力。此外，教師如能在學生進行習作前與他們共同制定評估準則，更可促使學生主動學習。教師可嘗試與學生一同想出評估的細節，就每個表現水平制定具體內容及解釋。當學生有機會參與這個過程時，他們便會更深入了解怎樣才能達到預期目標，以取得較理想的成果。

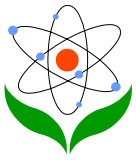
一般來說，一套有效的評分準則能清楚界定各項表現標準，作為溝通的工具，向學生給予回饋，從而推動學生學習。

總結上述各項分析，本文所分析的科學評估習作示例均具備三大特點。第一，以香港的情況來說，評估課業應配合教育改革的方向，鼓勵發展學生的科學知識及學習技巧。第二，多採用另類的評估模式。第三，評估課業應附以評分準則，達到可信、有效的評估目的。

方法

本文認為教師其實亦曾使用有效的評估課業，而不是以「虧欠模式」假設教師的教學方法基本上有問題。研究結果由參與「寰宇學校計劃」(香港網頁: www.ied.edu.hk/cric/saw/; 國際網頁: www.edc.org/CCT/saw2000/)的教師提供。這項計劃由美國基礎教育議會發起，共有其他九個國家／地區參與，當中包括香港特別行政區。參與計劃的教師須提交小四、中二及中四級*的科學習作樣本，並就每套習作樣本提供評分準則及習作背景資料。每套習作包含三份高、中、低水平的樣本。每個習作個案均會上載於計劃專屬網頁，供參與教師討論。另外，本文亦旨在提供實證，探討不同國家的教師曾採用的有效評估示例，以引發教師就科學科評估課業的問題作出專業討論。

*由於研究小組基於數個既定課題收集學生科學習作，而各國學生對有關課題的學習程度亦有所不同；故此，教師所提交的習作樣本亦可能屬於小五、中一或中三程度。

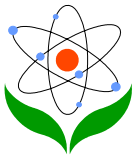


結果及討論

作者於「寰宇學校計劃」選取來自不同國家的六份習作(包括澳洲、美國、葡萄牙、法國)，並根據上文提及的四個評估課業層面作出分析。以下將總結習作的評估形式、學習技巧或科學知識、評分準則的運用、以及訂立評估課業的過程。表一總括了六份學生習作樣本的資料分析。個別習作的詳細資料將於下文逐一探討。

表一 六份習作示例的資料總結

習作名稱	級別	應用的科學知識	學習技巧及態度	選取的評估形式	評分準則的運用
動物研究——生態與環境的適應	小四	<ul style="list-style-type: none">動物對環境的適應棲息地	<ul style="list-style-type: none">從書本及互聯網搜集資料觀察記錄繪圖分類作出假設愛護動物的態度	<ul style="list-style-type: none">習作資料夾繪圖	<ul style="list-style-type: none">事先派發予學生由教師進行評估
消費者研究	中三	進行科學實驗的方法	<ul style="list-style-type: none">分析數據撰寫報告撰寫雜誌稿運用資訊科技控制變因處理數據	<ul style="list-style-type: none">雜誌稿信函以圖表組織概念	<ul style="list-style-type: none">只評核分析、表達及運用資訊科技的技巧以評語的方式給予回饋由教師進行評估
太陽系	中四	太陽系	<ul style="list-style-type: none">模型製作創作力組織能力	<ul style="list-style-type: none">模型描述報告	<ul style="list-style-type: none">不同比重的評分準則由教師進行評估
接駁電路：串聯與並聯電路	小四	串聯與並聯電路	<ul style="list-style-type: none">製作及測試模型展示結果繪圖描述能量的轉換	<ul style="list-style-type: none">模型繪圖	<ul style="list-style-type: none">評分準則以文字描述向學生解釋習作要求由教師進行評估
電子的流動	中二	<ul style="list-style-type: none">電子電運用符號	<ul style="list-style-type: none">創意寫作繪圖空間概念想像力認識能量的轉換	簡圖	<ul style="list-style-type: none">分為四個表現水平繪圖及寫作分別以兩套準則評核預先派發給學生由教師進行評估
繪畫配子	中二	配子的結構	<ul style="list-style-type: none">觀察繪圖反思	繪圖	<ul style="list-style-type: none">學生可利用教師的評分準則作自我評估由學生自評及教師進行評估



示例一： 動物研究——生態與環境的適應

級別： 小學四年級

評估形式： 此習作展示了習作資料夾及圖像運用的方法。教師要求學生為非洲矮種青蛙、招潮蟹及蝸牛建立及維持自然的棲息環境。學生在繪畫溫氏圖及比較動物的適應能力這類簡圖時，均考驗了他們的認知思考能力。

學習技巧：

透過維持動物的棲息地，可培育學生一些重要的技巧及態度，例如設法保護動物、與人交流經驗、並在飼養動物及記錄觀察日誌時，養成努力不懈、自給自足的態度。

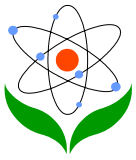
學生可透過寫作練習學會記錄觀察的重要技巧（於問題一及二描述及繪畫動物棲息地）、分類（於問題三繪畫溫氏圖、問題四的比較圖表）、推論（於問題五把動物棲息地與自然環境聯繫起來）、以及培養愛護動物的態度（問題六）。

評分準則的運用： 學生在進行習作前獲發評分準則，每項習作表現均配合教師對學生的學習期望。學生有機會重交習作。教師使用評分準則評核習作時，每個等級都包含五個要點，與寫作練習第一至五題相對應。教師沒有為問題六訂立任何評分標準，因為此問題要求學生給予個人意見。

示例一： 動物研究——生態與環境的適應

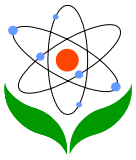
教師給予學生的指示：

學生進行了五星期的動物研究後，以此習作為最終評估。學生須建立及維持非洲矮種青蛙、招潮蟹、蝸牛的棲息地。教師鼓勵他們從書本、互聯網及百科全書搜集資料，並運用之前所學的筆記記錄技巧，定期觀察及詳細記錄有關資料。教師亦以透明膠片向學生解釋寫作部分的評分準則。



