

2019/20 學年

中國語文教育榮譽學士課程（五年全日制）（A5B060）

CHI 4655 學術研究 II

## 2004-2018 年小三級全港性系統評估 中國語文科閱讀卷題型分析

姓名：林健瑩

學生編號：

指導老師：廖先博士

繳交日期：2019 年 4 月 18 日

正文字數：9308 字

## 摘要

是次研究旨在分析 2004-2018 年小學三年級全港性系統評估 (TSA) 中國語文科閱讀試卷 (分卷一) 之文本選材特點、各題目類型分值比重和各題目的閱讀能力層次分值比重。研究發現, 隨 BCA 推行, 篇章數量維持在兩篇和 1200 字內的調整符合小三學生的基本能力。但篇章則只為個人用途而閱讀, 並且為連續性文本的記敘型文章, 限制評估的內容。另外, 各試題所選之類型符合小三學生能力, 而當局亦有意發展其他類型的題目。最後, 研究數據顯示歷年來部分試題未能圍繞小三閱讀範疇的基本能力出題。因此, 建議當局每年各分卷均設一篇實用文, 以及試題只考核複述、解釋和重整能力。望能有助日後 TSA 試題的編制, 推動 TSA 真正達到「促進教與學」。

關鍵字：全港性系統評估 (TSA)、試題分析、標準化測試

## 一、緒論

### 1.1 研究背景

全港性系統評估 (Territory-wide System Assessment, 下稱 TSA) 於 2004 年推行, 除了取代因過度操練而變質的香港學科測驗, 也為學校管理層和教育當局提供一系列的數據, 以了解校內學生在主要學習階段的學習表現是否達到基本的學科水平, 從而促進學與教 (香港教統局, 2000, 2004)。TSA 的評估制度在學與教的教育過程中是不可或缺的一環, 評估與教學是互為影響而有其存在的必要。通過評估, 教師和教育局能夠了解學校的教學質量和學生對基本能力的掌握程度。然而, 2015 年 TSA 掀起一場「爭取取消小三 TSA」風波, 小學三年級學生應否應考被受討論, 成為社會廣泛關注的教育議題, 更有言論指出 TSA 與原設初衷相互違背, 既影響日常教學, 也對學生造成學業壓力。

首先, TSA 試題脫離學生的基本能力, 形成操練的現象, 加重學生的學習負擔。羅天佑和梁操雅 (2017) 指出即便學生已達到該學習階段的學科基本能力, 也無法取得理想成績, 因部分試題並非圍繞基本能力而設, 其艱深程度已達到「不操不識」的地步。葉建源 (2016) 也揭示學生能透過操練相似的題目而提升成績; 亦有調查結果顯示: 現時已出現小一學生操練 TSA 的現象 (香港教育專業人員協會教育研究部, 2011)。除此之外, 部分教師反映小三學生的語文能力處於發展階段, 而 TSA 的評估內容甚為艱深, 把教學時間花在操練上實是剝削學生學習語文基礎知識的機會 (張壽洪, 2016)。

再者, Carless and Lam (2014) 指出標準化的測試評估有機會導致初小學生對學校教育形成焦慮, 說明了違背初衷的 TSA 令小三學生對學習產生負面的影響。而且, 韋惠英評論機械式的操練令學生易於對學習產生厭惡之感, 失去對學習的興趣, 特別對處於第一學習階段的小三學生 (香港教育專業人員協會, 2011)。學生透過不斷的操練才能達標, 實已違背了政府當初推行課改的原意及其教育理念。

葉建源和黃家樂 (2013) 批評 TSA 已成為一個壓迫孩子的政策, 整個小學教育被不必要的操練和壓力所吞噬, 對學生造成學業壓力, 演變成一項高風險評估, 失去了 TSA 原設的意義。事實上, 在西方國家也有一些像 TSA 般的教育措施, 如英國的標準評估考試 (Standard Assessment Tests, SAT)、美國的全國教育進度評審 (National Assessment of Education Progress, NAEP) 或歐洲的全國

學生讀寫及算數能力評估考試（National Assessment Program-Literacy，NAPLAN）等，但鮮有出現機械式的過度操練現象。因此，目前應立即檢討 TSA 試題設計和難度（葉建源，2013）。

有見及此，本研究的焦點將集中在小三 TSA 中文閱讀卷的試題設計上，先檢視並分析歷年試卷的評估框架、出題形式和題目難度，從而作進一步的討論，及就未來 TSA 的發展和評估設計提出相關建議，以助解目前的困局。

## 1.2 研究問題

在 2004-2018 年的小學三年級全港性系統評估，中國語文科閱讀試卷中：

- （一） 文本選材有何特點？
- （二） 各題目類型的分值比重是怎樣？
- （三） 各題目的閱讀能力層次分值比重是怎樣？

## 二、文獻回顧

### 2.1 全港性系統評估

2000 年香港全面推行教育改革，繼教育統籌委員會向政府提交《終身學習·全人學習 — 香港教育制度改革建議》後，課程發展議會發表《學會學習 — 課程發展路向》課程文件，當中提及其一改革重點為「幫助學生掌握中文、英文和數學三科的基本能力」，而全港性系統評估（Territory-wide System Assessment，簡稱 TSA）亦於 2004 年正式推行。TSA 源於基本能力評估（Basic Competency Assessment，簡稱 BCA）<sup>1</sup>中的「系統評估」（System assessment），屬一個標準參照的評估測試，由考試及評核局統籌，每年六月於全港統一執行。

