



1. 引言

1.1 全港性系統評估(Territory-wide System Assessment 或 "TSA")是在全港統一執行的評估，在各主要學習階段(即小三、小六及中三)終結時，釐定學生所能達到的基本能力水平。¹ 社會近年甚為關注學生及教師(尤其是小學師生)面對 TSA 的沉重壓力，為此政府推出了多項措施，包括於 2016 年在小三推行學校自願參與的"試行研究計劃"。

1.2 因應教育事務委員會的要求，資料研究組擬備了本資料摘要，研究選定海外地方所採用的學生能力評估制度。所選定的地方為澳洲、加拿大、南韓、英國² 及美國，各地均施行全國性評估制度以測試學生的學業水平，並藉此了解相關學生下列一項或多項的情況：個人測試成績、所屬學校的水平，以及身處的地區或國家整體學生的學習能力。本資料摘要所研究的評估制度分別為澳洲的全國讀寫及計算能力評估計劃("NAPLAN")；加拿大的加拿大全國學業評估計劃("PCAP")；南韓的全國教育成績評估("NAEA")；英國的全國課程評估(亦稱為 SAT 水平評估測驗)；以及美國的全國教育進展評估("NAEP")。

2. 香港的全港性系統評估

2.1 教育統籌委員會早於 2000 年建議推行基本能力評估，而 TSA 是基本能力評估的其中一環，³ 以評量學生在中、英、數 3 科的

¹ 根據 Education Bureau (2016)，基本能力指社會一般認為學生在不同學習領域課程中各主要學習階段必須掌握的能力和知識。

² 雖然英國 4 個構成國所採用的教育制度相若，但在兒童入學年齡以至學生接受評估測驗的類別等方面，仍有不少重大差異。為簡明起見，本資料摘要有關英國的研究範圍僅限於英格蘭一地。

³ 在基本能力評估下尚有"學生評估"一項，藉以為學校提供多一套評估工具。"學生評估"的特點包括：(a)備有網上中央評估庫；(b)設有網上評估活動；及(c)由電腦系統評核學生的表現，並提供即時評估報告，供教師參考。

表現。評估結果有助政府與學校了解學生在主要學習領域的水平，學校可依據評估結果改進教學，政府也可藉此為需要協助的學校提供針對性的支援。據政府所述，⁴ TSA 屬低風險的評估方案，既不會考核個別學生的表現，亦不會匯報學生的個人成績。

2.2 近年，部分持份者關注到 TSA 的試題越見艱深，以及政府在考慮資源分配時利用 TSA 的結果評估學校表現，⁵ 其影響所及，學校出現 TSA 過度操練的情況(尤以小三學生為甚)，加重師生們所面對的壓力。

2.3 政府認為過度操練或許歸咎於 TSA 的報告模式和學校處理成績的手法。⁶ 為回應社會人士的關注，政府於 2014 年推出多項改善措施，包括(a)不再向個別小學發放其學生中、英、數 3 科的基本能力達標率，以免校內以至學校之間互相比較；(b)將 TSA 從小學的"表現評量"⁷ 中剔除，以免社會人士誤會政府會 "懲罰"在 TSA 中表現稍遜的學校；⁸ 以及(c)延續小六 TSA 和"中一入學前香港學科測驗"隔年舉行的安排。⁹

2.4 政府在 2014 年推出改善措施後，於 2015 年 10 月底委託"基本能力評估及評估素養統籌委員會"就 TSA 進行全面檢討。政府根據委員會的建議，於 2016 年在小三推行學校自願參與的"試行研究計劃"。

2.5 "試行研究計劃"推展 4 項優化新元素：即(a)改良試卷及題目設計；(b)優化學校報告以照顧個別學校的需要，例如提供列明每道題目相對應的學習重點和評估重點的詳細資料分析報告；(c)加強為學校提供的專業支援措施；及(d)加入收集學生學習態度及動機的問卷調查。上述優化新元素旨在消除引致過度操練的誘因、提供對焦的服務以提升"學、教、評"的成效，以及加強學校及家長與各持份者的溝通和深化彼此的互信。

⁴ 請參閱 GovHK (2015b)。

⁵ 請參閱 Education Bureau (2016g)。

⁶ 據政府所述，一些教師(特別是小學教師)曾透露，學校數據曾被不適當使用，用在校內及學校之間互相比較，而此情況並不鮮見，為師生帶來壓力及加重他們的工作量／功課量。請參閱 Education Bureau (2014)。

⁷ "表現評量"有助學校檢討其工作成效，促進學校改善和持續發展。

⁸ 請參閱 Education Bureau (2014)。

⁹ 小六學生以往須在每年的 6 月及 7 月期間的數個星期內，應考兩個校內考試(即 TSA 與"中一入學前香港學科測驗")。為紓緩他們的考試壓力，政府由 2014 年起實行小六 TSA 與"中一入學前香港學科測驗"按年交替舉行的安排，即在單數年舉行小六 TSA，雙數年則舉行"中一入學前香港學科測驗"。

2.6 政府最近公布決定擴展"試行研究計劃"，在 2017 年推行"基本能力評估研究計劃"及將該計劃推展至全港小學。¹⁰ 按照政府的解釋，上述安排不是復考往日的 TSA。評估結果不會用作評核學校表現，而全港官立小學亦會率先不購買 TSA 補充練習，以釋除操練疑慮。

3. 選定海外地方的學生能力評估制度

3.1 澳洲、加拿大、南韓、英國及美國均推行全國性評估，對象是處於主要學習階段的學生，以便更能深入掌握他們所能達到的學業水平。評估所得的數據幫助政府及學校不斷改進日常課程，並加強為未能達標的學校及／或學生所提供的支援措施。下文各段重點介紹上述海外地方的學生能力評估制度，而有關比較各制度特色的摘要表載列於附表，隨後的附錄 I-VI 則詳述所選定海外地方的評估制度。

各地評估制度的特色及運作情況

3.2 是次研究的海外地方所施行的學生能力評估制度大致相若，但各制度之間仍存有不少差異，尤其是在評估對象、評估次數、評估的科目、所需準備和評估模式等方面。

評估對象

3.3 澳洲及美國揀選中、小學生參與全國性評估，而英國則只測試小學生。至於加拿大和南韓，兩地只評估中學生的學業水平；前者只以初中 8 年級的學生作為 PCAP 的評估對象，而後者豁免小學生接受 NAEP 評估，以豁免幼童接受評估測試。

3.4 澳洲、南韓及英國在決定應試學生的級別後，該級別的全體學生均須參與評估。美國和加拿大則採用兩階段的隨機抽樣方式揀選學生參加全國性評估，首先是從地區層面抽選應考學校，¹¹ 然後在每所應考學校的某一級別學生中抽選學生參與評估。事實上，美加兩地評估制度的設計，並非是要得出個別學生或學校的成績，因此無須要求每名學生參與評估。美國教育部表示以抽樣方式揀選學生應試，所得出的結果足以反映全國學生的整體表現。

¹⁰ 請參閱 Education Bureau (2017)。

¹¹ 美國從各州份揀選應考學校，而加拿大則從各省／地區揀選。

3.5 至於獲揀選的學生可否申請退出全國性評估，是次研究所涵蓋的海外地方做法各有不同。在美國，學生參與全國性評估屬於自願性質，獲隨機抽樣方式揀選的學生可決定是否參加。同樣地，澳洲的家長可基於宗教信仰或反對測試的理念要求子女退出全國性評估。與美國和澳洲不同，其他海外地方的全國性評估均屬強制性質，學生只可以基於身體機能障礙、學習上有嚴重困難或沒有足夠的作答語言能力等理由申請退出。

評估次數

3.6 除加拿大外，是次研究所涵蓋的海外地方均每年進行一次全國性評估。雖然英國於每年 5 月間舉行一年一度的全國性評估，但評估的日期會因應考生的年級不同而有別。教師須在 5 月份內為就讀主要階段 1¹² 最後一年的學生安排舉行 SAT 水平評估測驗，但確實的日期則由學校自行決定。學校可將測驗分開在數日甚至數星期舉行，讓學生盡可能在輕鬆的環境下進行測驗。然而，為就讀較高年級(主要階段 2 最後一年)的學生舉行的 SAT 水平評估測驗，必須在指定日期進行，而且通常指定在 5 月份某一個星期內舉行。

評估的科目及所需準備

3.7 是次研究的全國性評估所測試的科目，均着重測驗學生的讀寫及運算能力，因兩者皆是學習上不可或缺的基本能力。加拿大、英國及美國的全國性評估亦測試其他學科範疇的知識，例如科學素養(science literacy)。

3.8 至於學生是否須為應考全國性評估而作準備，澳洲、美國、南韓及加拿大負責評估的機構／政府部門均認為考生無須操練；澳洲的 NAPLAN、南韓的 NAEA 和美國的 NAEP 皆旨在測試學生在日常課堂所應掌握的知識和能力，而加拿大的 PCAP 並非根據任何特定課程出題。在澳洲，負責評估的機構亦建議家長將 NAPLAN 評估日視作校曆上的日常學校活動。相反，英國一些學校

¹² SAT 水平評估測驗的評估對象為就讀主要階段 1 及 2 最後一年的學生。英國的國定課程(national curriculum)分為 4 個主要階段：(a)主要階段 1：5 歲至 7 歲；(b)主要階段 2：7 歲至 11 歲；(c)主要階段 3：11 歲至 14 歲；及(d)主要階段 4：14 歲至 16 歲。小學屬主要階段 1(即 1 年級至 2 年級)及主要階段 2(即 3 年級至 6 年級)，中學則屬主要階段 3 及主要階段 4。

會在考試前為學生安排補課，以力求他們在 SAT 水平評估測驗中取得理想成績。操練的情況在應考主要階段 2 評估測驗的學生尤為普遍，因為該階段的評估測驗對學校來說屬高風險評估，傳媒會按公開的評估結果將考生所屬的學校排名，編製所謂學校"排行榜"(詳情請參閱 3.14 段)。

評估模式

3.9 在是次研究所涵蓋的海外地方，評估學生的方式以紙筆測驗(paper-based written test)為主，而筆試備有多類題型，包括多項選擇題及問答題。學生作答問答題時，須自行寫出答案，他們可被要求以短語(phrases)、兩三短句(sentences)以至數段文字(paragraphs)寫出答案。儘管如此，美國現正就若干科目試行電腦評估。澳洲亦計劃由 2017 年起，逐步將全國性評估的模式由紙筆測驗改為電腦化的網上測驗("NAPLAN Online")。

3.10 NAPLAN Online 的"度身訂造"測試功能，會因應學生的網上答題表現自動調節問題的難度，讓學生回答與其讀寫及運算能力相符的問題。¹³ 這項"度身訂造"的功能旨在讓所有考生均能各展所長，得到的評估結果更具代表性，而學生也更積極投入測驗。

3.11 是次研究所涵蓋的海外地方的另一特點，是其推行的全國性評估所定下的考生答卷時間限制不盡相同。除英國外，所有海外地方均要求考生在指定時間內完成作答。在英國，主要階段 1 評估測驗的時限無須嚴格執行，學校可自行決定學生完成測驗所需的時間，學生是否具有快速作答的能力，並非是評估所要測試量的一環。

匯報評估結果

3.12 是次研究所涵蓋的海外地方就匯報評估結果亦定下不同的機制；加拿大只在省份／地區及全國層面公開評估結果，而美國亦只會公布州份及全國的整體評估成績。美加兩地的做法是避免公眾利用個別學校的評估成績將他們排名，因而導致或驅使學校之間互相競爭，耗用課堂時間為學生準備評估測驗。

¹³ "度身訂造"功能會首先讓所有應考學生在網上回答一組相同的問題，他們作答的準確度會決定其後須作答的問題的難度。學生如在回答第一組相同的問題時答對較多，電腦會隨後為他們發出一組較為艱深的問題。若學生在回答第一組相同的問題時答對較少，電腦會為他們發出一組較為淺易的問題。

3.13 與加拿大和美國的情況不同，澳洲、英國和南韓提供的評估報告內容較為廣泛，列出學生、學校及地區／全國整體的評估表現。家長會從學校收到學生報告，載列其子女在全國性評估所取得的成績，並述明成績是否達到全國的最低／平均標準。在英國，家長亦同時收到由學校另行發出的教師評估報告，述明其子女在 4 核心科目(即寫作、閱讀、數學及科學)的學習進度。教師在進行教師評估時必須考慮學生日常的課堂表現，尤其是在書寫、實踐學習及言語表達方面的能力。教師評估與 SAT 水平評估測驗的結果在資料上互為補充，為學生的整體表現提供全面的評估。¹⁴