圖一：教師提供的評分準則（中文譯本）

<u>動物研究評分標準</u>	
<p style="text-align: center;">A</p> <ul style="list-style-type: none">◆ 確認出生態中需要的四種基本元素。◆ 提出一個動物生態的詳細圖解並加上標籤。◆ 列舉出兩種被觀察的動物的四種相同及四種不同的地方。◆ 確認及描述兩種被觀察的動物在行爲上及結構上的兩種適應方法。◆ 提出人類對生態環境影響的一個詳細解釋。	<p style="text-align: center;">B</p> <ul style="list-style-type: none">◆ 確認出生態中需要的四種基本元素。◆ 提出一個動物生態的圖解並加上標籤。◆ 列舉出兩種被觀察的動物的三種相同及三種不同的地方。◆ 確認及描述兩種被觀察的動物在行爲上的兩種適應及結構上的一種適應方法。◆ 提出人類對生態環境影響的一個解釋。
<p style="text-align: center;">C</p> <ul style="list-style-type: none">◆ 確認出生態中需要的二至三種基本元素。◆ 提出一個動物生態的圖解，沒有或有簡單的標籤。◆ 列舉出兩種被觀察的動物的兩種相同及不同的地方。◆ 確認及描述兩種被觀察的動物在行爲上或(及)結構上一至兩種適應方法。◆ 提出人類對生態環境影響的一個簡單解釋。	



習作樣本：

014
USA SC 2601 050 04 S 01
Name/# _____ Date 1/30/01

ANIMAL STUDIES

1. List the 4 basic needs of a habitat.
The 4 basic needs of a habitat are: food, water, space, and shelter.

2. Choose one of the three habitats (frog, crab or snail) that you and your team have been observing and illustrate it below. Be sure to label your diagram. Crab habitat

第一頁

014
USA SC 2601 050 04 S 01

3. Using the Venn diagram below, compare and contrast two (2) of your animals. (Frog/Crab/Snail)

第二頁

014
USA SC 2601 050 04 S 01

4. List two behavioral adaptations and two structural adaptations for two of your animals (Frog/Crab/Snail). Describe how these adaptations help the animals survive in their environments.

	List	Describe
Animal: Crab Behavioral adaptation	Eats Plankton Burrowes	Eats Plankton to survive burrowes for Shelter
Structural adaptation	Poppy Eyes Male has large claw	Poppy Eyes for seeing Male has large claw for protection
Animal: Frog Behavioral adaptation	Swims Fast Eats	Swims fast for Protection Eats for survival
Structural adaptation	Webbed Feet Long Tongue	Webbed feet for swimming Long tongue to catch food

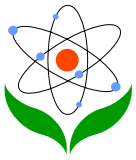
5. What organism has the greatest impact on the environment? Give at least three examples that show the important and unique way that this organism interacts with the environment.
The humans have the greatest impact on the environment. For one thing they pollute. Another thing is they start forest fires. And another reason is because we tear down trees.

第三頁

014
USA SC 2601 050 04 S 01

6. What was your favorite animal to observe? Why?
My favorite animal to observe was the crab. It was my favorite because they did neat things. I didn't know that our biggest crab could fit in the tiniest space. That's why my favorite animal was the crab.

第四頁



示例二： 消費者研究

級別： 中學三年級

評估形式：此習作是真實評估的例子。一份真實的評估習作應包含真實的任務、難題或專題，當中牽涉學生的判斷及創新意念，使學生自己動手完成有關習作，模擬現實世界的情況，包含一定的複雜性，並提供機會改進表現(Wiggins, 1998)。習作讓學生體驗實地考察的工作。學生須根據既定的科學家研究結果，以記者身份撰寫時事通訊形式的分析報告。

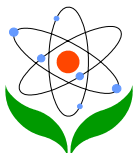
學習技巧：學生可學到在科學探究過程中控制變因的重要技巧，培養在分析所提供的證據時抱審慎、懷疑的態度。此習作亦加強了學生處理數據及資料的技巧、寫作能力、運用電腦及複雜的分析技巧。

評分準則的運用：學生於低年級時已學習有關科學探究的方法，亦嘗試設計及進行兩份相似性質的科學探究活動。雖然教師不會與學生一同進行習作，但亦會事前派發列明評分準則的習作指引及背景資料予學生。評分準則並無列明細節，教師在派發習作時會寫上評語。

示例二： 消費者研究

圖二： 教師提供的習作指引（中文譯本）

中三級消費者研究報告習作	
背景資料	
假設你是「聲音」雜誌的記者。該雜誌定期向消費者提供澳洲市場新產品的報告或資料。你收到美國"Conway"公司寄來的科學研究資料。該公司剛在澳洲市場上推出一系列的家居化學品及清潔劑。	
你的工作是為「聲音」雜誌分析附上的研究資料，並撰寫一份評論報告。	
學生必須完成以下各項工作：	
1. 為「聲音」雜誌撰寫文章——用簡單易明的方式（如圖表）去表達所提供的資料。（註：學生最好運用如"Excel"的電腦程式。）	
2. 以讀者的角度，批判性地評估實驗方法。 列舉出方法內的存在問題。	
3. 就所提供的產品資料，為讀者提供意見。（例如產品是否實惠；公司聲稱的產品效用是否正確。）	
4. 寫一封詳細的信給提供資料的科學家，建議兩項改良實驗的方法。	
學生是否準時遞交？ [是 / 否]	
評分項目	最高可得分數
實驗結果的表達方式（如圖 / 表）	3
實驗方法的分析（例如找出問題的數量）	5
為讀者提供的意見	2
改良實驗方法的質素	5
口頭報告、表達時的清晰度和原創性	3
電腦的運用	2
總分=_____	



背景資料：科學家的實驗方法（中文譯本）

“Shine”是最新研究出的衣物清潔劑。科學家發現，液體狀的“Shine”清潔劑比其他同類產品有高出百分之二十的清潔效能！此清潔劑可以應用在熱水或冷水中。它適合予手洗，對衣物和纖手都十分溫順。一點二公升的清潔劑只售價六元五角，極合乎經濟原則。