TSA 作為教改下一項全新的評估措施，起初設立的原意是希望協助學校管理階層了解學生在結束第一至第三學習階段時，中、英、數三個學科能力是否達到基本水平，從而改良教學策略和優化課程設計，並且能夠讓政府為有需要的學校提供適當的支援，進一步提升教學素質（香港教統局，2000，2004）。

另外，根據政府文件，TSA 不單是一套輔助學與教的工具，也是一個低風

---

<sup>1</sup> 基本能力評估（Basic Competency Assessment，簡稱 BCA）可分為學生評估（SA）與系統評估（TSA）兩部份。學生評估（SA）是在網上進行校內評估，主要幫助教師了解學生在三科主科（中文、英文和數學）的學校表現，例如他們的強或弱項和學習需要等，以改善學生間的學習差異和提升學習成效。

險的評估計劃，強調評估不是為篩選之用，學校報告內的資料和數據並不會用於比較學生和學校（香港教統局，2006）。雖然如此，侯傑泰和何穎欣（2008）指出 TSA 能達致監察學校之目的，但另一方面卻成為向學校問責的指標，影響恆常的教學。這是因為在 TSA 推行的初期恰是香港小學的「殺校潮」，學界一直傳出學生的 TSA 表現直接影響教育局對學校的評核。因此，學校之間視 TSA 為競爭指標，出現操練學生的現象。

隨後數年，不同的學者陸續指出操練對學生造成過大壓力（黃毅英、林智中、陳美英，2008；廖佩莉，2009），一度引起社會關注。面對學界和社會的聲音，教育局一直嘗試作出調整：2014 年推出優化措施、2015 年起隔年推行小六 TSA、2016 年試行基本能力評估研究計劃（Basic Competency Assessment，簡稱 BCA），以及 2017 年於全港學校推行 BCA。

## 2.2 TSA 試題設計越趨艱深

香港教育專業人員協會（2011）曾評論現時的 TSA 與因過度操練而被廢除的香港學科測驗並無差別。協會所進行的問卷調查結果亦顯示 73% 教師不認同普遍小三學生可於沒有操練的情況下應付 TSA 試題，67% 教師更認為小三 TSA 不應繼續推行（香港教育專業人員協會，2015）。此外，廖佩莉（2013）的研究也反映 77% 教師認為 TSA 的試題需要反覆操練，因學生並不善於回答高層次的問題。不少教師已表示 TSA 所評核的內容已經脫離學習階段的基本能力範疇（葉建源，2013）；學者黃毅英、林智中和陳美恩（2013）亦提及近年 TSA 試題已超出學生的基本能力，張壽洪（2016）更進一步指出 TSA 部分題目並不是根據基本能力指標而擬定。故此，TSA 出現操練現象的其一原因正是試題的設計和難度越加刁鑽與艱深，並非按照基本能力所制定。

然而，試題越見艱深並非於短時間內形成，當中涉及到歷年來 TSA 在推行過程中的嚴謹性與出題原則。葉建源（2016）將其成因歸納為兩項，一是「基本能力評估」和「促進學習的評估」概念之間的矛盾：當 TSA 除「基本能力」外，還強調「有效地促進學習的評估」時，擬題者則須把題目難度設定為一個「鐘形曲線」，才能分辨出不同學生的學習能力；二是考評局對 TSA 的兩大出題要求：不可重覆題目和不斷轉變題型，以作為學生評估（SA）的試題和杜絕操練。

### 2.3 閱讀測試的能力指標

相對其他範疇的語文能力，閱讀能力不僅在語文評估中佔着一個無可比擬的位置，更是學生在生活和社交中一項重要的基本能力。在閱讀評估中，不同的文章體裁所考核學生的閱讀能力有所不同；而試題設計亦應與文本遵循文題依屬的關係，即保持試題和閱讀材料間的密切關係。合適的試題設計和評分標準等評估資料更能有效地檢視學生的水平和學習成效。

在國際教育領域中，現時最具代表和影響力的國際閱讀測試有二，分別是由國際教育成就評價協會（IEA）主辦的「全球學生閱讀能力進展研究」

（PIRLS），以及經濟合作及發展組織（OECD）所策劃的「學生能力國際評估計劃」（PISA）（謝錫金等，2005；羅燕琴，2011）。廖先、祝新華（2010）認為 PISA 和 PIRLS 的更新和修訂反映出目前閱讀評估的最新發展模式和特點，而兩者的發展趨勢對目前的閱讀評估有所啟示。祝新華（2015）進一步補充兩項國際閱讀測試的能力指標和評估框架值得學界加以重視和借鑒，特別在教育質量評估方面。

### 2.4 六層次閱讀能力系統

認知是進行閱讀活動時閱讀者所需具備的基本能力，亦是閱讀能力中不可忽視的一環。在閱讀評估中，認知能力表現在各閱讀結果上，即由閱讀者理解對材料的字詞到內容的表層意思，繼而理解深層或隱含意思，至最後聯繫個人經驗和知識，對材料作出評價。Anderson, Krathwohl, Airasian, (2001) 在 Bloom (1956) 對學生學習的六個認知層次：識記、理解、應用、分析、綜合、評價的基礎上，新「增創」造層次；而祝新華（2003）概括國內外多個閱讀理論，發展出「閱讀能力系統」，當中包括六種由低至高層次的考核能力：複述、闡述、重整、伸展、評鑑和創新。

目前，這系統在香港和新加坡的語文課程、教學和評估範疇上得到實際的應用。而且，這亦是一測試題型系統，可用作設計中國或分析閱讀試卷，為中國語文科之課程規劃和教學設計等研究領域提供了具參考價值的依據（祝新華，2015）。加上，此系統與 PIRLS、PISA 等國際閱讀測試的能力評核框架是相對一致的，值得以此系統作為分析的工具。