3.14 此外，澳洲、英國及南韓會公布個別考生所屬的學校在全國性評估的表現，傳媒多利用相關的數據為學校排名，編製學校"排行榜"。學校或因此受壓而採取因"試"施教的方針，將應考科目的操練置於其他學習範疇之上，以求在評估測驗中表現出色，在學校"排行榜"取得較高的排名。在英國，教育部亦按學校在 SAT 水平評估測驗的表現編製"學校成績與水平列表"(School Achievement and Attainment Tables)。英國的全國視學機構 — 教育及兒童服務與技能標準局("英國標準局") (Office for Standards in Education, Children's Services and Skills) — 為學校確立視學評鑒時，會參考該學校在"學校成績與水平列表"排名的高低。若英國標準局認為某學校的整體教學成效未如理想時，有關督學可有權決定是否對該校採取"特殊措施"或發出"改善通知書"。英國將個別學校在全國性評估結果公開的做法，可對參加評估的學校造成壓力，尤其是"學校成績與水平列表"排名較低的一群。

公眾意見及評估制度的改善措施

3.15 在英國，全國性評估制度一向被批評為對學童造成過多的壓力，尤其是對應考主要階段 2 評估測驗的學生而言。事實上，按學校在 SAT 水平評估測驗的表現為他們排名，並將排名結果公布的做法，造成學校因"試"施教，令學童壓力問題更趨嚴重。SAT 水平評估測驗在 2016 年 5 月引進了新的測試模式，惹來更多的批評。¹⁵

¹⁴ 主要階段 1／主要階段 2 評估測驗只顯示學生於某一測試日在某一科的選定範疇所能達到的學業水平，而教師評估則是教師對學生整個學年在該科目整體表現的評鑒。

¹⁵ 英格蘭於 1991 年首次引進全國性評估，而相關制度隨後經歷數次改動。最近一次改動在 2014 年 9 月發生，英國政府為所有政府資助學校頒佈全新的國定課程，而按全新的國定課程而出題的首屆評估於 2016 年舉行。

3.16 英國的小學生從 2016 年起每年應考按新模式設計的 SAT 水平評估測驗，題目較往年艱深。¹⁶ 這一改動引起了部份老師和學生家長的抗議和抵制。更令人關注的是，約半數參加 2016 年主要階段 2 評估測驗的學生未能取得達標的分數。因應上述的發展，下議院教育委員會在 2016 年 9 月宣布就全國性評估進行調查，其中的重點包括審查(a)舉行 SAT 水平評估測驗的目的，以及新的評估制度是否達到預期的目的；(b)在小學階段施行評估測驗的利弊，以及(c)隨後所須採取的改善措施。下議院教育委員會已與主要持份者舉行了 3 場聽證會，但仍未就 SAT 水平評估測驗完成總結報告。

3.17 同樣地，南韓的學生能力評估測驗亦遭受批評，傳媒按學校在評估測驗的表現編製學校"排行榜"，做法備受質疑，認為此舉導致學校因"試"施教，並加劇學校之間的競爭，以及加重學生所承受的壓力。為改善上述情況，南韓政府於 2010 年將評估測試的時間由 70 分鐘縮短至 60 分鐘，又於 2013 年推行"愉快教育：讓學生實現夢想與發揮才華"政策，廢除在小學階段進行全國性評估，而在中學繼續施行的評估亦減少評估科目，以進一步紓緩學生的壓力。在澳洲，NAPLAN 被批評鼓勵學校因"試"施教，其作為評估工具的成效亦受輿論質疑，尤其是 NAPLAN 的評估沒有包括教師評估在內，難以全面反映學生的整體表現。

3.18 美國和加拿大推行低風險的全國性評估制度，所受到的批評較少。與英國、南韓及澳洲不同，美加兩地的輿論較少批評全國性評估為學校及學生所帶來的壓力。有關加拿大的 PCAP 的批評多集中於 PCAP 的設計，而針對美國的 NAEP 所提出的批評，則關乎 NAEP 按學生的表現水平(achievement level)匯報評估成績的成效。

¹⁶ 根據英國政府所述，新推行的國定課程的設計，可以與世界上最好的教育制度所推行的媲美，故此學童應達到較高的學業水平。提高對學童的要求是希望他們離開小學時，已能掌握閱讀、寫作和運算的能力，在中學求學階段以至成年期發揮他們的潛能。請參閱 Standards & Testing Agency (2016d)。

表 —— 選定地方的學生能力評估制度

	香港	澳洲	加拿大	南韓	英國	美國
制度特色						
制度名稱	<ul style="list-style-type: none"> 全港性系統評估 ("TSA")。 由 2017 年起，小三學生的評估將會以命名為 "基本能力評估研究計劃" 進行。 	<ul style="list-style-type: none"> 全國讀寫及計算能力評估計劃 (National Assessment Program – Literacy and Numeracy) ("NAPLAN")。 	<ul style="list-style-type: none"> 加拿大全國學業評估計劃 (Pan-Canadian Assessment Program) ("PCAP")。 	<ul style="list-style-type: none"> 全國教育成績評估 (National Assessment of Educational Achievement) ("NAEA")。 	<ul style="list-style-type: none"> 全國課程評估，亦稱為 SAT (Standard Assessment Tests) 水平評估測驗。 	<ul style="list-style-type: none"> 全國教育進展評估 (National Assessment of Education Progress) ("NAEP")，包括主要評估及長期趨勢評估。
施行年份	<ul style="list-style-type: none"> 2004 年。 	<ul style="list-style-type: none"> 2008 年。 	<ul style="list-style-type: none"> 2007 年。 	<ul style="list-style-type: none"> 2000 年。 	<ul style="list-style-type: none"> 1991 年。 	<ul style="list-style-type: none"> 1969 年。
評估對象	<ul style="list-style-type: none"> 小三學生(約 8 歲)； 小六學生(約 11 歲)； 及中三學生(約 14 歲)。 	<ul style="list-style-type: none"> 中小學學生。 3 年級學生(7 至 8 歲)；5 年級學生(9 至 10 歲)； 7 年級學生(11 至 12 歲)；及 9 年級學生(13 至 14 歲)。 	<ul style="list-style-type: none"> 初中 8 年級學生(13 至 14 歲)。 	<ul style="list-style-type: none"> 中學 9 年級學生(約 14 歲)及 11 年級學生(約 16 歲)。 	<ul style="list-style-type: none"> 就讀主要階段 1 (5 至 7 歲)及主要階段 2 (7 至 11 歲)最後一年的小學生。 	<ul style="list-style-type: none"> NAEP 主要評估：小學 4 年級學生(約 9 歲)；及中學 8 年級(約 13 歲)和 12 年級學生(約 17 至 18 歲)。 NAEP 長期趨勢評估：年滿 9 歲、13 歲及 17 歲的學生。

表 —— 選定地方的學生能力評估制度(續)

	香港	澳洲	加拿大	南韓	英國	美國
制度特色(續)						
全級或抽樣應試	<ul style="list-style-type: none"> 指定級別的所有學生。 	<ul style="list-style-type: none"> 指定級別的所有學生。 	<ul style="list-style-type: none"> 隨機抽選學校和學生應試。 	<ul style="list-style-type: none"> 指定級別的所有學生。 	<ul style="list-style-type: none"> 指定級別的所有學生。 	<ul style="list-style-type: none"> 隨機抽選學校和學生應試。
強制或自願參與評估	<ul style="list-style-type: none"> 強制參與。⁽¹⁾ 	<ul style="list-style-type: none"> 自願參與。家長可以要求子女退出。 	<ul style="list-style-type: none"> 強制參與。⁽¹⁾ 	<ul style="list-style-type: none"> 強制參與。⁽¹⁾ 	<ul style="list-style-type: none"> 強制參與。⁽¹⁾ 	<ul style="list-style-type: none"> 自願參與。獲抽選的學生可決定是否參加。
評估次數	<ul style="list-style-type: none"> 小三及中三：每年評估； 小六：隔年評估。 	<ul style="list-style-type: none"> 每年評估。 	<ul style="list-style-type: none"> 每 3 年評估一次。 	<ul style="list-style-type: none"> 每年評估。 	<ul style="list-style-type: none"> 每年評估。 	<ul style="list-style-type: none"> NAEP 主要評估：每年舉行。 NAEP 長期趨勢評估：每 4 年舉行。
評估的運作情況						
評估的科目	<ul style="list-style-type: none"> 中文、英文及數學。 	<ul style="list-style-type: none"> 閱讀、寫作、語言慣例 (即拼寫、語法及標點符號)及計算。 	<ul style="list-style-type: none"> 閱讀、數學及科學。 	<ul style="list-style-type: none"> 韓文、數學及英文。 	<ul style="list-style-type: none"> 主要階段 1 及 2 測試閱讀及數學。 主要階段 2 還測試(a)語法、標點符號及拼寫和(b)科學。 	<ul style="list-style-type: none"> 主要為閱讀、寫作、數學及科學。
評估的時間	<ul style="list-style-type: none"> 各卷別的時間不同。 	<ul style="list-style-type: none"> 評估限時 40 分鐘至 65 分鐘。 	<ul style="list-style-type: none"> 限時 90 分鐘。 	<ul style="list-style-type: none"> 每項測驗限時 60 分鐘。 	<ul style="list-style-type: none"> 時限因試卷而異。 主要階段 1 的評估時限無須嚴格執行。⁽²⁾ 	<ul style="list-style-type: none"> 限時 60 分鐘。

註：(1) 考生只可以基於身體機能障礙、學習上有嚴重困難或沒有足夠的作答語言能力等理由才可申請退出。

(2) 學校可自行決定學生完成評估測驗所需的時間，學生是否具有快速作答的能力，並非是評估所要測試的一環。

表 —— 選定地方的學生能力評估制度(續)

	香港	澳洲	加拿大	南韓	英國	美國
評估的運作情況(續)						
評估模式	<ul style="list-style-type: none"> 多項選擇題及開放式提答。 	<ul style="list-style-type: none"> 多項選擇題、简答题及開放式提答。 NAPLAN 將逐步由紙筆測驗改為電腦化的網上測驗(NAPLAN Online⁽³⁾)。 NAPLAN Online 會因應學生在網上的初段答題表現自動調節隨後發出問題的難度，讓學生回答與其學業水平相符的問題。 	<ul style="list-style-type: none"> 選擇式應答題(多項選擇題、是非題等題型)及問答題。⁽⁴⁾ 	<ul style="list-style-type: none"> 七至八成題型屬多項選擇題，其他題型包括問答題。 	<ul style="list-style-type: none"> 多類題型，包括多項選擇題、是非題、配對題和開放式提答。 除評估成績外，教師還須提供評估報告，述明所教學生在核心科目的學習進度。 	<ul style="list-style-type: none"> 多項選擇題及開放式提答。
所需準備	<ul style="list-style-type: none"> 據政府所述，學生應可透過日常的學習獲取基本能力，沒有必要為 TSA 進行額外操練，或購買額外的 TSA 練習試題。 	<ul style="list-style-type: none"> 無需準備。負責評估的機構建議家長應將評估日視作日常學校活動。 	<ul style="list-style-type: none"> 無需準備。評估沒有根據任何特定課程出題。 	<ul style="list-style-type: none"> 無需準備。用作評估的試題設計，只着重測試考生在日常課堂中所應掌握的知識和能力。 	<ul style="list-style-type: none"> 一些學校會在評估前為學生安排補課。 	<ul style="list-style-type: none"> 無需準備。用作評估的試題設計，只着重測試考生在日常課堂中所應掌握的知識和能力。

註：(3) 請參閱第 3.10 段以了解 NAPLAN Online 的運作詳情。

(4) 學生作答問答題時須自行寫出答案(例如短語、短句及或數段文字)。

表 —— 選定地方的學生能力評估制度(續)

