

探討達成不同貼文類型適用的圖像替代文字描述方式研究

A study of the description of alt text suitable for different types of posts

¹ 蔡涵如 ² 何俊亨

¹ 國立成功大學工業設計研究所 | 碩士 | abby963852741@gmail.com

² 國立成功大學工業設計研究所 | 副教授 | hoch@gs.ncku.edu.tw

¹ Han-Ju Tsai ² Chun-Heng Ho

¹ Department of Industrial Design, National Cheng Kung University, Master, abby963852741@gmail.com

² Department of Industrial Design, National Cheng Kung University, Associate Professor, hoch@gs.ncku.edu.tw

引用本文：

蔡涵如、何俊亨 (2021)。探討達成不同貼文類型適用的圖像替代文字描述方式研究。《感性學報》，9(1)，4-28。

近年來社群軟體的使用率提高，圖像、影像逐漸取代文字為主要的社交物件。視障者也是社群軟體的高度使用者，當視障者在使用社群軟體時，會為沒有替代文字的圖片感到困擾。臉書近年推出自動替代文字系統，該系統仍然有許多可改善的空間。本研究探討另一項長期做為視障者提供視覺資訊的口述影像服務後，提出臉書貼文中的圖像替代文字應該因應貼文類型調整敘述方式。研究發現，由於相繼型的貼文在文字內容中，解釋圖像資訊的比例較少，因此貼文適合使用特定的圖像替代文字。而說明型貼文的文字內容中，已解釋許多圖像資訊，因此可以使用特定或抽象的圖像替代文字。以研究結果彙整成建議圖像替代文字產出流程圖，以提供未來相關研究或者相關服務參考。

關鍵詞：替代文字、口述影像、螢幕翻譯、可用性、圖像索引

Recently, images and videos have replaced text as the main social objects on the social software. When the visually impaired use social software, they will be troubled by pictures without alt text. Facebook has launched an automatic alt text system, which still has a lot of room for improvement. After exploring the Audio Description service that provides visual information for the visually impaired, this research proposes that the alt text of posts should be adjusted according to the types of posts. This research found that since successive posts spend a smaller proportion of the text explaining the image information, the posts are suitable for specific alt text. The text of the illustrated post contains a lot of image information, so these posts are appropriate for specific or abstract alt text. The research results are compiled into a flowchart of the different ways of producing alt text provides a reference for future related research or the development of related services.

Keywords: Alt text, Audio Description, Accessibility, Image Index

一 緒論

"A picture is worth 1000 words? Maybe. But the audio describer might say that a few well-chosen words conjure vivid and lasting images." - Joel Snyder

1.1 研究背景與動機

由於網路的普及、智慧型手機和其他行動裝置的迅速發展，促使社群軟體擁有高普及率 (Kemp, 2019)。社群軟體中的社交物件 (Social object) 也越來越多樣化，圖像、影像逐漸取代文字成為主要的社交物件。根據臉書的官方統計資料，人們每一天在 Facebook、Instagram、Messenger 和 WhatsApp 與超過 20 億張照片產生互動 (Facebook, 2017)。而許多視障者也是社群軟體的高度使用者 (Wu & Adamic, 2014)。近年來社群軟體中圖像、影像資訊量大增，使得社群軟體的可用性降低。在早期推特的社交物件還只有文字時，推特被視為有高度可用性的社群軟體，但是當現在圖片大量充斥於推文的情況下，推特對於視障人士的可用性正逐漸降低 (Morris et al., 2016)。雖然社群軟體上的圖像對視障人士來說較難以觸及，但是有研究指出，視障使用者對於社群軟體上的圖像資訊有著高度參與動機 (Voykinska, Azenkot, Wu, & Leshed, 2016)。

目前雖然已經有許多無障礙輔助閱覽工具，但是仍然存在一些問題是輔助工具無法解決的，如當視障者接觸到圖像內容沒有替代的說明文字，會使視障使用者感到挫折和困擾 (Lazar, Allen, Kleinman, & Malarkey, 2007)。為了解決視障使用者在社群軟體上遇見的窘境，臉書開發了自動替代文字系統 (Automatic Alt-Text)，此功能結合了人工智慧和替代文字語法，可為圖像內容自動生成替代文字 (Facebook, 2018)。該功能的發布確實提升了許多視障人士與臉書軟體上圖像內容互動的體驗，也讓許多視障使用者感到驚艷 (Wu, Wieland, Farivar, & Schiller, 2017)。原先臉書可以產生的圖像替代文字還是相當的簡單，例如「圖像可能包含三個人，微笑，在戶外」(Facebook, 2016a, 2016b, 2019)。到了 2021 年臉書有針對自動替代文字再進行改版升級，目前進步成可以除了可以辨識物件之外，還可以辨識相片人物動作，還有些許地標，除此之外也會根據相片中的物件面積比例判定照片中的主角物件為何，提升了自動提袋文字敘述的準確性 (Facebook, 2021)。以下圖 1 為目前 2021 年最新版本的臉書自動圖像替代文字範例。