### 三、研究方法

#### 3.1 研究樣本

本研究以 2004 年至 2015 年的全港性系統評估、2016 年試行研究計劃、2017 年基本能力評估研究計劃，以及 2018 年的全港性系統評估之小學三年級中國語文科閱讀試卷（分卷一）作為研究樣本，合共 15 份，研究者將從香港考試及考評局的網頁下載歷年試卷。是次的研究樣本覆蓋了全港性系統評估、試行基本能力評估研究計劃和基本能力評估研究計劃歷年的閱讀評估試卷，具代表性。

#### 3.2 研究方法與步驟

##### 3.2.1 針對研究問題一：文本選材方面有何特點？

本研究先記錄各年小學三年級 TSA 中國語文科閱讀試卷（分卷一）篇章的數量和文本的內容摘要，然後再就篇章的「情景」、「形式」、「類型」和「字數」，四部分進行分析。

首先，PISA2009（OECD，2009a）將篇章情景區分為「個人目的」、「公共目的」、「工作目的」和「教育目的」四種，不同的篇章情景具有不同的閱讀目的（參見表 1），而此框架乃作為閱讀分類和文本選擇的基礎。

表 1 篇章情景

篇章情景	說明	文本例子
1. 為個人用途而閱讀	為滿足個人的興趣、 休閒消遣的閱讀	個人信件、小說、傳記
2. 為公共用途而閱讀	為參與社會活動的閱讀	官方文件、公眾事務資訊
3. 為工作用途而閱讀	為完成某項任務或工作	研究、調查
4. 為教育用途而閱讀	為較大規模的學習任務 而取得資訊	教科書、參考書

再而，根據 PISA2009（OECD，2009a）篇章形式共有四類，分別是連續性文本、非連續性文本、混合文本與組合文本，而歷年 TSA 並未出現組合文本，故是次研究將集中分析篇章屬連續性文本、非連續性文本或混合文本（參見表 2），不同形式的閱讀篇章考核學生不同的閱讀能力。

表 2 篇章形式

篇章形式	說明	例子
1. 連續性文本	由句子組成的段落，由段落組成的文章	節、章、書
2. 非連續性文本	運用不同的方式呈現信息	表格、圖形
3. 混合文本	包含連續性文本與非連續性文本	電郵、論壇

另外，篇章類型可分為「描寫」、「記敘」、「說明」、「議論」，以及「交易型文本」。「交易型文本」即是用作達到某種特定的目的，例如安排會議、索取物件或相約朋友，一般以信件、郵件或博客等形式出現。

最後，研究者將統計每年試卷中的閱讀篇章總字數。

### 3.2.2 針對研究問題二：各題目類型的分值比重是怎樣？

縱觀各年的閱讀卷，試題的類型可分為四類型：填充題、多項選擇題、簡答題，以及其他題型（包括排序題、圖表題、配對題和判斷題）。鑒於小三學生正處於第一學習階段，不論語文能力或書寫能力皆處於發展的階段，故而部分題型被簡化，導致出現題型混淆的情況，研究者將加以為其作出定義：首先，填充題為要求學生在一句出現漏空的句子中填上合適的字詞。另外，短答題指試題出現一問句後並提供一句答句供學生回答。

其後，研究者將分析每道試題的題目類型，並統計每年該類型題目的賦分、比重和題量。

### 3.2.3 針對研究問題三：各題目的閱讀能力層次分值比重是怎樣？

從 15 份研究樣本的試題中，根據祝新華的六層次閱讀能力系統以表格的形式進行分析，然後統計每年該層次的賦分和其比重。

## 四、研究結果與分析

## 4.1 文本選材的特點

表 3 篇章數量與文本選材

年份	數量	文本內容摘要		年份	數量	文本內容摘要	
2018	2	篇章一	山羊醫生	2010	3	篇章一	小蜜蜂迷路
		篇章二	妹妹在蝸牛身上學到靠自己，不依賴別人			篇章二	參觀雪糕廠
2017	2	篇章一	阿花令美好學校成為森林的樂園	2009	3	實用文一	宣傳單張：天文台展覽
		篇章二	弟弟解說水熊蟲的習性			篇章一	童軍售賣獎券
2016	2	篇章一	白白豬整蛋糕（上）	2008	3	篇章二	家兒對玉英的態度由生氣到原諒
		篇章二	白白豬整蛋糕（下）			實用文一	宣傳單張：武術匯演
2015	3	篇章一	快樂谷音樂會	2007	4	篇章一	大海兄弟吵架
		篇章二	體育課集體活動學到「團結就是力量」			篇章二	風車發電
		實用文一	宣傳單張：《小星星》徵文			實用文一	書信：問候表哥近況，並邀請他到花卉展覽
2014	3	篇章一	大威哈奇王子宇宙（上）	2006	3	篇章一	在學校的園圃種番茄
		篇章二	大威哈奇王子宇宙（下）			篇章二	下雨撐傘：晴晴對別人外表的偏見
		實用文一	電郵：哈奇與大威分享生活近況			實用文一	宣傳單張：閱讀獎勵計劃
2013	3	篇章一	快樂農莊：從照顧幼苗領悟到爸媽的偉大	2005	4	實用文二	書信：邀請外公外婆出席頒獎典禮，並共進午膳
		篇章二	鱷魚蛋唔被當作鳥蛋			篇章一	無空調日與怕熱的志強
		實用文一	報章：繽紛蝴蝶迎暑假			篇章二	林村許願樹
2012	3	篇章一	內心善良的怪獸阿餅	2004	3	實用文	邀請卡：邀請表姐參加步行籌款
		篇章二	魔法豹先生			篇章一	與鄰座小燕的誤會
		實用文一	書信：向表姐表示想念及交代近況			篇章二	回想學騎自行車的過程
2011	3	篇章一	森林大王選舉	2004	3	實用文一	邀請卡：邀請表哥觀看花式跳繩比賽
		篇章二	文龍減肥			實用文二	賀卡：祝賀表哥奪網球比賽冠軍
		實用文一	網頁：兒童快樂城			篇章一	小猫小胖
						實用文一	書信：告知楊老師獲得畫畫比賽冠軍和會到法國探訪她
						實用文二	賀卡：祝賀成績欠佳的志明考獲全班第五名