	香港	澳洲	加拿大	南韓	英國	美國
匯報評估結果						
量度考生表現的指標	<ul style="list-style-type: none"> 中、英、數 3 科每科的全港整體學生的基本能力達標率。 	<ul style="list-style-type: none"> 考生的成績按 5 個全國考評量表評核。 	<ul style="list-style-type: none"> 考生的成績按：(a) 整體平均分及 (b) 學生達到指定表現水平的百分比顯示。 	<ul style="list-style-type: none"> 學生在 NAEA 每科測驗中的分數按 4 級"學業成績水平"評核。 	<ul style="list-style-type: none"> 測驗分數會轉換成"等級分數" (scaled score)，而 100 分則定為平均學生達標分數。 	<ul style="list-style-type: none"> NAEP 所評估的科目成績以 (a) 整體學生的平均標準得分及 (b) 學生達到或超越某個能力水平的百分比顯示。
披露評估結果						
學生層面	X	✓	X	✓	✓	X
學校層面	✓	✓	X	✓	✓	X
全港(地區)／全國層面	✓	✓	✓	✓	✓	✓
比較評估結果	<ul style="list-style-type: none"> 參與評估的學校獲提供數據，載列該校學生的基本能力達標率。^{(5)·(6)} 	<ul style="list-style-type: none"> 學生、學校、州及領地之間互相比較。 傳媒按評估結果將學校排名，編製學校"排行榜"。 	<ul style="list-style-type: none"> 各省份／地區可與其他省份／地區及全國平均水平互相比較。 	<ul style="list-style-type: none"> 學校以至各道／廣域市之間互相比較。 傳媒按評估結果將學校排名，編製學校"排行榜"。 	<ul style="list-style-type: none"> 主要階段 1 的評估結果不會送交政府。 學校及各地區之間可就主要階段 2 的測驗結果作比較。⁽⁷⁾ 	<ul style="list-style-type: none"> 各州份及地區之間互相比較。
主要目的	<ul style="list-style-type: none"> 透過評估促進學習。 	<ul style="list-style-type: none"> 評估。 	<ul style="list-style-type: none"> 評估。 	<ul style="list-style-type: none"> 評估。 	<ul style="list-style-type: none"> 評估及問責。⁽⁸⁾ 	<ul style="list-style-type: none"> 評估。

註：(5) 由 2014 年起，當局只向個別中學公布該校學生在中、英、數 3 科達到基本能力水平學生的人數及百分比。

(6) 市民只可下載全港的 TSA 評估報告，當中只載述整體 TSA 結果摘要。

(7) 主要階段 2 評估測驗由校外人士負責批改試卷，而教育部會按學校及地區分項在其"學校成績與水平列表"登載評估結果。

(8) 在英國，學校須就其學生在 SAT 水平評估測驗問責。考生的表現會影響其就讀學校在"學校成績與水平列表"中的排名，而英國標準局(全國視學機構)在確立其視學評鑒時會參考相關學校在"學校成績與水平列表"中的排名。

表 —— 選定地方的學生能力評估制度(續)

	香港	澳洲	加拿大	南韓	英國	美國
公眾的接受程度						
公眾意見	<ul style="list-style-type: none"> 學生、教師及學校為應付 TSA 而面對沉重壓力的境況備受關注。 	<ul style="list-style-type: none"> 學校"排行榜"可能對學校造成壓力，令學校因"試"施教。 	<ul style="list-style-type: none"> PCAP 屬低風險評估，或令學生失卻參與評估的動機，以及難以認真對待測試，積極投入。 	<ul style="list-style-type: none"> 公開 NAEA 成績的做法帶來反效果，例如學校會因"試"施教，以及在學校以至學生之間造成競爭。 	<ul style="list-style-type: none"> 輿論指 SAT 水平評估測驗對學童造成沉重壓力。主要階段 2 的評估測驗屬高風險測試，更導致學校因"試"施教。 	<ul style="list-style-type: none"> NAEP 的評估報告按表現水平區分學生的學習能力。有輿論認為表現水平的劃分過於武斷，未必能反映學生的實際表現。
檢討制度	<ul style="list-style-type: none"> 在 2016 年，當局試行"試行研究計劃"，同時推出多項改進措施(例如改良題目的設計)。 	<ul style="list-style-type: none"> 逐步由紙筆測驗改為電腦化的網上測驗。 	<ul style="list-style-type: none"> 聯邦政府並未就 PCAP 進行任何檢討。 	<ul style="list-style-type: none"> 廢除在小學階段進行的全國性評估，而在中學繼續施行的評估亦減少評估科目及縮短測試時間。 	<ul style="list-style-type: none"> 下議院教育委員會 (Education Committee) 現就全國性評估進行調查。 	<ul style="list-style-type: none"> 全國教育統計中心⁽⁹⁾ 曾檢討 NAEP，建議減少以表現水平衡量學生表現，而應着重其他評估數據，例如整體學生的平均標準得分。

註：(9) 美國的全國教育統計中心是負責收集及分析美國及其他國家的教育數據的主要聯邦機構。

香港的學生能力評估制度

(a) 制度特色		
描述	簡介	詳情
制度名稱	全港性系統評估("TSA")。	<ul style="list-style-type: none"> TSA 是基本能力評估的其中一環，在小三、小六及中三 3 個學習階段施行。 由 2017 年起，小三學生的評估將會以命名為"基本能力評估研究計劃"施行。
施行年份	2004 年。	<ul style="list-style-type: none"> TSA 於 2004 年首先在小三實施，其後在 2005 年及 2006 年分別推展至小六及中三。
評估對象	小三、小六及中三學生。	<ul style="list-style-type: none"> 本港學童約 6 歲開始接受六年制的小學教育，升中後接受 6 年中學教育(即 3 年初中及 3 年高中)。
制度目的	以評估促進學習。 ¹	<ul style="list-style-type: none"> TSA 的主要目的是：(a)為政府及學校提供學生在主要學習範疇的能力水平資料；(b)為教師提供正面回饋的依據，以提升學與教的效能；及(c)加強學校運用評估資料的能力，並推廣促進學習的評估文化。
評估次數	<ul style="list-style-type: none"> 小三及中三：每年評估。 小六：隔年評估。 	<ul style="list-style-type: none"> 自 2014 年起，小六 TSA 逢單數年舉行。²
強制或自願參與評估	接受政府資助的學校均須舉行評估。	<ul style="list-style-type: none"> 學校如確定學生因其障礙而不能參與評估，可豁免學生參加部份或全部的評估。 2016 年，小三 TSA 以"試行研究計劃"(Tryout Study)形式進行，有超過 50 多間不同類別的小學自願參與。

¹ 據政府所述，TSA 旨在為學校提供學生表現的客觀數據，推行促進學習的評估。

² 然而，學校在雙數年可因應情況自願參與評估。

香港的學生能力評估制度

(b) 制度的運作情況		
描述	簡介	詳情
負責評估的機構	香港考試及評核局。	<ul style="list-style-type: none"> 香港考試及評核局屬獨立運作、自負盈虧的法定機構，負責在香港舉辦公開考試。
評估的科目	中文、英文及數學。	<ul style="list-style-type: none"> 中文科和英文科再細分為閱讀、寫作、聆聽及說話評估；中文科另設"視聽資訊"一卷。
評估的時間	各卷別的評估時間不同。	<ul style="list-style-type: none"> 舉例而言，2016年小三TSA英文科讀寫試卷限時30分鐘，而數學科試卷限時40分鐘。
評估模式	多項選擇題及開放式提答。	<ul style="list-style-type: none"> 主要以紙筆形式作答，中、英文科還兼考說話評估。 2016年的評估加入收集學生學習態度及動機的問卷調查。³
所需準備	沒有。據政府所述，學生應可透過日常的學習獲取基本能力，沒有必要為TSA進行額外操練，更無需額外購買TSA練習試題。	<ul style="list-style-type: none"> TSA歷屆評估試卷可供下載使用。 學生可購買由出版商出版的補充練習，而據報道，學校曾安排補課及使用模擬試卷操練學生應付TSA。

³ 問卷調查是一項新元素，旨在收集學生的非學業數據(例如參與課外活動的時間、學習興趣及動機等數據)，為學校提供更多資料，從而(a)加深了解影響學習表現的因素；及(b)進一步協助學生學習。

香港的學生能力評估制度

(c) 匯報評估的結果		
描述	簡介	詳情
量度考生表現的指標	中、英、數 3 科每科的全港整體學生基本能力達標率。 ⁴	<ul style="list-style-type: none"> 舉例而言，2015 年全港小三學生的中、英文及數學科基本能力達標率分別為 86.4%、80.4% 及 87.6%。
評估報告	<ul style="list-style-type: none"> 每個評估年度均公開全港報告，述明整體 TSA 結果摘要。 參與評估的學校可獲該校的學校報告。 	<ul style="list-style-type: none"> 在學校層面，TSA 為學校提供學校報告，報告不僅載列該校學生在中、英、數 3 科的整體達標率，還提供題項分析及其他補充資料，包括各題的答對率及學生在各項測試基本能力題目的表現等。
比較評估結果	學校不應互相比較。	<ul style="list-style-type: none"> 不會公布個別學生所得的分數。 為免學校在校內及與其他學校作不適當的比較，由 2014 年起，當局不再向個別小學提供學生在中、英、數 3 科的基本能力達標率。
運用評估結果	政府和學校利用評估數據優化教學安排及協助學生學習。	<ul style="list-style-type: none"> 在全港層面，評估數據有助政府檢視教育政策、提供資源和訂定支援措施及專業培訓的方向等。 在學校層面，評估數據有助教師了解整體學生的強弱項，從而制訂改善學與教效能的計劃。

⁴ 當局在首年舉行 TSA 後即成立了顧問／評審小組，釐定中、英、數 3 科的基本能力水平，該等水平至今未變。

香港的學生能力評估制度

(d) 公眾對制度的接受程度	
描述	重點
公眾意見	<ul style="list-style-type: none"> 近年，TSA 成為備受關注的議題，在社會上引起廣泛討論。部分持份者深切關注到學校為 TSA 過度操練，對學生和教師造成壓力，尤以小學的情況為甚。學校過度操練的成因，有意見認為 TSA 的題目越來越艱深，加上評估結果在校內及與其他學校作不適當的比較所致。 為回應社會人士的關注，政府於 2014 年推出多項改善措施，包括(a)不再向各小學發放其學生的基本能力達標率，以免有所比較；及(b)TSA 成績不再列入小學的"學校表現評量"。
檢討制度	<ul style="list-style-type: none"> "基本能力評估及評估素養統籌委員會"於 2015 年年底全面檢討 TSA 的施行情況。 當局根據委員會的建議，於 2016 年在小三推行"試行研究計劃"。該計劃推展 4 個新元素：即(a)改良試卷及題目設計；(b)優化學校報告以照顧個別學校的需要；(c)加強為學校提供的專業支援措施；及(d)加入收集學生學習態度及動機的問卷調查。 政府進一步解釋，2016 試行研究計劃的題目整體較以往淺易，對準小三學生的基本能力要求。至於學校報告，當局向學校提供 4 款評估報告以供選擇：現行版、精簡版、整合版及資料分析版報告。⁵ 學校可按校情及科本情況，選擇獲取任何一款或多款報告。
最新發展	<ul style="list-style-type: none"> 政府於 2017 年 1 月公布，"試行研究計劃"將會以命名為"基本能力評估研究計劃"施行，在 2017 年推展至全港小學。 政府表示，有關安排不是復考往日的 TSA，而 2016 年"試行研究計劃"推展的新元素將會在 2017 年的評估中保留下來。

⁵ 精簡版學校報告只提供個別學校在各道題目的表現數據，當中並無全港整體數據。整合版學校報告則為一份基本能力題組綜合報告，附有整體學生表現示例。此外，資料分析報告提供每道題目相對應的學習重點、基本能力、提問用意，以及多項選擇題的選項分析。請參閱 Education Bureau (2016f)。

澳洲的學生能力評估制度

(a) 制度特色		
描述	簡介	詳情
制度名稱	全國讀寫及計算能力評估計劃 (National Assessment Program—Literacy and Numeracy) ("NAPLAN")。	<ul style="list-style-type: none"> NAPLAN 在全國施行，旨在評估所有就讀 3 年級、5 年級、7 年級及 9 年級學生在閱讀、寫作、語言慣例(即拼寫、語法及標點符號)和計算能力範疇的知識和能力。NAPLAN 屬全國考評計劃 (National Assessment Program)的一環。⁶
施行年份	2008 年。	<ul style="list-style-type: none"> NAPLAN 於 2008 年開始實施，旨在加強收集、統一評估及匯報全國學生的讀寫及計算能力的表現數據。在此之前，各個州份和領地各自實施有關評估學生讀寫及計算能力的測試。
評估對象	所有 3 年級、5 年級、7 年級及 9 年級學生。	<ul style="list-style-type: none"> 澳洲各個州份和領地的教育制度不盡相同。一般而言，學童接受 6 至 7 年的小學教育，由 1 年級至 6 或 7 年級。中學的修業期則為 3 或 4 年，由 7 至 10 年級或 8 至 10 年級；而高中的修業期為兩年，由 11 年級至 12 年級。
評估目的	NAPLAN 旨在提供學生在讀寫和計算方面的成績數據以作評估。	<ul style="list-style-type: none"> NAPLAN 為教師及其他主要持份者提供一項客觀的量度工具，定期評估學生的學業成績，讓他們就有關學生的成績與全國平均值及澳洲其他同年學生的成績作出比較。
評估次數	每年評估。	<ul style="list-style-type: none"> 評估逢每年 5 月舉行。
強制或自願參與評估	自願參與。家長可以宗教信仰或反對測驗的理念為理由要求讓其子女退出評估。	<ul style="list-style-type: none"> 來自非英語背景的學生或殘疾學生可獲豁免。