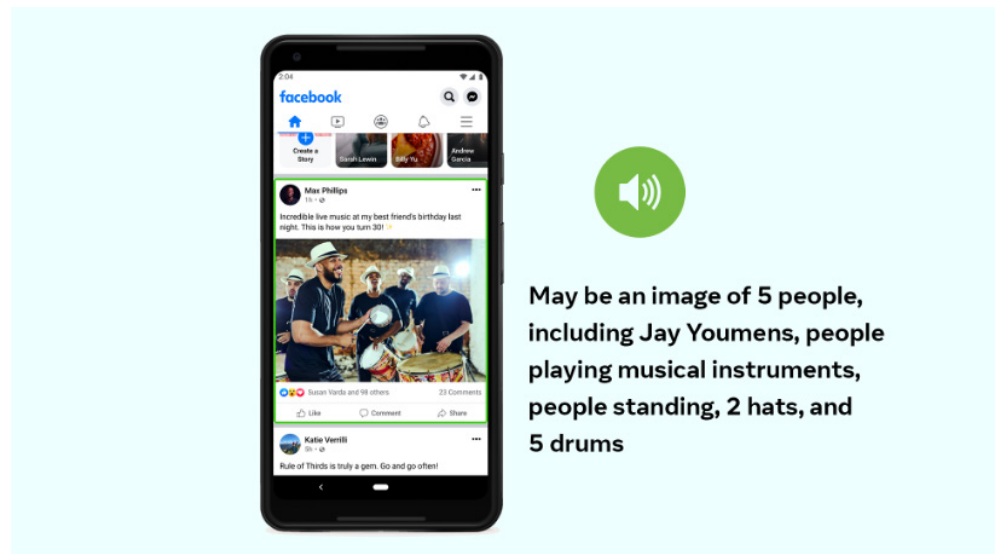


圖 1 | 圖像替代文字範例

本研究探討了另一項常作為為視障人士提供視覺資訊的口述影像（Audio Description）服務。簡單來說，口述影像就是把「看見」的「說」出來，對任何視覺屬性之傳播媒介訊息的口語轉述，都屬口述影像（口述影像服務網，2019）。目前已經有通用的指導規則可以參考，以幫助口述影像人員製作口述影像內容（趙雅麗，2002）。但是目前口述影像的通用指導規則大多是以客觀、準的方式傳達視覺資訊為首要目標，視障使用者的體驗感受為次要目標的現象。近年有一些學者開始研究起口述影像的使用者體驗（Walczak & Fryer, 2017）。有研究者開始提倡，應該要轉而以更有戲劇張力、主觀的口語敘事（Audio Narrative）代替口述影像，以提高使用者體驗（Kruger, 2010）。也還有學者堅持應該要繼續維持以客觀、單純的方式敘述，以避免偏頗的資訊傳達（Caro, 2016）。因此現在對於口述影像的製作原則，產生眾說紛紜的局面。然而以上任一說法並不是絕對的最佳解答，適合的敘述方式取決於需要被口述影像「翻譯」或者「再現」的影像內容（趙雅麗，2002）。

1.2 研究目的

而圖像替代文字目前正往客觀且提高準確性的積極改正，然而如同口述影像相關研究一樣，我們無法斷定只要提高內容準確性，就能給視障使用者帶來較佳體驗。因此本研究以口述影像研究為借鑒，提出圖像替代文字的研究方向，以探討不同臉書貼文所對應適合的「圖像替代文字」敘述方式為目的。本研究使用 Jaimes 和 Chang 提出對圖像索引的金字塔概念模型作為產出不同敘述方式的替代文字參考，

以模型中提出的一般 (Generic)、特定 (Specific)、抽象 (Abstract) 的概念，為本研究的貼文樣本分別產出一般、特定、抽象的圖像替代文字以搭配不同新聞內容種類與不同圖文相互關係，以了解對於不同貼文內容，是否會有適合不同的圖像替代文字敘述方式 (Jaimes & Chang, 1999)。

二 研究方法

本研究的實驗方法為問卷調查法，請受測者觀看不同內容的臉書貼文樣本後，再聆聽不同敘述方式的圖像替代文字樣本。請受測者根據不同貼文對三種圖像替代文字的適合度排名。以了解對應不同貼文所適合的圖像替代文字敘述方式。

2.1 搜集貼文樣本

在臧國仁、蔡琰著作的《敘事傳播》中，提及新聞媒體的報導也需要因應內容種類調整適合的報導方式，因此確立新聞媒體有報導手法的多樣性 (臧國仁、蔡琰，2017)。故本研究根據紐約時報與英國廣播公司分類新聞的方式，選定國際、商業、科技、科學、故事、娛樂與藝術、健康、運動、旅遊、風格十種內容作為本研究貼文選定範疇 (BBC, 2020; Times, 2020)。除了內容種類之外，本研究認為貼文中的圖文關係也會影響替代文字適合的敘述方式，在《解讀影像》著作中，作者 Kress 和 Leeuwen 介紹並引用 Barthes 將圖文關係分類的兩種方式，分別為圖文相繼與圖文說明。圖文相繼指的是當文字與影像兩者在傳遞資訊時，兩者相互接力並且共同完整一訊息的意義，影像傳遞的資訊內容不與文字內容相互重複；圖文說明指的是當文字與影像兩者在傳遞資訊時，兩者重複說明同一訊息，只是使用不同種符號來呈現 (Barthes, 1993; Kress & VanLeeuwen, 1996)。根據以上條件，最終選訂十種新聞內容與兩種圖文關係相互搭配的二十則新聞貼文為本研究實驗樣本。圖 2 為圖文相繼關係的貼文樣本，圖 3 為圖文說明關係的貼文樣本。



2020 春夏男裝周期間有一件大事，就是 6 月 20 日在巴黎大皇宮舉辦的卡爾拉格斐紀念會。

這場活動由 Chanel、Fendi 和 Karl Lagerfeld 三個品牌共同主辦，由歌劇名導 Robert Carsen 監製整體視覺，她曾經執導過描述卡爾和香奈兒女士之間情感的改編劇作《Platée》，當時演出卡爾的演員是法國男中音 Edwin Crossley-Mercer。

這次沒人扮演卡爾，而是由他自己真實的故事背景和創作過的作品、在好友心中留下的印象來向所有喜愛他的人告別。他說過最討厭葬禮，那會讓他「寧願死掉」，所以這場活動應該也會以一種不哀傷的方式進行，可惜全世界只有 2500 人可以參加。如果他在場，應該還是會以這種臭屁的模樣為自己超精彩的人生謝幕吧！

圖 2 | 圖文相繼貼文樣本



約莫 7 年前，英國時尚金童 Jonathan Anderson 執掌 Loewe，憑藉一只 Puzzle 包重新定義這個向來低調的西班牙老牌，這款包也橫空出世要生 IT Bag 高度，成為眾多女孩每年犒賞自己的口袋名單。設計師從品牌歷史資料庫中一只斑駁碎裂的皮包獲得靈感，應用 40 張嚴選皮革與超越物件形狀影像收納限制的高深拼接工藝創造這款名包，獨特的延展性帶來更多使用方式（肩背、斜背、手提、手拿包等），這麼好用的東西當然不能只有女孩獨享，大尺寸款式想當然適合男孩們擁有，品牌堪稱龜毛的選皮標準（千分之四）完全體現在 Puzzle 的質感。

Puzzle Large 淺焦糖色小牛皮肩背提包 by Loewe

Photographs by Mau Lin ()

