

科學圖文閱讀眼動研究之系統回顧*

王孜甯 簡郁芬

國立臺灣師範大學
教育心理與輔導學系

本研究之目的在於系統性地回顧近 30 年來（1990 年—2020 年 5 月）科學圖文閱讀的眼動研究，綜整文獻上科學圖文閱讀的歷程與樣態。579 篇英文及 33 篇中文論文經標題摘要與全文檢視兩階段篩選，共 51 篇（含 55 個研究）納入本回顧，以質性綜整後歸納為兩大主題：認知歷程與介入設計。認知歷程的研究主要探討讀者（如年齡、能力等）及閱讀材料（如難度、順序等）對閱讀歷程的影響；介入設計的研究主要是材料操弄（如空間接近原則、信號原則等）與教學介入（如圖文閱讀策略、眼動楷模示範等）。研究結果顯示，在科學文本中添加圖像通常會影響讀者的視覺行為，但未必保證好的學習表現。整體而言讀者主要以文字驅動閱讀行為，而歷程中重視圖像以及圖文參照的讀者，通常閱讀表現較佳；材料特徵經常與讀者特徵有交互作用；部分讀者特徵因子（如能力與先備知識）可能對眼動型態與學習表現產生調節效果。教學介入亦會影響讀者的眼動型態，如：增加圖文整合、投注更多認知資源解碼圖片，進而提升閱讀理解。而材料操弄的效果不定，通常有信號、物理整合的材料會增加凝視與參照，且學習效果較好，但仍需考量讀者特徵（如年齡）以及不同研究操弄的細節。最後，本文亦針對未來研究方向與實務應用提供建議。

關鍵詞：系統回顧、科學圖文閱讀、眼球追蹤、閱讀歷程

* 1. 通訊作者：簡郁芬，jianyucin@ntnu.edu.tw。

2. 本文為柯華蕓教授提議與組織「眼動與閱讀的回顧性研究」系列論文的其中一篇，謹以此文紀念和感謝柯教授。本文獲科技部「年輕學者養成計畫——哥倫布計畫」（MOST109-2636-H-003-003-，MOST110-2636-H-003-003-）和教育部高等教育深耕計畫核定之國立臺灣師範大學學習科學跨國頂尖研究中心經費支持，一併致謝。同時感謝審查委員的修改建議，使本文更加完善。

在科學教科書或科普文章中，圖文並置是常見的呈現形式，文本用來描述知識概念，圖像則帶有視覺與空間訊息，能描繪出文本較難呈現的概念關係、物體結構和發展歷程，兩者聯合透過多元表徵來傳達複雜的科學概念（簡郁苓、吳昭容，2012；Carney & Levin, 2002）。儘管視覺圖像的使用增加，但相對於文本理解，讀圖理解的研究仍屬起步階段（Guo, Zhang, et al., 2020）。目前科學圖文閱讀的實徵研究主要探討學習成效，研究結果莫衷一是，圖片效果仍屬分歧。有些研究發現在文本添加圖像的學習效果優於純文本，例如 Mayer 針對此一現象提出「多媒體原則」（multimedia principle），強調圖像帶來的正向學習成效（Mayer, 2005, 2014）；但有些研究發現圖像效果不明顯，甚至是負向效果（McTigue, 2009; Segers et al., 2008）。當不一致的證據浮現時，顯然必須再考慮更細微的問題：什麼原因造成學習成效結果不一致？其背後隱藏的圖文閱讀歷程為何？還有哪些調節因子會帶來閱讀歷程的改變？

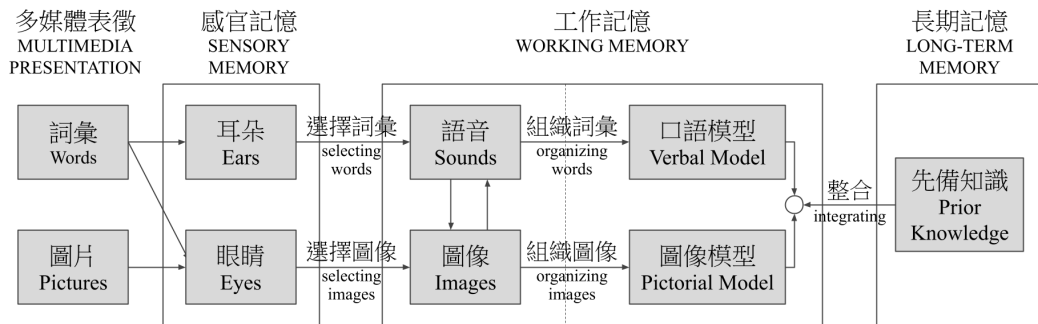
近二十年來，研究者針對圖文閱讀的認知機制提出一些理論模式，主要為 Mayer 的多媒體學習認知理論（cognitive theory of multimedia learning, CTML）（Mayer, 2005, 2014）及 Schnotz 的圖文理解整合模式（integrated model of text and picture comprehension, ITPC）（Schnotz & Bannert, 2003; Schnotz et al., 2014），兩者都認為讀者在進行圖文閱讀時是否能夠形成連貫的心理表徵，關鍵在於能否適切地參照與整合文字表徵與圖片表徵。儘管過去已有數篇針對圖文閱讀之文獻回顧，但發表年代已超過十年（Carney & Levin, 2002; Phillips et al., 2010; Vekiri, 2002），恐怕無法反映新近圖文閱讀研究的樣貌；此外，目前有關圖文閱讀的回顧性文獻多側重學習成效的討論（Carney & Levin, 2002; Guo, Zhang, et al., 2020; Phillips et al., 2010; Schnotz, 2014），然而僅透過離線（off-line）的成效測量無法獲取圖文處理與參照的資訊，若能進一步探究線上（on-line）即時的歷程資料，就更有機會了解讀者實際上如何閱讀，進而一窺閱讀歷程的黑盒子，探究讀者是否能讀好科學圖文的可能原因。目前已有一些科學圖文閱讀的眼動實徵，但迄今沒有系統性回顧來耙梳科學圖文閱讀在不同條件與介入操弄下的眼動型態，及其反映的閱讀歷程與學習表現之間的關聯。因此，本研究之目的在於系統性地回顧近 30 年來科學圖文閱讀的眼動研究，綜整文獻上科學圖文閱讀的歷程與樣態。

（一）圖文閱讀的理論基礎

同前所述，目前圖文閱讀研究大多以 Mayer 的多媒體學習認知理論（以下通稱 CTML）（Mayer, 2005, 2014）及 Schnotz 的圖文理解整合模式（以下通稱 ITPC）（Schnotz & Bannert, 2003; Schnotz et al., 2014）做為理論基礎；這兩個理論背後的假設都來自雙碼理論（dual coding theory）（Paivio, 1986），認為人類認知系統是由語言系統和圖像系統這兩種不同但彼此關聯的子系統組成。在圖像系統中，圖像的表徵形式類比於我們知覺到的形式，屬於類比式編碼（analog codes）；而語言系統中，對詞語和概念形成的表徵形式為符號性編碼（symbolic codes），與被表徵的對象之間並不具有知覺上的相似性。文句等語言訊息通常只能由語言系統進行處理與編碼；圖片訊息則可能透過圖像與語言系統進行處理與編碼（Paivio, 1986）。

Mayer（2005）的 CTML 架構如圖 1，包含三大假定。第一，「雙重通道」，人類訊息處理系統是由兩個獨立對稱的通道組成：聽覺—口語通道（auditory-verbal channel）以及視覺—圖像通道（visual-pictorial channel），前者用以處理聽覺輸入與口語表徵，後者則處理視覺輸入與圖像表徵。第二，「有限容量」，每一個子系統在工作記憶中運作的容量都是有限的。第三，「主動處理」，人會主動以有意義的方式建構知識，在多媒體學習中，主動處理需要五種認知歷程：選擇相關詞彙／圖片、將選擇的詞彙／圖像組織為口語／圖像模型、最後將兩種模型與先備知識整合為連貫的心理表徵。根據 CTML 與實徵結果，學習者從圖片與文字的學習效果優於只有文字的材料，Mayer 認為圖像有利於高層次的學習，能夠幫助讀者形成概念的心智模型，且工作記憶中對於訊息的心理保留與運作是有容量限制的，當材料中同時提供圖片與文字，能夠讓訊息透過口語和圖像雙重通道進行處理與編碼，降低認知超載（cognitive overload）的情況，進而促進學習。

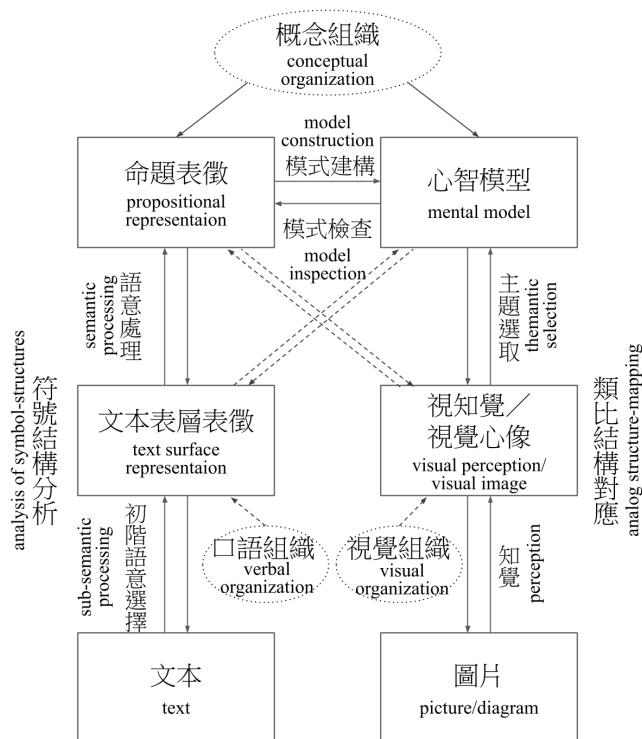
圖 1
多媒體學習認知理論 (CTML)



註：翻譯自“Cognitive theory of multimedia learning,” by R. E. Mayer, 2005, In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (p. 37). Copyright 2021 by the Cambridge University Press.

另一個廣為圖文理解相關研究採納的模式為 ITPC (Schnotz & Bannert, 2003; Schnotz et al., 2014)，架構如圖 2。此模式將表徵區分為描述性 (descriptive) 與描繪性 (depictive) 兩個路徑。描述性路徑包括外在文本、文本表層結構的內在心理表徵、以及文本語意內容的命題表徵，屬於符號處理的問題；至於描繪性路徑則包括外在圖片、圖片的內在視知覺或心像、以及圖片中呈現主題的心智模型，屬於因描繪性表徵之間類比關係而進行的結構對應問題。在文本理解中，讀者形成文本表層結構的心理表徵，觸發了概念組織歷程，從而產生了連貫的語意內容命題表徵，並建立所描述主題訊息的心智模型。至於在圖片理解中，讀者透過圖形實體的辨識和區辨等知覺處理，建立圖形的視覺心理表徵，產生的視知覺被假定保留在工作記憶的圖像部分，接著讀者透過語意處理來建立圖片中所呈現主題的心智模型和命題表徵。心智模型建構是透過基模驅動的對應過程所引導：將圖形實體對應到心智模型實體，並將空間關係對應到語意關係，屬於視覺空間關係系統和語意關係系統之間的類比結構對應歷程。與 CTML 不同的是，ITPC 模式中只會建構出一個心智模型，假定文本和圖片在本質上是不同形式的表徵，需要不同形式的認知處理，並因不同目的而具有不同用途 (Schnotz & Wagner, 2018)。

圖 2
圖文理解整合模式 (ITPC)



註：翻譯自“Construction and interference in learning from multiple representations,” by W. Schnotz and M. Bannert, 2003, *Learning and Instruction*, 13, p. 145 ([https://doi.org/10.1016/S0959-4752\(02\)00017-8](https://doi.org/10.1016/S0959-4752(02)00017-8)). Copyright 2021 by the Pergamon.