⁶ 全國考評計劃是一項量度工具，包括 NAPLAN 和抽樣評估 (sample assessments)，用以評估學生的表現是否符合重要的教育成效指標。只有少數 6 年級和 10 年級的學生須參與抽樣評估，而抽樣評估每年進行，以 3 年為一周期，分別測驗學生在 (a)"科學素養"；(b)"公民及公民權"；及 (c)"資訊及通訊科技素養"方面所掌握的知識和能力。舉例而言，抽樣評估於 2003 年開始施行，首先就"科學素養"進行評估，隨後於 2004 年及 2005 年施行的抽樣評估則分別就"公民及公民權"及"資訊及通訊科技素養"進行評估。

澳洲的學生能力評估制度

(b) 制度的運作情況		
描述	簡介	詳情
負責評估的機構	澳洲全國課程評核和報告局 (Australian Curriculum, Assessment and Reporting Authority) ("ACARA")。	<ul style="list-style-type: none"> ACARA 為獨立的法定機構，負責制訂國定課程 (national curriculum)、推行全國性評估，以及收集和公布學校在評估中表現的數據，以支援學生學習。
評估的科目	4 個範疇：即 (a) 閱讀；(b) 寫作；(c) 語言慣例 (即拼寫、語法及標點符號)；及 (d) 計算能力。	<ul style="list-style-type: none"> NAPLAN 旨在評核學生修讀國定課程後所能掌握的讀寫及計算能力。
評估的時間	NAPLAN 限時約 40 分鐘至 65 分鐘。	<ul style="list-style-type: none"> NAPLAN 為期 3 天，以筆試形式進行。 閱讀評估時限：3 年級 45 分鐘；5 年級 50 分鐘；7 年級及 9 年級 65 分鐘。 寫作評估時限：所有年級均限時 40 分鐘。 語言慣例評估時限：3 年級及 5 年級 40 分鐘；7 年級及 9 年級 45 分鐘。 計算能力評估時限：3 年級 45 分鐘；5 年級 50 分鐘；7 年級及 9 年級 60 分鐘。
評估模式	<ul style="list-style-type: none"> 筆試包括多項選擇題、简答题及開放式提答。 NAPLAN 將逐步由紙筆測驗改為電腦化的網上測驗 ("NAPLAN Online")。NAPLAN Online 的 "度身訂造" 測試功能，會因應學生的網上答題表現自動調節問題的難度，讓學生回答與其學習能力相符的問題。 	<ul style="list-style-type: none"> NAPLAN 將由 2017 年起於 3 年內逐步轉為在網上測試學生。 NAPLAN Online 會首先讓所有應考學生在網上回答一組相同的問題，他們作答的準確度會決定其後須作答的問題的難度。學生如在回答第一組相同的問題時答對較多，電腦會隨後為他們發出一組較為艱深的問題。若學生在回答第一組相同的問題時答對較少，電腦會為他們發出一組較為淺易的問題。⁷

⁷ NAPLAN Online 的題目將有別於紙筆測驗題目。NAPLAN Online 新增多個不同的題目形式和互動功能，例如考生可以點按文字 (clickable text)、拖放文字 (drag and drop) 及聽寫等方式作答。

澳洲的學生能力評估制度

(b) 制度的運作情況(續)		
描述	簡介	詳情
所需準備	<ul style="list-style-type: none"> 歷屆試卷和例題登載於 ACARA 的網站，學生可下載以便熟習在測驗中須回答的題型。 ACARA 表示，NAPLAN 評核學生修讀國定課程後所能掌握的讀寫及計算技巧，他們無須過度操練。家長應把 NAPLAN 的評估日視作校曆上的日常學校活動。 	<ul style="list-style-type: none"> 應考 NAPLAN 的學生可購買商業機構所提供的備考服務或練習(例如測驗試題／答案和考題分析)。然而，ACARA 建議學生需評估該等服務或練習是否適合他們才選擇購買。
(c) 匯報評估結果		
描述	簡介	詳情
量度考生表現的指標	<ul style="list-style-type: none"> 應考 NAPLAN 的學生的成績按 5 個全國考評量表(national achievement scale)評核，分別顯示考生在閱讀、寫作、計算、拼寫和標點符號等範疇的能力。 每個全國考評量表均分有 10 個等級，等級的高低代表經 NAPLAN 評估的 3 年級至 9 年級學生所能掌握的知識和能力的多寡。每個全國考評量表均附有全國最低標準以作比較。⁸ ACARA 按個別考生在某個範疇所取得的分數("原始分數")轉換成"等級分數"(scaled score)，⁹ 然後按"等級分數"的高低歸入就該範疇訂定的全國考評量表內的其中一個等級。學生的"等級分數"若等於或高於某等級的達標分數，便會歸入該等級。 	<ul style="list-style-type: none"> 就每個年級而言，學生報告只載列 10 個等級的其中 6 個，詳情如下： <ol style="list-style-type: none"> 3 年級學生的成績都會顯示在等級 1 到等級 6 的範圍，全國最低標準為等級 2； 5 年級學生的成績都會顯示在等級 3 到等級 8 的範圍，全國最低標準為等級 4； 7 年級學生的成績都會顯示在等級 4 到等級 9 的範圍，全國最低標準為等級 5；及 9 年級學生的成績都會顯示在等級 5 到等級 10 的範圍，全國最低標準為等級 6。

⁸ 就每個年級所定的全國考評量表內的第二低等級，代表該年級學生應該達到的全國最低標準。全國最低標準是經各方同意可以接受的最低知識和能力標準。如學生達不到此標準，將難以在學校學習上有所長進。

⁹ 由於不同年份的試卷的難度可能有輕微差異，將"原始分數"轉化為"等級分數"，以確保在不同年份進行的評估可資比較。

澳洲的學生能力評估制度

(c) 匯報評估結果(續)		
描述	簡介	詳情
評估報告	ACARA 提供學生、學校及全國整體的評估表現。	<ul style="list-style-type: none"> 學生成績報告載有個別學生的 NAPLAN 成績，學校的平均成績則公開刊登在 My School 網站。¹⁰ 載有全國的讀寫和計算能力水平資料的全國報告在 ACARA 的網站發布。
比較評估結果	學生成績報告、"My School"網站及全國報告的設計可供學生、學校、州及領地之間互相比較。	<ul style="list-style-type: none"> 學生成績報告可讓個別學生將自己的成績與同校其他學生的成績、全國平均成績、全國最低標準及學校平均成績(只適用於某些州和領地)作比較。 學校可利用 "My School" 網站就其學生所取得的成績，與取錄社會及教育背景相若學生的學校，以及澳洲所有學校作一比較。 全國報告載列每個年級及每項科目範疇的評核結果，並按州／領地、性別、原住民身份、以及父母職業和教育程度分項列出。
運用評估結果	NAPLAN 有助監察學生的學習進度；以及讓學校了解其教學課程的強弱項。	<ul style="list-style-type: none"> 如某所學校的學生的讀寫及計算能力被評為低於平均水平，該校會獲政府分配額外撥款，協助他們提升學生的讀寫及計運的能力。

¹⁰ "My School" 網站提供學校的概況，包括學校課程、學術表現、教職員數目、經費來源、學生背景、收生情況及上課出席率等資料。

澳洲的學生能力評估制度

(d) 公眾對制度的接受程度	
描述	重點
公眾意見	<ul style="list-style-type: none"> 部分持份者認為 NAPLAN 已達到所定的目標，但卻有意見質疑公開 NAPLAN 成績的做法，特別是傳媒按評估結果將學校排名，編製所謂學校"排行榜"(league tables)，所帶來的反效果包括(a)對排名較低的學校造成壓力；及(b)鼓勵學校因"試"施教，例如把讀寫及計算能力的操練置於其他學習範疇之上，以及過分強調達標而忽略了創意和分析技巧的培育。 輿論亦質疑 NAPLAN 作為分析評估工具的成效，因為 NAPLAN 並不包括教師的評估(teacher assessments)，未能全面反映學生的表現。
檢討制度	<ul style="list-style-type: none"> 澳洲參議院曾於 2010 年就 NAPLAN 的施行及公布評估結果的安排進行調查，並於 2013 年就 NAPLAN 的成效進行另一項調查。調查工作完成後發表的報告提出多項建議，包括 ACARA 應(a)探討如何協助成績稍遜的學生；以及(b)為 3 年級至 10 年級每級學生每年進行測驗，以便更準確監察學生的表現¹¹。
已經實行或計劃作出的修訂／改善(如有的話)	<ul style="list-style-type: none"> 澳洲政府計劃於 2017 年推出 NAPLAN Online，旨在： <ul style="list-style-type: none"> (a) 確保能力較高的學生可接受適合他們程度的測試，而能力稍遜的學生亦有機會盡展所長； (b) 擴闊評估範圍；及 (c) 為教師和學校提供更確切及詳細的資料，加深其了解學生應考 NAPLAN 的能力。

¹¹ 政府不同意此項建議，認為進一步擴大評估對象並非是為家長及教師提供更詳盡資料的最佳方案。

加拿大的學生能力評估制度

(a) 制度特色		
描述	簡介	詳情
制度名稱	加拿大全國學業評估計劃 (Pan-Canadian Assessment Program) ("PCAP")。	<ul style="list-style-type: none"> PCAP 在全國施行，用以評估就讀 8 年級的中學生在閱讀、數學及科學 3 學科所能掌握的知識和能力，評估是以作為加拿大官方語言的英語及法語進行。
施行年份	2007 年。	<ul style="list-style-type: none"> 加拿大在 2007 年推出 PCAP，以取代學業成績指標計劃 (School Achievement Indicators Program)，目的是配合課程更動及各省份／地區日漸重視國際評估的趨勢¹²，以及藉 PCAP 為閱讀、數學及科學這 3 核心科目進行評估測驗。
評估對象	所有初中 8 年級學生，不論年齡，均須應考。	<ul style="list-style-type: none"> 在加拿大，中小學教育的修業期一般為 12 年，由 1 年級至 12 年級。在不同的省份和地區，由小學升讀中學的年級各所不同，但學童通常 6 歲(1 年級)入學，11 歲時完成小學課程(6 年級)，然後升讀初中(7 年級至 9 年級)。
評估目的	對學習的評估。	<ul style="list-style-type: none"> 在加拿大，學生因所處的省份／地區不同而所修讀的課程和接受的評估有異。然而，所有加拿大學生均須修讀閱讀、數學及科學 3 科，而 PCAP 旨在評核全國各地學生在這 3 個共通學科所能掌握的知識和能力，從而得出可資比較的數據，以評估處於同一學習階段的學生的學業水平。
評估次數	每 3 年一次。	<ul style="list-style-type: none"> PCAP 每 3 年進行一次，剛於 2016 年舉行。
強制或自願參與評估	強制參與。	<ul style="list-style-type: none"> PCAP 使用兩階段隨機抽樣方式從 8 年級學生中抽選參與評估，首先是從省份／地區隨機揀選應考學校，然後在每所應考學校的 8 年級學生中隨機抽選學生參與評估。 學生只可基於(a)身體障礙；(b)智障；及(c)沒有足夠的英語或法語(PCAP 的作答語言)能力等原因，才可獲豁免參與 PCAP。

¹² PCAP 的設計是要與國際學生評估計劃 (Programme for International Student Assessment 或簡稱 PISA) 所採用的模式吻合。PISA 是由經濟合作及發展組織 (Organisation for Economic Co-operation and Development) 舉辦的國際調查，每 3 年施測一次，旨在測量世界各地 15 歲學生於閱讀、數學及科學這 3 個領域上所能掌握的知識及能力，而每次調查都會輪流針對上述領域當中的一科進行詳細測試。