“Shine”與市面上其他一些清潔劑的比較。測試結果如下：

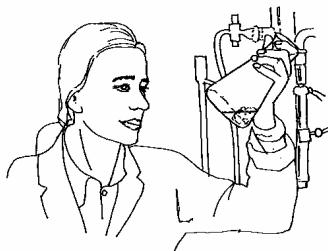
測試一：「清潔效能」



三樣名爲“Omo”、“Dynamo”、“Surf”的清潔劑與“Shine”作比較。實驗中，把一杯清潔劑加入盛載「少量」衣物的洗衣機。經過乾水的清洗過程之後，消費評審員比較結果。(本公司向每評審員支付每小時十五元的酬勞。)

產品名稱	“Omo”	\$8	750毫升	「乾淨」
	“Dynamo”	\$9	2公升	「完全地乾淨」
	“Surf”	\$7.50	1.2公升	「十分乾淨」
	“Shine”	\$6.50	1.2公升	「驚人地乾淨」

測試二：「溫順度」



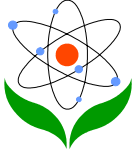
首先測試各清潔劑對衣物的溫順度，然後測試對手部皮膚的溫順度。

Scum太太、Sludge先生、Slime小姐及Sublime女士分別用不同的清潔劑清洗白襯衫。他們每天清洗，直到襯衫有明顯的磨損跡象爲止。

Scum太太是一位九十歲，很健壯的家庭主婦。她報告說，用“Omo”清洗過的衣服能經歷至少一百零五次的循環清洗，衣物的鈕扣才會脫落。她載滿衣物的洗衣機從來沒有壞過，不過她的手出了一點皺紋。

Sludge先生已經獨居兩年。他光顧市內名叫“Sticky Fingers”的自助洗衣店，用“Surf”每天兩次清洗他最喜愛的、夜總會交際用的襯衫。很不幸地，在清洗過只有三十次之後，清潔劑已損壞了「泡泡袖」的針線口。他每天戴著膠手套洗衣服。

Sublime女士使用“Shine”清潔劑。她舊式的洗衣機仍然操作如常，而襯衫到目前爲止還沒有損壞的跡象。換句話說，襯衫能通過起碼一百五十次的清洗考驗。她的皮膚甚至逃離了以往牛皮癬病的痛苦。



習作樣本：

Voice

Volume 7, Issue 6 October 2000

SHINE WITH YOUR CLOTHES

Amazing results guaranteed.
The latest development in clothes washing products.

Written by

I have just received some information from the American company called Conway. They have just recently made their products available on the Australian market.

The new product they have released is called *Shine*. It is the latest development in clothes washing products. Scientific tests have shown that *Shine* cleans 20% better.

I was given two tests that were done on this product. *Shine* and three others products were used in the testing but *Voice* is unable to publish what the other products were called. The first test looks at the washing effectiveness. And the second, at the gentleness shown on clothes and hands.

Four people were asked to each test a detergent. After they tested their product they told us what they

The other side to the story.
Is it all what it really seems?

Perhaps all is not what it seems. If you have already read the main article you would have seen that *Shine* showed some fantastic results.

But if you take a closer look at the method used and the way things were done it doesn't seem so rosy.

In the first test done, there was a lot of varying factors. A capful of each detergent was placed in the washing machine. The caps would have all been a different size. And if the back of the

第一頁

Continued from page 1

Cost of Products

Product	Price (\$ per 100 mL)
Shine	0.54
Brand 1	1.06
Brand 2	0.45
Brand 3	0.63

2 was also a good product. It showed that the clothes came up very clean. However the superior standard made by *Shine* can in no way compare.

On the left I have included a graph which shows you a price of each of these products. Brand 1 is by far the most expensive. Brand 2 is slightly cheaper than *Shine*. Nine cents cheaper per hundred mL. This can surely not win over in the factor of which one is better. *Shine's* results are wondrous, and you can not put a price on bright clean clothes.

The next test tested how gentle the products were on clothes and also on hands.

The tests were done on white shirts and were washed daily until there were obvious signs of damage.

Tester	Product tested	Gentleness on clothes	Gentleness on hands
Mrs. Scum*	Brand 1	Lasted 105 washes until buttons came off.	Hands looked slightly wrinkled.
Mr. Sludge*	Brand 3	Stitching in sleeves wore out after only 30 washes.	Gloves used.
Ms Sublime*	<i>Shine</i>	No signs of damage. Tested right through to end of test. 150 washes.	Psoriasis on hands gone.

Brand 1 was good in gentleness on clothes but seemed to fail on the gentleness on hands. Brand 3 did not give very satisfying results. After only thirty washes it seemed to wear out. Brand 3 was not tested for gentleness on hands. Once again *Shine* showed tremendously great results, lasting up until the end of the testing. It may even hold the cure for Psoriasis!♦

第二頁

Shine *Amazing* added to a small load of

continued on page 2

As you can see above *Shine* showed some fantastic

Rachael's Recommendations

Shine is said to clean 20% better. And is said to be amazing. It is a great price and is in a good-sized 1.2L bottle. It is not too small that you run out quickly but not too big either.

The shirts that were tested with *Shine* did not show any signs of damage. And *Shine* was gentle on hands. If it really will get rid of Psoriasis, I can't tell you. But I guess it is worth a try.

There did seem to be a lot of errors in the way the method was done though, and I myself wrote away to the scientist who provided me with the information and told him about how I thought it could be improved.

But can we believe what we see? This time I'll let you make up your own mind on that one. ♦

Next Week: A review on the new computer game out. Called *Deddasgart* but known on the American market as *"Dedda"*.

Continued from page 1

washing. Each load would undoubtedly be a different size. No two people would agree on the size of a "small" load.

How long did each cycle go for? How were they dried? Was it hot or cold water used? What sort of machine was used to wash them? Each of these things would contribute to the result given in the end.

And how can four different people describe what the shirts look like. Each person would say something different from the same shirt let alone different ones.

In the second test there was also a lot of errors. White shirts were washed daily until obvious signs of damage were shown. But each detergent tested a different type of white shirt. They all need to be the same. They need to be the same size, age, thickness, brand and have the same sort of stitching. The same type of washing machine needs to be used and the same length of time in the cycle, and when drying.

How can one person wash their shirts twice a day, or wear rubber gloves on a test on your hands? If buttons came off or stitching came undone it would be the shirt not the detergent used which made that happen. And I certainly don't see how someone's favorite night clubbing shirts could all be white.

And with Mrs. Scum*. Of course her hands are going to look "slightly" wrinkled. She's ninety years old!