#Loewe #JonathanAnderson #LoewePuzzle
#LoewePuzzle

#Mensfashion #BAG #ITBAG

圖 3 | 圖文說明貼文樣本

2.2 製作貼文的圖像替代文字樣本

本研究參考文獻探討的 Jaimes 和 Chang 提出的金字塔概念模型作為選取圖像中概念的方式 (Jaimes & Chang, 1999; Jörgensen, Jaimes, Benitez, & Chang, 2001; Shatford, 1986)。金字塔概念模型將語義級別分為六個層級，分別為一般物件、一般場景、特定物件、特定場景、抽象物件、抽象場景，模型圖示如圖 4。

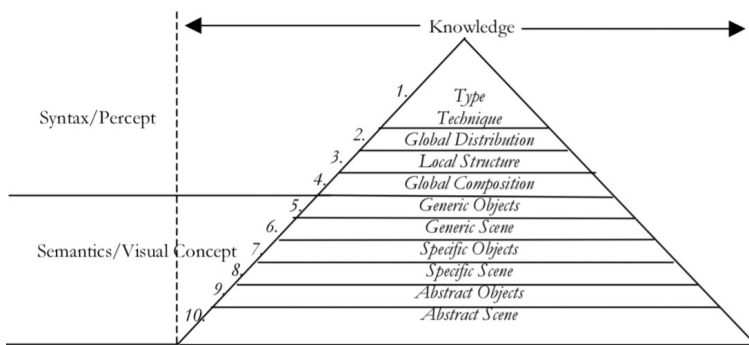


圖 4 | 金字塔概念模型

本研究以該模型對國際相繼貼文樣本如圖 5 索引得出表 1 的索引概念們。接著再進一步組織概念形成完整的語句，概念順序安排採用對視障者較容易完形的從大概念到小概念的報讀順序 (翁煌德, 2019)。最終得到表 1 中三種圖像替代文字，分別為一般、特定、抽象的圖像替代文字。並且使用與 Google 翻譯一樣的聲音模擬器錄製語句，做為本研究圖像替代文字的樣本。



圖 5 | 貼文樣本

表 1 | 不同層級的索引概念與圖像替代文字

層級	索引概念	圖像替代文字
一般	一個黑種人、一個黃種人、四隻旗子、花朵、麥克風、壁畫、西裝、握手、室內	在室內，有一個黑種人在左方，與一個黃種人在右方，兩人面對面握手，身後有黃色和紅色的旗幟交錯排列。
特定	索羅門群島外交部長、中國外交部長、索羅門群島國旗、中國國旗、蓮花樣式壁畫、花束、外交會議室	在會議室中，索羅門群島外交部長與中國外交部長兩人面對面握手，身後有兩國國家的國旗交錯並列。
抽象	合作、成交、首長、微笑、會議、外交、正式、莊重的場合	在莊重的外交會議中，代表兩國外交的首長握手以示雙方的建交與合作。

2.3 圖像替代文字實驗

實驗採用線上問卷進行實驗與資料搜集，請受測者自行選擇適合的場域並且使

用自己的電腦作為填卷工具，實驗受測者為三十三位聽力無受損的台灣人。整份實驗問卷包含二十個題組，一題組中有一則貼文、三段圖像替代文字音訊、排序題。

以下為實驗流程：

- (1) 步驟一、觀看一則圖文並茂的新聞貼文。
- (2) 步驟二、聆聽三段圖像替代文字。
- (3) 步驟三、受測者以整則貼文(圖+文)的意義為單位，為三段圖像替代文字對該貼文的匹配程度排名。

三 結果分析與討論

本研究使用常用來分析排序資料的 Friedman 檢定作為分析工具，並且採用 Bonferroni 事後檢定。

3.1 Friedman 檢定結果

以下表 2 為每個題組進行 Friedman 檢定的結果，由表格中我們可以了解到，每題組中的圖像替代文字在排名上是有顯著差異的 ($p < .05$)，但是無法得知三種圖像替代文字之間的詳細排名差異，因此再進一步進行事後檢定。

表 2 | 二十題題組 Friedman 分析結果

變項	相繼				說明			
	P-value	G	S	A	P-value	G	S	A
國際	.001	1.00	2.79	2.21	.001	1.12	2.39	2.48
商業	.015	1.70	2.39	1.91	.001	1.94	2.55	1.52
科技	.001	1.33	2.24	2.42	.001	1.33	2.30	2.36
科學	.001	1.42	2.30	2.27	.001	1.27	2.33	2.39
故事	.001	1.36	2.09	2.55	.003	1.64	2.45	1.91
娛樂	.001	1.12	2.39	2.48	.001	1.18	2.06	2.76
健康	.001	1.39	2.52	2.09	.035	1.67	2.03	2.30
運動	.001	1.18	1.97	2.85	.001	1.15	2.24	2.61
旅遊	.001	1.55	2.48	1.97	.001	1.39	2.45	2.15
風格	.001	1.24	2.58	2.18	.001	1.30	2.24	2.45

註解：G = Generic; S = Specific; A = Abstract。

3.2 Bonferroni 事後檢定結果

將每題組進行事後檢定之後，得到了詳細的排名情況，總共可以歸納出以下若干種情況。第一種結果為三個圖像替代文字之間，兩兩的排名比較皆有顯著的差異，因此可以得知三者比較下，完整的圖像替代文字排名。第二種結果為三個圖像替代文字之間，只有單一圖像替代文字的排名明顯高於其他兩個圖像替代文字，也就是特定圖像替代文字顯著大於抽象與一般圖像替代文字。第三種結果為三個圖像替代文字之間，只有單一圖像替代文字的排名明顯低於其他兩個圖像替代文字，也就是一般圖像替代文字顯著小於特地與抽象替代文字。最後的第四種結果為三種圖像替代文字之間，只有兩種圖像替代文字的排名之間有顯著差異，第三種圖像替代文字與這兩種圖像替代文字的排名皆無明顯差異，只有抽象圖像替代文字大於一般圖像替代文字。