加拿大的學生能力評估制度

(b) 制度的運作情況		
描述	簡介	詳情
負責評估的機構	教育部長議會 (Council of Ministers of Education) ("CMEC")。	<ul style="list-style-type: none"> CMEC 於 1967 年設立，成員包括各省份及地區負責教育的部長。CMEC 旨在提供平台，讓各教育部長討論及推行彼此關注的項目及政策措施，包括評估加拿大學生的知識及能力。
評估的科目	閱讀、數學及科學。	<ul style="list-style-type: none"> 每次 PCAP 評估都是就閱讀、數學及科學其中的一科進行詳細測試(主考科)，而以其餘兩科作為輔考科。¹³ 在每次 PCAP 評估中，3 個測驗科目會輪流交替成為主考科和輔考科。¹⁴ 此外，每次 PCAP 的主考科與兩年後舉行的 PISA 的主考科均為同一科目。¹⁵
評估的時間	90 分鐘。	<ul style="list-style-type: none"> 測驗限時 90 分鐘。學生如需更多時間，將獲給予額外最多 30 分鐘完成測驗。 所有學生完成答題後，必須填寫一份關於其興趣、學習態度和課堂活動的問卷，填答問卷之時間為 30 分鐘。¹⁶

¹³ 與輔考科比較，主考科的評估更為詳細，參與測試的學生人數亦較多。

¹⁴ 舉例而言，2007 年的主考科是閱讀、2010 年和 2013 年的主考科分別是數學和科學。到了 2016 年的 PCAP，閱讀再次成為主考科，而數學及科學則為輔考科。

¹⁵ 第四屆 PCAP 在 2016 年進行，以閱讀作為主考科。預期在 2016 年應考 PCAP 的 8 年級學生中，大部分會在 15 歲時參與 2018 年的 PISA 評估。由於 2018 年的 PISA 亦會以閱讀作為主考科，加拿大當局遂可就學生在 2016 年 PCAP 的表現與他們在 2018 年 PISA 的表現作一比較。

¹⁶ 除應試的學生外，教師和校長亦須各自填答一份提供背景資料的問卷。教育工作者、研究員及政策制訂的官員可根據問卷數據，研究學生學習能力、情境因素、學習態度與教學策略相互之間的關係。

加拿大的學生能力評估制度

(b) 制度的運作情況(續)		
描述	簡介	詳情
評估模式	<ul style="list-style-type: none"> 筆試包括選擇題及問答題。 閱讀測驗或包括要求學生閱讀一篇短文，然後討論他們對該篇短文的體會。 	<ul style="list-style-type: none"> 在選擇題中，學生獲提供多個選擇，從中選取其答案。題型包括多項選擇題、方格勾選題(check-box)及是非題。 至於問答題，學生須自行寫出答案，他們可被要求以短語(phrases)、兩三短句(sentences)以至數段文字(paragraphs)答題。
所需準備	考生無需準備。	<ul style="list-style-type: none"> 根據 CMEC，PCAP 沒有根據任何特定課程出題，故此學生無須為 PCAP 備考，日常的課堂教學應已為他們作好充分準備應試。
(c) 匯報評估結果		
描述	簡介	詳情
量度考生表現的指標	<ul style="list-style-type: none"> 就主考科而言，學生成績按以下方式匯報：(a)主考科整體平均分及學生達到指定表現水平(performance level)的百分比；以及(b)主考科下各分項的整體平均分。(a)、(b)兩項的成績亦再按省份／地區、作答語言及考生性別區分。 至於其餘兩科輔考科，學生成績僅按省份／地區、作答語言及考生性別區分，匯報整體平均分。 	<ul style="list-style-type: none"> 主考科的評估項目較多，故可再按分項匯報成績。舉例而言，2013 年的 PCAP 評估了學生在 4 個細項(科學本質、生命科學、物理科學及地球科學)的表現。 CMEC 制訂表現水平把考生的分數轉化為 PCAP 所須量度的知識和能力水平。¹⁷ 學生的 PCAP 分數若等於或高於某表現水平的達標分數，該學生的表現可視為已達到該水平。

¹⁷ 在 2013 年的 PCAP 中，應考科學一科的學生成績按 4 級表現水平顯示，第 4 級屬最高級，而第 1 級則視為可以接受或達到"基本水平"(即 8 年級學生普遍均能達到的水平)。

加拿大的學生能力評估制度

(c) 匯報評估結果(續)		
描述	簡介	詳情
評估報告	CMEC 只在省份／地區及全國層面公開評估結果。	<ul style="list-style-type: none"> CMEC 並未有編制有關個別學校／學生的評估報告。
比較評估結果	各省份／地區可與其他省份／地區及全國平均水平作比較。	<ul style="list-style-type: none"> CMEC 強調評估結果不是用作反映個別學生的學習表現，亦不是供學校用於相互比較。
運用評估結果	<ul style="list-style-type: none"> 在加拿大，不同的省份／地區實施不同的教育制度。當局推行 PCAP 以便在閱讀、數學及科學 3 科上創建一套共通評核基準，即使全國的教育制度不同，亦有互相比較的基礎。 由於 PCAP 將各省份／地區納入一套共通評核基準內，即使彼此的教育政策可能有別，也能互相比較。 	<ul style="list-style-type: none"> 各省份／地區可根據其在 PCAP 取得的評估結果，與其他省份／地區、全加以至國際的評估結果比較，從而掌握境內 8 年級學生的學業水平，並以此為基礎檢視課程成效及其他教育制度的環節。 PCAP 亦旨在量度剛升上初中的小學生所能掌握的知識和能力，尤其是在核心科目範疇的表現。當局可藉此得悉學童接受正規教育後數年間所取得的教學成果，如有不足之處亦可在他們中學畢業前彌補過來。
(d) 公眾對制度的接受程度		
描述	重點	
公眾意見	<ul style="list-style-type: none"> 由於並非所有學生均須參與 PCAP 評估，抽樣調查可能難以真正反映不同省份／地區學生的學習表現。 另一個關注點是 PCAP 屬低風險評估，或令學生失卻參與評估的動機，以及難以認真對待 PCAP，積極投入。 	
檢討制度	<ul style="list-style-type: none"> 聯邦政府近年並無就 PCAP 進行任何檢討。 	

南韓的學生能力評估制度

(a) 制度特色		
描述	簡介	詳情
制度名稱	全國教育成績評估(National Assessment of Educational Achievement)("NAEA")。	<ul style="list-style-type: none"> NAEA 評估所有 9 年級及 11 年級學生在國定課程中韓文、數學及英文科所能掌握的知識和能力。
施行年份	2000 年。	<ul style="list-style-type: none"> 有關 NAEA 的發展大綱於 1998 年制訂，倡議推行測試以評估國定課程的成效和監察學生的學習進度。NAEA 經兩年時間落實，在 2000 年開始施行。
評估對象	所有 9 年級及 11 年級的學生。	<ul style="list-style-type: none"> 南韓的中小學教育包括 6 年小學(1 年級至 6 年級)、3 年初中(7 年級至 9 年級)及 3 年高中(10 年級至 12 年級)。 NAEA 最初以隨機抽樣方式挑選 6 年級、9 年級及 10 年級學生應試，但自 2009 年起 NAEA 改為測試全國的 6 年級、9 年級及 10 年級學生。擴闊 NAEA 的評估對象旨出更能識別學業成績稍遜的學生，在學習上向他們提供適切的支援，力求所有學生的學業表現不會低於全國最低標準。 當局其後改革 NAEA 以減輕學生的壓力，作出下列兩項的安排：(a)由 2010 年起由 11 年級學生取代 10 年級學生應考；及(b)在 2013 年豁免 6 年級學生參加 NAEA 評估。
評估目的	NAEA 旨在提供學生在讀寫和運算方面的成績數據。	<ul style="list-style-type: none"> NAEA 利用評估所提供的數據以評核全國的教育事業及國定課程的實施是否適合學生的需求。
評估次數	每年評估。	<ul style="list-style-type: none"> 在每年 6 月／7 月間進行測試。
強制或自願參與評估	強制參與。	<ul style="list-style-type: none"> 所有 9 年級及 11 年級學生均必須參加評估，惟就讀"不符合評估資格學校"的學生除外。¹⁸

¹⁸ "不符合評估資格學校"指並非施教國定課程的學校，其中包括開辦自設課程的職業學校及專科學校。

南韓的學生能力評估制度

(b) 制度的運作情況		
描述	簡介	詳情
負責評估的機構	韓國教育課程評鑑院 (Korea Institute for Curriculum and Evaluation)("KICE")。	<ul style="list-style-type: none"> KICE 是政府資助從事教育研究的機構，負責制訂國定課程及舉行評估和測試。
評估的科目	韓文、數學及英文。	<ul style="list-style-type: none"> 在 2013 年，評估科目由 5 科(韓文、數學、英文、科學及社會)減至現時的 3 科(韓文、數學及英文)，以減輕應考 NAEA 學生的壓力。¹⁹
評估的時間	每科測驗限時 60 分鐘。	<ul style="list-style-type: none"> 在 2010 年，按科目劃分的測驗時間由 70 分鐘縮短至 60 分鐘。 所有學生亦須填答一份問卷，限時 20 分鐘。
評估模式	兼具紙筆測驗和問卷。	<ul style="list-style-type: none"> 每科測驗包括 30 至 40 道題目，當中七至八成屬多項選擇題，其他題型包括問答題。 學生均須填答關於他們背景資料的問卷。²⁰
所需準備	無需準備。但學生及家長可取閱歷屆試卷。	<ul style="list-style-type: none"> NAEA 旨在評核學生修讀國定課程後所能掌握的語文及數學的知識及能力。
(c) 匯報評估結果		
描述	簡介	詳情
量度考生表現的指標	NAEA 旨在評估學生對國定課程所涵蓋內容的掌握程度。學生在 NAEA 每科測驗中所取得的分數按 4 級"學業成績水平"評核。	<ul style="list-style-type: none"> 4 級"學業成績水平"分別是"優異"、"良好"、"合格"及"低於合格水平"。全國最低標準訂於"合格"水平。 學生的 NAEA 分數若等於或高於某水平的達標分數，該學生的表現可視為已達到該水平。

¹⁹ 自 2013 年起，當局取消強制所有 9 年級及 11 年級學生應考科學及社會科的測試，現時以隨機抽樣方式挑選若干名 9 年級學生參與該兩科的測試。

²⁰ NAEA 亦向校長發出問卷，收集有關他們學校的背景資料。

南韓的學生能力評估制度

(c) 匯報評估結果(續)		
描述	簡介	詳情
評估報告	KICE 提供學生、學校及全國整體的評估表現。	<ul style="list-style-type: none"> 學生成績報告載有個別學生的 NAEA 成績。 個別學校的 NAEA 平均成績上載於學校資料披露系統 (School Information Disclosure System), ²¹ 而載有全國平均成績的全國報告則在 KICE 負責管理的網站披露。
比較評估結果	除學校之間比較外, 各道/廣域市亦互相比較。	<ul style="list-style-type: none"> 法例訂明個別學校須在學校資料披露系統披露與 NAEA 相關的資料, 例如學生應試率及他們在各級學業成績水平所佔的百分比。 全國報告載列每科的評估結果, 並按考生性別及學業成績水平各級的百分比再分項。
運用評估結果	為當局提供數據以評估全國教育事業發展的情況。	<ul style="list-style-type: none"> 自 2013 年起, 政府不再按 NAEA 成績向學校問責, 例如學生的 NAEA 成績高低再不成為政府出資向學校提供支援措施所要考慮的因素。²² 政府改為實施了一項名為 "夢想學校"(Do-Dream schools) 的新措施, 根據教師所作的建議、心理評估及其他評核來識別學業成績稍遜的學生, 為他們提供支援, 而非以 NAEA 的評估結果作為依據。 除學習支援外, "夢想學校" 透過提供個人化支援、輔導及課後支援等服務, 全方位協助學業成績稍遜的學生, 以提高其學習動機、自尊及社交技巧。
(d) 公眾對制度的接受程度		
描述	重點	
公眾意見	<ul style="list-style-type: none"> 教師工會、家長及其他持份者質疑當局公開 NAEA 成績的做法, 此舉帶來不少反效果, 例如對排名較低的學校造成壓力、引發學校之間的競爭, 以及驅使學校採取因"試"施教的方針。 	
檢討制度	<ul style="list-style-type: none"> 教育部於 2013 年推行 "愉快教育: 讓學生實現夢想與發揮才華" 的政策, 目的是 (a) 促進學生身心均衡發展; (b) 聚焦於為學生減壓; 及 (c) 提供充實的學習機會。據此政策實施的優化措施, 包括上文所述在小學階段廢除進行 NAEA; 以及學校的 NAEA 成績高低再不成為政府出資提供支援措施所要考慮的因素。 	