(二) 可能影響科學圖文閱讀的調節因子

多媒體原則指出，圖文材料的學習效果應該優於純文本（Butcher, 2014），然而 Guo、McTigue 等人（2020）的回顧中發現，圖像輔助並非總是能強化閱讀理解表現。為何多媒體原則並未獲得實徵研究上穩定一致的強韌證據？Vekiri（2002）綜整相關研究指出，讀者特徵以及材料特徵都會影響圖像對學習的效果。

年齡是讀者特徵中常被關注的變項之一。Mayer 團隊所進行的研究幾乎都以大學生為對象（Mayer, 2014），而大學生通常是較有能力、有技巧的讀者；McTigue（2009）及 Coleman 等人（2018）懷疑 CTML 是否同樣適用於年紀較小的讀者，結果發現添加圖像無法提升六年級兒童在物理文本的閱讀表現（McTigue, 2009），也未提升四年級兒童在生物與天文材料的名詞記憶和閱讀理解（Coleman et al., 2018），顯示多媒體原則很可能無法從大學讀者完全複製到年紀較輕的讀者（Segers et al., 2008），後設分析結果也呈現這樣的趨勢，從大學讀者得到多媒體原則的證據比中小學讀者更為強韌（Guo, Zhang, et al., 2020）。

就能力因子來看，各研究納入的能力概念不盡相同，包括智力、閱讀能力、空間能力等。過去研究顯示，高智力兒童比低智力兒童更能以視覺圖像做為閱讀輔助（Hannus & Hyönä, 1999; Parkhurst & Dwyer, 1983）；Höfler（2010）的後設分析發現，高空間能力者比低空間能力者更能從視覺圖像中獲益。然而有部分研究者持相反意見，認為高能力讀者會自行建構心智模型，視覺圖像對他們來說可能是冗餘甚至是干擾訊息，反而對低能力讀者的幫助較大（Schnotz, 2014）。儘管如此，多數文獻仍傾向於支持圖像對於高能力讀者更具優勢的說法（Mayer, 2014）。而圖像對於年紀較小、智力或空間能力較弱的讀者而言效果較為有限，原因可能在於其不懂得將注意力放在圖像上、

圖像解碼能力較弱、經常誤解圖像規則、或是缺乏有效獲取圖像中重要訊息的策略 (Coleman et al., 2018; Norman, 2012)。

在某些研究中，能力與領域知識、先備知識的概念重疊混用，例如 Scheiter 等人 (2018) 以「認知先驗條件」綜合了能力與領域知識的概念，Hegarty 與 Just (1993) 的「機械能力」實際上概念偏向於領域知識的測量 (Vekiri, 2002)；部分回顧性文獻也將領域知識和先備知識合併討論 (Phillips et al., 2010; Renkl & Scheiter, 2017; Vekiri, 2002)。在先備知識因子裡，不同觀點爭論得更為激烈 (Phillips et al., 2010; Vekiri, 2002)，其中一派研究者認為，高先備知識者能從圖文材料中學習更多 (Hegarty & Just, 1993)，Renkl 與 Scheiter (2017) 的回顧持此觀點，認為高先備知識者通常較有策略，能花費較少心力成功整合視覺與口語訊息 (Vekiri, 2002)，而低先備知識者很難自行建立圖像與其他表徵之間的連結。然而這派觀點似乎缺乏新近實徵研究的證據 (Phillips et al., 2010; Vekiri, 2002)。另一派研究者 Kalyuga 累積實徵研究後指出，對新手讀者 (低先備知識者) 來說有益的訊息對於專家讀者 (高先備知識者) 而言反而可能變得冗餘。圖像能提供額外的心智模型建構路徑，幫助低先備知識者建立他們自己很難形成的心智模型 (Schnotz, 2014)，但對於高先備知識者來說卻可能產生專家反轉效果 (expertise reversal effect) (Kalyuga et al., 2003)，不但無助於學習，甚至造成負向影響 (Mayer, 2014; Mayer & Gallini, 1990)。Vekiri (2002) 認為，兩派研究結果的差異可能與先備知識的測量，以及先備知識與圖像複雜度之間的交互作用有關。

至於材料特徵方面，Carney 與 Levin (2002) 根據圖像在文本處理中的功能做區分：表徵性圖像 (representational pictures) 用來指涉部分或全部的文本內容概念；解釋性圖像 (interpretational pictures) 有助於澄清困難文本，例如以泵系統表示血壓；裝飾性圖像 (decorational or decorative pictures) 的功能則是在感官上吸引注意力，通常無助於學習，甚至可能造成分心、遺漏重要的相關訊息，反而產生負向影響 (Harp & Mayer, 1998; Phillips et al., 2010)，尤其生手讀者更容易如此 (Renkl & Scheiter, 2017)。Guo、McTigue 等人 (2020) 認為圖像可能帶有不同的訊息複雜度，對理解會產生不同程度的挑戰 (Guo, Zhang, et al., 2020)，通常僅在讀者能以最少的認知處理來解釋和整合訊息時才有益 (Vekiri, 2002)。

部分研究著眼於材料操弄，CTML 以實徵研究為基礎提出多個多媒體教學原則，例如信號原則 (signaling or cueing principle) 是指材料中添加醒目提示來引導讀者注意相關元素或凸顯組織結構；空間接近原則 (spatial contiguity principle) 是指將相對應的文字與圖片放在接近的位置，會比分散放置時學習效果更好 (Mayer, 2014)。Renkl 與 Scheiter (2017) 及 Guo、McTigue 等人 (2020) 則針對教學介入進行回顧，這些介入包含三大面向：1. 提供教學，包括讀寫能力教學、圖像規則教學、領域特定教學等；2. 事前訓練，即在學習前提供與內容相關的訊息，包括眼動楷模示範、注意力分佈訓練、後設認知訓練等；3. 提示或策略，包括以提示促發學習策略、教學回饋或引導、自我解釋與後設認知策略等。目前研究顯示，教學介入大多能促進學習，但累積研究數量並不多。

整體而言，讀者特徵 (例如年齡、能力和先備知識)、材料特徵 (例如圖像類型、難度) 與介入都會影響圖文閱讀的效果。教學介入效果也經常受到其他讀者特徵調節 (Krebs et al., 2019; Mason et al., 2016; Scheiter & Eitel, 2015)，讀者特徵與材料特徵有時候也會產生交互作用，例如低先備知識者閱讀複雜材料時的圖像效果尤其明顯 (Carney & Levin, 2002)，但也特別容易受到誘人細節干擾 (Renkl & Scheiter, 2017)，這些都是在探討科學圖文閱讀時需要留意的。

(三) 過去圖文閱讀相關回顧性文獻的脈絡

過去有關圖文閱讀的回顧性文獻，焦點從探討視覺圖像的本質、讀者特徵，轉移至教學介入 (Guo, McTigue, et al., 2020)，以下摘述幾篇，並立基於過去研究，開展本篇回顧的重要概念。Carney 與 Levin (2002) 回顧 1990 年至 2002 年間發表的實徵研究，以量化方式探討「為什麼」和「何時」圖像有效，結果發現具有特定功能的圖像可以不同程度地提升學習。Phillips 等人 (2010) 系統性回顧了 1936 年至 2009 年的 247 篇實徵研究與討論文章，聚焦於視覺圖像的定義以及良好圖像的構成要素，最後建議視覺圖像輔助需與課程目標有關，圖像應做為文本的補充，而非直接替代，且圖像的內容比形式 (例如顏色、線條、逼真度等) 更為重要。然而 Carney 與 Levin 及 Phillips 等人的回顧都聚焦於材料特徵，較少考慮讀者變項與教學介入帶來的影響。

Vekiri (2002) 系統性回顧 1990 年至 2000 年的文獻，在不同的理論架構下檢驗圖像在學習中的作用，結果發現材料特徵和讀者特徵都會影響圖像對學習的效果。接著要問的是：怎樣的情況下圖像會輔助或阻礙學習？Renkl 與 Scheiter (2017) 界定出讀者在圖像學習時可能面臨四種困難：第一，讀者偏誤，指讀者傾向於閱讀較習慣的文本而忽略圖像，且容易對多媒體學習的理解程度感到過度自信；第二，先備知識與技能對學習會產生影響；第三，對視覺圖像的注意力分佈問題，例如注意力放錯重點；第四，視覺圖像與其他表徵之間的整合問題。此外，Renkl 與 Scheiter 還提供一些優化學習的輔助措施。然而，這份回顧並未系統性的考慮調節因子，也未呈現可複製的文獻搜尋與篩選程序。Guo、Zhang 等人 (2020) 認為在提供輔助前，需先系統性地探究「應該教哪種類型的圖片？」和「應該要教誰？」，他們透過後設分析回顧 1985 年至 2018 年的 36 篇文章 (39 個研究)，試圖了解靜態圖像對於閱讀理解的成效，以及圖像效果受到年級、圖像類型、測量題型與文本類型等變項調節的程度。結果發現圖像具有中度正向效果，年級的調節效果未達顯著，而單純的圖像比混合圖像 (使用兩類以上的圖像) 更有效，開放問答與混合測量比是非題表現更好。但這份後設分析排除許多更細緻的歷程資料 (例如眼動、放聲思考)，Guo、McTigue 等人 (2020) 則進行另外一項系統回顧，以 K-12 學生為對象，納入方法學上更為全面的研究。結果顯示圖像可以增強學生高層次的學習 (即分析、評估、應用和創造)，但單純納入圖像不能保證正向學習效果，建議教師應設計教學活動來促進高層次的圖像學習。

儘管上述幾個回顧各有獨特之處，但也形成新的研究需求。具體來說，Vekiri (2002)、Carney 與 Levin (2002) 及 Phillips 等人 (2010) 的發表年代已超過十年以上，恐怕限制了研究範疇，無法符應現代圖文閱讀型態的多樣性；目前有關圖文閱讀的回顧性文獻大多側重於學習成效的討論 (Carney & Levin, 2002; Guo, Zhang, et al., 2020; Phillips et al., 2010; Schnotz, 2014)，缺乏針對閱讀歷程的探討，少數回顧納入零星眼動研究做為效果解釋的歷程證據 (Guo, McTigue, et al., 2020; Renkl & Scheiter, 2017)，但焦點皆非透過眼動文獻探究閱讀歷程；因此本研究希望填補文獻上的缺口，針對科學圖文閱讀的眼動研究進行系統性回顧。

(四) 透過眼球追蹤技術探究閱讀歷程

眼球追蹤 (eye-tracking) 是指透過儀器追蹤眼球的移動，並將這些移動與刺激做連結，讓研究者了解個體將視覺注意力放在刺激的哪個位置、持續多久、有著怎樣的順序模式 (Holmqvist et al., 2011)。根據 Just 與 Carpenter (1980) 的立即性假設 (immediacy assumption) 與眼心假設 (eye-mind assumption)，立即性假設是指讀者在遇到文本時會立刻嘗試進行不同層次的語言處理，不會延遲而是立即發生；眼心假設則認為視覺資訊呈現時，個體的眼動資料能對應個體知覺與認知歷程之間的關聯，只要對單詞進行處理，眼球就會持續凝視該位置，凝視時間反映的就是處理時間。透過眼球追蹤通常可以取得兩項重要的測量資料：凝視 (fixation) 與跳移 (saccade)。凝視指的是眼睛幾乎靜止，僅有微小顫動，能從刺激中擷取訊息，也因此凝視的位置與持續時間能指涉出哪些訊息被觸及以及個體在處理該訊息的程度有多深，一個凝視點的持續時間通常在 100 毫秒到 500 毫秒之間 (Rayner, 1998)；跳移則是不同凝視之間的快速眼動，視覺注意力從一個焦點移動至另一個焦點，我們通常無法吸收其中的視覺訊息 (Rayner, 1998)。目前已有許多研究使用眼球追蹤技術探究各種材料的閱讀歷程，包括短文 (官英華, 2020)、笑話 (呂昕頤等人, 2019)、圖文 (Eitel, 2016)……等。分析眼球軌跡的指標龐雜將近百種，除了以整體角度檢視眼動資料，許多研究者會進一步以「感興趣區域」(area of interest, AOI)，例如以單詞、句子、區域為單位，或「感興趣時段」(period of interest, POI)，例如以問題解決前、中、後為觀察時間，切割出更細的分析單位 (陳學志等人, 2010)。本研究參考過去文獻 (陳學志等人, 2010; 簡郁苓、吳昭容, 2012; Alemdag & Cagiltay, 2018; Hyönä et al., 2003; Mason, Pluchino, Tornatora, & Aries, 2013) 及 Lai 等人 (2013) 的分類尺度 (即時間、空間、次數)，綜整為 14 個常見的圖文閱讀眼動測量指標¹：

1. 時間尺度指標

帶有時間性的眼動指標，通常用來回答認知處理上有關「何時」以及「多久」的問題。科學圖

文眼動研究常用的時間尺度指標有八項：

(1) 總凝視時間 (**Total fixation duration, TFD**)。落在 AOI 的凝視點時間加總，通常反映閱讀時認知處理的程度，凝視時間越長，代表讀者需要花費越多認知資源處理該區域的訊息 (陳學志等人, 2010; 簡郁芩、吳昭容, 2012)。

(2) 平均凝視時間 (**Mean fixation duration, MFD**)。將 AOI 中的總凝視時間除以凝視點個數，即每個凝視點平均的處理時間，平均凝視時間越長，通常反映解碼該 AOI 的訊息 (如：文字或圖片) 越困難 (Jian & Ko, 2017; Jian et al., 2019)。