And lastly how can anyone really tell when signs of wear or damage are showing. Perhaps one of them forgot to look at the stitching or the collar. You can not have four different people looking for signs of damage each person will not see what another sees. ♦

第三頁

VOICE
832 Collins St
Melbourne 3000
Australia

October 7, 2000

Dr Richard Harvey
CONWAY
763 Hudson Rd
Phoenix Arizona
USA

Dear Dr Richard Harvey,

I am a journalist from the Australian magazine "Voice". Your product *Shine* has recently been released onto the Australian market. I received some scientific information from your company Conway on the product. I appreciate the information that your company sent to me.

Our magazine reviews products that are on the Australian market. After the review I did on *Shine*, I thought I would write back to you and give you some feedback as to what I thought of the method used. Enclosed with this letter you will find a copy of the article I wrote about the product.

As you can see in the article, I have come across numerous errors. I rewrote what I thought would be an adequate method. You will find it on page two.

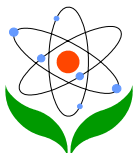
I hope you will find it to be helpful, and that you truly do consider using my new and improved method. Thankyou.

Sincerely yours,

Journalist
Voice
RAR

Please note: In magazine article *Product 1- "Omo"*
Product 2- "Dynamo"
Product 3- "Surf"

第四頁



Test 1

“Washing effectiveness”

In the lab compare the four products. If this is not possible only allow one person to compare them all. Use the amount specified on the back of the package added to 6 white shirts of the same brand. Repeat for the other three products using same machine, water temperature and wash time. After they are all dry, ask only one person to compare all of the shirts against the other brands.

Test 2

“Gentleness”

Ask four groups of three people to wash all of the white shirts (which are of the same brand and make) only once a day, in the same machine. Making a total of 12 people. All of the people must use 6 shirts in the wash how the instructions on the back of the detergent advise. Have the same length of cycle time and the same method of drying. Have another three people checking for any sign of obvious damage daily on every single shirt and make sure it was just not the shirt poorly made. Ask the twelve people how gentle they believe the detergent was on their hands. Gather all of the information and average out what they say. That way you will see a clearer indication of how gentle the detergent really is on people hands.

第五頁

教師的評估結果：

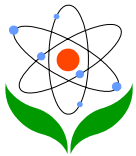
評分項目	教師評語	得分
實驗結果的表達方式（如圖 / 表）	各方面都表現出色	3/3
實驗方法的分析（例如找出問題的數量）	文章內提及實驗出現的問題數目在全班前幾名	5/5
為讀者提供的意見	所提供的改良意見能針對報告內的弊端	2/2
改良實驗方法的質素	惟一缺點，是沒有描述一部分實驗方法	4/5
口頭報告、表達時的清晰度和原創性	文字流暢，雜誌設計見心思	3/3
電腦的運用	懂得利用 Excel 程式處理文字和圖表	2/2

示例三： 太陽系

級別： 中學四年級

評估形式：此習作是探討太陽系及地球內部結構的專題研習，主要包括模型製作及報告兩種評估形式。

學習技巧：習作以小組形式進行。教師鼓勵學生多表達個人意見，同時亦要尊重別人的意見。教師只提供有限度的指引及監察，以培養學生的責任感。習作鼓勵學生發揮創意，訓練組織技巧，這兩項亦是教師評分的項目。



評分準則的運用：

評分指引所列的各項評分準則，都再細分為不同類別，每項準則所佔分數比重不同。雖然教師可為各項準則定下不同評分比重，但如能在學生進行習作前，把習作的重點通知學生，則會更為理想。此習作較著重學生在「運用理論的準確性」表現，所佔分數較「報告」一項高出一倍，而習作的「整體表現」及「創造力」，則只佔總分 20%。學生知道有關的分數比例，則可適當分配時間，尋找合適的資料，而不會把時間花在美化模型的外觀。不過，使用評分指引來評核學生表現並非一定「萬無一失」，因為學生及家長有時並不認同教師的評分。例如他們會質疑「為甚麼學生在某項準則上只拿到 12 分，而不是 13 分？」因此，分數的比重只是用來顯示不同層面的重要性，而實際的分數其實並不重要。

示例三： 太陽系

教師給予學生的指示：

學生可利用五個月時間來製作模型。習作共分三個階段：

第一階段：搜集有關太陽系和地球內部結構的資料，完成一份書面報告，介紹太陽系的模型、描述模型結構及相關現象、列出所用的材料、模型的草圖、以及提交下一階段習作的日期。

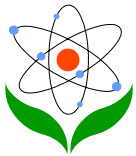
第二階段：教師與學生商議特定時間，共同討論製作模型及進行報告時所遇的問題。

第三階段：學生公開匯報已完成的工作，由教師根據評分紙評估學生的表現。

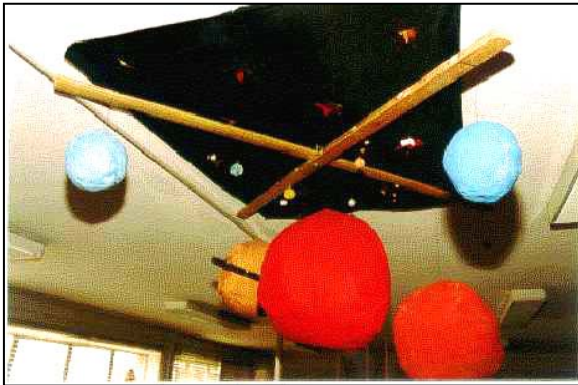
圖三：

習作樣本：

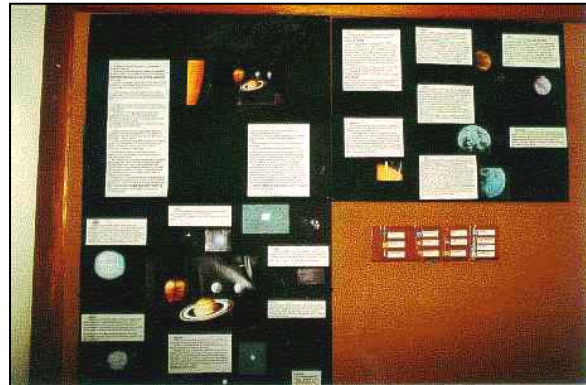




模型放大圖：

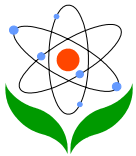


學生描述太陽系的各個星球：



教師的評估結果（中文譯本）：

<u>立體模型——太陽系</u>			評分紙		
準則		最高得分	獲得的分數		
繳交日期		3	3		
創造力		15	10		
報告	確認	作者	3	3	
		工作	4	4	
	副標題	是/ 否	5	5	
		科學 準確性	10	10	
	整體表現		5	5	
運用理論的準確性	模型製作		15	10	
	科學準確性		15	10	
	懂得運用其他顏色作強調		10	10	
	填寫的資料		15	10	
總分		100	80		