²¹ 學校資料披露系統是一個由教育部負責管理的網站, 為公眾提供有關學校的概況, 例如校內設施、收生率及就讀學生的成績。

²² 在 2013 年之前, 南韓政府要求學校就其學生在 NAEA 的表現問責, 並透過各項措施(例如加設補習班), 以支援 NAEA 成績欠佳的學生, 提升他們的學習表現。

英國的學生能力評估制度²³

(a) 制度特色		
描述	簡介	詳情
制度名稱	全國課程評估 (National curriculum assessment)。	<ul style="list-style-type: none"> 全國課程評估，亦稱為 SAT (Standard Assessment Tests)，指按法例為位處英格蘭的小學進行的評估。
施行年份	1991 年。	<ul style="list-style-type: none"> 英格蘭於 1991 年首次引進全國性評估，而相關制度隨後經歷數次改動。最近一次改動在 2014 年 9 月發生，英國政府為所有政府資助學校頒佈全新的國定課程，²⁴ 而按全新的國定課程而出題的首屆評估於 2016 年舉行。 據英國政府所述，就 2016 年開始進行的評估而言，其要求較之前為高，而期望學生所能達到的水平亦較高，其用意是評估學童是否已作好準備接受中學教育。
評估對象	就讀主要階段 1 及 2 最後一年的小學生。	<ul style="list-style-type: none"> 在英格蘭，國定課程分為 4 個主要階段：(a) 主要階段 1：5 歲至 7 歲 (1 年級至 2 年級)；(b) 主要階段 2：7 歲至 11 歲 (3 年級至 6 年級)；(c) 主要階段 3：11 歲至 14 歲 (7 年級至 9 年級)；及 (d) 主要階段 4：14 歲至 16 歲 (10 年級至 11 年級)。小學屬主要階段 1 及主要階段 2，中學則屬主要階段 3 及主要階段 4。
評估目的	提供數據以評估學生的學習表現，並就此向學校問責。	<ul style="list-style-type: none"> 主要階段 1 的評估測驗旨在量度幼童所達到的學業水平，以了解他們自就學以來的學習進展，而主要階段 2 的評估測驗則核實學生在多大程度上已準備好接受中學教育。 主要階段 2 的評估測驗亦旨在提供數據，以比較不同學校的表現，並藉此推動學校改善求進。

²³ 雖然英國 4 個構成國所採用的教育制度相若，但在學童入學年齡以至學生們接受評估測驗的類別等方面，仍存有不少重大差異。為簡明起見，本資料摘要有關英國的研究範圍僅限於英格蘭一地。

²⁴ 國定課程的涵蓋範圍包括英格蘭所有年齡介乎 5 歲至 16 歲的學生。國定課程旨在劃一中小學的授課內容，不但指明須予教授的科目，亦訂明修讀這些科目的學生應達到的水平。根據法例，政府資助學校有責任依循國定的課程授課，但其他類別的學校，例如公辦民營學校 (academies) 及私立學校，則不在此限。公辦民營學校為公帑資助的獨立民營學校，而私立學校則沒有政府資助，學生須繳交學費。

英國的學生能力評估制度

(a) 制度特色(續)		
描述	簡介	詳情
評估次數	每年評估。	<ul style="list-style-type: none"> • 評估通常於每年 5 月舉行。 • 教師須在 5 月份內安排舉行主要階段 1 的評估測驗，但確實的日期則由學校自行決定。學校可將測驗分開在數日甚至數星期舉行，讓學生盡可能在輕鬆的環境下進行評估測驗。 • 與主要階段 1 的評估測驗相反，主要階段 2 的評估測驗必須在指定日期進行，而且通常指定在 5 月份某一個星期內舉行。
強制或自願參與評估	強制參與。	<ul style="list-style-type: none"> • 所有政府資助學校均必須參與評估測驗。其他類別的學校可自行選擇是否參與。 • 校長可基於學生的個別情況而豁免他們接受評估測驗，例如學生有嚴重學習困難，或因重病而缺課多日。
(b) 制度的運作情況		
描述	簡介	詳情
負責評估的機構	標準及考試局(The Standard and Testing Agency)。	<ul style="list-style-type: none"> • 標準及考試局是教育部轄下的行政機關，負責為 14 歲(即主要階段 3 最後一年的學生)或以下的學童設計評估測驗，並舉行有關測驗。
評估的科目	<ul style="list-style-type: none"> • 主要階段 1：閱讀及數學。 • 主要階段 2：(a)閱讀；(b)語法、標點符號及拼寫；(c)數學；及(d)科學。 	<ul style="list-style-type: none"> • 在主要階段 1 結束時亦設有語法、標點符號及拼寫評估測驗，但學校可選擇是否舉行該項測驗。 • 主要階段 2 的科學評估測驗每兩年舉行一次，並將於 2018 年舉行。科學評估測驗屬隨機抽樣的測試，即只會抽選某一百分比的學校舉行測驗。

英國的學生能力評估制度

(b) 制度的運作情況(續)		
描述	簡介	詳情
評估的時間	<ul style="list-style-type: none"> 2017 年舉行的主要階段 1 評估測驗時限：閱讀(1 小時 10 分鐘)；及數學(55 分鐘)。 2017 年舉行的主要階段 2 評估測驗時限：閱讀(1 小時)；語法、標點符號及拼寫(1 小時)；及數學(1 小時 50 分鐘)。²⁵ 	<ul style="list-style-type: none"> 與主要階段 2 評估測驗不同，主要階段 1 評估測驗的時限無須嚴格執行。學校可自行決定學生完成評估測驗的所需時間，學生是否具有快速作答的能力，並非須予評估的其中一環。
評估模式	<ul style="list-style-type: none"> 兼具筆試及教師評估兩部分，而筆試採用多種不同題型。²⁶ 	<ul style="list-style-type: none"> 在主要階段 1 及主要階段 2 結束時，教師須提交教師評估報告，述明所教學生在 4 個核心科目(即寫作、閱讀、數學及科學)的學習進度。²⁷ 教師的評估報告，連同主要階段 1/主要階段 2 評估測驗結果，應可全面反映學生的學習表現。²⁸ 教師在進行教師評估時必須考慮學生日常的課堂表現，尤其是在書寫、實踐學習及言語表達方面的能力。教師在主要階段 1 結束時就學生所作的教師評估，亦必須考慮主要階段 1 的評估測驗結果。
所需準備	歷屆試卷及例題可在政府網站下載使用。	<ul style="list-style-type: none"> 應考學生可購買商業機構所提供的備考服務及/或練習，而部分學校亦會在考試前為學生(尤其是須參加主要階段 2 評估測驗的學生)安排補課。

²⁵ 主要階段 2 科學科的評估測驗每兩年舉行一次，該測驗不會在 2017 年舉行。

²⁶ 舉例而言，主要階段 1 閱讀試的題型包括多項選擇題、是非題、配對題、短答題及開放式提答。主要階段 1 數學試的題型包括多項選擇題、配對題、是非題、繪製圖表及繪畫圖形，而部分題目亦要求應考學生列出或解釋運算過程。

²⁷ 舉例而言，修畢主要階段 1 課程的學生會收到教師在寫作、閱讀及數學科所作的評估，評估結果以下述 3 個水平的其中一個顯示：(a)邁向預期水平；(b)達到預期水平；及(c)達到較高水平。然而，主要階段 1 科學科只有一個標準：是否達到預期水平。

²⁸ 主要階段 1/主要階段 2 評估測驗只顯示學生於某一測試日在某一科的選定範疇所達到的學業水平，而教師評估則是教師對學生整個學年在該科目的整體表現的評鑒。

英國的學生能力評估制度

(c) 匯報評估結果		
描述	簡介	詳情
量度考生表現的指標	學生在主要階段 1／主要階段 2 評估測驗中取得的分數("原始分數")，經調整後轉換成"等級分數"(scaled score)， ²⁹ 而 100 分則視為平均學生達標的分數。 ³⁰	<ul style="list-style-type: none"> 在夏季學期臨近結束時，家長會收到一份報告，載列其子女所取得的"原始分數"、經調整後的"等級分數"，以及所得的"等級分數"是否達到全國平均標準。家長亦會收到由學校另行發出的教師評估報告。
評估報告	<ul style="list-style-type: none"> 主要階段 1 評估測驗結果只供家長查閱。 主要階段 2 評估測驗提供學生、學校及整個地區的評估結果。 	<ul style="list-style-type: none"> 主要階段 1 評估測驗由校內教師負責批改試卷。學校不會公布測驗結果，亦不會將測驗結果送交政府，但家長可查閱子女所考取的成績。 主要階段 2 評估測驗由校外人士負批改試卷，家長可查閱子女所考取的成績，而教育部亦會按學校及地區層面(local-authority area)分項在教育部網站的"學校成績與水平列表"登載測驗結果。
比較評估結果	透過"學校成績與水平列表"比較不同學校在主要階段 2 評估測驗中取得的成績。	<ul style="list-style-type: none"> "學校成績與水平列表"的設計令學校可與(a)收生情況相若的學校及(b)其他位處英格蘭的學校在主要階段 2 評估測驗中所取得的成績作一比較。
運用評估結果	<ul style="list-style-type: none"> 主要階段 1 及主要階段 2 的評估有助教師掌握學生所達到的學業水平；如有需要的話，教師會為學生訂定適當的支援措施，提升他們的學業成績。 主要階段 2 評估測驗結果亦會用於編製"學校成績與水平列表"，藉以令學校為其學生的學業水平及學習進度問責。 	<ul style="list-style-type: none"> "學校成績與水平列表"旨在協助"英國教育及兒童服務與技能標準局"(Office for Standards in Education, Children's Services and Skills)進行視學工作。³¹ 英國"教育及兒童服務與技能標準局"在為其學校確立視學評鑒時，會研究該學校在列表的排名高低。若該局認為某學校的整體教學成效未如理想，有關督學必須決定，應對該校採取"特殊措施"，還是向其發出"改善通知書"。 家長為子女選擇學校時亦會參考"學校成績與水平列表"。

²⁹ 由於不同年份的試卷的難度可能有輕微差異，將"原始分數"轉化為"等級分數"，以確保在不同年份進行的評估可資比較。

³⁰ 學生得分低於 100 分表示他們需要在學習方面取得支援，得分超過 100 分則表示他們較同齡學生達到較高的學業水平。

³¹ "英國教育及兒童服務與技能標準局"是一個非部長級的政府部門，負責督導及規管(a)照顧兒童及青少年的服務和(b)為不同年齡人士提供教育及技能學習的服務。

英國的學生能力評估制度

(d) 公眾對制度的接受程度	
描述	重點
公眾意見	<ul style="list-style-type: none"> 全國課程評估備受家長、教師工會及其他持份者批評，指該評估為學童造成過多壓力，尤其是在主要階段 1 終結時所進行的評估，可能令兒童就學不久已經充滿挫敗感。 輿論亦質疑主要階段 2 評估測驗定位為高風險測驗的設計，令教學過份聚焦於測驗內容，因為學校都希望在測驗中取得優異成績，令學校在"學校成績與水平列表"中取得較高的排名。 SAT 水平評估測驗在 2016 年 5 月引進了新的測試模式，惹來更多的批評。英國的小學生從 2016 年起每年應考按新模式設計的 SAT 水平評估測驗，題目較往年艱深³²，這一改動引起了部份老師和學生家長的抗議和抵制。更令人關注的是，約半數參加 2016 年主要階段 2 評估測驗的學生未能取得達標的分數。
檢討制度	<ul style="list-style-type: none"> 在 2016 年 9 月，下議院教育委員會宣布就全國課程評估進行調查，其中的重點包括審查 (a)舉行 SAT 水平評估測驗的目的，以及新的評估制度是否達到預期的目的；(b)在小學階段參加評估測驗的利弊，以及(c)隨後所須採取的改善措施。教育委員會已與主要持份者舉行了 3 場聽證會，但仍未就 SAT 水平評估測驗完成總結報告。

³² 根據英國政府所述，新推行的國定課程的設計，可以與世界上最好的教育制度所推行的媲美，故此學童應達到較高的學業水平。提高對學童的要求是希望他們離開小學時，能掌握閱讀、寫作和運算的能力，在中學求學階段以至成年期發揮他們的潛能。請參閱 Standards & Testing Agency (2016d)。