(3) 總閱讀時間 (**Total reading time**)。在某個 AOI 中停留的所有時間加總，包含凝視與跳移的時間 (Lai et al., 2013)，反映的認知歷程與總凝視時間類似。

(4) 首次注視時間 (**First-pass fixation duration**)。凝視點第一次進入某個 AOI 一直到離開該 AOI 期間的所有凝視點時間加總，通常反映初始處理，如詞彙觸接後的語意處理 (Hyönä et al., 2003; Mason, Pluchino, Tornatora, & Aries, 2013)。

(5) 再次注視時間／回視時間 (**Revisited/Second-pass fixation duration**)。離開某個 AOI 後再次進入該 AOI 的凝視點時間加總，通常反映晚期處理，亦即理解整合，或有意圖的解決閱讀的不理解 (Jian et al., 2019; Mason, Tornatora, et al., 2013)。

(6) 佔總凝視時間比率 (**Percentage of fixation duration**)。其他凝視時間指標 (例如：某個 AOI 的總凝視時間、首次注視時間或再次注視時間等) 除以學習期間的總凝視時間，通常反映讀者的選擇性注意力分配 (簡郁芩、吳昭容, 2012; Alemdag & Cagiltay, 2018; Mason, Pluchino, Tornatora, & Aries, 2013)。

(7) 首次凝視前花費時間 (**Time to first fixation on AOIs**)。從閱讀開始或目標刺激出現後花費多少時間，讀者的凝視才第一次進入該 AOI，此指標通常反映的是該區域吸引讀者選擇性注意的程度 (Alemdag & Cagiltay, 2018)，越具吸引力則首次凝視前花費時間越短 (陳學志等人, 2010; Scheiter & Eitel, 2015)。

(8) 平均每單位凝視時間 (**Mean fixation duration per unit**)。將 AOI 內的凝視時間除以某個單位數量，例如字數、圖像數、單位面積等，藉此可以比較不同面積的凝視時間 (Mason, Pluchino, Tornatora, & Aries, 2013; Zander et al., 2017)。

2. 空間尺度指標

空間尺度上的眼動指標涉及位置、距離、方向、順序、凝視或跳移在空間上的配置，通常可以回答認知處理上有關「在哪裡」以及「如何」的問題。Lai 等人 (2013) 採納 Liversedge 與 Findlay (2000) 的觀點，認為這類測量可能與選擇性認知處理有關；然而 Alemdag 與 Cagiltay (2018) 則持不同觀點，認為掃視路徑可能反映整合層次的認知歷程，凝視位置則反映組織層次的認知歷程。科學圖文眼動研究常用的空間尺度指標有三項：

(1) 掃視路徑 (**Scanpath pattern**)。眼球移動順序形成的模式，反映出閱讀路徑 (簡郁芩、吳昭容, 2012; Lai et al., 2013)。

(2) 平均跳移距離 (**Mean saccade length**)。平均每兩個接續凝視點間的距離，通常反映訊息的密度程度。平均跳移距離越短，可能訊息密度越大，需要較高程度的認知處理 (簡郁芩、吳昭容, 2012; Jian et al., 2013)。

(3) 凝視位置 (**Fixation position**)。凝視點落在的位置，反映認知處理的位置 (Chuang & Liu, 2012; Lai et al., 2013)。

3. 次數尺度指標

基於次數或頻率來測量眼動情況的指標，通常用來反映視覺材料的重要性。科學圖文眼動研究常用的次數尺度指標有三項：

（1）總凝視次數（**Total fixation count, TFC**）。落在 AOI 中的凝視點個數加總，通常凝視次數越多反映更高程度的認知處理，與總凝視時間有高相關（Eitel, 2016; Schnotz & Wagner, 2018）。

（2）區域間跳移次數（**Transition between AOIs**）。讀者的凝視在 AOI 之間移動的次數加總，通常反映閱讀時段落或區域間相互參照整合的情況（簡郁芩、吳昭容，2012；Mason, Pluchino, Tornatora, & Aries, 2013）。然而，簡郁芩與吳昭容（2012）指出，若區域間跳移次數頻繁，也可能有兩種不同的解釋，一方面反映讀者參照頻繁而成功整合材料，另一方面也可能反映讀者因整合困難而持續嘗試。

（3）回視次數（**Regression count**）。凝視點離開某個 AOI 後再次進入該 AOI 的次數加總，通常反映的也是晚期處理，當讀者閱讀到不理解的内容，例如讀者閱讀能力較弱、文章的連貫性低或難度較高時，其回視次數以及再次注視時間通常都會增加（陳學志等人，2010；Mason et al., 2016; Schüler, 2017）。

值得一提的是，每篇研究使用的眼動指標內涵與名稱不盡相同，有時候會出現「異名同義」的情況，例如：Revisits、Revisited fixation count 與 Regression count 在各自研究的定義上都表示凝視點離開某個 AOI 後再次返回該 AOI 的次數；另一種情況是「同名異義」，例如：Dwell time 在部分研究中指的是總凝視時間（陳學志等人，2010；Stárková et al., 2019），但在其他研究指的是總閱讀時間（Höffler et al., 2017; Ketzer-Nöltge et al., 2019），因此必須更仔細地審視每個研究中眼動指標的定義。

（五）本研究

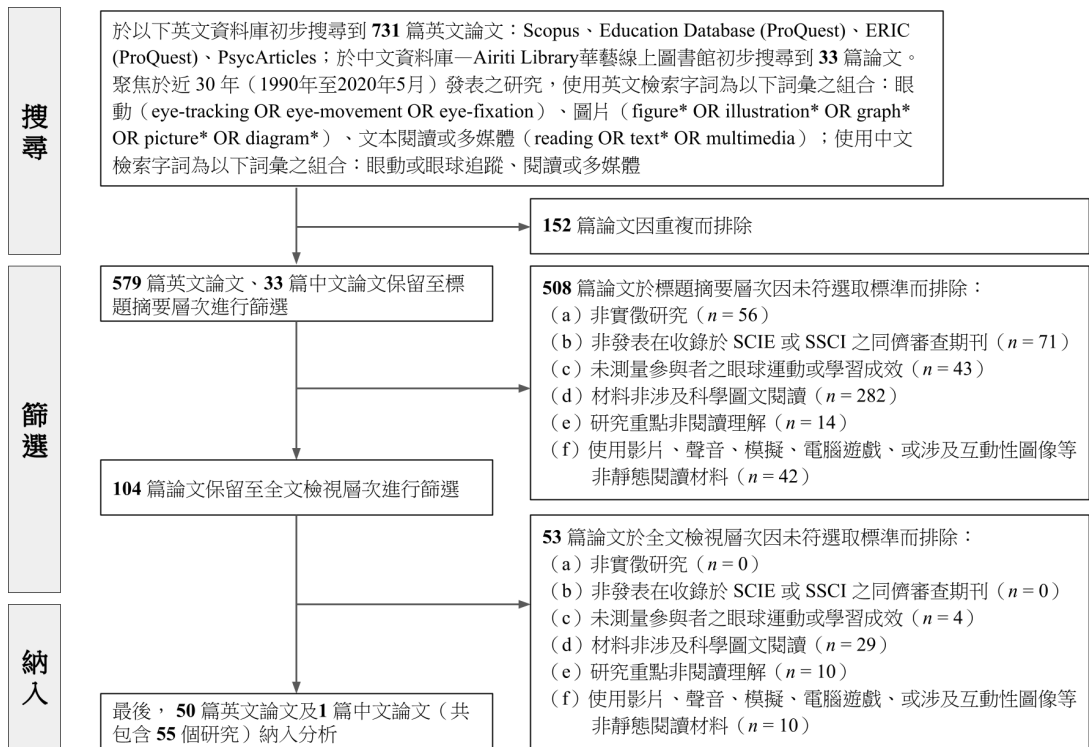
從上述說明可得知，眼動指標能反映出不同於學習表現的歷程資料，但目前仍缺乏針對此一範疇的回顧性文獻，因此，本研究之目的在於系統性回顧近 30 年來科學圖文閱讀的眼動研究，試圖回答以下研究問題：

1. 讀者在科學圖文閱讀的眼動型態為何？與學習表現之間有何關聯？
2. 怎樣的條件會影響以及如何影響讀者的科學圖文閱讀歷程及表現？
3. 操弄介入是否影響以及如何影響讀者的科學圖文閱讀歷程及表現？

方法

本研究參考 Guo、McTigue 等人（2020）發表於《Educational Psychology Review》之文獻回顧方法進行，文獻搜尋與選取歷程如圖 3 所示。

圖 3
文獻搜尋與選取歷程



（一）文獻搜尋

本分析聚焦於 1990 年至 2020 年 5 月發表之實徵研究。英文論文於以下資料庫進行搜尋：Scopus、Education Database (ProQuest)、ERIC (ProQuest)、PsycArticles，檢索字詞為以下詞彙之布林邏輯組合：眼動（eye-tracking OR eye-movement OR eye-fixation）、圖像（figure* OR illustration* OR graph* OR picture* OR diagram*）、閱讀或多媒體（reading OR text* OR multimedia），產生 731 篇文章，排除重複文章後有 579 篇。中文論文於 Airiti Library 華藝線上圖書館搜尋，檢索詞為「眼動或眼球追蹤」及「閱讀或多媒體」，產生 33 篇論文。

（二）選取標準

篩選與編碼歷程有兩階段：標題摘要層次與全文檢視層次。首先，我們使用以下標準檢視所有文章的標題與摘要：1. 實徵研究；2. 發表在收錄於 SCIE、SSCI、TSSCI 或 THCI 之同儕審查期刊；3. 測量參與者之眼球運動與學習成效；4. 材料涉及圖文閱讀，且內容範疇為科學；5. 主要聚焦於閱讀理解，而非問題解決或推理任務；6. 材料的圖像部分為研究者提供之靜態圖像；使用影片、聲音、模擬、電腦遊戲、或涉及互動性圖像（例如由讀者自行繪製圖像）等研究將被排除。英文論文在標題摘要層次篩選後，保留 103 篇文章，接著在閱讀論文內文後，依照上述六標準再進一步篩選出 50 篇英文論文。中文論文於兩階段篩選後保留一篇。最終中英文論文共有 51 篇（涵蓋 55 個研究）納入分析。

（三）分析與詮釋階段

本回顧使用歸納派典分析（inductive paradigmatic analysis; Polkinghorne, 1995）來處理最終納入的 55 個研究。歸納派典分析能夠從內容中耙梳並呈現出不同類別的主題，使我們能夠以具邏輯性的方式納入並綜整各種研究設計的實徵研究；這些主題並未由研究者預先確定，而是以歸納的方式在分析過程中逐步浮現，從而使我們能夠掌握所回顧研究的複雜特徵，並將其歸類為更廣泛的主題（Guo, McTigue, et al., 2020）。本研究參考 Guo、McTigue 等人（2020）以歸納派典分析進行系統性回顧的方法，首先從回顧的研究中提取描述性訊息（例如：參與者特徵、研究設計、使用眼動指標、研究發現等，示例如表 1）；接著在確定關鍵資訊與主要研究發現後，將描述性訊息進行暫定編碼；接下來比較各研究所編代碼的相同與相異之處，並將研究進行主題分類，把屬性相似的研究放在一起以回應研究問題；最後根據不同主題類別來呈現研究發現。

表 1
個別研究之質性描述與編碼示例

研究	參與者	材料領域	研究設計	眼動指標	主要研究發現
Johnson & Mayer (2012, exp.1)	大學生 $N = 48$	一篇圖文 （機械）	受試者間：材料 操弄（圖文分散 組 vs. 圖文整合 組）	區域間跳移次數 （及比率）、總 凝視時間比率	（1）整合組遷移表現優於分散組，保留亦較高但未顯著。（2）整合組有更多整合移動與相應移動，但圖像凝視無差異。兩組的文區凝視比例皆高於圖區。

註：因篇幅限制，完整內容請見補充資料二，連結：http://bit.ly/ESM2_2021

結果與討論

（一）研究特徵描述

本節以每十年分群，根據研究參與者、材料內容、發表期刊出處、眼動儀與取樣率、眼動指標等向度進行研究特徵的綜整²。

1. 研究參與者

科學圖文閱讀眼動研究大多以大學生為對象（ $n = 32$ ），近七年來以小學後期（ $n = 14$ ）和中學前期（ $n = 10$ ）為參與者的研究逐漸增加，主要來自 Jian、Mason 及 Schnotz 等三組研究團隊。學前兒童、小學前期、中學後期與成人的眼動研究甚少（參閱補充資料三 A）。若有研究橫跨兩種年齡類別（ $n = 6$ ）則皆計入。

2. 材料內容

超過半數的材料領域涉及生物／醫學（ $n = 34$ ），其次為機械／物理（ $n = 13$ ）、地球科學／地理（ $n = 10$ ）、心理學／認知神經科學（ $n = 5$ ）及化學（ $n = 3$ ）等（參閱補充資料三 B）。若有研究涉及兩種以上的領域類別（ $n = 8$ ）則皆計入。