由上述表 3 可見歷年 TSA 閱讀卷的篇章數量和每篇閱讀材料的內容摘要。

首先，在篇章數量方面。TSA 剛推行的首五年（2004 年至 2008 年），篇章數量最高達四篇（包括兩篇文章和兩篇實用文）；在 2008 年後，則穩定維持在三篇的數量（包括兩篇文章和一篇實用文）；到了 2016 年 BCA 推行後，篇章數量降至兩篇。

其次，在文本內容選材方面，歷年的篇章主要以日常生活事件或動物的生活為主題。另外，在 2013 年或之前，試卷內的各篇文章皆為獨立篇章，即篇章與篇章之間沒有人物或情節的關聯與承接，而近年（2014 年和 2016 年）試卷內的篇章出現上集和下集的現象，篇章之間屬承上啟下的關係。

再而，在 2004 年至 2008 年的實用文篇章皆為傳統紙本形式的書信、賀卡或邀請卡；到了 2009 年或之後，實用文的選材始現不同的新形式，有更貼近的生活的宣傳單張，也有配合信息年代的電子閱讀，例如電郵和網頁等。這樣的轉變能夠更全面和真實地考核學生閱讀生活化文類的基礎閱讀能力。但可惜的是，2016 年起 TSA 僅在其中一份分卷設實用文的考核，其餘兩份分卷則被刪減該部分。

表 4 篇章的情景、形式和類型

年份	篇章情景	篇章形式	篇章類型
2018	個人	連續性	記敘
2017	個人	連續性	記敘
2016	個人	連續性	記敘
2015	個人+公共	連續性+混合	記敘+交易型
2014	個人	連續性+混合	記敘+交易型
2013	個人+公共	連續性+混合	記敘+交易型
2012	個人	連續性	記敘+交易型
2011	個人+公共	連續性+混合	記敘+交易型
2010	個人+公共	連續性+混合	記敘+交易型
2009	個人+公共	連續性+混合	記敘+交易型
2008	個人	連續性	記敘+交易型
2007	個人+公共	連續性+混合	記敘+交易型
2006	個人	連續性	記敘+交易型
2005	個人	連續性	記敘+交易型
2004	個人	連續性	記敘+交易型

從上述表 4 整體而言，TSA 各年的篇章主要圍繞在為個人目的和公共目的而閱讀的情景，當中出現連續性和混合文本的篇章形式，而篇章類型則有記敘型和交易型的。而記敘型的篇章符合 Chall 所提出的閱讀發展階段，適合小三學生閱讀敘事性文體的閱讀能力發展特徵（McKenna & Stahl, 2003）。

此外，歷年來試卷中的篇章情景、形式和類型存有三種的規律模式。在 BCA 尚未推行前，文本選材是「為個人用途而閱讀的記敘型和交易型連續性文章」以及「為個人和公共而閱讀的記敘型和交易型的連續性文本與混合文本」。但自 BCA 實施後，文本選材劃一改成「為個人用途而閱讀的記敘型連續性文章」，這樣單一的特定規律限制了評估的框架和評核內容。



圖 1 2004-2018 年小三級 TSA 中國語文科閱讀卷（卷一）總字數

在上述的圖 1 中可先剔除首年的數據，因當年閱讀卷尚未獨立成卷，與聆聽卷合卷，故評估內容較短，字數亦較少，未具參考價值。剔除後可見，整個線形圖呈現出一個與「鐘形曲線」相似的形狀。曲線的前端（2005 年）和末端（2018 年）與同一的水平上，表示總字數的數量相若，為 900 多字；而曲線中段高出的位置（2010 年和 2011 年）代表總字數曾最高達 1500 字。除此之外，這樣的「鐘形曲線」趨勢其實也代表着試卷的難度，評估時間只有二十五分鐘，文章字數的多與少，影響着學生的閱讀、審題和作答速度，增加了試卷的難度。

繼而，從圖中「先升後降」的趨勢可見，2011 年前的篇章字數呈遞升的趨勢，2011 年後呈遞降的趨勢。這種降幅在 2012 年、2014 年和 2016 年尤其顯著，而此現象是受到當時社會事件所影響：香港教育專業人員協會（下稱教協）於 2011 年、2013 年和 2015 年就 TSA 進行了三次的教師意見調查，其報告所提出的建議受政府關注，政府並作出相關的回應。可是，這些改善措施未能解決 TSA 的異化問題，促使 TSA 在 2015 年爆發成為一項社會議題，引起社會廣泛關注，並曾在立法會召開公聽會。因此，當局在 2016 年作出檢討措施，字數隨即下調，並限制在 1200 字的範圍內。而 PIRLS 為了制定一個更適合小四學生閱讀水平的測試，自 2011 年起把文本長度調整至 800 字，測試內容合共兩篇 1600 字（祝新華，2015）。與之比較，經調適後的 TSA 文本字數與國際的閱讀