表 3 | 統計結果歸納

三個圖像替代文字排名有相互顯著差異		單一圖像替代文字對其他兩個圖像替代文字排名有顯著差異				只知道兩種圖像替代文字之間的排名差異	
		得知最前排名		得知最後排名			
貼文	排名	貼文	排名	貼文	排名	貼文	排名
國際相繼	S>A>G	商業相繼	S>A=G	國際說明	A=S>G	健康說明	A>G
健康相繼	S>A>G	商業說明	S>A=G	科技相繼	A=S>G		
風格相繼	S>A>G	故事說明	S>A=G	科技說明	A=S>G		
故事相繼	A>S>G	旅遊相繼	S>A=G	科學相繼	A=S>G		
娛樂說明	A>S>G			科學說明	A=S>G		
運動相繼	A>S>G			娛樂相繼	A=S>G		
運動說明	A>S>G			旅遊說明	A=S>G		
				風格說明	A=S>G		

註解：G = Generic; S = Specific; A = Abstract

3.3 歸納同種排名結果的貼文

最後歸納成只根據被選出第一名匹配的圖像替代文字結果來討論。以下表 4 分別歸納統整出哪些貼文樣本對於特定 (Specific)、抽象 (Abstract) 的圖像替代文字為第一名匹配。由於有些貼文樣本的統計結果並沒有顯示出何種圖像替代文字為第一名匹配，只有顯示出最後一名匹配的樣本，因此也將這個結果歸納討論。

表 4 | 名次結果歸納

特定 (Specific) 第一名	抽象 (Abstract) 第一名	一般 (Generic) 最後一名
國際 / 相繼	故事 / 相繼	國際 / 說明
商業 / 相繼	娛樂 / 說明	科技 / 相繼
商業 / 說明	運動 / 相繼	科技 / 說明
故事 / 說明	運動 / 說明	科學 / 相繼
健康 / 相繼	健康 / 說明	科學 / 說明
旅遊 / 相繼		娛樂 / 相繼
風格 / 相繼		旅遊 / 說明
		風格 / 說明

本研究觀察表 4 以統整結果，歸納出影響圖像替代文字的方式有三種，分別為圖文關係為主要影響因子、內容種類為主要影響因子、圖文關係與內容種類交互影響。以下圖 6 為分析脈絡圖供參考。

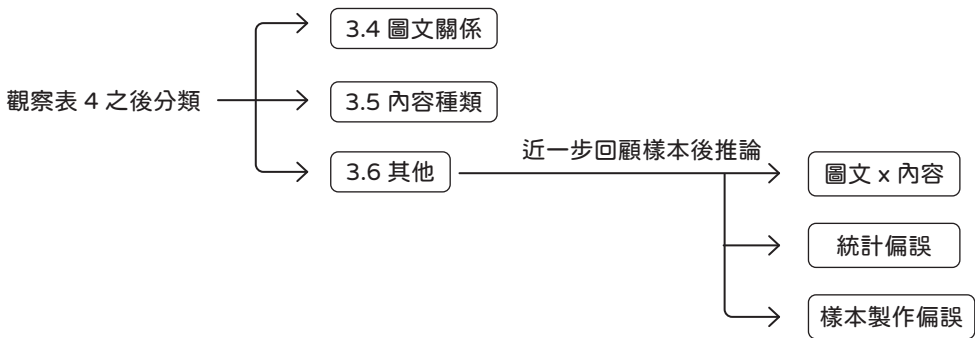


圖 6 | 分析討論脈絡圖

3.4 圖文關係為主要影響圖像替代文字的因子

根據表 4，有三種內容種類，分別為國際、旅遊、風格，對於圖文關係變化時有一樣的結果，這三種內容種類中如果是相繼型的貼文，都適用特定 (Specific) 的圖像替代文字，而如果是說明型的貼文時，可同時適用特定或抽象 (Abstract) 的圖像替代文字。因此推論在圖文關係中，相繼型的貼文適用特定的圖像替代文字，說明型的貼文可以使用特定或抽象的圖像替代文字。

根據文獻探討以及實驗規劃中的對於圖文相互關係的解說，圖文相繼的關係是圖與文傳遞的資訊不重疊，而圖文說明的關係是圖與文傳遞的資訊重疊。對於上述三種內容種類，國際、旅遊、風格的貼文來說，相繼型貼文適合特定的圖像替代文字，說明型貼文適合特定或抽象的圖像替代文字。本研究推測圖文關係影響造成此結果的原因，可能是因為圖文相繼型的貼文中，文字並沒有詳細的解釋圖像內容，因此

較適合使用特定的圖像替代文字；而圖文說明型的貼文，由於貼文中的文字內容有較多解釋圖像的資訊，因此圖文說明型的貼文就不一定只能使用特定的圖像替代文字，也可以使用抽象的圖像替代文字。圖 7 為圖文關係與圖像替代文字敘述方式的關係圖。

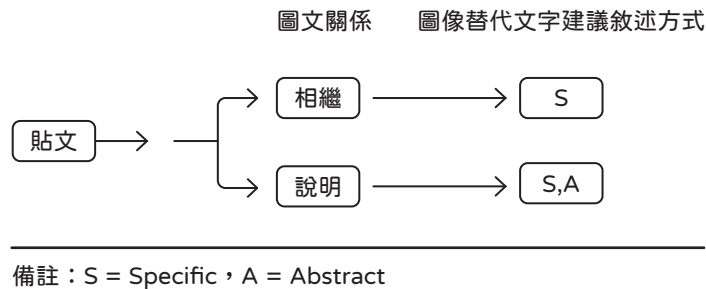


圖 7 | 圖文關係與圖像替代文字敘述方式關係圖

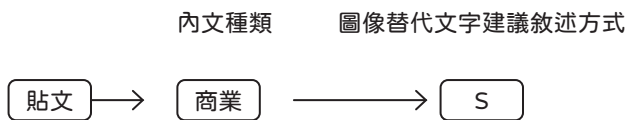
3.5 內容種類為主要影響圖像替代文字的因子

依照 3.4 小節的推論來說，實驗結果應該要是趨向所有內容種類的相繼貼文都是特定 (Specific) 第一名的結果，說明貼文則是可以使用特定或者抽象 (Abstract) 的替代文字。但是並非所有的實驗都是如此，因此除了圖文關係影響之外，貼文的內容種類也有對圖像替代文字適合的敘述方式造成影響，而且不同種類也有不同的影響，以下小節，將會依照貼文內容種類對於圖像替代文字的影響逐一討論。接下來的小節會討論為何除了國際、旅遊、風格之外，其他內容種類並不會依照圖文關係影響圖像替代文字：相繼型貼文適合特定、說明型貼文適合特定或者抽象。