3. 發表期刊出處

發表於《Learning and Instruction》最多（ $n = 6$ ），接續為《Computers & Education》次之（ $n = 5$ ）、《Journal of Computer Assisted Learning》（ $n = 4$ ）、《Computers in Human Behavior》（ $n = 4$ ）和《Applied Cognitive Psychology》（ $n = 3$ ），其他心理、閱讀、教育科技、科學教育等相關期刊皆一或二篇（參閱補充資料三 C）。

4. 眼動儀

整體而言圖文閱讀眼動研究幾乎都是使用 SMI、EyeLink 與 Tobii 等三大廠牌的眼動儀（參閱補充資料三 D），主要型號為 SMI RED 250（ $n = 10$ ）、EyeLink 1000（ $n = 10$ ）、Tobii T120（ $n = 9$ ）。特別注意的是，眼動儀型號與研究發表團隊有很大的關係，同一團隊發表的研究通常會使用同一款眼動儀。

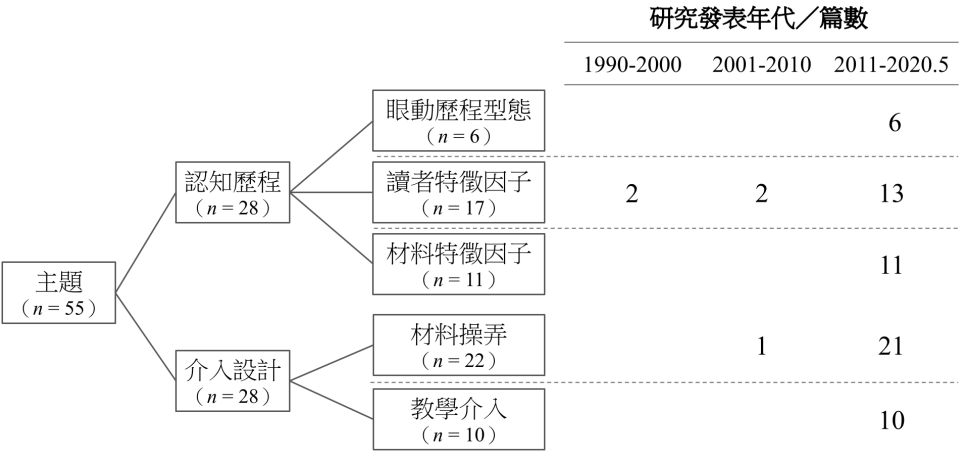
5. 眼動指標

圖文閱讀的眼動研究最常使用的眼動指標為總凝視時間（或比率）與區域間跳移次數（或比率），分別有 41 篇及 36 篇，其次為總凝視次數（ $n = 19$ ）、平均凝視時間（ $n = 12$ ）、總閱讀時間（ $n = 12$ ）；亦有研究將閱讀處理區分為初始處理與晚期處理，即首次注視時間（ $n = 12$ ）、再次注視時間（ $n = 11$ ）以及回視次數（ $n = 6$ ）；整體而言，使用掃視路徑（ $n = 5$ ）、平均跳移距離（ $n = 4$ ）、凝視位置（ $n = 2$ ）等空間尺度指標的研究很少，與 Lai 等人（2013）及 Alemdag 與 Cagiltay（2018）的發現類似（參閱補充資料三 E）。

（二）研究綜整

經歸納派典分析浮現兩大主題：認知歷程與介入設計，主題階層架構如圖 4，研究綜整結果摘要如表 2。首先，在認知歷程主題中以「眼動歷程型態」回應研究問題一：「讀者在科學圖文閱讀的眼動型態為何？與學習表現之間有何關聯？」再以「讀者特徵因子」與「材料特徵因子」回應研究問題二：「怎樣的條件會影響以及如何影響讀者的科學圖文閱讀歷程及表現？」而材料特徵因子經常與讀者特徵因子產生交互作用，因此一併呈現。接著在介入設計主題中以「材料操弄」與「教學介入」等兩個子主題來回應研究問題三：「操弄介入是否影響以及如何影響讀者的科學圖文閱讀歷程及表現？」由圖 4 可知，早期的數篇圖文閱讀眼動研究多在探究認知歷程，近 10 年來的介入設計研究較過去更為盛行。

圖 4
主題階層架構



註：由於部分研究涵蓋兩個以上的主題，例如研究中同時探討讀者特徵因子與材料特徵因子，因此各子主題合計數量會超過上層主題之研究總數。

表 2
研究綜整主要結果摘要表

研究主題	眼動型態	學習效果
認知歷程		
眼動歷程型態	讀者主要以文字驅動而非圖像驅動閱讀行為，圖像通常做為鷹架輔助	重視圖像及圖文參照者通常表現較佳
讀者特徵因子	圖像檢視 (1) 高年級圖像檢視多於低年級 (2) 高能力圖像檢視多於低能力 (3) 不同先備知識者在圖像檢視的多寡未有一致趨勢 (4) 視覺者圖像檢視多於語文者	(1) 高年級表現優於低年級 (2) 高能力表現優於低能力 (3) 添加圖像有利於先備知識高或低者的表現未有一致趨勢 (4) 不同認知風格者的表現未有一致趨勢
	文字檢視 (1) 低先備知識者文字檢視多於高先備知識者 (2) 語文者文字檢視多於視覺者	
	圖文參照 (1) 高年級圖文參照多於低年級 (2) 高能力圖文參照多於低能力 (3) 不同先備知識者在圖文參照的多寡未有一致趨勢 (4) 不同認知風格者在圖文參照的多寡可能無差異	
	(1) 讀者會因應文章難度而有不同的注意力分配：高能力在困難文章投注較多心力、低能力在中低難度文章投注較多心力 (2) 高能力與高年級會隨著項目難度增加而增加圖像檢視與圖像一項目間的參照 (3) 重讀時文字檢視減少、圖像檢視增加 (4) 不一致訊息讓一開始閱讀時有更多的凝視、回視與參照	(1) 在中低難度材料的表現優於困難材料 (2) 在重讀的表現優於初讀 (3) 讀者表現未受到材料中不一致訊息的影響
介入設計		
材料操弄	(1) 相較於無信號的區域，讀者更快觀看有信號的區域，且可能在低先備知識者的效果更大 (2) 讀者在解釋性圖像的檢視多於裝飾性圖像 (3) 空間整合時的圖文參照多於空間分散時	(1) 在有信號材料的表現優於無信號材料，但年輕讀者不一定有此趨勢 (2) 空間接近效果視文字是否必要而定，非必要文字不會產生空間接近效果 (3) 在不同類型圖像的表現未有一致趨勢
教學介入	教學組在圖像檢視、圖文參照多於無教學組	教學組表現優於無教學組，且在低能力與低先備知識者的教學效果更大

註：本表呈現回顧研究之結果趨勢，需特別留意各主題涵蓋的研究數量可能較少，以及其他矛盾的研究結果。

1. 認知歷程

(1) 眼動歷程型態。整體而言讀者主要以文字驅動 (text-driven) 的方式閱讀科學圖文，在文區的凝視明顯較多 (簡郁芬、吳昭容, 2012; Hannus & Hyönä, 1999; Hung, 2014)，Hegarty 與 Just (1993) 發現讀者在閱讀機械相關圖文時會以累進的方式重複閱讀文本，在基於文本的層次上整合訊息，然後藉助圖像精緻化心智模型。Schnotz 的研究團隊提出更為完整的概念，進行一系列的眼動研究驗證其 ITPC 模式，結果發現在初始心智模型建構 (initial mental model construction) 時，讀者的確傾向於以文本驅動多過於圖片驅動的方式進行閱讀，圖像僅充當讀者的外部鷹架，支持整體

連貫形成策略；一旦初始心智模型建構後，讀者在適應心智模型精緻化階段（adaptive mental model elaboration）似乎傾向於圖片驅動多過於文本驅動，會選擇性地依照需求處理圖像，支持任務特定訊息選擇策略（Schnotz et al., 2014; Schnotz & Wagner, 2018; Zhao et al., 2020）。

此外，讀者面對科學圖文有不同類型的閱讀模式與策略類型，然而從讀者眼動行為進行的分群結果並不相同。Mason、Tornatora 等人（2013）找出四年級讀者的三種閱讀模式分別為低度整合者（14%）、中度整合者（39%）、高度整合者（47%）；Jian 等人（2019）則找出六年級讀者採用四種閱讀策略，包括初步全面檢視者（21%）、淺層處理者（58%）、文字主導者（12%）與圖像主導者（9%）；Jian 與 Wu（2015）從大學讀者的眼動資料找出圖文參照者、圖像優先者、文字優先者與只讀文字者等四種閱讀類型；Scheiter 等人（2019）透過集群分析將大學讀者的閱讀模式分為三群，其中集群 3 的人數過少，因此僅比較集群 1 和集群 2。集群 1 在圖片和文字凝視時間、次數與圖文參照次數明顯高於集群 2，較接近 Mason、Tornatora 等人的高度整合者以及 Jian 與 Wu 的圖文參照者；集群 2 的圖片和文字凝視時間、次數與圖文參照次數都是最少的，接近 Jian 等人的淺層處理者與 Mason、Tornatora 等人的低度整合者。整體而言，以上分群結果大多牽涉圖像處理、文字處理或圖文整合在程度上的差異，通常低度處理者較少注意圖像，整合移動也較少；高度處理者在初始處理時特別注意圖像，也重視晚期的圖文處理與圖文整合。

研究亦發現閱讀歷程中重視圖像處理以及圖文參照的讀者，通常閱讀表現較佳（Hung, 2014; Jian, 2017; Jian et al., 2019; Jian & Wu, 2015; Scheiter et al., 2019; Zhao et al., 2020），尤其帶有標籤的圖像能促進文字解碼初步處理的語意資訊（Jian & Wu, 2015）。若再將閱讀分為問題未呈現階段與問題回答階段，表現越好的學生在問題未呈現階段投注越低的時間在文本，但在問題回答階段投注越多的時間在圖像（Zhao et al., 2020）。呼應過去研究，讀者經常跳過或僅略讀圖像，但當他們閱讀圖像時，從事的是更高層次的認知活動（Cromley et al., 2010）。高度整合者在立即學習表現優於低度整合者（Mason, Tornatora, et al., 2013），也有較好的遷移表現（Scheiter et al., 2019）。

綜上所述，眼動實徵結果大致支持 CTML 及 ITPC 背後的雙碼理論假設（Paivio, 1986），亦即人類認知系統由語言系統和圖像系統這兩套子系統組成。一般而言，讀者傾向於以文本驅動多過於圖片驅動的方式進行閱讀，但卻可能有不同的閱讀模式，牽涉到圖文處理與整合程度的高低。通常閱讀歷程中重視圖像處理以及圖文參照的讀者，閱讀表現較佳；然而目前研究仍無法確切回應閱讀表現與讀者初始處理或晚期處理之間的關聯性。

（2）讀者特徵因子。研究中納入的讀者特徵因子主要包括：年級或年齡、學校屬性、能力、先備知識或領域知識、認知風格等。一般而言，高年級讀者比年輕讀者更重視圖像以及圖文參照，在閱讀時較會同時使用文本和圖像兩種表徵，通常也有較好的學習表現（Jian, 2016; Schnotz & Wagner, 2018）；文本處理會隨著年級增加，這可能與閱讀能力的提升有關，但圖像處理則否（Schnotz & Wagner, 2018）。少數兩則針對老年人的研究發現，老年人初步很少看圖，傾向於在閱讀時以及完成文本閱讀後對圖片有較長一點的凝視（Liu et al., 2009），整體而言較少注意圖像，在圖像理解與圖文整合似乎較有困難，圖像理解的表現也較差（Bol et al., 2016; Liu et al., 2009），但文本理解的正確度優於年輕人（Liu et al., 2009）。老年人的閱讀模式很可能不同於年輕人，然而由於相關研究數量極少，有待後續研究的驗證。

在學校屬性（即學術導向與非學術導向）的部分，讀者在初始心智模型建構的文本與圖像處理並未因學校屬性而改變，但學術導向學生在材料初次呈現時有較高的圖文參照次數比例，且在後續呈現時，圖文參照次數比例的下降幅度比非學術導向學生更大（Schnotz & Wagner, 2018）。年級與學校屬性上的差異很可能與能力有關，一般而言高年級與學術導向的學校會是能力較佳的學習者（Zhao et al., 2020），高年級相對於低年級的眼動型態與學習表現，類似於學術導向學校相對於非學術導向學校的趨勢（Schnotz & Wagner, 2018）。