評估相若，符合小三學生的閱讀能力，表示 TSA 在字數方面已重新針對學生的基本能力。

#### 4.2 各題目類型的分值比重

表 5 題目類型的分值比重

	填充題			多項選擇題			簡答題			其他題型 排序、圖表、配對、判斷			總計	
	賦分	比重	題量	賦分	比重	題量	賦分	比重	題量	賦分	比重	題量	賦分	題量
2018	5	25%	5	9	45%	9	3	15%	3	3	15%	3	20	20
2017	4	18%	4	12	55%	12	0	0%	0	6	27%	4	22	20
2016	4	18%	4	11	50%	11	3	14%	2	4	18%	3	22	20
2015	4	15%	4	13	48%	13	2	7%	2	8	30%	4	27	23
2014	4	14%	4	18	64%	14	3	11%	2	3	11%	3	28	23
2013	8	28%	6	16	55%	13	1	3%	1	4	14%	3	29	23
2012	4	16%	4	18	72%	14	1	4%	1	2	8%	2	25	21
2011	5	19%	5	15	56%	11	5	19%	4	2	7%	2	27	22
2010	5	21%	5	13	54%	11	2	8%	2	4	17%	3	24	21
2009	8	26%	6	16	52%	12	2	6%	1	5	16%	3	31	22
2008	7	28%	6	13	52%	11	1	4%	1	4	16%	3	25	21
2007	7	28%	6	13	52%	13	4	16%	2	1	4%	1	25	22
2006	5	21%	5	11	46%	11	6	25%	4	2	8%	2	24	22
2005	6	24%	6	12	48%	12	5	20%	4	2	8%	2	25	24
2004	3	21%	3	5	36%	5	5	36%	4	1	7%	1	14	13
總計	79	21%	73	195	53%	172	43	12%	33	51	14%	39	368	317

從上表 5 的數據分析可見，多項選擇題為分值比重最高的題型，佔整體的 53%，其次是 21% 的填充題，而簡答題和其他題型的比重則相若，各佔 12% 和 14%。而多項選擇題和填充題共佔全部題型的 70%，這兩類型的題目具客觀評分和簡易作答的特點。由此可見，當局在設計試題時考慮到小三學生的書寫能力和書寫速度，為避免學生在作答時花大量時間於書寫答案上，從而影響評核其閱讀能力，因此選擇答題時間較短的選擇題和填充題為試卷的主要題型。

再者，祝新華（2015）曾說明每類型的題目皆有其重點考核的閱讀能力，當中指出多項選擇題和填充題主要考核學生的複述和解釋能力，屬低層次認知能力。加上 TSA 本屬於一個標準化測驗，旨在評估學生在結束學習階段時是否達到該學科的基本能力，故此，多項選擇題和填充題佔頗高的分值比重與 TSA 的原設意義相合。

另外，由各年的出題量和賦分量之趨勢而看，多項選擇題和填充題每年的升幅相對穩定，但值得注意的是簡答題和其他題型則較為飄忽。簡答題由 2007

年起均持下降的趨勢，甚至在 2017 年的試卷中沒有出現過此題型，這樣的現象亦是被當時社會上出現 TSA 試題過深的聲音所影響；而其他題型則持上升的趨勢，所佔的分值比重曾在 2015-2017 年間高於填充題。可見，當局有意在試題設計上有意發展其他類型的題目。

### 4.3 各題目的閱讀能力層次分值比重

表 6 題目的閱讀能力層次分值比重

	1.複述			2.解釋			3.重整			4.伸展			5.評鑒			6.創意			7.其他			總計	
	賦分	比重	題量	賦分	比重	題量	賦分	比重	題量	賦分	比重	題量	賦分	比重	題量	賦分	比重	題量	賦分	比重	題量	賦分	題量
2018	5	25%	5	5	25%	5	9	45%	9	1	5%	1	0	0%	0	0	0%	0	0	0%	0	20	20
2017	7	32%	5	5	23%	5	10	45%	10	0	0%	0	0	0%	0	0	0%	0	0	0%	0	22	20
2016	8	36%	6	5	23%	5	7	32%	7	2	9%	2	0	0%	0	0	0%	0	0	0%	0	22	20
2015	8	30%	4	6	22%	6	10	37%	10	2	7%	2	0	0%	0	0	0%	0	1	4%	1	27	23
2014	7	25%	5	4	14%	4	15	54%	12	1	4%	1	0	0%	0	1	4%	1	0	0%	0	28	23
2013	13	45%	9	4	14%	4	11	38%	9	1	3%	1	0	0%	0	0	0%	0	0	0%	0	29	23
2012	4	16%	3	5	20%	5	11	44%	10	4	16%	2	0	0%	0	0	0%	0	1	4%	1	25	21
2011	10	37%	7	5	19%	5	10	37%	8	2	7%	2	0	0%	0	0	0%	0	0	0%	0	27	22
2010	9	38%	8	5	21%	5	8	33%	6	2	8%	2	0	0%	0	0	0%	0	0	0%	0	24	21
2009	7	23%	4	5	16%	4	14	45%	10	5	16%	4	0	0%	0	0	0%	0	0	0%	0	31	22
2008	7	28%	4	5	20%	5	9	36%	8	3	12%	3	0	0%	0	0	0%	0	1	4%	1	25	21
2007	5	20%	3	7	28%	6	9	36%	9	4	16%	4	0	0%	0	0	0%	0	0	0%	0	25	22
2006	6	25%	5	4	17%	4	7	29%	6	5	21%	5	0	0%	0	1	4%	1	1	4%	1	24	22
2005	6	24%	5	6	24%	6	6	24%	6	6	24%	6	0	0%	0	1	4%	1	0	0%	0	25	24
2004	3	21%	2	4	29%	4	6	43%	6	0	0%	0	0	0%	0	1	7%	1	0	0%	0	14	13
總計	105	29%	75	75	20%	73	142	39%	126	38	10%	35	0	0%	0	4	1%	4	4	1%	4	368	317