示例四： 接駁電路：串聯與並聯電路

級別： 小學四年級

評估形式： 此習作包含繪圖及電路模型製作兩部分。學生可利用所提供的材料，自行選擇製作串聯或並聯電路，然後繪畫電路圖。最後，學生須向教師解釋模型所涉及的能量轉換。

學習技巧： 透過練習及進行實驗，可訓練學生以下各種學習技巧：製作及測試模型、展示探究結果、按照實物繪圖、描述能量的轉變。

評分準則的運用：

學生可多次重交習作，直至達到超出要求的表現。教師事先已派發評分指引，並向全體學生清楚解釋對他們的學習期望。

教師訂立了一套完整的描述性評分指引，指出及解釋習作內三項評估範疇中四個不同層面的要求。總分達到 8-9 分的習作視為超出要求；5-7 分為符合水平；0-4 則未能達到要求。

Enger & Yager (2001) 指出，每項評分標準應只描述一項要求，否則對一些只達到部分要求的表現，則較難作出評分。教師可考慮把三項評估項目細分為不同類別，例如把評核電路圖練習的標籤、著色及準確程度分為三個評分細項。

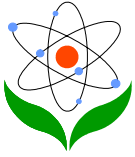
示例四： 接駁電路：串聯與並聯電路

教師給予學生的指示（中文譯本）：

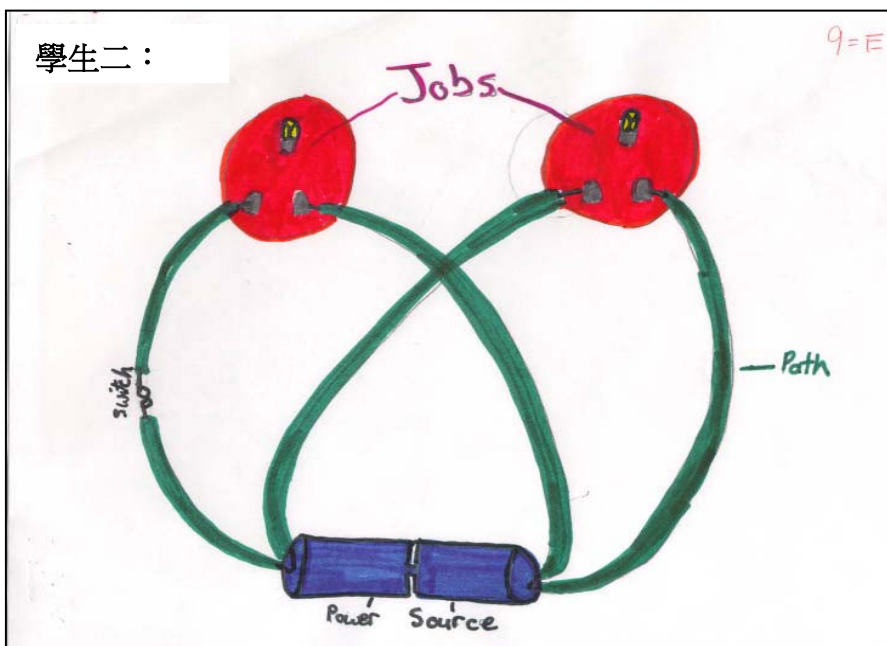
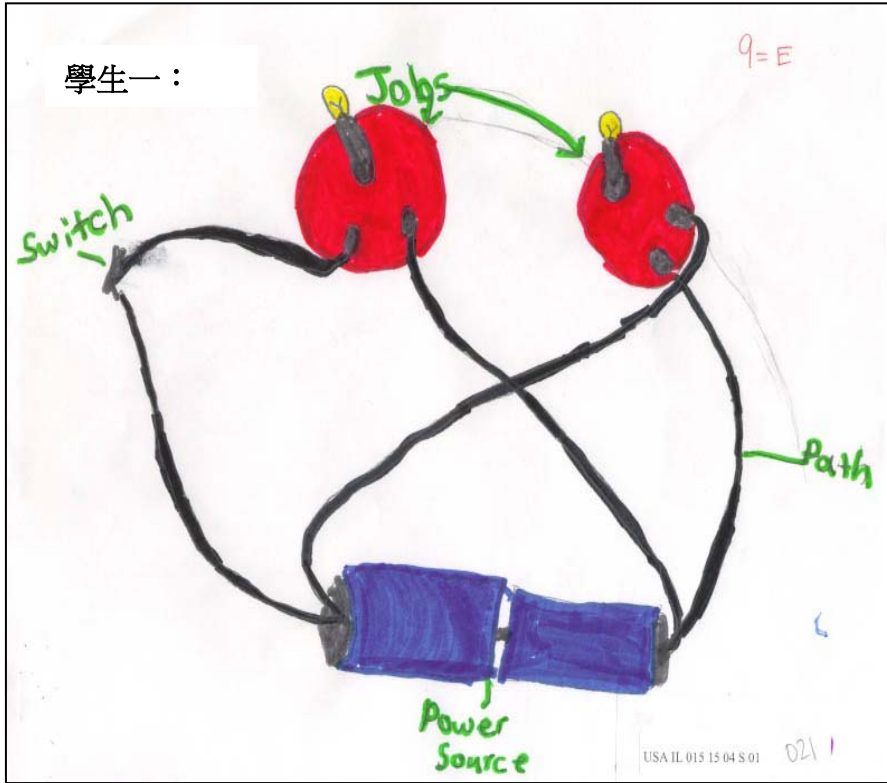
學生二人一組，利用所提供的材料合力製作一個串聯或並聯電路模型，然後個別按模型繪畫電路圖，並向教師解釋模型所涉及的能量轉換。教師利用評分指引（參閱習作樣本後頁）解釋對每項工作的要求。