3.5.1 商業內容種類對於圖像替代文字敘述方式的影響

在商業內容種類中，無論是何種圖文相互關係的貼文，最適合的替代文字都是特定的圖像替代文字，本研究回顧了商業相繼、商業說明的貼文。兩篇報導的內容都是使用特定的物件來傳達抽象的意義，例如：某特定知名餐廳的人潮可以反應當天的股市，某特定規模大小的外送公司在疫情來到的過年後復工，前者是使用特定餐廳的人潮（特定）來反應經濟的走向（抽象），後者是使用特定外送公司復工（特定）來反應經濟復甦（抽象）。兩者都傾向使用具體的事件讓人體會到經濟的變動，而在細部探究搭配的特定的圖像替代文字中，由於特定的替代文字能很具體的描繪出圖像的內容，而讓聽眾聽到具體的描繪方式，能更想像出畫面樣貌，因為特定、

具體的關鍵字在聽眾腦中已經有種既定的印象，此印象不一定是視覺的印象，有可能是該關鍵詞在該時空背景下的特殊印象，如商業相繼報導透過某知名餐廳的人潮反應股市漲跌，必定要是那指標性的餐廳才能代表這個經濟起伏，因為該餐廳在那時空背景下的地位意義。還有如商業說明型的貼文透過報導指標性的外送公司復工，代表著經濟復甦。這也就是本研究推論為何對於商業內容種類的貼文，較適合使用特定圖像替代文字的原因。以下圖 8 為商業內容貼文與圖像替代文字敘述方式的關係圖。



備註：S = Specific

圖 8 | 商業內容與圖像替代文字敘述關係圖

3.5.2 運動內容種類對於圖像替代文字敘述方式的影響

運動內容的貼文樣本無論是搭配相繼、說明的圖文關係，樣本的統計結果分析皆是抽象的圖像替代文字為最適合的圖像替代文字。本研究回顧了運動相繼、運動說明的貼文之後發現，前者的報導內容為知名退休籃球員殞落的消息，並且搭配一張拍得非常具有戲劇張力的背影離場照片；後者報導內容是知名網球選手宣布退休的消息，並且搭配一張該網球選手年幼時打網球的照片。在運動類別的報導中，文字就可以將屬於特定內的容範疇報導完整，如運動選手的比賽數據，但是如果談論關於運動精神層面這類抽象的事物，就很難單以文字報導呈現，若有文配圖會有更好的呈現，因此對於運動類別的報導，在搭配的图片方面就不太需要使用呈現特定內容的資訊，多會是為了呈現抽象事物，故搭配的圖像替代文字也會較適合使用抽象的圖像替代文字。本研究挑選的兩則運動報導的內容雖然都有提及特定的資訊內容，如兩位退役的運動選手的輝煌戰績，但是兩篇報導主旨是在講告別、別離這些抽象且感性的主題，故推論是這個報導的主旨為較抽象的概念。因此造成抽象的圖像替代文字為最適合的圖像替代文字。以下圖 9 為運動內容貼文與圖像替代文字敘述方式的關係圖。

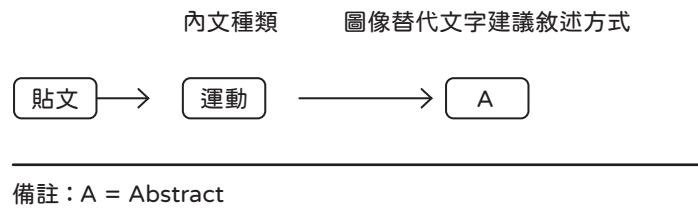


圖 9 | 運動內容與圖像替代文字敘述關係圖

3.5.3 科技和科學內容種類對於圖像替代文字敘述方式的影響

科技、科學內容種類的相繼貼文中，貼文報導只是在單純傳遞知識訊息，並沒有傳達具象知識之後更深層的意義。除此之外，雖然對於相繼貼文來說，文字與圖片呈現資訊不重複，可能造就圖像內容非常的重要需要詳細解說。但在對這兩種類型的相繼貼文來說，文字資訊已經完整的傳達了該貼文想要傳達的特定知識，而圖片只是搭配示意，由此可見，由於圖片在貼文中的重要性不大，因此這兩種類型的相繼貼文結果會偏向說明型的貼文結果，也就是同時適合特定和抽象的圖像替代文字。以下圖 10 為科技、科學內容貼文與圖像替代文字敘述方式的關係圖。

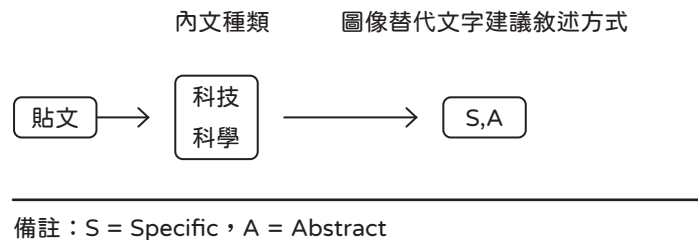


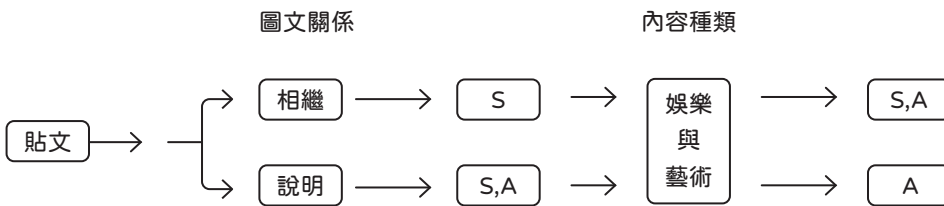
圖 10 | 科技、科學內容與圖像替代文字敘述關係圖

3.6 其他因子影響

對於貼文適用的替代文字敘述方式，除了會單獨受到圖文關係的影響之外，也會單獨受到內容種類的影響。但是除了單獨影響之外，圖文關係與內容種類也會發生交互影響，也有其他因素所造成的影響，接下來的討論是本研究根據圖文關係與內容種類的交互影響造成貼文適用圖像替代文字的結果，以及其他非實驗變因的因素所造成的結果。