從上述表 6 的數據分析可見，TSA 試題覆蓋了五層的閱讀能力，當中考核最多的是重整題，分值比重達 39%；隨後是複述題和解釋題，各佔 29%和 20%；再而是佔 10%和 1%的伸展題和創意題。當中值得注意的地方是：有 1% 題目並不屬於考核學生的閱讀能力，而是考核其語文知識和數學能力，例如

2006 年的第 19 題和 2015 年的第 21 題等（參見圖 2 和圖 3）。

我邀請你一同參加這次活動。活動結束後，我們會到康樂廣場附近的歡樂酒樓喝茶，順便和你預祝生日。

小敏  
2006 年 6 月 25 日

19. 如果為小敏補寫自稱，應在文中 \_\_\_\_\_ 的部分填寫下列哪一項？

A. 妹妹  
B. 學妹  
C. 表妹  
D. 堂妹

評卷員專用

(36)

圖 2 2006 年第 19 題

21. 根據以下月曆，哪一天會出版 9 月第一期《小星星》？

第一期《小星星》在 9 月 \_\_\_\_\_ 日出版。

圖 3 2015 年第 21 題

2006 年第 19 題考核學生是否掌握實用文中的書信格式，要求學生在灰色方格內提上適當的稱謂，此屬考核學生的語文知識，超出「理解簡單實用文」的基本能力。而 2015 年第 21 題則要求學生先理解宣傳單張（實用文）中「星光出版社逢星期三出版的兒童刊物《小星星》」一句，繼而須配合題目和其所提供之月曆，方可得出正確答案，但是看月曆屬於數學科有關年月日概念的能力，亦是另一脫離閱讀能力的試題。

再者，屬低層次閱讀能力的複述題和解釋題佔整體的 49%，中階層次的重整題和伸展題佔 49%，高層次的閱讀能力題目亦佔 1%。然而，TSA 原設為一標準化的測試，旨在評估學生對學科基本能力的掌握。故此，小三 TSA 的試題應對照第一階段的基本能力。研究者嘗試製作一表，以檢視兩者的相互的對應關係（參見表 7）。

表 7 第一學習階段閱讀範疇基本能力與閱讀能力層次的對應

基本能力 \ 閱讀能力層次	複述	解釋	重整	伸展	評鑒	創意
能認讀一般閱讀材料中的常用字	✓					
能理解所學詞語		✓				
能理解簡淺敘述性文字的段意及段落關係			✓			
能概略理解篇章中簡淺的順敘/倒敘事件		✓				
能理解簡單的實用文	✓	✓				

從上述的對照表可見，第一學習階段的閱讀基本能力大多停留在低層次的複述題、解釋題和重整題，而伸展、評鑒和創意能力並不能呼應與之提及的基本能力。因此，歷年來共有 12% 的題目脫離了基本能力。

再而，重整題歷年來的升幅亦值得留意（參見表 6）：由最低的 2005 年 24% 上升至最高的 2014 年 54%，雖近兩年下調至 45%，但其比重亦過高，因表 7 中重整題僅對應其中一項的基本能力。

## 五、結論

### 5.1 文本選材的特點

根據研究結果，推行 BCA 的前後文本選材具不同的特點。在 BCA 尚未推行前文本數量最高為四篇，閱讀材料總字數為 1500 字。而篇章為個人和公共用途而閱讀的，當中出現的形式包括連續性文本和混合文本的記敘型和交易型文章。隨着 BCA 的推行，篇章數量將至兩篇，閱讀材料總字數維持在 1200 字內。而篇章則只為個人用途而閱讀，當中出現的形式只有連續性文本的記敘型文章。這樣的改變限制了評估的內容，以及未能全面考核學生是否已掌握「能理解簡單的實用文」的基本能力。然而，篇章數量維持在兩篇和 1200 字內的調整符合小三學生的基本能力。

### 5.2 各題目類型的分值比重

上文的研究結果分析顯示，各題目類型的分值比重由高至低，分別是多項選擇題（53%）、填充題（21%）、其他題型（14%）、簡答題（12%）。而考評局選擇以多項選擇題和填充題為主要題型乃考慮到小三學生的書寫能力和作答時間，令評估數據能夠更準確地反映出學生的閱讀能力是否已達基本能力。此外，其他題型的佔分比重持上升的趨勢表示當局有意在未來的試卷中發展不同類型的題目，有助舒緩學生因操練而能應付千篇一律的 TSA 題目（廖佩莉，2013）之情況。

### 5.3 各題目的閱讀能力層次分值比重

從是次研究數據可見，歷年來的 TSA 試題不但涵蓋了複述（29%）、解釋（20%）、重整（39%）、伸展（10%）和創意（1%）的題目，還出現考核學生語文知識和數學能力的題目（1%）。TSA 屬一標準化測試，題目若根據基本能力而設只需包含複述、解釋和重整題。因此，過去的 TSA 試題並沒有圍繞基本能力而出題，故而導致學生即使達到基本能力亦不懂作答 TSA 的題目。