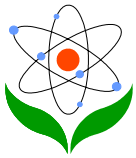
圖四： 模型樣本：





學生根據上述模型繪畫的電路圖：





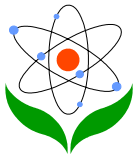
教師的評估結果：

電路設計				
請圈出你所示範的電路種類		串聯電路		並聯電路
	0	1	2	3
繪畫電路圖	沒有付出努力或不求甚解。	只有某部份的電路圖繪畫正確。	電路圖包含了電路應有的部份但不能完全正確地繪圖。	電路圖包含了電路應有的部份及完全正確地繪圖。
創作電路模型	沒有付出努力或不求甚解。	只有某部份的線路連接正確。	電路模型包含了電路應有的部份及學生能夠解釋模型的運作程序。	電路模型包含了電路應有的部份及學生能夠解釋模型的運作原理。
解釋能量種類的轉變	沒有理解。	學生只解釋了一種能量的種類。	學生解釋了在模型中所有的能量種類。	學生解釋了在模型中所有能量種類的轉換。

超出要求：8-9 符合要求：5-7 不符合要求：0-4

* 學生的串聯及並聯電路模型一定要達到超出要求的水平，才能獲得甲等成績。

組員一得分：9 (超出要求)
組員二得分：9 (超出要求)



示例五： 電子的流動

級別： 中學二年級

評估形式：此習作包括繪畫電路圖及創意寫作兩部分，前者要求學生繪畫評分指引所列的所有元素，後者要求學生運用想像力，假設自己是一粒名叫「哈利」的電子，解釋有關電壓、電阻器和能量這幾個概念之間的關係。

學習技巧：此習作設計顧及個別學生的多元智慧發展，如創意寫作練習能提升邏輯數學思維及語言技巧，而繪畫電路圖則可訓練學生以圖像顯示空間感（繪畫觀察圖），憑空想像，然後把想像轉化為物理符號，找出事物的模式（展示串聯及並聯電路都包含相同的基本元素），了解元素之間的轉化情形（兩種電路的電子流動路徑）。

評分準則的運用：

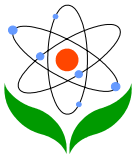
教師分別針對繪畫電路圖及創意寫作兩部分訂立了兩套不同的評分指引。學生須在沒有教師協助下完成習作，但教師會於事前派發評分指引，並講解習作指示。評分指引所列的各項要求，便成為學生習作表現的指標。

其實，教師亦可於每個評分項目加入描述性的準則。創意寫作部分的評分是以四種表現水平來評估，包括「優異」、「良好」、「一般」及「沒有提交」。雖然教師在評分時就每項準則寫下意見，但其實學生最好於事前已得悉每個等級的特定要求，例如教師可考慮在評分指引「討論所用的能量種類」一欄，列明「優異」及「良好」等級之不同要求。教師可指明拿取「優異」等級的學生須列明最少三種所用的能量，而「良好」等級的學生，則只列出其中兩種。

示例五： 電子的流動

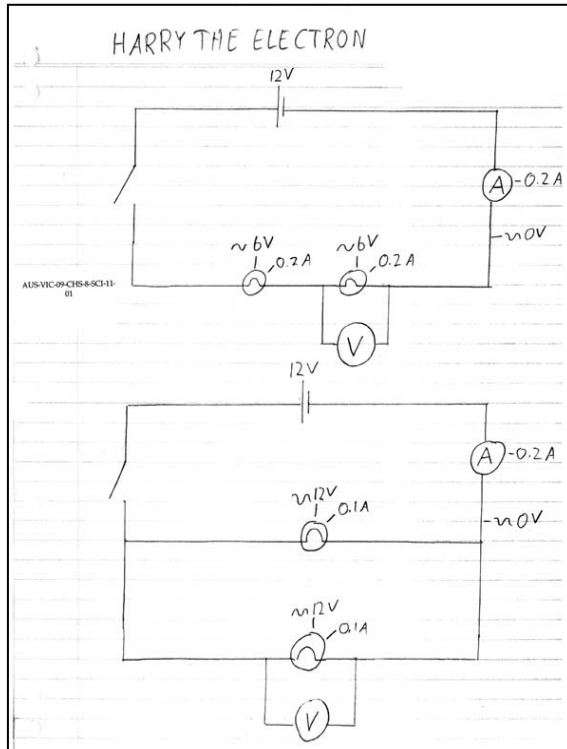
教師給予學生的指示（中文譯本）：

學生須根據電路模型準確繪畫電路圖，當中必須包括評分指引所列的各個主要配件。當學生以「電子的流動」創作故事時，可根據評分準則作為指引。教師表明習作重點，是展示對科學概念的理解及創意的發揮。



圖五：

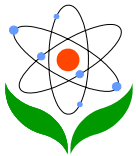
學生習作樣本：



HARRY THE ELECTRON

Harry and his mates are all electrons. At 11:00 Harry and his mates had a rude awakening when Tom connected a series circuit with 2 light globes, a switch, a Voltmeter, a Ammeter and the 12V Power Pack were Harry is sleeping. When Tom turns the Power Pack on Harry and his mates get up and go to the Power Pack Check in desk. This is were Harry gets his energy from. After the Check in desk he enters a torpedo chamber were a chemical reaction takes place and sends him flying around the circuit. "Why go around this circuit" asked one of Harry's friends. "I want to go around this dangerous circuit because I am in love with the Positive side of the power pack and the only way I can get to it is around the circuit," said Harry. As soon as Harry is released out of the Power Pack his movement is restricted, but he knows what worse is to come. These are called by the electrons, those damn things that get in the way" but are known by every one else as resistors. The first structure he passes through is the copper wire. Harry enjoys this freeway because there is little resistance. Then Harry runs into a funny regulator thing called an ammeter. This counts all the electrons that pass through. When Harry went through the number was 0.2. Harry doesn't mind this because it has little resistance. Both the Wire, Freeway and the Regulator, Ammeter's doors were arranged so that Harry and his mates could pass through without many collisions with other atoms. A few things in the hard work through the Mountains of resistors. These Mountains are commonly known as light bulbs. The atoms in these light bulbs hate Harry and his mates so they try their best to annoy them and get in their way. As Harry comes up to this he can see another light bulb in his way so he thinks to himself he only way I'm going to get past this is to use less energy on each. So Harry charges, on the first globe he has many collisions but makes it through using only 6V to get through. But on the second light globe he uses only 6V again but he finds Voltmeter Gym. In this gym he was wish to see how much energy he had lost in the light globe. All the energy Harry lost in the light globes come out as heat and light. With the little energy he had left he had a strange run home. But All of a sudden the switch was closed right in front of Harry. Harry looked at the switch and decided not to jump the gap. If Harry had alot more