在能力因子的部分，儘管 Mason、Tornatora 等人（2013）指出空間能力並未與任何眼動指標相關，但整體而言眼動文獻傾向支持圖像對於高能力者更為有利的說法，高智力及高閱讀能力者在閱讀處理上稍微更具策略，在文、圖相關部分花費了相對更多時間，尤其在閱讀困難文章的初始階段，注意力會回到文本前一段，且在兩張圖像之間來回凝視，低能力者較少觀察到此情況，在閱讀初始階段即遭遇困難（Hannus & Hyönä, 1999; Jian & Ko, 2017）。Mason、Tornatora 等人（2015）亦發

現閱讀能力能預測圖像回憶與遷移表現。但不論能力如何，基本上仍是文本驅動的方式，讀圖時間甚少（Hannus & Hyönä, 1999; Jian & Ko, 2017）。

在先備知識的部分，Morrow 等人（2012）發現高知識者傾向於在讀完文本後大量讀圖，低知識者則傾向於在整段閱讀中看圖。Hegarty 與 Just（1993）發現低機械領域知識者在句子的部分有更多重讀、圖文跳移與圖像的區域檢視（即解碼兩、三個部件間的關係，這相對於整合更多部件間關係的全域檢視），因此認為高低機械領域知識者的理解歷程，主要差異在於整合文本句子的訊息與圖像的區域表徵建構。然而有些研究卻發現，先備知識很可能與跳移次數、文到圖再次注視時間、回視次數有正相關（Ho et al., 2014; Mason, Tornatora, et al., 2013），或者無法提供太多先備知識、理解表現與整體圖文眼動指標有關的證據（Morrow et al., 2012），Scheiter 等人（2018）就發現認知先驗條件（涵蓋能力與領域知識的綜合分數）並未影響讀者文字與圖片的閱讀時間以及圖文跳移次數。此外，多數考慮先備知識的研究並非以先備知識做為研究關心的自變項，而是做為研究控制（例如，Johnson & Mayer, 2012; Koć-Januchta et al., 2017; Korbach et al., 2020; Krebs et al., 2019; Lin et al., 2017; Ozelik et al., 2009）。

認知風格為另一個可能影響讀者視覺行為的因子。語文者（verbalizer）傾向於依賴文本、有更多的文字間跳移，視覺者（visualizer）則傾向於依賴圖像、有更多的圖片間跳移，然而兩者在圖文間跳移次數並無差異（Höfler et al., 2017; Koć-Januchta et al., 2017）。Höfler 等人（2017）進一步將視覺者區分出兩種類型後，發現空間視覺者（spatial visualizer）在圖文上並無明確的偏好，未若物體視覺者（object visualizer）那般注意圖像。Liu（2018）關注的議題則為場地依賴性，他發現場地依賴者的凝視次數、跳移距離與回視次數皆比場地獨立者更高，呈現出較為全面的搜尋模式。在學習成效上，視覺者與場地獨立者似乎較有優勢（Koć-Januchta et al., 2017; Liu, 2018），但另一研究並未發現不同認知風格者在視覺行為上的差異會影響學習成效（Höfler et al., 2017）。在圖文閱讀的眼動研究中探討認知風格的文獻甚少，未能獲得一致的結果傾向。

綜上所述，年級較高的讀者比年輕讀者有更多的圖像檢視及圖文參照，通常也有更好的學習表現；類似的趨勢反映在學術導向學校相對於非學術導向學校，以及高能力者相對於低能力者。然而能力因子涵蓋內容甚廣，又經常與領域知識、先備知識的概念重疊，研究結果較為混雜。在先備知識上，低先備知識者的文字檢視多於高先備知識者，但在圖像檢視以及圖文參照上未有一致趨勢，添加圖像有利於何者的表現也未有定論。在認知風格上，視覺者的圖像檢視較多，語文者的文字檢視較多，不同認知風格者在圖文參照的多寡上可能無差異，學習表現則未呈現一致趨勢。

（3）材料特徵因子及其與讀者特徵的交互作用。材料難度、閱讀階段順序與內容一致性是較為常見的材料特徵因子。Hegarty 與 Just（1993）發現讀者在較複雜的機械材料有較多的全域檢視（整合許多部件之間的關係），但在閱讀不同複雜度的材料時的區域檢視（對兩個或三個部件之間的關係進行編碼）則無差異，顯示讀者能回應高複雜度材料的整合需求。Jian 與 Ko（2017）發現文章難度與閱讀能力在學習表現與眼動行為產生交互作用，比起簡單文章的處理，四年級高能力讀者投注較多時間心力處理困難文章；相反的，四年級低能力讀者讀困難文章很快就放棄，反而願意多花時間仔細閱讀與之能力較相當的中低難度文章。

Schnotz 研究團隊一系列的眼動研究中同時呈現圖文材料與題目，但題目由易到難依序呈現，結果發現高年級與學術導向學生會根據項目難度增加而調整策略，例如花費較少的注意力處理文本、更密集的處理圖像，圖像與項目間的參照次數亦隨著增加，也的確有更好的學習表現，但低年級與非學術導向學生的眼動並未呈現此調整性（Hochpöchler et al., 2013; Schnotz et al., 2014）。由上述研究可知，眼動資料或許能提供自我調整閱讀的證據。Rop 等人（2018b）將閱讀材料區分為前三頁與後三頁，眼動資料發現，讀者在高任務經驗的後三頁會降低對非必要文字的注意；Eitel（2016）則發現給予大學讀者第二次閱讀與測驗的機會時，增加了在圖片上相對於文本上的注意力（即文區凝視次數下降、圖區平均凝視時間增加），也與學習成效的進步有關，顯示重複閱讀與重測有助於降低多媒體學習中對文本的過度依賴，進而幫助表現；首讀較差的表現與重複閱讀時更長的閱讀時間有關，支持「根據任務需求調整」（Adaptation to task demands）的假設。

有些研究者關心讀者如何處理材料中訊息的不一致，結果發現不一致訊息提高了讀者的訊息處理需求，的確會影響其視覺行為（Mudrick et al., 2019; Scheiter & Eitel, 2015; Schüler, 2017, 2019），

例如有更多的凝視、回視與參照 (Schüler, 2017, 2019)，但可能僅在閱讀的一開始引起注意，後來讀者更大量地注意相對應的其他訊息 (Scheiter & Eitel, 2015)，甚至僅有 22% 的參與者記得材料中帶有不一致訊息 (Schüler, 2017)。然而閱讀帶有不一致訊息的讀者和一般讀者在學習成效上大致沒有差異，似乎未產生負向影響 (Scheiter & Eitel, 2015; Schüler, 2017, 2019)。Mudrick 等人 (2019) 則進一步發現，圖文不一致可能較會影響讀者的後設監控與推論理解表現，但文本內的不一致卻影響不大；眼動資料可看出，讀者在圖文不一致的材料會有更多基於事件的圖文間跳移，以及基於時間的圖像長時間凝視到文本凝視，顯示讀者已偵測出圖文間的不一致訊息並會來回確認。整體而言，不一致訊息會影響讀者視覺行為，但未必對閱讀表現產生傷害。

綜上所述，讀者會因應文章或題目難度而有不同的注意力分配，而年級、能力與學校屬性可能在其中扮演調節角色。高能力者在困難文章投注較多心力，低能力者在中低難度文章投注較多心力；高能力與高年級會隨著項目難度提高而增加圖像檢視與「圖像—項目」間的參照；第二次閱讀會減少文字檢視、增加圖像檢視，顯示出讀者的調整性。材料中出現不一致訊息時，會讓讀者一開始有更多的凝視、回視與參照，但可能並未影響閱讀表現。

2. 介入設計

(1) 材料操弄。包含 19 篇論文中的 22 個研究，大多以大學生或成人為研究對象，主要包括以下三種類型：第一，信號 vs. 無信號 ($n = 10$)：材料是否以色彩、箭頭、標籤等方式提示重點，即信號原則的操弄；第二，空間整合 vs. 分散 ($n = 5$)：是否將相關的圖文訊息放在較接近的位置，即空間接近原則的操弄；第三，圖像或文字敘述類型 ($n = 8$)：例如將圖像區分為認知圖或情感圖、裝飾性誘人細節或教學圖、詳細圖或簡化圖、個人化或傳統型等。儘管有部分研究 ($n = 7$) 操弄多媒體原則（即圖文與純文本的比較），但這些研究的焦點大多放在其他變項，例如年齡 (Bol et al., 2016; Liu et al., 2009)、不一致訊息 (Schüler, 2019)、領域知識與能力 (Meppelink & Bol, 2015)；或是做為其他操弄的效果對照，例如醒目標示 (Mason, Pluchino, & Tornatora, 2013)、圖像類型 (Mason, Pluchino, Tornatora, & Aries, 2013)、重複閱讀與重測 (Eitel, 2016)；此外，圖文材料與純文本材料的眼動指標較難直接比較，因此將相關研究納入其他主題討論。

a. 信號 vs. 無信號。部分研究顯示，信號標示有助於圖文整合表現 (Scheiter & Eitel, 2015) 或遷移表現 (Mason, Pluchino, & Tornatora, 2013; Ozcelik et al., 2009)，通常讀者能夠更快（即首次凝視前花費時間少）且更密集地（即凝視次數或平均凝視時間高）注視信號元素 (Ozcelik et al., 2009; Scheiter & Eitel, 2015)，幫助重要圖像與文本間的晚期處理 (Mason, Pluchino, & Tornatora, 2013)。另外亦發現部分效果，信號僅影響了低先備知識者的注視行為，會較早觀看圖像，進而支持學習表現，但不影響高先備知識者的眼動行為與學習表現 (Richter & Scheiter, 2019)。甚至 Jian (2019) 發現信號組與無信號組在各眼動指標皆無差異，且信號組的學習表現並未優於無信號組，因此懷疑信號原則本身很可能無法類推至六年級的年輕讀者。除了年齡因素以外，推測造成研究結果分歧的原因之一可能也與信號的樣態有關。相較於其他研究以粗體或彩色標示重點或專業術語 (Ozcelik et al., 2009; Richter & Scheiter, 2019; Scheiter & Eitel, 2015)，或是圖像上添加箭頭 (簡郁苓、吳昭容, 2012; Jian & Wu, 2016) 或標籤圖例 (Johnson & Mayer, 2012; Mason, Pluchino, & Tornatora, 2013; Scheiter & Eitel, 2015)，Jian 的信號是在文本中標示出「圖 1」、「圖 2」字樣，或許醒目程度較低，因此效果有限。這樣的推測有賴後續研究驗證。信號標示搭配其他操弄，例如物理位置上的整合 (Johnson & Mayer, 2012) 或教學介入 (Jian, 2019)，通常在閱讀時會有更多的圖文跳移整合，也會明顯地比控制組有更好的學習表現。

b. 物理空間整合 vs. 分散。物理空間上的整合會促進讀者的圖文跳移 (Johnson & Mayer, 2012; Rop et al., 2018a, 2018b)，當整合的文字為機械運作說明時，空間整合有助於學習表現 (Johnson & Mayer, 2012, exp. 2)；但若將整合形式與圖例形式做比較，整合組的學習表現並未優於圖例組，眼動型態也呈現更為複雜的「圖例—文本」和「圖例—圖像」路徑 (Johnson & Mayer, 2012, exp. 3)；當整合的文字為非必要文字（即文字與圖像傳遞冗餘訊息）時，空間整合無助於學習 (Rop et al., 2018b)，甚至反而會阻礙學習 (Rop et al., 2018a)。

c. 圖像或文字敘述類型。一般而言圖像訊息量會影響讀者投注的心力，例如 Hung (2014) 在小樣本的六年級學生中發現訊息量最低的裝飾性圖像比表徵性和解釋性圖像獲得最少注意（凝視時間最短、跳移距離最長）；Lin 等人 (2017) 發現詳細圖像的讀者比簡化圖像的讀者在圖像總閱讀時間、比例以及圖像晚期處理各指標都來得高，且在閱讀甫開始時，詳細圖像更快能吸引讀者注意；Stárková 等人 (2019) 發現擬人圖只在每頁閱讀最初兩秒時，比一般示意圖吸引較多的注意，但不影響整體閱讀的投入程度。有些研究從不同讀者特徵來探究圖像類型的差異，Bol 等人 (2016) 發現中壯年人對認知圖的注意力比老年人更多，並且回憶更多的訊息，但在情感圖則無此趨勢；Jian (2019) 從閱讀模式著眼，發現初步全面檢視者在解釋性圖像比表徵性圖像有更多整體凝視以及晚期處理，圖像主導者則有更多初始處理，至於淺層處理者及文字主導者在處理兩種圖片上並無差異。Zander 等人 (2017) 是唯一操弄個人化語言的研究，閱讀個人化語言的讀者有較多的文區凝視、較少的圖區凝視和圖文跳移次數。整體研究顯示，圖像或文字敘述類型會影響讀者的眼動行為，但可能不足以影響學習成效以及主觀情意層面的測量 (Lin et al., 2017; Stárková et al., 2019; Zander et al., 2017)。Lin 等人進一步發現，在簡化圖條件下，如果讀者閱讀簡化圖片且重新讀圖的時間比例越高，理解表現越好，然而在詳細圖條件下，眼動與理解測量無相關，顯示眼動指標與學習表現的關係有可能受到圖像類型的調節。