再而，歷年來重整題持上升的趨勢，最高佔整體的 54%，近兩年亦佔整體的 45%。若此情況持續出現，TSA 便會演變成常模參照測試。當擬題者有見學生於去年能應付一些較艱深的題目，就會在來年進一步相對地提高試卷的難度，以適時更新對學生學習能力的掌握，達致「促進學習的評估」。這樣無疑把去年的試卷作為對照組，用以擬定下一年度試題的難度。這是往後 TSA 發展值得注意的地方。

## 六、建議與啟示

### 6.1 每年各分卷均設一篇實用文

政府就回應社會上反對小三 TSA 的聲音，經一輪檢視後，對閱讀卷的文本作出了相關的修訂，當中包括調整篇章數量，由三篇減至兩篇，而且為避免實用文在閱讀卷的比重過大，只在其中一份閱讀分卷設實用文（教育局，2016）。以本研究的研究樣本為例，小三考生只需考核兩篇連續性文章，這樣的做法未能評估所有完成第一學習階段的學生是否已達「能理解簡單的實用文」的基本能力。

小三學生的閱讀能力處於學習閱讀（learning to read）（謝錫金，林偉業、林裕康和羅嘉怡，2005），他們除需要掌握閱讀敘事性的閱讀能力外，閱讀實用文的基本能力也不可忽視。實用文是一種生活化的文本，既源於生活，又被普遍地運用於生活，相當具實用性。加上，PISA 的研究者認為非連續性文本或組合文本所要求的閱讀能力是學生日後投身社會、實現個人奮鬥目標必需的閱讀素養。故而，實用文不僅是一種文本，更是一種生存方法。學生掌握閱讀這類型文本的基本能力有利他們往後適應社會生活和參與社會生活，這亦是其所在之價值。因此，建議當局於每年各閱讀分卷均設實用文一篇。

### 6.2 試題只考核複述、解釋和重整能力

上文的研究結果三顯示出歷年來 TSA 的試題皆涵蓋不同難度，當中部分試題不單遠離基本能力，考核學生相對較高層次的閱讀能力，甚至脫離閱讀能力，考核學生的語文知識和數學能力，完全違背了 TSA「評估整體學生在完成小學三年級時，中文科的基本能力」之原設目的。

有見及此，建議當局應根據第一學習階段閱讀範疇的基本能力設計試題，使致每年的 TSA 是一個真正的標準能力測試，旨檢視學生對基本能力的掌握；而試題所考核的閱讀能力應圍繞在複述、解釋和重整能力，使每年的試題難度劃一保持在相約的水平，減低學生因試題過難而操練的現象，進一步確保 TSA 的題目不會再度出現猶如「跨欄」中的「欄」（葉建源，2016），刻意設定在中等的高度，讓應考的一半考生跨過，另一半考生則不能跨過。

### 6.3 對教學的啟示

教師應善用每年考評局於 TSA 試後向各學校提供學校報告。報告書中提供

不同的數據，並分析學生的能力表現。從這些數據與分析中，教師能夠了解不同學生的能力水平。以小三中國語文科的閱讀卷為例，報告書中顯示出學生在各試題的得分情況，讓教師能夠進一步得悉學生對基本閱讀能力的掌握程度，以及反映出學生在閱讀不同文本和作答不同類型試題的強弱。透過這些數據資訊，教師能夠從中反思課程的規劃和課堂設計，並且對教學作針對性的改進，望以能更有效地幫助學生改善及提高其閱讀能力，藉此促進學與教。

另外，作為一名準語文教師，應當明白機械式的操練非一顯著提升學生語文能力的方法。操之過渡，操之過急，反而適得其反，令學生在低年班時已對學習語文失去興趣，甚至抗拒和逃避學習語文。因此，教師需自我審視。定時審視自己的教學，因應學生的能力不時改變教學策略，令學生於恆常的課堂學習上已能掌握各基本能力。

## 七、結語

Lam, R. (2018) 表示 TSA 呈現出一種沒有贏家的教育生態：教師工作壓力大，學生未能愉快學習，父母也常與子女因溫習問題而發生爭執。TSA 推行至今已踏入第十六年，一路走來原來已經進入了一個「三輸」的局面，相信這並非當局昔日推行 TSA 的原意，以及所預計之結果。因此，還望是次研究所分析歷年 TSA 試題的數據能有助日後 TSA 試題的編制，推動 TSA 真正達到「促進教與學」，而非礙於教師教學，學生慣於操練僵化的試題。

最後，是次研究只選取小三 TSA 中國語文科閱讀卷每年的其中一份分卷，未有分析剩下的兩份分卷，未能分析 TSA 歷年來所有的評估文本與試題，寄望未來的研究能作進一步的分析。

## 八、參考文獻

### 8.1 小三級全港性系統評估試卷

香港考試及評核局（2004）：2004 年全港性系統評估小學三年級中國語文聆聽及閱讀評估，檢自

<https://www.bca.hkeaa.edu.hk/web/Common/res/2004priPaper/chi/C1.pdf>

香港考試及評核局（2005）：2005 年全港性系統評估小學三年級中國語文閱讀評估，檢自 <https://www.bca.hkeaa.edu.hk/web/Common/res/2005priPaper/chi/2005-3CR1.pdf>

香港考試及評核局（2006）：2006 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2006priPaper/P3Chi/2006\\_TSA\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2006priPaper/P3Chi/2006_TSA_3CR1.pdf)

香港考試及評核局（2007）：2007 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2007priPaper/P3Chi/2007\\_TSA\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2007priPaper/P3Chi/2007_TSA_3CR1.pdf)

香港考試及評核局（2008）：2008 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2008priPaper/P3Chi/2008\\_TSA\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2008priPaper/P3Chi/2008_TSA_3CR1.pdf)