3.6.1 娛樂與藝術內容種類與圖文關係對圖像替代文字的影響

藝術娛樂就是傳遞抽象主旨較偏向抽象的內容，如八卦新聞、介紹藝術品，較適合使用許多抽象且富有渲染力的詞彙組成圖像替代文字。依照原先推論的結果，圖文相繼應該會是特定第一名，但是會因為主旨抽象導向特定和抽象都適合作為該圖像替代文字，而依照原先推論應該會同時適合特定和抽象圖文說明貼文卻變成抽象第一名，形成兩則貼文適用的圖像替代文字的結果，同時往抽象偏移的情況。會造成此結果就是因為該內容種類的主旨大多偏向抽象的內容，並且此情況下是適用抽象的詞彙來形容抽象主旨，因此原本推測只試用特定敘述的相繼型貼文，變成特定與抽象都適用，而原本推測特定與抽象敘述都適用的說明型貼文，變成抽象更為適合。以下圖 11 為娛樂藝術內容貼文和圖文關係與圖像替代文字敘述方式的關係圖。



備註：S = Specific，A = Abstract

圖 11 | 娛樂內容和圖文關係與圖像替代文字敘述方式關係圖

3.6.2 故事內容種類與圖文關係對圖像替代文字的影響

對於故事內容中，相繼型的貼文沒有依照原先推論圖文相互關係影響是特定第一名，反而是抽象第一名，推測原因是因為，該內容種類與本次運動種類貼文選擇到類似貼文，因而有類似的結果，文字充滿了許多抽象又富有情緒張力的詞彙，並且都是搭配有特殊處理過的照片，並不是實際事件發生當下的側拍照片。因此故事相繼的貼文較適合使用抽象第一名的結果。以下圖 12 為故事內容貼文與圖像替代文字敘述方式的關係圖。



備註：A = Abstract

圖 12 | 故事內容與圖像替代文字敘述方式關係圖

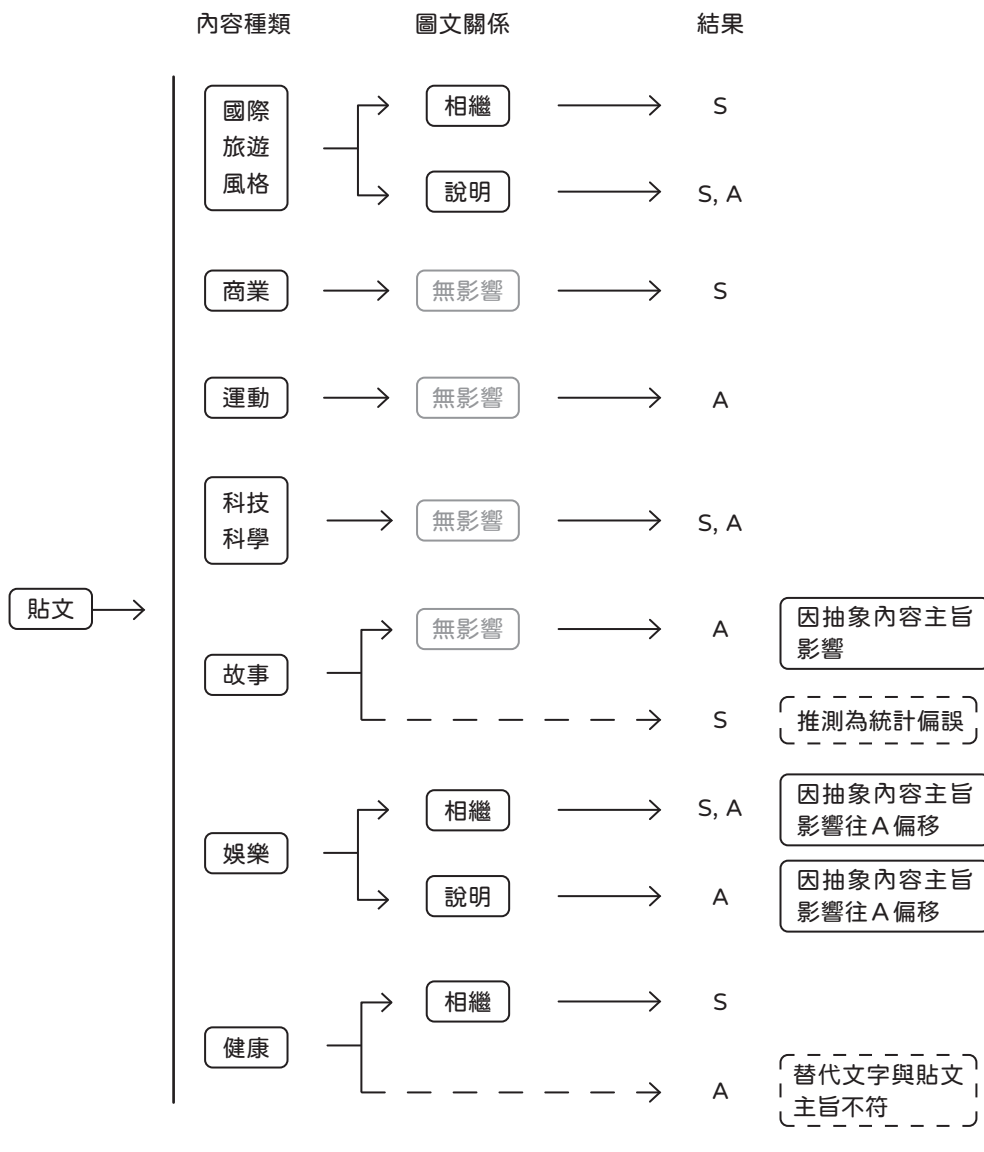
而故事說明的貼文，沒有符合 3.4 小節提及的圖文關係影響的結果，也就是適合特定和抽象，反而是單獨適合特定，本研究推測該統計結果的原因可能為統計上的偏誤，會這樣推論的原因是因為該貼文中的文字內容已經非常詳細的解說了圖像內的資訊，因此該貼文應該可以適用特定和抽象的圖像替代文字，但是統計結果並不如此，此結果的偏誤，有待留給後續研究探討。