(2) 教學介入。此主題的 10 個研究皆在 2015 年後發表，主要來自 Jian、Mason 與 Scheiter 等三個研究團隊，研究參與者分別為小學後期（四和六年級）、中學前期（七年級）以及大學生，材料則分別為中文、義大利文與德文。教學介入的內容最大宗為眼動楷模示範 (eye-movement modeling examples, 以下通稱 EMME)，共有五個研究，指的是在閱讀前播放優秀讀者在閱讀該材料時的眼動軌跡做為示範引導 (Krebs et al., 2019; Mason, Pluchino, et al., 2015; Mason et al., 2016; Mason et al., 2017; Scheiter et al., 2018)；另有三個研究為閱讀策略教學，例如教導讀者讀圖或進行圖文整合 (Jian, 2018, 2019; Rop et al., 2018a)；尚有研究探討執行意向 (implementation intentions; Stalbovs et al., 2015) 與追描手勢 (tracing gestures; Korbach et al., 2020) 的教學。執行意向指的是「如果—那麼」計畫 (“if-then” plans)，假定採取行動的機會 (例如：「如果我已經讀完一段段落」) 有助於實現學習目標的行動 (例如：「那麼我會搜尋圖片中相對應的訊息」)，以便在遇到機會時自動執行該行動；追描手勢則是以肢體 (此處為手指) 指出並追蹤材料來幫助閱讀。多數研究指出，教學介入會增加讀者的圖片檢視以及圖文整合的嘗試 (Jian, 2018, 2019; Korbach et al., 2020; Mason et al., 2016; Mason et al., 2017; Scheiter & Eitel, 2015; Stalbovs et al., 2015)。Rop 等人 (2018a) 發現整合教學組比圖文分散組花費較少時間閱讀非必要文字，也比圖文整合組在非必要的文字及相關圖之間有較少的跳移，或許因為 Rop 等人操弄的內容為非必要文字，該研究並未找到整合教學組的優勢。Mason、Pluchino 等人 (2015) 的相關分析顯示，從相關文字到相關圖片的凝視時間與回憶、遷移表現有關；Stalbovs 等人 (2015) 指出，圖文跳移次數能預測學習成效；Mason 等人 (2017) 甚至發現 EMME 教學對各項學習表現的效果受到眼動指標 (即從文到圖或圖到文的再次注視時間、圖文跳移次數) 的完全中介或部分中介。

此外，閱讀理解能力和領域知識對教學介入可能有調節效果：低閱讀能力者若能觀察眼動楷模會比沒有觀察眼動楷模者有更好的事實知識與遷移表現；高閱讀能力者的表現則不受能否觀察眼動楷模而影響 (Mason et al., 2016)；低領域知識者中，EMME 組在圖片凝視時間明顯長於無楷模組，但在高領域知識者則無此差異 (Krebs et al., 2019)。Krebs 等人 (2019) 進一步發現調節中介效果，在低領域知識讀者中，若告知參與者所觀摩的眼動來自同儕 (同儕楷模) 能促進更長的圖片凝視時間，進而有利於理解表現，若告知參與者所觀摩的眼動來自表現優異者 (勝任楷模) 則無此效果。Korbach 等人 (2020) 的追描策略教學研究亦發現，視覺注意力的跳移確實可以解釋對辨認測驗表現的正向影響。

整體而言，教學介入通常會影響讀者的眼動型態，很可能透過圖文參照與圖片檢視等相關視覺行為為中介，進而提升學習成效 (Korbach et al., 2020; Mason et al., 2017)，而閱讀能力和領域知識可能是其中的調節變項 (Krebs et al., 2019; Mason et al., 2016)。

綜合討論

儘管 Mayer 的「多媒體原則」植基於不少實徵研究證據，但在特定條件下（例如，年輕或低能力讀者）未必獲致同樣強韌的證據，且我們對於讀者實際上如何聯合處理文字與圖像的理解仍然相當缺乏。學習成效資料顯示，高年級與高能力者似乎從圖像中獲益更多，眼動資料呈現出高年級與高能力者更加重視圖像檢視與圖文參照，亦即有較高的圖像凝視時間比例、圖區平均凝視時間、圖文區域間跳移次數等。對照 CTML (Mayer, 2005, 2014) 可發現，高年級與高能力者很可能透過認知系統中的口語及圖像雙重路徑，帶來更多的記憶容量或心智模型數量，可以解釋學習成效的提升；年輕或低能力讀者的圖像檢視與圖文參照較少，很可能未有效地選擇詞彙與圖片等多元表徵，亦缺乏語音與圖像的參照與組織，造成口語模型與圖像模型的建構困難，進而無法帶來學習成效的進步。由此可知，單純添加圖像無法保證更好的學習成效，必須讓讀者認知系統的雙重通道皆能有效運作才是學習提升的關鍵。

在 Schnotz 的 ITPC (Schnotz & Bannert, 2003; Schnotz et al., 2014) 中，圖文閱讀的好處來自於文本處理和圖像處理之間固有的不對稱性，使讀者可以結合兩種形式表徵的特定優勢 (Schnotz & Wagner, 2018)。眼動資料顯示，在閱讀階段開始時，讀者似乎主要將心力投注於文本理解上，由文本引導讀者對主題的概念分析，因此注意力在文本和圖片之間頻繁跳移是為了在口語和圖像訊息之間建立交叉參照。學術導向以及高年級讀者在此階段的確比非學術導向和低年級讀者有更高的圖文參照次數比例 (Schnotz & Wagner, 2018)。一旦初始心智模型建構後，讀者主要使用圖片來調整心理模型以適應任務需求，因此造成「圖像一項目間的參照」會比「文字一項目間的參照」更多 (Schnotz & Wagner, 2018)，或是在第二次閱讀時增加了對圖片相對於文本的注意，這也與學習成效的進步有關 (Eitel, 2016)。簡而言之，文本提供概念引導，而圖像則依照需求充當外部認知工具。這樣的結果證實了 ITPC 中文本處理和圖像處理之間不對稱性的假定，亦即文本和圖片因著不同的目的具有不同的功能。

本研究在方法上使用系統回顧，帶來四點研究限制：第一，因建立了嚴格的篩選標準，儘管文獻搜尋與選取歷程清楚明確，但也可能限制了回顧的範疇；第二，部分實徵研究使用小樣本或是單一測量，其品質較為受限，而綜整中有些主題的結果來自的研究數量較少，因此在解釋的時候必須特別小心；第三，實徵研究在發表時本身可能帶有些許偏誤，例如教學介入有效果之研究可能較無效果之研究更容易發表至期刊，進而膨脹了介入的有效性，但就研究內部結果來看，眼動歷程與學習成效之間的連結仍具參考性，惟在解讀學習成效時需特別謹慎；第四，本綜整著眼於科學圖文材料，因此結果在其他內容領域的類推性較為受限。

本回顧之綜整結果顯示此領域尚有許多研究缺口。在方法論上有三點建議。第一，近六成的科學圖文閱讀眼動研究來自大學生參與者資料，而小學後期和中學前期為參與者的研究約占四成，但主要來自三組研究團隊，因此，亟待其他眼動研究者投入探討非大學生參與者的科學圖文閱讀，包括學前兒童、K-12 學生和成人。第二，在選用的眼動指標方面，相較於時間和次數尺度指標，空間尺度指標極少被選用，然而這類指標例如掃視路徑、凝視位置等通常能夠呈現出注意力的配置與順序，建議未來研究可增加空間尺度眼動指標的探討。第三，目前研究在看待眼動指標高低及其反映的認知歷程可能會有不同的解讀，例如區域間跳移次數高究竟反映的是讀者因掌握閱讀技巧而能進行成功整合，亦或因理解困難而增加整合嘗試？眼動資料的推論解釋並不容易，建議後續研究可以考慮增加質性訪談，例如閱讀後，播放參與者的眼動資料請參與者說明自己的思考歷程。

在研究變項的建議有四，首先，本回顧許多主題下的研究結果未有一致的傾向（例如能力、先備知識），或者累積研究數量少（例如認知風格），或是不同研究操弄的細節差異大（例如圖像類型眾多複雜），這些都需要後續更細緻地深究。其次，能力與先備知識可能是其他操弄的調節因子，有些介入操弄對於低能力者、低先備知識者會產生影響，在高能力者和高先備知識者未必有效果。例如：高工作記憶者比低工作記憶者不容易受到誘人細節影響，較會忽略誘人細節圖片並聚焦於文本上 (Sanchez & Wiley, 2006)；Richter 與 Scheiter (2019) 發現信號只對低先備知識者產生學習支持，而高先備知識者則不受影響。亦尚有其他可能的調節變項，例如年齡、材料內容領域帶來的影響，在本回顧中因果積篇數過少而未成為綜整主題，建議未來研究將這些可能的調節變項納入探討，並在累積充分的研究後進一步彙整分析。第三，目前少有研究區分閱讀初始處理與晚期處理，建議

未來研究可再探討閱讀表現與兩者的關聯性。第四，目前研究發現讀者會隨著材料難度或任務經驗增加而調整閱讀模式，然而多數研究僅讓讀者進行一次性的閱讀，建議後續研究可以藉由操弄材料難度、重複閱讀、增加任務經驗等，透過眼動技術持續探究讀者的自我調整歷程。

最後，在教育實務上提供三點建議。第一，鼓勵讀者在閱讀科學圖文時掌握圖像訊息以及圖文之間的參照，有效使用兩種表徵形式。第二，單純操弄材料儘管可能改變讀者的眼動型態，但未必能提升學習成效，建議教學者可以考慮再透過教學介入輔助，引導讀者掌握讀圖與圖文參照的技巧。第三，進行介入設計時要特別留意個別差異（例如不同能力與先備知識）讀者的介入效果並不一致，可能有不同的需求。除此之外，在研究呈現上建議使用的眼動指標名詞應前後一致，並呈現出明確的定義與計算方式，使其他研究者能夠更容易複製或回顧該研究。

註釋

¹ 各主要文獻之眼動指標對照表如補充資料一，連結：http://bit.ly/ESM1_2021

² 各項研究特徵描述摘要表如補充資料三，連結：http://bit.ly/ESM3_2021

參考文獻

* 為納入本回顧之論文

呂昕頤、詹雨臻、陳學志（2019）：〈笑話中歧義與推論歷程之眼動分析〉。《教育心理學報》，50，587–609。[Lu, H.-I., Chan, Y.-C., & Chen, H.-C. (2019). Ambiguity and inference processing in verbal jokes: Analyses of eye movement. *Bulletin of Educational Psychology*, 50, 587–609.] [https://doi.org/10.6251/BEP.201906_50\(4\).0002](https://doi.org/10.6251/BEP.201906_50(4).0002)

官英華（2020）：〈語音轉碼對臺灣大學生閱讀歷程的影響：來自眼動實驗的證據〉。《教育心理學報》，52，459–488。[Guan, Y.-H. (2020). The effects of phonological recoding on Taiwanese students' reading processes: Evidence from eye-tracking experiments. *Bulletin of Educational Psychology*, 52, 459–488.] [https://doi.org/10.6251/BEP.202012_52\(2\).0010](https://doi.org/10.6251/BEP.202012_52(2).0010)

陳學志、賴惠德、邱發忠（2010）：〈眼球追蹤技術在學習與教育上的應用〉。《教育科學研究期刊》，55（4），39–68。[Chen, H.-C., Lai, H.-D., & Chiu, F.-C. (2010). Eye tracking technology for learning and education. *Journal of Research in Education Sciences*, 55(4), 39–68.] <https://doi.org/10.3966/2073753X2010125504002>

* 簡郁芩、吳昭容（2012）：〈以眼動型態和閱讀測驗表現探討箭頭在科學圖文閱讀中的圖示效果〉。《中華心理學刊》，54，385–402。[Jian, Y.-C., & Wu, C.-J. (2012). The effect of arrows in an illustration when reading scientific text: Evidence from eye movements and reading tests. *Chinese Journal of Psychology*, 54, 385–402.] <https://doi.org/10.6129/cjp.2012.5403.07>

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, 125, 413–428. <https://doi.org/10.1016/j.compedu.2018.06.023>

* Bol, N., van Weert, J. C. M., Loos, E. F., Romano Bergstrom, J. C., Bolle, S., & Smets, E. M. A. (2016). How are online health messages processed? Using eye tracking to predict recall of information in younger and older adults. *Journal of Health Communication*, 21, 387–396. <https://doi.org/10.1080/10810730.2015.1080327>