香港考試及評核局（2009）：2009 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2009priPaper/P3Chi/2009\\_TSA\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2009priPaper/P3Chi/2009_TSA_3CR1.pdf)

香港考試及評核局（2010）：2010 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2010priPaper/P3Chi/2010\\_TSA\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2010priPaper/P3Chi/2010_TSA_3CR1.pdf)

香港考試及評核局（2011）：2011 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2011priPaper/P3Chi/2011\\_TSA\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2011priPaper/P3Chi/2011_TSA_3CR1.pdf)

香港考試及評核局（2012）：2012 年全港性系統評估小學三年級中國語文閱讀

評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2012priPaper/P3Chi/2012\\_TSA\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2012priPaper/P3Chi/2012_TSA_3CR1.pdf)

香港考試及評核局（2013）：2013 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2013priPaper/P3Chi/TSA2013\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2013priPaper/P3Chi/TSA2013_3CR1.pdf)

香港考試及評核局（2014）：2014 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2014priPaper/P3Chi/TSA2014\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2014priPaper/P3Chi/TSA2014_3CR1.pdf)

香港考試及評核局（2015）：2015 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2015priPaper/P3Chi/TSA2015\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2015priPaper/P3Chi/TSA2015_3CR1.pdf)

香港考試及評核局（2016）：2016 年試行研究計劃小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2016priPaper/P3Chi/TryoutStudy2016\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2016priPaper/P3Chi/TryoutStudy2016_3CR1.pdf)

香港考試及評核局（2017）：2017 年基本能力評估研究計劃小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2017priPaper/P3Chi/BCA2017\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2017priPaper/P3Chi/BCA2017_3CR1.pdf)

香港考試及評核局（2018）：2018 年全港性系統評估小學三年級中國語文閱讀評估，檢自

[https://www.bca.hkeaa.edu.hk/web/Common/res/2018priPaper/P3Chi/TSA2018\\_3CR1.pdf](https://www.bca.hkeaa.edu.hk/web/Common/res/2018priPaper/P3Chi/TSA2018_3CR1.pdf)

## 8.2 課程文件

香港教統局（2000）：香港教育制度改革建議，香港，香港印務局。

教育統籌委員會（2000）：終身學習·全人學習 — 香港教育制度改革建議，香港，香港印務署。

香港教統局（2004）：教育改革進展性評估（三），香港，香港印務局。

課程發展議會（2004）：中國語文課程指引（小一至小六），香港，香港印務署。

香港教統局（2006）：教育改革進展性報告（四），香港，香港印務署。

### 8.3 書籍

Anderson, L. W., Krathwohl, D. R., Airasian, P. W. (2011). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives*. New York: Longman.

Bloom, B. S. (1956). *Taxonomy of education objectives: The classification of education goals, Handbook I: Cognitive domain*. New York: Longman.

羅天佑和梁操雅（2017）：香港考評文化的承與變——從強調篩選到反映能力，香港，商務印書館有限公司。

羅燕琴、岑紹基、林偉業和鍾嶺崇（2011）：《香港中國語文課程新路向：學習與評估》，香港，香港大學出版社。

謝錫金、林偉業、林裕康和羅嘉怡（2005）：《兒童閱讀能力進展：香港與國際比較》，香港，香港大學出版社。

葉建源和黃家樂（2013）：全港性系統評估(TSA)評論文集，香港，香港教育專業人員協會和香港數學教育學會。

葉建源（2016）：全港性系統評估（TSA）政策實施研究，香港，葉建源議員辦事處。

祝新華（2015）：促進學習的閱讀評估，北京，人民教育出版社。

### 8.4 期刊論文

Carless, D., & Lam, R. (2014). *The examined life: perspectives of lower primary school students in Hong Kong*. *Education 3 - 13: International Journal of Primary, Elementary and Early Years Education*, 42(3), 313 - 329.

Lam, R. (2018). *Testing, Drilling and Learning: What Purpose Does the Grade 3 Territory-Wide System Assessment Serve?* *Asia Pacific Education Review*, 19(3), 363-374. <https://doi.org/10.1007/s12564-018-9523-z> (SSCI)

侯傑泰、何穎欣（2008）：學習回饋及系統檢查：香港的經驗。《教育研究與發展期刊》，4（4），1-18。

黃毅英、林智中、陳美恩（2013）：「基本能力」還是「基本」嗎？，《全港性系統評估(TSA)》，香港，香港教育專業人員協會。

- 廖佩莉（2009）：析論香港小學中國語文科教師為學生準備「全港系統性評估」（TSA）的策略，*教育研究與發展期刊*，5（4），109-128。
- 廖佩莉（2013）：倒流效應：香港全港性系統評估(TSA)對小學中國語文科教師的影響，《*教育研究月刊*》，228，86-102。
- 廖先，祝新華（2010）：從國際閱讀項目的最近發展探討閱讀評估策略，《*Global Education 《全球教育展望》(CSSCI journal)*》，12，53-59。
- 香港教育專業人員協會教育研究部（2011）：TSA 流弊未除，打倒評核怪獸，《*全港性系統評估(TSA)*》，香港，香港教育專業人員協會。
- 香港教育專業人員協會教育研究部（2011）：談 TSA 衍生的操練文化 —專訪教育顧問韋惠英，檢自 <https://www.hkptu.org/15068>。
- 張壽洪（2016）：小學語文教師對全港性系統評估的認識和意見，輯於余虹、何文勝編，《*全球語境下漢語文教育的建設和發展*》，（頁 270-281），北京，開明出版社。
- 祝新華（2003）：以閱讀認知理論為基礎探討理解能力的層次，*中國語文測試*，（4）：15-21。