energy he could jump the gap in the form of a spark. This is where Harry wished he was a bolt of lightning. If he was he could jump any gap. After he had gone around the circuit Harry had a chat with some of his friends and found out that they went through the same experiences he went through. Later in the day the circuit was changed to a parallel circuit and Harry went around it again. He did the same things except for this time he come across a junction. He decided to go straight ahead but his mate went left so Harry felt lonely. This time Harry had only 1 light globe to go through so he used 12V instead of 6V releasing more light and heat on the light globe.



教師的評估結果 (中文譯本):

47/48 A+

中二年級科學科 – 創意寫作的評分工作紙

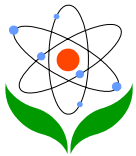
姓名:

電路繪圖

準則	在兩個電路中的位置及符號都正確	只有一個電路正確	在兩個電路中的位置及符號不正確	沒有提交
電路的設計	✓			
12 伏特的電池	✓			
開關制	✓			
電燈泡	✓			
安培計	✓			
伏特計	✓			
電壓的讀數	✓			
電流的讀數	✓			
	8			
	× 3	× 2	× 1	× 0
總成績	24			

創意寫作

準則	優異	良好	一般	沒有提交
能量來源及電勢差的討論	✓			
討論所用的能量種類	✓			
電阻及有關結構的討論	✓			
電流的討論	✓			
討論開關擊的運作	✓			
討論開關擊的限制	✓			
創作意念	✓			
表達方式		✓		
	7	1		
	× 3	× 2	× 1	× 0
總成績	21	2		



示例六： 觀察記錄圖

級別： 中學二年級

評估形式： 此習作包括兩部分：繪畫配子的特徵及填寫自我評估表格。

學習技巧： 學生已學習繪畫生物圖的技巧，所用的自我評估方法能推動學生反思自己的思考方法，是主動改善學習的重要。

評分準則的運用：

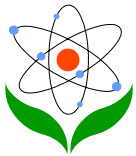
此習作的自我評估表分教師及學生評估兩欄。自我評估表是最快最方便的方法，讓學生根據既定的準則自行分析習作表現。不過，大部分教師都會擔心這種評估方法的可信性，但研究顯示，當學生有機會參與評估，而學習目標又在評估前及評估期明清楚界定時，學生普遍都會在自我評估及同儕互評如實反映(Black and Dylan, 1998)。

如教師與學生使用相同的評估標準，學生就能比較自己與教師的評分。為使評估課業更公平、更有效，可在評估表上加上空白欄，以便填寫意見。更理想的做法，是讓達不到所有準則的學生於限期內重交習作。另一項值得注意的地方，是此評估表以「成功」及「失敗」為評分等級，對某些學生來說或會較難接受。教師可考慮以較正面的用字，如「達到要求」及「仍須改進」取而代之。

示例六： 觀察記錄圖

教師給予學生的指示（中文譯本）：

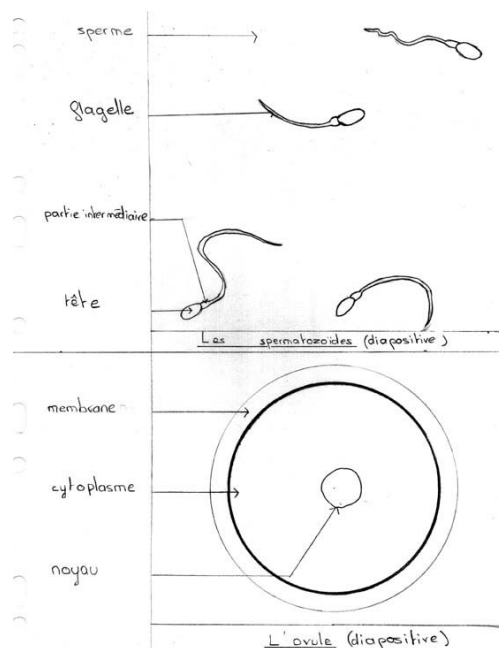
學生須根據以顯微鏡觀察精子和卵子的結果，如實繪畫觀察記錄圖。習作由學生個別進行，並可根據以下評分指引繪圖。學生可於限期前重交或分段提交已作修改的習作。

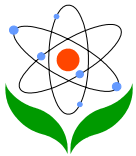


圖六： 教師提供的評分指引（中文譯本）：

	姓名： 班級：	自我評估		老師評估	
		成功	失敗	成功	失敗
工作紙的報告	用尺沿標籤欄的垂直邊界繪畫一條直線				
	整齊，用鉛筆橫行書寫				
	小心				
繪圖技巧	平均地使用繪圖框				
	置中繪圖				
	整齊，細緻及一筆過繪畫				
	整體的外形正確，適當的繪圖比例				
	繪畫正確的細胞內部結構				
標籤欄	用尺繪畫一條指示箭咀來加上標籤，不可與其他指示箭咀相交				
	指示箭咀的尾部須與適當的標籤相連				
	標籤填寫在標籤欄內，一個接一個垂直地填上				
	標籤內容須以細楷字母填寫				
	正確及完整的標籤				
	在圖的下方寫上準確的標題，並加上底線				
	列出觀察時的情況				