- Butcher, K. R. (2014). The multimedia principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 174–205). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.010>
- Carney, R. N., & Levin, J. R. (2002). Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14, 5–26. <https://doi.org/10.1023/A:1013176309260>
- *Chuang, H.-H., & Liu, H.-C. (2012). Effects of different multimedia presentations on viewers' information-processing activities measured by eye-tracking technology. *Journal of Science Education and Technology*, 21, 276–286. <https://doi.org/10.1007/s10956-011-9316-1>
- Coleman, J. M., McTigue, E. M., & Dantzler, J. A. (2018). What makes a diagram easy or hard? The impact of diagram design on fourth-grade students' comprehension of science texts. *The Elementary School Journal*, 119, 122–151. <https://doi.org/10.1086/698819>
- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, 35, 59–74. <https://doi.org/10.1016/j.cedpsych.2009.10.002>
- *Eitel, A. (2016). How repeated studying and testing affects multimedia learning: Evidence for adaptation to task demands. *Learning and Instruction*, 41, 70–84. <https://doi.org/10.1016/j.learninstruc.2015.10.003>
- Guo, D., McTigue, E. M., Matthews, S. D., & Zimmer, W. (2020). The impact of visual displays on learning across the disciplines: A systematic review. *Educational Psychology Review*, 32, 627–656. <https://doi.org/10.1007/s10648-020-09523-3>
- Guo, D., Zhang, S., Wright, K. L., & McTigue, E. M. (2020). Do you get the picture? A meta-analysis of the effect of graphics on reading comprehension. *AERA Open*, 6(1). Advance online publication. <https://doi.org/10.1177/2332858420901696>
- *Hannus, M., & Hyönä, J. (1999). Utilization of illustrations during learning of science textbook passages among low- and high-ability children. *Contemporary Educational Psychology*, 24, 95–123. <https://doi.org/10.1006/ceps.1998.0987>
- Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, 90, 414–434. <https://doi.org/10.1037/0022-0663.90.3.414>
- *Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717–742. <https://doi.org/10.1006/jmla.1993.1036>
- Ho, H. N. J., Tsai, M.-J., Wang, C.-Y., & Tsai, C.-C. (2014). Prior knowledge and online inquiry-based science reading: Evidence from eye tracking. *International Journal of Science and Mathematics Education*, 12, 525–554. <https://doi.org/10.1007/s10763-013-9489-6>
- *Hochpöchler, U., Schnotz, W., Rasch, T., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2013). Dynamics of mental model construction from text and graphics. *European Journal of Psychology of Education*, 28, 1105–1126. <https://doi.org/10.1007/s10212-012-0156-z>
- Höffler, T. N. (2010). Spatial ability: Its influence on learning with visualizations—a meta-analytic review.

- Educational Psychology Review*, 22, 245–269. <https://doi.org/10.1007/s10648-010-9126-7>
- *Höfler, T. N., Koć-Januchta, M., & Leutner, D. (2017). More evidence for three types of cognitive style: Validating the object-spatial imagery and verbal questionnaire using eye tracking when learning with texts and pictures. *Applied Cognitive Psychology*, 31, 109–115. <https://doi.org/10.1002/acp.3300>
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- *Hung, Y.-N. (2014). “What are you looking at?” An eye movement exploration in science text reading. *International Journal of Science and Mathematics Education*, 12, 241–260. <https://doi.org/10.1007/s10763-013-9406-z>
- Hyönä, J., Lorch, R. F., Jr., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Elsevier. <https://doi.org/10.1016/B978-044451020-4/50018-9>
- *Jian, Y.-C. (2016). Fourth graders' cognitive processes and learning strategies for reading illustrated biology texts: Eye movement measurements. *Reading Research Quarterly*, 51, 93–109. <https://doi.org/10.1002/rrq.125>
- *Jian, Y.-C. (2017). Eye-movement patterns and reader characteristics of students with good and poor performance when reading scientific text with diagrams. *Reading and Writing*, 30, 1447–1472. <https://doi.org/10.1007/s11145-017-9732-6>
- *Jian, Y.-C. (2018). Reading instructions influence cognitive processes of illustrated text reading not subject perception: An eye-tracking study. *Frontiers in Psychology*, 9, Article 2263. <https://doi.org/10.3389/fpsyg.2018.02263>
- *Jian, Y.-C. (2019). Reading instructions facilitate signaling effect on science text for young readers: An eye-movement study. *International Journal of Science and Mathematics Education*, 17, 503–522. <https://doi.org/10.1007/s10763-018-9878-y>
- Jian, Y.-C., Chen, M.-L., & Ko, H.-W. (2013). Context effects in processing of Chinese academic words: An eye-tracking investigation. *Reading Research Quarterly*, 48, 403–413. <https://doi.org/10.1002/rrq.56>
- *Jian, Y.-C., & Ko, H.-W. (2017). Influences of text difficulty and reading ability on learning illustrated science texts for children: An eye movement study. *Computers & Education*, 113, 263–279. <https://doi.org/10.1016/j.compedu.2017.06.002>
- *Jian, Y.-C., Su, J.-H., & Hsiao, Y.-R. (2019). Differentiated processing strategies for science reading among sixth-grade students: Exploration of eye movements using cluster analysis. *Computers & Education*, 142, Article 103652. <https://doi.org/10.1016/j.compedu.2019.103652>
- *Jian, Y.-C., & Wu, C.-J. (2015). Using eye tracking to investigate semantic and spatial representations of scientific diagrams during text-diagram integration. *Journal of Science Education and Technology*, 24, 43–55. <https://doi.org/10.1007/s10956-014-9519-3>
- *Jian, Y.-C., & Wu, C.-J. (2016). The function of diagram with numbered arrows and text in helping readers construct kinematic representations: Evidenced from eye movements and reading tests. *Computers in Human Behavior*, 61, 622–632. <https://doi.org/10.1016/j.chb.2016.03.063>

- *Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, 18, 178–191. <https://doi.org/10.1037/a0026923>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist*, 38, 23–31. https://doi.org/10.1207/S15326985EP3801_4
- *Ketzer-Nöltge, A., Schweppe, J., & Rummer, R. (2019). Is the seductive details effect moderated by mood? An eye-tracking study. *Applied Cognitive Psychology*, 33, 62–70. <https://doi.org/10.1002/acp.3487>
- *Koć-Januchta, M., Höffler, T., Thoma, G.-B., Precht, H., & Leutner, D. (2017). Visualizers versus verbalizers: Effects of cognitive style on learning with texts and pictures – An eye-tracking study. *Computers in Human Behavior*, 68, 170–179. <https://doi.org/10.1016/j.chb.2016.11.028>
- *Korbach, A., Ginns, P., Brünken, R., & Park, B. (2020). Should learners use their hands for learning? Results from an eye-tracking study. *Journal of Computer Assisted Learning*, 36, 102–113. <https://doi.org/10.1111/jcal.12396>
- *Krebs, M. C., Schüller, A., & Scheiter, K. (2019). Just follow my eyes: The influence of model-observer similarity on Eye Movement Modeling Examples. *Learning and Instruction*, 61, 126–137. <https://doi.org/10.1016/j.learninstruc.2018.10.005>
- Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, 10, 90–115. <https://doi.org/10.1016/j.edurev.2013.10.001>
- *Lin, Y. Y., Holmqvist, K., Miyoshi, K., & Ashida, H. (2017). Effects of detailed illustrations on science learning: An eye-tracking study. *Instructional Science*, 45, 557–581. <https://doi.org/10.1007/s11251-017-9417-1>
- *Liu, C.-J., Kemper, S., & McDowd, J. (2009). The use of illustration to improve older adults' comprehension of health-related information: Is it helpful? *Patient Education and Counseling*, 76, 283–288. <https://doi.org/10.1016/j.pec.2009.01.013>
- *Liu, H.-C. (2018). Investigating the impact of cognitive style on multimedia learners' understanding and visual search patterns: An eye-tracking approach. *Journal of Educational Computing Research*, 55, 1053–1068. <https://doi.org/10.1177/0735633117697020>
- Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, 4, 6–14. [https://doi.org/10.1016/S1364-6613\(99\)01418-7](https://doi.org/10.1016/S1364-6613(99)01418-7)
- *Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of picture labeling on science text processing and learning: Evidence from eye movements. *Reading Research Quarterly*, 48, 199–214. <https://doi.org/10.1002/rrq.41>
- *Mason, L., Pluchino, P., & Tornatora, M. C. (2015). Eye-movement modeling of integrative reading of an illustrated text: Effects on processing and learning. *Contemporary Educational Psychology*, 41,

- 172–187. <https://doi.org/10.1016/j.cedpsych.2015.01.004>
- *Mason, L., Pluchino, P., & Tornatora, M. C. (2016). Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. *British Journal of Educational Technology*, 47, 1083–1095. <https://doi.org/10.1111/bjet.12271>
- *Mason, L., Pluchino, P., Tornatora, M. C., & Ariasi, N. (2013). An eye-tracking study of learning from science text with concrete and abstract illustrations. *The Journal of Experimental Education*, 81, 356–384. <https://doi.org/10.1080/00220973.2012.727885>
- *Mason, L., Scheiter, K., & Tornatora, M. C. (2017). Using eye movements to model the sequence of text–picture processing for multimedia comprehension. *Journal of Computer Assisted Learning*, 33, 443–460. <https://doi.org/10.1111/jcal.12191>
- *Mason, L., Tornatora, M. C., & Pluchino, P. (2013). Do fourth graders integrate text and picture in processing and learning from an illustrated science text? Evidence from eye-movement patterns. *Computers & Education*, 60, 95–109. <https://doi.org/10.1016/j.compedu.2012.07.011>
- *Mason, L., Tornatora, M. C., & Pluchino, P. (2015). Integrative processing of verbal and graphical information during re-reading predicts learning from illustrated text: An eye-movement study. *Reading and Writing*, 28, 851–872. <https://doi.org/10.1007/s11145-015-9552-5>
- Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press.
- Mayer, R. E. (2014). *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press.
- Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, 82, 715–726. <https://doi.org/10.1037/0022-0663.82.4.715>
- McTigue, E. M. (2009). Does multimedia learning theory extend to middle-school students? *Contemporary Educational Psychology*, 34, 143–153. <https://doi.org/10.1016/j.cedpsych.2008.12.003>
- *Meppelink, C. S., & Bol, N. (2015). Exploring the role of health literacy on attention to and recall of text-illustrated health information: An eye-tracking study. *Computers in Human Behavior*, 48, 87–93. <https://doi.org/10.1016/j.chb.2015.01.027>
- *Morrow, D., D'andrea, L., Stine-Morrow, E. A. L., Shake, M., Bertel, S., Chin, J., Kopren, K., Gao, X., Conner-Garcia, T., Graumlich, J., & Murray, M. (2012). Comprehension of multimedia health information among older adults with chronic illness. *Visual Communication*, 11, 347–362. <https://doi.org/10.1177/1470357212446413>
- *Mudrick, N. V., Azevedo, R., & Taub, M. (2019). Integrating metacognitive judgments and eye movements using sequential pattern mining to understand processes underlying multimedia learning. *Computers in Human Behavior*, 96, 223–234. <https://doi.org/10.1016/j.chb.2018.06.028>
- Norman, R. R. (2012). Reading the graphics: What is the relationship between graphical reading processes and student comprehension? *Reading and Writing*, 25, 739–774. <https://doi.org/10.1007/s11145-011-9298-7>
- *Ozcelik, E., Karakus, T., Kursun, E., & Cagiltay, K. (2009). An eye-tracking study of how color coding

- affects multimedia learning. *Computers & Education*, 53, 445–453. <https://doi.org/10.1016/j.compedu.2009.03.002>
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford University Press.
- Parkhurst, P. E., & Dwyer, F. M. (1983). An experimental assessment of students' IQ level and their ability to profit from visualized instruction. *Journal of Instructional Psychology*, 10, 9–20.
- Phillips, L. M., Norris, S. P., & Macnab, J. S. (2010). The concept of visualization. In L. M. Phillips, S. P. Norris, & J. S. Macnab (Eds.), *Visualization in mathematics, reading and science education* (Vol. 5, pp. 19–34). Springer Science & Business Media.
- Polkinghorne, D. E. (1995). Narrative configuration in qualitative analysis. *International Journal of Qualitative Studies in Education*, 8, 5–23. <https://doi.org/10.1080/0951839950080103>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Renkl, A., & Scheiter, K. (2017). Studying visual displays: How to instructionally support learning. *Educational Psychology Review*, 29, 599–621. <https://doi.org/10.1007/s10648-015-9340-4>
- *Richter, J., & Scheiter, K. (2019). Studying the expertise reversal of the multimedia signaling effect at a process level: Evidence from eye tracking. *Instructional Science*, 47, 627–658. <https://doi.org/10.1007/s11251-019-09492-3>
- *Rop, G., Schöler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2018a). The effect of layout and pacing on learning from diagrams with unnecessary text. *Applied Cognitive Psychology*, 32, 610–621. <https://doi.org/10.1002/acp.3445>
- *Rop, G., Schöler, A., Verkoeijen, P. P. J. L., Scheiter, K., & Van Gog, T. (2018b). Effects of task experience and layout on learning from text and pictures with or without unnecessary picture descriptions. *Journal of Computer Assisted Learning*, 34, 458–470. <https://doi.org/10.1111/jcal.12287>
- *Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, 34, 344–355. <https://doi.org/10.3758/BF03193412>
- *Scheiter, K., & Eitel, A. (2015). Signals foster multimedia learning by supporting integration of highlighted text and diagram elements. *Learning and Instruction*, 36, 11–26. <https://doi.org/10.1016/j.learninstruc.2014.11.002>
- *Scheiter, K., Schubert, C., & Schöler, A. (2018). Self-regulated learning from illustrated text: Eye movement modelling to support use and regulation of cognitive processes during learning from multimedia. *British Journal of Educational Psychology*, 88, 80–94. <https://doi.org/10.1111/bjep.12175>
- *Scheiter, K., Schubert, C., Schöler, A., Schmidt, H., Zimmermann, G., Wassermann, B., Krebs, M. C., & Eder, T. (2019). Adaptive multimedia: Using gaze-contingent instructional guidance to provide personalized processing support. *Computers & Education*, 139, 31–47. <https://doi.org/10.1016/j.compedu.2019.05.005>
- Schnotz, W. (2014). Integrated model of text and picture comprehension. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 72–103). Cambridge University Press.