3.6.3 健康內容種內與圖文關係對圖像替代文字的影響

健康相繼型貼文的統計結果符合圖文關係中，相繼適合使用特定的圖像替代文字的結果。但是健康說明型貼文沒有符合圖文關係的影響，因此再次回顧貼文內容探討該結果的原因。對於健康說明的貼文來說，該貼文文字本身雖然有傳遞圖像資訊，但文字與圖像的主旨如果分開來看是不相同的，而本研究的認為該貼文的主旨是透過文字部分的內容來傳遞的。該貼文文字內容在宣導懷孕時養寵物的注意事項，但是圖像資訊是一對懷胎的夫妻與一隻小狗，小狗摸著孕婦的肚子，有透露出抽象的情感意義。但是該貼文文字部分並沒有要傳遞小狗摸著孕肚的抽象情感，只是非常仔細的說明懷孕養寵物的注意事項，該圖像比較像是隨意找來的示意圖，但是因為圖像本身傳達的情感又過於強烈，又因為製作該圖像的抽象替代文字時，著重於小狗摸著孕肚的抽象意象，該抽象圖像替代文字與貼文主旨並沒有直接的關係，但是卻有與圖像有強烈的連結，所以就會使受測者傾向選擇該抽象的圖像替代文字，因此本研究認為此部分是因為製作替代文字時的偏誤。

3.7 總結

將以上 3.4 到 3.6 小節講述的所有分析討論歸納為一個貼文的圖像替代文字製作流程圖，如下圖 13，提供未來製作圖像替代文字之建議與改善。



備註：S = Specific, A = Abstract

圖 13 | 結果統整

以上統整圖可作為未來產生圖像替代文字的建議，在貼文的內容種類中，如果內容種類屬於商業、運動、科技、科學、故事就會是以內容的種類影響為重，因此建議根據該內容種類來判定適合的圖像替代文字敘述方式。對商業內容種類的貼文中，適合特定的圖像替代文字，因為必須表現出特定資訊為何，才能代替背後的抽象意義，但是該特定抽象意義為客觀的抽象意義如股市與經濟走向，並非為主觀的情感抽象意義。對於運動內容種類的貼文，特定的資訊如賽事數據適合使用文字來呈現，圖像多為了呈現抽象意義，而運動貼文呈現的抽象意義通常是富有情感的形象意義，如努力不懈的精神、毅力、傳奇的運動員人生。對於科技、科學內容種類

的貼文，由於內容是客觀的知識傳遞，鮮少有關於主觀情感的內容傳遞，加上文字內容就足以傳遞整則貼文的內容，因此可以適用特定或者抽象的圖像替代文字。對於故事內容種類的貼文，內容主旨多為偏向富有情感的抽象內容傳遞，因此適合使用抽象的圖像替代文字。

而對於國際、旅遊、風格這三類貼文，只有圖文關係會對圖像替代文字的敘述方式造成影響，圖文關係與內容種類不會交互影響。因此對於這三種種類的貼文來說，只要是相繼型的貼文都適合使用特定的圖像替代文字，而說明型的貼文都既適合特定也適合抽象的圖像替代文字。

對於娛樂與藝術貼文來說，同時受到內容種類與圖文關係的交互影響，因內容種類較偏向主觀且富有戲劇張力的抽象概念，因此影響了原先圖文關係會導致的結果，而將圖文關係導致的結果往抽象偏移了。

最後對於健康內容種類的貼文來說，相繼型的貼文有符合圖文關係導致的結果，但是對於說明型貼文來說，圖文關係的影像並沒有發揮作用，經過再次回顧貼文內容以及圖像替代文字樣本之後，確認造成此結果的原因可能是因為，雖然抽象的圖像替代文字與該圖片情境非常符合，但是並不與貼文文字內容的主旨相符合，因此算是製作圖像替代文字的偏誤，此部分有待未來研究繼續探討。

3.8 與先前相關研究結果討論

先前 Hollink 的圖像描述、搜尋研究的結果指出，受測者最常使用一般 (Generic) 類別的概念作為搜尋或者描述的概念，只有在針對新聞報紙圖片或者是真實攝影照片時才會是特定 (Specific) 最常被使用，該研究結果有部分內容與本研究結果相同，也就是新聞圖片或是真實圖片會時常使用特定 (Specific) 類別的。對此現象 Hollink 也有在自己研究中對 Armitage 的圖書館圖像搜尋研究結果，以後者研究涵蓋的圖像樣本為實際照片為由，故也有該類似結果之說的推論。

然而，Hollink 的研究實驗是要求受測者觀看一段文字內容，執行幫該文字內容找尋搭配插圖的任務，在找尋插圖前，先在內心想像該插圖應該會是什麼樣貌，接著直接憑空描述心中想像的插圖，再去執行搜尋的動作。由於該研究任務（憑空想像搭配圖像再進行描述與搜尋）取向與本研究任務（評斷描述樣本與圖片相符程度）取向不相同，因此這可能是本研究與先前相關研究有些許不同的地方。除此之外，由於該研究搭配的三段文字樣本分別取自於小說以及童書還有一則國際新聞報導，

取自於小說以及童書中的文字片段中，完全沒有提及特定和抽象的物件或是場景，只有一般的物件與場景，因此在對於小說以及童書的想像插圖描述與搜尋時，才會大量使用一般層級的概念，偶爾使用抽象層級的概念，幾乎不使用特定層級的概念。而國際新聞報導文字內容大量提及特定種類的物件或是場景的詞彙，因此受測者在執行插圖描述與搜尋時，才會大量使用特定層級的概念詞彙（Hollink, Schreiber, Wielinga, & Worrying, 2004）。而這個結果與 Armitage 和 Enser 在 1996 年時，對圖書館中進行圖像資料搜尋引擎建置時的研究有類似的結果，在該研究結果也顯示特定（Specific）層級是最長被使用的級別，Hollink 在針對該文獻結果與自己本身的研究結果一起討論時，也是認為由於在圖書館的圖像很多是歷史照片或者是電影海報等真實攝影照片，因此造就 Armitage 和 Enser 的研究結果，也就是特定（Specific）級別最常作為描述搜尋圖片的類（Armitage & Enser, 1997）。而本研究搜集的文字樣本全部皆為新聞媒體內容，性質與 Hollink 的研究中的小說與兒童讀物的文字樣本非常不同。故這可能是 Hollink 研究部分結果與本研究結果相異的原因。