美國的學生能力評估制度

(a) 制度的主要特點		
描述	簡介	詳情
制度名稱	全國教育進展評估 (National Assessment of Education Progress) ("NAEP")。	<ul style="list-style-type: none"> • NAEP 又稱為 "全國成績表" (Nation's Report Card)，用以評估中小學生的學習表現。NAEP 分為 (a) NAEP 主要評估 (NAEP main assessment) 和 (b) NAEP 長期趨勢評估 (NAEP long-term trend assessment)。前者的評估對象為 4 年級、8 年級及 12 年級的學生，而後者的評估對象則為年滿 9 歲、13 歲及 17 歲的學生。 • 當局藉 NAEP 主要評估和 NAEP 長期趨勢評估收集和公布有關全國學生的學習表現數據，而 NAEP 主要評估更會在州份及地區層面收集及公布關於若干評估科目的數據。
施行年份	1969 年。	<ul style="list-style-type: none"> • 進行全國性學生能力評估的構思在 1960 年代初浮現，當局在 1964 年開始規劃推行 NAEP，而首次全國性評估於 1969 年展開。州份層面的 NAEP 評估報告於 1990 年以試行方式向外公布，其後成為常設安排，至於地區層面的 NAEP 評估報告則於 2002 年開始向外公布，但相關的評估只會在一些大城市進行。
評估對象	<ul style="list-style-type: none"> • NAEP 主要評估：4 年級、8 年級及 12 年級的學生。 • NAEP 長期趨勢評估：年屆 9 歲、13 歲及 17 歲的學生。 	<ul style="list-style-type: none"> • 在美國，兒童約 6 歲入讀小學 1 年級。約 5 至 6 年後，他們接受初中及高中教育 (初中通常是 3 年，由 6 年級至 8 年級；高中通常是 4 年，由 9 年級至 12 年級)。學生高中畢業後，可升讀學院 (college) 或大學 (university)。
評估目的	評估美國整體學生的學業水平和學習進度。	<ul style="list-style-type: none"> • NAEP 的兩大目標分別為：(a) 比較各州份學生的學業水平；及 (b) 監察學生的學習進度。為達致上述目標，有關當局在全國抽選具有代表性的學生參加 NAEP 主要評估或 NAEP 長期趨勢評估。

美國的學生能力評估制度

(a) 制度的主要特點(續)		
描述	簡介	詳情
評估次數	<ul style="list-style-type: none"> • NAEP 主要評估：每年 1 次。 • NAEP 長期趨勢評估：每 4 年 1 次。 	<ul style="list-style-type: none"> • NAEP 主要評估於每年 1 月至 3 月期間進行。 • NAEP 長期趨勢評估每 4 年進行 1 次：年滿 17 歲的學生在春季進行測試，年滿 13 歲及 9 歲的學生則分別在秋季和冬季應考。
強制或自願參與評估	學生和學校可選擇是否參加評估。	<ul style="list-style-type: none"> • 學校及學生皆以隨機抽樣的形式被揀選接受評估。³³ 然而，接受指定教育撥款的州份 (Title 1 Funding) 則必須參加為 4 年級及 8 年級學生而設的 NAEP 閱讀及數學科評估，但獲抽選的學生仍可選擇是否參加。
(b) 制度的運作情況		
描述	簡介	詳情
負責評估的機構	全國教育統計中心 (National Center for Education Statistics)。	<ul style="list-style-type: none"> • 全國教育統計中心設置於國家教育科學研究院 (Institute of Education Sciences) 下，屬教育部 (Department of Education) 的執行機構，為收集及分析美國及其他國家的教育數據而設的主要聯邦機構。
評估的科目	涵蓋多個科目，包括數學、閱讀、寫作及科學	<ul style="list-style-type: none"> • 就 NAEP 主要評估而言，科目範疇每年不同。最經常進行評估的科目為數學、閱讀、科學及寫作。³⁴ 其他科目如藝術、公民教育、經濟、地理、美國史等，亦不時被選為評估科目。雖然評估科目眾多，但每名獲抽選的學生只會應考其中一科。 • 就 NAEP 長期趨勢評估而言，評估科目則為閱讀及數學兩科。某些題目在歷屆的評估重複出現，以確保在不同年份進行的評估可資比較。

³³ 獲抽選舉行 NAEP 的學校須具代表性，反映其所屬州份的學生人口特徵及校區分佈結構。為確保獲抽中參加評估的學生來自大規模及小規模的學校的比例適當，大規模學校獲抽中的機會較小規模學校為高。在每所獲選抽中的學校，再以隨機抽選方式選取學生參加評估。

³⁴ 當局會每兩年評估 4 年級及 8 年級學生在閱讀及數學方面的表現。

美國的學生能力評估制度

(b) 制度的運作情況(續)		
描述	簡介	詳情
評估的時間	學生一般須在 60 分鐘內完成作答。	<ul style="list-style-type: none"> 在評估時，學生可隨時終止作答，或略過作答部份問題。
評估模式	多項選擇題及開放式提答。學生亦須填答一份問卷。	<ul style="list-style-type: none"> 評估於日常上學時間進行。 在 2016 年，NAEP 的數學及閱讀評估試行以附連鍵盤、手寫筆及耳筒的平板電腦進行。部分題目採用多媒體材料，例如影音材料。 學生所填答的問卷旨在收集他們的背景、校內外的學習機會，以及其學習經歷的資料。³⁵
所需準備	學生無需準備。然而，全國教育統計中心於每次評估結束後，都會公開部分 NAEP 試題，學生可登入中心網站下載有關試題。	<ul style="list-style-type: none"> 全國教育統計中心認為學生無須為 NAEP 作特別準備，用作評估的試題設計，只着重測試考生在 4 年級、8 年級及 12 年級的日常課堂中所應掌握的知識。

³⁵ 有關當局亦向教師及學校行政人員發出問卷，以收集有關教師培訓、教學方式及學校政策方面的資料。

美國的學生能力評估制度

(c) 匯報評估結果		
描述	簡介	詳情
量度考生表現的指標	NAEP 所評估的科目成績以整體學生的平均標準得分 (average scale scores) 和學生達到或超越某個能力水平的百分比顯示。	<ul style="list-style-type: none"> • 平均標準得分顯示學生在評估中的表現。一般而言，NAEP 評估科目的得分介乎 0 分至 500 分之間。 • 全國教育統計中心將能力水平分為基礎(basic)、精熟(proficient)及進階(advanced)3 級水平，並按學生在 NAEP 的表現定出他們是否能達到或超越某一級的能力水平。
評估報告	全國教育統計中心在全國層面公開評估結果，披露全國的整體成績，而部分評估亦會公布州份及地區在 NAEP 所取得的成績。	<ul style="list-style-type: none"> • 當局不會公布個別學校／學生的評估結果。 • 當局亦不會每年發放州份層面的評估結果。舉例而言，4 年級及 8 年級數學及閱讀科的州份整體成績每兩年公布一次，亦只會包括公立學校的資料。 • 在若干大城市(例如紐約市及波士頓)，4 年級及 8 年級數學及閱讀科會有地區層面的整體成績，而有關數據只會每兩年公布一次。
比較評估結果	評估數據可供州份及地區之間互相比較。	<ul style="list-style-type: none"> • NAEP 不計算個別學生的得分。然而，有關當局會按學生的人口特徵分項載列評估結果，例如按學生的種族／性別分類的分項數字。
運用評估結果	政府利用評估結果監察學生的學習進度，並就此制訂改進美國教育制度的措施。	<ul style="list-style-type: none"> • 學校、家長、政策制訂者及研究人員均有利用評估數據。

美國的學生能力評估制度

(d) 公眾對制度的接受程度	
描述	重點
公眾意見	<ul style="list-style-type: none"> 雖然 NAEP 自 1969 年開始運作至今，並號稱為學生能力評估的"黃金標準"，³⁶ 但有輿論認為將能力水平劃分為基本、精熟及優異三級的做法，過於主觀武斷，缺乏有力的科學論據支持。 目前全國教育統計中心在設定能力水平時，首先要釐訂不同等級水平中，考生最少需要獲取的分數，而每個等級的最低分數稱為臨界得分。由於 NAEP 的能力水平按臨界得分釐定，部份持份者認為有可能令學生評估結果有欠準確。
檢討制度	<ul style="list-style-type: none"> 美國政府問責辦公室 (Government Accountability Office) 於 1990 年代及 2000 年代就 NAEP 曾進行數次檢討，內容包括 NAEP 訂定學生能力水平的方式、進行評估所涉及的費用及殘疾學生參加 NAEP 的情況。全國教育統計中心近年成立了一個小組，負責討論、研究及構思 NAEP 未來日子的發展路向，並提出多項建議。 小組提出了多項建議，包括減少以能力水平衡量學生表現，以及在報告能力水平結果時應更清晰闡明所提供資料的局限性。小組亦建議應更着重其他評估數據，例如整體學生的平均標準得分和學生表現差距的標準差 (standard deviation)。 此外，有意見提議須重新設計 NAEP 以提供個別學生／學校的評估結果，小組認為這提議並不可行，因為會大幅增加整項評估的複雜性和進行評估的費用。此外，各州份現時進行的其他評估已可提供有關個別學生／學校的數據。

³⁶ 根據全國教育統計中心所述，NAEP 經常被奉為學生能力評估的"黃金標準"，因為 NAEP 是由全美最熟悉評估機制及設計考試內容的專家、教育專家及教師共同研發。請參閱 An Introduction to NAEP (2016)。

參考資料

澳洲

1. *Australian Curriculum, Assessment and Reporting Authority*. (2017) Available from: <http://www.acara.edu.au/> [Accessed March 2017].
2. Australian Government. (2011) *Australian Government Response to the Senate Education, and Employment and Workplace Relations References Committee Report on the Administration and Reporting of NAPLAN Testing*.
3. Australian Government. (2014) *Australian Government Response to the Senate Education and Employment References Committee Report – Effectiveness of the National Assessment Program – Literacy and Numeracy*.
4. Australian Government. (2017) *Australian Education System*. Available from: <https://www.studyinaustralia.gov.au/global/australian-education/education-system> [Accessed March 2017].
5. *My School*. (2017) Available from: <http://myschool.edu.au/> [Accessed March 2017].
6. National Assessment Program. (2017) *The National Assessment Program*. Available from: <https://www.nap.edu.au/> [Accessed March 2017].
7. The Senate, Education and Employment References Committee. (2014) *Final Report on the Effectiveness of the National Assessment Program – Literacy and Numeracy*. Available from: http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Education_and_Employment/Naplan13 [Accessed March 2017].
8. The Senate, Education, Employment and Workplace Relations References Committee. (2010) *Final Report on the Administration and Reporting of NAPLAN Testing*. Available from: http://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Education_Employment_and_Workplace_Relations/Completed_inquiries/2010-13/naplan/index [Accessed March 2017].

加拿大

9. Andrews, C. et al. (2007) *Compulsory assessment systems in the INCA countries: Thematic Probe*. Qualifications and Curriculum Authority. Available from: <https://www.publications.parliament.uk/pa/cm200607/cmselect/cmmeduski/memo/test&ass/ucm3102paper5.pdf> [Accessed March 2017].
10. Council of Ministers of Education. (2007) *What is the Pan-Canadian Assessment Program*. Available from: <http://www.cmec.ca/docs/pcap/pcap2007/chapter1.pdf> [Accessed March 2017].
11. Council of Ministers of Education. (2014) *PCAP 2013: Report on the Pan-Canadian Assessment of Science, Reading, and Mathematics*. Available from: <http://www.cmec.ca/Publications/Lists/Publications/Attachments/337/PCAP-2013-Public-Report-EN.pdf> [Accessed March 2017].
12. Council of Ministers of Education. (2016a) *PCAP 2016 Assessment Framework*. Available from: <http://www.cmec.ca/docs/pcap/pcap2016/PCAP-2016-Assessment-Framework.pdf> [Accessed March 2017].
13. Council of Ministers of Education. (2016b) *PCAP 2016 Reading, Mathematics, and Science Assessment: Handbook for Schools*. Available from: http://www.cmec.ca/docs/pcap/pcap2016/PCAP2016_Handbook_Schools_EN.pdf [Accessed March 2017].
14. Council of Ministers of Education. (2017) *Overview: Pan-Canadian Assessment Program (PCAP)*. Available from: <http://www.cmec.ca/131/Programs-and-Initiatives/Assessment/Overview/index.html> [Accessed March 2017].
15. Ho, S.C. (2013) *Asia-Pacific Education System Review Series: Student Learning Assessment*. United Nations Educational Scientific and Cultural Organisation & Hong Kong Institute of Educational Research. Available from: <http://unesdoc.unesco.org/images/0021/002178/217816E.pdf> [Accessed March 2017].
16. Klinger, D. et al. (2016) Chapter 9: Canada: the Intersection of International Achievement Testing and Educational Policy Development. In: Volante, L. (ed.) *The Intersection of International Achievement Testing and Educational Policy*. Routledge.