- Schnotz, W., & Bannert, M. (2003). Construction and interference in learning from multiple representation. *Learning and Instruction, 13*, 141–156. [https://doi.org/10.1016/S0959-4752\(02\)00017-8](https://doi.org/10.1016/S0959-4752(02)00017-8)
- *Schnotz, W., Ludewig, U., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2014). Strategy shifts during learning from texts and pictures. *Journal of Educational Psychology, 106*, 974–989. <https://doi.org/10.1037/a0037054>
- *Schnotz, W., & Wagner, I. (2018). Construction and elaboration of mental models through strategic conjoint processing of text and pictures. *Journal of Educational Psychology, 110*, 850–863. <https://doi.org/10.1037/edu0000246>
- *Schüler, A. (2017). Investigating gaze behavior during processing of inconsistent text-picture information: Evidence for text-picture integration. *Learning and Instruction, 49*, 218–231. <https://doi.org/10.1016/j.learninstruc.2017.03.001>
- *Schüler, A. (2019). The integration of information in a digital, multi-modal learning environment. *Learning and Instruction, 59*, 76–87. <https://doi.org/10.1016/j.learninstruc.2017.12.005>
- Segers, E., Verhoeven, L., & Hulstijn-Hendrikse, N. (2008). Cognitive processes in children's multimedia text learning. *Applied Cognitive Psychology, 22*, 375–387. <https://doi.org/10.1002/acp.1413>
- *Stalbovs, K., Scheiter, K., & Gerjets, P. (2015). Implementation intentions during multimedia learning: Using if-then plans to facilitate cognitive processing. *Learning and Instruction, 35*, 1–15. <https://doi.org/10.1016/j.learninstruc.2014.09.002>
- *Stárková, T., Lukavský, J., Javora, O., & Brom, C. (2019). Anthropomorphisms in multimedia learning: Attract attention but do not enhance learning? *Journal of Computer Assisted Learning, 35*, 555–568. <https://doi.org/10.1111/jcal.12359>
- Vekiri, I. (2002). What is the value of graphical displays in learning? *Educational Psychology Review, 14*, 261–312. <https://doi.org/10.1023/A:1016064429161>
- *Zander, S., Wetzel, S., Kühl, T., & Bertel, S. (2017). Underlying processes of an inverted personalization effect in multimedia learning – An eye-tracking study. *Frontiers in Psychology, 8*, Article 2202. <https://doi.org/10.3389/fpsyg.2017.02202>
- *Zhao, F., Schnotz, W., Wagner, I., & Gaschler, R. (2020). Texts and pictures serve different functions in conjoint mental model construction and adaptation. *Memory & Cognition, 48*, 69–82. <https://doi.org/10.3758/s13421-019-00962-0>

收稿日期：2020 年 12 月 30 日

一稿修訂日期：2021 年 01 月 02 日

二稿修訂日期：2021 年 03 月 11 日

三稿修訂日期：2021 年 03 月 31 日

四稿修訂日期：2021 年 04 月 09 日

接受刊登日期：2021 年 04 月 14 日

Bulletin of Educational Psychology, 2022, 53(4), 773–800
National Taiwan Normal University, Taipei, Taiwan, R. O. C.

A Systematic Review of Eye-Tracking Studies on Text-Diagram Science Reading

Tzu-Ning Wang

Yu-Cin Jian

Department of Educational Psychology and Counseling,
National Taiwan Normal University

Text and diagrams are frequently presented together in science textbooks and popular science articles. Text is used to describe concepts, and diagrams contain visual and spatial information depicting conceptual relationships, object structures, and developmental processes. Two primary theoretical models have been widely adopted by researchers in the domain of text–diagram comprehension: Mayer’s cognitive theory of multimedia learning (CTML; Mayer, 2005, 2014) and Schnotz’s integrated model of text and picture comprehension (ITPC; Schnotz & Bannert, 2003; Schnotz et al., 2014). In contrast to findings of benefits associated with Mayer’s “multimedia principle” (Mayer, 2005, 2014), several studies have documented neutral or negative learning effects from studying images (McTigue, 2009; Segers et al., 2008). Previous literature reviews on text–diagram reading have either been published more than 10 years ago (Carney & Levin, 2002; Phillips et al., 2010; Vekiri, 2002) or mostly focused on offline outcome measures (Carney & Levin, 2002; Guo et al., 2020; Phillips et al., 2010; Schnotz, 2014), which may not reflect the complexity of text–diagram processing. Empirical evidence of eye movements during text–diagram reading has been obtained in several studies, but no systematic review has been conducted to synthesize these findings.

This systematic review was designed to synthesize the empirical research findings of eye-tracking studies in the domain of text–diagram science reading over the past 30 years. The three specific research questions were the following:

a. *What are the readers’ eye movement patterns, and what is the relationship between these patterns and learning performance in text–diagram science reading?*

b. *How and under what conditions do potential influencing factors affect readers’ eye movement patterns and performance in text–diagram science reading?*

c. *Can interventions affect eye movement patterns and performance in text–diagram science reading?*

A systematic literature review was conducted using a methodological three-step process (Guo et al., 2020).

a. *Identifying and searching articles* published between January 1990 and May 2020 in Scopus, Education Database (ProQuest), ERIC (ProQuest), PsycArticle, and Airiti Library (Chinese) by using a combination of multiple key terms: eye-tracking (eye-tracking OR eye-movement OR eye-fixation), diagram (figure* OR illustration* OR graph* OR picture* OR diagram*), and reading (reading OR text* OR multimedia). After duplicates were excluded, 579 English-language articles and 33 Chinese-language articles were identified during the initial search.

b. *Screening and coding* studies using predetermined selection criteria: (a) examined empirically; (b) published in a peer-reviewed SCIE, SSCI, TSSCI, or THCI journal; (c) measured participants’ eye movements and learning outcomes; (d) involved text–diagram reading in the science domain; (e) primarily focused on reading comprehension (studies involving problem-solving or reasoning tasks were excluded); (f) used static visual displays as materials (studies using video, audio, simulation, computer games, and interactive diagrams were excluded). After abstract and full-text screening, 50 English-language articles (including 54 studies) and one Chinese-language article were retained for inclusion in the analysis.

c. *Analyzing* the included studies and *interpreting* findings using inductive paradigmatic analysis (Polkinghorne, 1995).

Based on Guo et al. (2020), categories were not predetermined by the researchers but rather were identified in an inductive manner. After descriptive information was extracted from the included studies and tentatively coded, we compared the similarities and differences of the codes and categorized them into themes to answer the research questions. Finally, the research findings were presented on the basis of these different themes.

Of the 55 studies, 32 were conducted with college students. Regarding content area, more than half of the studies involved biology and medicine ($n = 34$), and mechanics and physics ($n = 13$) and earth sciences and geography ($n = 10$) represented the second and third most common content areas. Articles were most commonly sourced from the journals *Learning and Instruction* ($n = 6$) and *Computers & Education* ($n = 5$). SMI RED 250 ($n = 10$), EyeLink 1000 ($n = 10$), and Tobii T120 ($n = 9$) were the most common models of eye-trackers used. The most commonly used eye movement indicators were total fixation duration ($n = 41$) and the transition between areas of interest ($n = 36$).

The qualitative synthesis of the studies was organized according to two major themes: cognitive processes and interventions.

a. *Cognitive processes.* (1) *Eye movement patterns.* The eye movement data support dual-coding theory (Paivio, 1986), which is the theoretical basis of CTML and ITPC and suggests that the human cognitive system is composed of two sets of subsystems: verbal and pictorial. Generally, reading is text-driven rather than diagram-driven, and readers may have different reading patterns that involve varying degrees of diagram processing and text–diagram integration. Those who more often inspect diagrams and refer to text and diagrams during reading usually performed better. (2) *Reader characteristics.* Higher-level readers were more likely to utilize diagram inspection and text–diagram referencing tactics compared with lower-level readers, and they generally had better learning performance; this applied to students in both academic-oriented schools (compared to non-academic-oriented schools) and high-ability students (compared to low-ability students). People with low prior knowledge (PK) tended to inspect more text than those with high PK, but no consistent pattern was discernible in diagram inspections, text–diagram referencing, and learning performance. In terms of cognitive style, visualizers fixated more on diagrams, and verbalizers fixated more on text, but the findings regarding learning performance were mixed. (3) *Material characteristics and their interactions with reader characteristics.* High-ability readers put more effort into difficult articles, and low-ability readers put more effort into articles with medium-low difficulty. High-ability and senior-grade readers increased diagram inspection and “diagram–item” references as the difficulty of the task increased. Readers in the second cycle of reading decreased text inspection and increased diagram inspection, reflecting adjustment by the reader. When inconsistent information appeared in material, readers typically performed more inspections, reinspections, and transitions in the beginning, but this did not affect reading performance.

b. *Interventions.* (1) *Material manipulation.* First, readers tended to fixate on signaled sections (i.e., key points were highlighted via colors, labels, and arrows) faster than on nonsignaled sections, and this effect was greater for readers with low PK. Readers performed better on signaled materials than on nonsignaled materials, but this pattern did not show among younger readers. Second, readers inspected explanatory diagrams more than decorative diagrams; however, the effect on performance by adding different types of diagrams was unclear. Third, readers transitioned more between text and diagrams when the two were physically integrated than when they were separated. The spatial contiguity effect on performance depended on whether the information was necessary. (2) *Instructional interventions.* Overall, instructional intervention usually affected readers’ eye movement patterns and was likely to enhance learning performance through mediators of visual behavior such as text–diagram references and diagram inspection. Reading ability and domain knowledge may be moderating variables.

Our results demonstrated that the inclusion of diagrams in science texts usually affects readers’ visual behavior but does not guarantee positive learning performance. The key to learning improvement is the effective operation of the dual channels of a reader’s cognitive system. Some reader characteristics (such as ability and PK) may play moderating roles in eye movement and learning performance. Several gaps in eye-tracking research on text–diagram science reading were identified in this review. First, more research studies should be conducted with preschool children, K-12 students, and adults. Second, future research studies should include spatial-scale eye movement indicators and distinguish between initial and late processing indicators. Third, other process measures, such as think-aloud protocols, should be included to help interpret eye movement data. Fourth, future research should focus on potential moderators (such as ability, PK, age, and content area) or the factors that contain complex research manipulations and were underrepresented in the studies in our review (such as cognitive style and type of

diagram). Additionally, it is recommended that future research further explore readers' self-regulatory processes by using eye-tracking technology. Finally, the findings of this study suggest that instructors should also consider individual differences and modify their instructions to help readers master the skills of reading diagrams and text–diagram references rather than focusing exclusively on the course content.

First, the strict screening criteria used in this study may have limited the scope of this review. Second, some empirical studies were based on small samples or a single measurement, which may have limited the findings, and the data pertaining to some themes in the review came from a small number of studies; therefore, special care must be taken when interpreting the results. Third, the empirical research itself may have been somewhat biased when published, which may have inflated the effectiveness of the interventions. Fourth, this review focused on the domain of scientific reading materials, and consequently, the results may not be generalized to other content areas.

Keywords: systematic review, text-diagram science reading, eye-tracking, reading process