習作樣本（原文為法文）：





綜合分析習作示例

評估準則的運用

透過分析本文的習作示例，在運用評分準則方面，建議教師可考慮以下要點：

應事先派發有關評分準則予學生參考，這可作為溝通工具（例一），亦可作自評之用（例六）；

把學生的表現水平劃分為三或四個等級（例五）；

列出不同評分準則的內容，例如：科學概念、表達能力等，並標示各項準則所佔的比重（例二及三）；

利用文字詳細描述各項準則的要求／標準（例四）；

除提供準則外，亦加強對學生的回饋，教師評語對幫助學生自我完善及積極學習方面是十分重要的（例二）。

多元化的評估形式

從表一可見，用作評估學生的形式是可以相當多樣化的，這些形式包括：習作資料夾、繪圖、雜誌稿信函、以圖表組織概念、模型、描述、報告及簡圖等。

除使學生更容易掌握科學概念外，利用不同類型的評估形式，更可進一步發展學生學習科學的技巧及態度，例如：

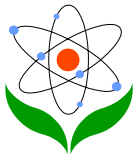
科學過程技巧：觀察、記錄、繪圖、分類、作出假設、分析數據、控制變因、處理數據、製作及測試模型、認識及描述能量的轉換等。

一般學習技巧：從書本及互聯網搜集資料、撰寫報告、撰寫雜誌稿、運用資訊科技、模型製作、創作力、組織能力、展示結果、繪圖、創作寫作、空間概念、想像力及反思等。

態度：愛護動物的態度。

總結

上述六個示例符合了最近所鼓吹的評估方向，亦展示了各種另類評估形式，包括：習作資料夾、雜誌稿、撰寫信函、繪圖、創意寫作、模型製作等，各示例亦提供評分指引，用以交待教師對學生的習作要求和給予學生的習作指示。雖然訂立評估工作的過程未能在習作示例中展示出來，但其中一位教師（示例六）



亦嘗試加入學生自評部分，讓師生使用同一套評分標準。這種評估方法亦可成為學生學習過程的一部分，讓他們深入了解教師對自己的期望，同時亦能培養自我評估的技巧。

如考慮香港的實際情況，這些外國示例其實亦可讓本地教師作為借鏡。每個示例都評估了一連串的學習技巧，評分指引的內容與培養這些學習技巧的目標亦屬一致。然而，建立科學知識是學習的一部分，學習技巧的培養也應建基於此。當香港的教師正尋找測驗考試以外的評估模式之際，上述示例相信有助他們了解如何使用平日較少採用的評估方法。

文中討論的示例是教師專業發展的實用參考材料。教師可從中了解其他教師的科學教學方法。由於所引例子均為實例，教師嘗試將之引入自己的教學過程時，定能有更大的信心。縱使這些習作並不能切合特定學校或課堂需要，然而教師必可舉一反三，按學生的學習需要作出改動。這種鼓勵教師作專業發展的取向較為正面，相信較容易為教師接受。「寰宇學校計劃」希望能透過網上討論習作示例，讓教師互相交流科學評估及教學經驗。

參考文獻：

Baird, H. (1997). *Performance Assessment for Science Teachers*. Utah: The Utah State Office of Education.

[Online]. Available: <http://www.usoe.k12.ut.us/curr/science/Perform/past1.htm>

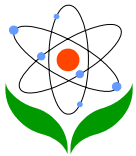
Baron, J. B. (1991) *Performance Assessment: Blurring the Edges of Assessment, Curriculum, and Instruction*. In G. Kulm, and S. M. Malcom. (Eds.). *Science Assessment in the Service of Reform*. USA: American Association for the Advancement of Science.

Black P. (1998, December). *Formative Assessment: Raising Standards Inside the Classroom*, *School Science Review*, **80**(291), 39-46.

Black P. and Dylan W. (1998, October). *Inside the Black Box: Raising Standards Through Classroom Assessment*, *Phi Delta Kappan*, **80**(2), 139-148.

Bloom, B. S., Englehart, N. D., Furst, E. J., Hill, W. H. & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: Handbook I. Cognitive domain*. New York: McCay.

Chicago Board of Education. (2000). *Chicago Public Schools Instructional Intranet Assessment Page*. [Online]. Available: <http://intranet.cps.k12.il.us/Assessments>



Curriculum Development Council. (2002). *Basic Education Curriculum Guide: Building on strengths*. Hong Kong: Government Printing Department.

Doran, R., Chan, F. & Tamir, P. (1998). *Science Educator's Guide to Assessment*. Arlington: National Science Teachers Association.

Dougherty, M.J. (1997). Formative Assessment. *Science Teacher*, **64**(6), 29-33.

Enger, S. K. and Yager, R. E. (2001). *Assessing Student Understanding in Science: A Standards-Based K-12 Handbook*. California: Corwin Press, Inc.

Erickson, H. L. (2001). *Stirring the Head, Heart, and Soul: Redefining Curriculum and Instruction*. California: Corwin Press, Inc.

Giuseppe, M.D., Bence, L., Pedrett, E., Goodnough, K., Hodson, D. and Mylchreest, L. (2001). Developing Assessment Tools for Improving the Quality of Student Performance. A presentation at the annual meeting of the National Association for Research in Science Teaching, St. Louis, Missouri, March 2001.

Haetrel, E. H. (1991) Form and Function in Assessing Science Education. In G. Kulm and S. M. Malcom. (Eds.). *Science Assessment in the Service of Reform*. USA: American Association for the Advancement of Science.

Jackie D., Jenny M., Keith V., Margaret H., Doug J., and Cherry A. (1998). *CSF: Making consistent judgments in Science*. Australia: Department of Education, Victoria.

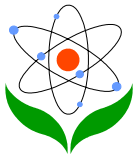
Jones, R. W. (1998). Overcoming an all-too-common problem: Solving the mismatch between science curricula and assessment. *Australian Science Teachers' Journal*, **44**(3).

Kulm, G. (1991) The Control of Assessment. In G. Kulm, and S. M. Malcom (Eds.). *Science Assessment in the Service of Reform*. USA: American Association for the Advancement of Science.

Marsh, C. and Morris, P. (1991). *Curriculum Development in East Asia*. London: Falmer.

Meng, E. (1990) What Research Says...About Appropriate Methods of Assessment. *Science and Children*, **28**(1), 42-45.

Raizen, S. A. and Kaser, J. S. (1989) Assessing Science Learning in Elementary School: Why, What, and How? *Phi Delta Kappan*, **70**(9), 718-722.



Sadler, D.R. (1998). Formative Assessment: revisiting the territory. *Assessment in Education*, 5(1), 77-84.

Torff, B. (1997) *Multiple Intelligence and Assessment*. USA: SkyLight Training and Publishing, Inc.