17. Organisation for Economic Co-operation and Development. (2015) *Education Policy Outlook: Canada*. Available from: <http://www.oecd.org/education/policyoutlook.htm> [Accessed March 2017].
18. Rabbny, R. et al. (2014) Mining Large Scale Data from National Educational Achievement Tests. A Case Study. In: *Workshop on Data Mining for Educational Assessment and Feedback (ASSESS 2014)*. 24 August. New York. Available from: http://www.aspiringminds.com/pages/assess/2014/camera_ready/poster/rabbany_etal.pdf [Accessed March 2017].
19. Statistics Canada. (2012) *Handbook for the Pan-Canadian Education Indicators Program*. Available from: http://publications.gc.ca/collections/collection_2012/statcan/81-582-g/81-582-g2012001-eng.pdf [Accessed March 2017].

香港

20. Education Bureau. (2014) *Review on the Implementation of Territory-wide System Assessment*. Available from: <http://www.legco.gov.hk/yr13-14/english/panels/ed/papers/ed0113cb4-284-3-e.pdf> [Accessed March 2017].
21. Education Bureau. (2016a) *Issues Related to the Suspension, Continuation or Abolition of the Territory-wide System Assessment*. Available from: <http://www.legco.gov.hk/yr15-16/english/panels/ed/papers/ed20160111cb4-435-1-e.pdf> [Accessed March 2017].
22. Education Bureau. (2016b) *Key Performance Measures (with effect from the 2016/17 school year)*. Available from: <http://www.edb.gov.hk/en/sch-admin/sch-quality-assurance/performance-indicators/kpm/> [Accessed March 2017].
23. Education Bureau. (2016c) *Latest Development on the Review of the Territory-wide System Assessment*. Available from: <http://www.legco.gov.hk/yr16-17/english/panels/ed/papers/ed20161212cb4-262-2-e.pdf> [Accessed March 2017].

24. Education Bureau. (2016d) *Review of the Territory-wide System Assessment*. Available from: <http://www.legco.gov.hk/yr15-16/english/panels/ed/papers/ed20160322cb4-729-3-e.pdf> [Accessed March 2017].
25. Education Bureau. (2016e) *Territory-wide System Assessment 2016 (Primary 6) Quick Guide*. Available from: http://www.bca.hkeaa.edu.hk/web/TSA/en/2016QuickGuideP6/QG_P6_Part1_Eng.pdf [Accessed March 2017].
26. Education Bureau. (2016f) *The Coordinating Committee on Basic Competency Assessment and Assessment Literacy Report on 2016 Tryout Study (Primary 3)*. Available from: <http://www.edb.gov.hk/attachment/en/curriculum-development/tsa/2016fullreport-e.pdf> [Accessed March 2017].
27. Education Bureau. (2016g) *The Coordinating Committee on Basic Competency Assessment and Assessment Literacy Report on Review of the Territory-wide System Assessment*. Available from: <http://www.edb.gov.hk/attachment/en/curriculum-development/tsa/fullreport.pdf> [Accessed March 2017].
28. Education Bureau. (2017) *Basic Competency Assessment Research Study*. Circular Memorandum No. 20/2017. Available from: <http://www.bca.hkeaa.edu.hk/web/TSA/zh/pdf/EDBCM17020E.pdf> [Accessed March 2017].
29. GovHK. (2015a) *On The Record TSA for learning, not drilling*. Available from: http://www.news.gov.hk/en/record/html/2015/06/20150625_192145.lin.shtml [Accessed March 2017].
30. GovHK. (2015b) *Press Releases, LCQ3: Territory-wide System Assessment*. Available from: <http://www.info.gov.hk/gia/general/201511/04/P201511040532.htm> [Accessed March 2017].
31. GovHK. (2016) *TSA proposal submitted*. Available from: http://www.news.gov.hk/en/categories/school_work/html/2016/02/20160204_151403.shtml [Accessed March 2017].

32. GovHK. (2017a) *Press Releases, EDB announces arrangements for Basic Competency Assessment Research Study*. Available from: <http://www.info.gov.hk/gia/general/201701/23/P2017012300338.htm> [Accessed March 2017].
33. GovHK. (2017b) *Trial test extends to all schools*. Available from: http://www.news.gov.hk/en/categories/school_work/html/2017/01/20170123_123643.shtml [Accessed March 2017].
34. *Hong Kong Examinations and Assessment Authority*. (2017a) Available from: <http://www.bca.hkeaa.edu.hk/web/en/Introduction.html> [Accessed March 2017].
35. Hong Kong Examinations and Assessment Authority. (2017b) *TSA, Frequently Asked Questions (FAQ)*. Available from: <http://www.bca.hkeaa.edu.hk/web/TSA/en/Faq.html> [Accessed March 2017].
36. Hong Kong Examinations and Assessment Authority. (2017c) *TSA Report*. Available from: <http://www.bca.hkeaa.edu.hk/web/TSA/en/PriTsaReport.html> [Accessed March 2017].
37. Legislative Council Secretariat. (2016) *Issues related to the implementation of Territory-wide System Assessment*. Updated background brief submitted to the Panel on Education of the Legislative Council. LC Paper No. CB(4)729/15-16(04).

南韓

38. Chung, P. & Hong, W. (2015) Transitions in education policy and interventions for underachievers in South Korea. *KEDI Journal of Educational Policy*, vol. 12, no.2, December, pp.151-171.
39. Korea Institute for Curriculum and Evaluation. (2015a) *Results of NAEA 2013: Cognitive/Affective Characteristics and Trends*.
40. Korea Institute for Curriculum and Evaluation. (2015b) *Results of the 2014 NAEA: Trends in Academic Achievement in Middle Schools*.

41. Korea Institute for Curriculum and Evaluation. (2017a) Available from: <http://www.kice.re.kr/> [Accessed March 2017].
42. Korea Institute for Curriculum and Evaluation. (2017b) *Information Service for National Assessment of Educational Achievement*. Available from: <https://naea.kice.re.kr/> [Accessed March 2017].
43. Ministry of Education. (2017) *The School Information Disclosure System*. Available from: <http://www.schoolinfo.go.kr/> [Accessed March 2017].
44. Organisation for Economic Cooperation and Development. (2016) *Education Policy Outlook Korea*. Available from: <http://www.oecd.org/edu/Education-Policy-Outlook-Korea.pdf> [Accessed March 2017].

英國

45. BBC. (2016a) *Sats protest. Parents speak out over testing*. 3 May. Available from: <http://www.bbc.com/news/education-36192219> [Accessed March 2017].
46. BBC. (2016b) *Teachers' union calls for ballot on primary test boycott*. 27 March. Available from: <http://www.bbc.com/news/education-35904242> [Accessed March 2017].
47. GOV.UK. (2016) *The national curriculum*. Available from: <https://www.gov.uk/national-curriculum/overview> [Accessed March 2017].
48. Hall, K & Øzerk, K. (2008) *Primary Curriculum and assessment: England and other countries*. Research Survey 3/1, University of Cambridge. Available from: <http://cprtrust.org.uk/wp-content/uploads/2014/06/research-survey-3-1.pdf> [Accessed March 2017].
49. House of Commons. (2016) *Hansard: Debate on SATs Results*. 12 July. Available from: <https://hansard.parliament.uk/commons/2016-07-12/debates/3D8FA64F-F38D-4A74-8864-F644097AD952/SatsResults> [Accessed March 2017].

50. Roberts, N. (2016) *Briefing Paper Number 06798: The school curriculum and SATs in England : Reforms since 2010*. 10 November. Available from: <http://researchbriefings.files.parliament.uk/documents/SN06798/SN06798.pdf> [Accessed March 2017].
51. Standards & Testing Agency. (2016a) *2017 key stage 1: assessment and reporting arrangements (ARA)*. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/561720/2017_KS1_ARA_v1.0.pdf [Accessed March 2017].
52. Standards & Testing Agency. (2016b) *2017 key stage 2: assessment and reporting arrangements (ARA)*. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/561723/2017_KS2_ARA_v1.0.pdf [Accessed March 2017].
53. Standards & Testing Agency. (2016c) *Information for parents: 2016 national curriculum tests at the end of key stages 1 and 2*. Available from: <https://www.gov.uk/government/publications/2016-national-curriculum-tests-for-key-stages-1-and-2-information-for-parents> [Accessed March 2017].
54. Standards & Testing Agency. (2016d) *Information for parents: Results from the 2016 national curriculum assessments at the end of key stage 2*. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/534658/Information_for_parents_-_2016_NCT_results_at_the_end_of_key_stage_2.pdf [Accessed March 2017].
55. TheSchoolRun. (2017a) *Changes to KS1 SATs in 2017. what parents need to know*. Available from: <http://www.theschoolrun.com/KS1-SATs-in-2017> [Accessed March 2017].
56. TheSchoolRun. (2017b) *Changes to KS2 SATs in 2017. what parents need to know*. Available from: <http://www.theschoolrun.com/key-stage-2-SATs-2017> [Accessed March 2017].
57. UK Parliament. (2016) *Primary assessment inquiry launched*. 23 September. Available from: <https://www.parliament.uk/business/committees/committees-a-z/commons-select/education-committee/news-parliament-2015/primary-assessment-launch-16-17/> [Accessed March 2017].

美國

58. Center for Public Education. (2006) *Score wars: What to make of state v. NAEP tests*. Available from: <http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/The-proficiency-debate-At-a-glance/Score-wars-What-to-make-of-state-v-NAEP-tests.html> [Accessed March 2017].
59. DC.Gov. (undated) *NAEP Frequently Asked Questions (FAQs)*. Available from: <https://osse.dc.gov/page/naep-frequently-asked-questions-faqs> [Accessed March 2017].
60. Fair Test. (undated) *Is NAEP a Fair and Valid Benchmark?* Available from: <http://www.fairtest.org/naep-fair-and-valid-benchmark> [Accessed March 2017].
61. National Center for Education Statistics. (2010) *An Introduction to NAEP*. Available from: <https://nces.ed.gov/nationsreportcard/pdf/parents/2010468.pdf> [Accessed March 2017].
62. National Center for Education Statistics. (2012a) *NAEP Achievement Levels*. Available from: <https://nces.ed.gov/nationsreportcard/achievement.aspx> [Accessed March 2017].
63. National Center for Education Statistics. (2012b) *NAEP: Looking Ahead – Leading Assessment into the Future*. Available from: https://nces.ed.gov/nationsreportcard/pdf/Future_of_NAEP_Panel_White_Paper.pdf [Accessed March 2017].
64. National Center for Education Statistics. (2016a) *Digitally Based Assessments (DBA)*. Available from: <https://nces.ed.gov/nationsreportcard/dba/> [Accessed March 2017].
65. National Center for Education Statistics. (2016b) *Frequently Asked Questions*. Available from: <https://nces.ed.gov/nationsreportcard/faq.aspx> [Accessed March 2017].
66. National Center for Education Statistics. (2017) *Parents' Frequently Asked Questions About NAEP*. Available from: <https://nces.ed.gov/nationsreportcard/parents/faq.aspx> [Accessed March 2017].

67. The Washington Post. (2011) *NAEP: A flawed benchmark producing the same old story*. Available from: https://www.washingtonpost.com/bl ogs/answer-sheet/post/naep-a-flawed-benchmark-producing-the-same-old-story/2011/11/03/gIQAbnonmM_blog.html?utm_term=.aa9eac473c15 [Accessed March 2017].

立法會秘書處
資訊服務部
資料研究組
吳穎瑜及劉奕君
2017年3月9日
電話：2871 2122

資料摘要為立法會議員及立法會轄下委員會而編製，它們並非法律或其他專業意見，亦不應以該等資料摘要作為上述意見。資料摘要的版權由立法會行政管理委員會("行政管理委員會")所擁有。行政管理委員會准許任何人士複製資料摘要作非商業用途，惟有關複製必須準確及不會對立法會構成負面影響，並須註明出處為立法會秘書處資料研究組，而且須將一份複製文本送交立法會圖書館備存。本資料摘要的文件編號為 IN06/16-17。