除了本研究與 Hollink 的研究在文字樣本上有明顯的差異之外，研究實驗的任務取向也影響很大，本研究認為在 Hollink 的研究中，由於並不是針對一個存在的圖像進行描述，而是要憑空想像適合搭配文字的圖像，對那張虛構的圖像進行描述，並且搜尋，先撇開搭配文字內容大多散佈於何種層級概念中，這樣實驗結果會指向大量使用一般（Generic）是可想而知，因為是在對一個不確定的圖像描述，並且使用這些詞彙去搜尋合適的圖片，使用一般（Generic）層級的描述詞去搜尋較可以搜尋到多樣化的圖片，才可以多樣比較找出合適搭配插圖。故這也可能是造成該結果部分結果與本研究結果相異之原因。

四 結論

本章將先對本研究做結論，接續討論本研究的研究貢獻，最後針對本研究的不足的地方討論，以給予未來相關的研究建議。

4.1 研究結論

隨著社群軟體的發達，越來越多人成為社群軟體的使用者，社群軟體也佔據了人們的生活一大部分。除了社群軟體的崛起以外，社群軟體上的社交物件也有許多

的轉變，逐漸從文字轉變到圖像、影像為主，全球的社交軟體中一天當中發佈的照片可數以億計，當許多以視覺媒介傳遞訊息的社交物件出現後，便會降低社群軟體的可用性，讓社群軟體上的視障用戶無法觸及這些視覺媒介傳遞的資訊內容。雖然臉書在近年有開發了一項自動替代文字的系統，使用圖像資料庫訓練人工智慧辨識圖像，以提供臉書視障使用者在閱覽臉書貼文時可以搭配自動的圖像替代文字功能，但是目前該系統對於每種貼文以及圖像產出的敘述架構都非常簡單以及類似，例如如下，圖像中可能包含一人、天空、大自然。也就是屬於金字塔概念模型中一般 (Generic) 資訊內容的圖像替代文字。

本研究以另一項時常為視障者提供圖像資訊的服務——口述影像為借鏡，提出圖像替代文字的敘述方式應該根據該整貼文想要傳達的目的與意義進行調整，方能準確傳遞該訊息的意義給視障使用者，也能提高視障者的臉書使用體驗。本研究參考許多研究者針對圖像資料庫研究中的圖像索引模型為圖像替代文字多樣敘述方式的參考依據，最終以 Jaimes 和 Chang 提出的金字塔概念模型 (Jaimes & Chang, 1999)，根據該模型為每則貼文樣本的圖像產生分別為一般 (Generic)、特定 (Specific)、抽象 (Abstract) 的圖像替代文字。所以本研究試想了解不同類別的報導貼文與不同的圖文相互關係對於圖像替代文字的影響，找到對應內容類別與圖文相互關係和圖像替代文字的關係。

研究結果發現，綜觀所有樣本的分析結果，特定 (Specific) 的圖像替代文字最常被認為最適合的圖像替代文字，此結果與先前探討 Jaimes 和 Chang 的金字塔概念模型相關文獻時，Hollink et al. (2004) 執行的圖像描述、搜尋研究中，有類似的結果，也就是針對新聞照片、真實攝影照片，受測者傾向使用「特定層級」的概念作來描述。除此之外，本研究還有發現受測者也時常認為抽象 (Abstract) 的圖像替代文字為最適合作為貼文的圖像替代文字，並且數量與特定的圖像替代文字差距不大，這也是先前研究沒有提到的現象，本研究推測由於貼文內容種類所導致的目的意義不同或者圖文的關係影響，而讓受測者認為讓抽象是最合適的圖像替代文字。故貼文內容種類與圖文關係確實會影響圖像替代文字合適的敘述方式。而實驗結果顯示最適合的兩種圖像替代文字的敘述方式也與目前臉書使用的方式不相同。除此之外，本研究也有歸納出內容種類對圖像替代文字影響的原因跟方式，以及圖文關係對於圖像替代文字影響的規則。

4.2 研究貢獻

先前研究中指出，對於新聞類或是真實拍攝的照片，人們認為特定的圖像替代文字是最適合的。本研究除了發現了類似的結果，還有更近一步的發現，也就是當新聞內容種類不同時，或者圖文關係變化時，會影響圖像替代文字適合的敘述方式，因而造成抽象的替代文字也時常被認為是合適的圖像替代文字。但是這兩個影響因子並不是對每則貼文都有完全一致的影響，圖文關係影響會因為內容種類影響削弱，甚至消失，只剩下內容種類的影響。如下：

- (1) 對於商業內容種類的貼文，無論圖文關係如何，建議使用特定的圖像替代文字。
- (2) 對於運動、故事內容種類的貼文，無論圖文關係如何，建議使用抽象的圖像替代文字。
- (3) 對於科學、科技內容種類的貼文，無論圖文關係如何，可以使用特定或者抽象的圖像替代文字。
- (4) 對於娛樂藝術內容種類的貼文，圖文關係與內容種類交互影響下，娛樂相繼型貼文，可以使用特定或者抽象的圖像替代文字；娛樂說明型貼文，建議使用抽象的圖像替代文字。

以上是單受貼文內容種類影響的結果，除了上述內容類別之外的貼文，大多遵守圖文關係單獨影響的規則。然而，針對本研究未探討的貼文類型而言，由此研究結果觀察到的應用通則大致為：

由於相繼型的貼文在文字內容中，解釋圖像資訊的比例較少，因此貼文中的圖像適合使用特定的圖像替代文字。說明型貼文的文字內容中，已經解釋許多圖像資訊，因此可以使用特定或者抽象的圖像替代文字。

4.3 研究侷限與後續相關研究建議

本次研究在討論結果分析時，才發現挑選健康說明型貼文的樣本時，沒有考慮到該貼文本身的圖文符合度，因此導致雖然圖像替代文字與圖像符合，但是與貼文本身主旨不符合的狀況，以至於受測者在選擇圖像替代文字的時候受到影響，而與預期結果不符，建議未來相關研究可以探討圖文主旨符合度對圖像替代文字適合敘述的影響。

由於本次研究中，運動、故事內容種類的皆是因為圖像內容與文字內容非常抽象，圖像內容與文字內容抽象與類別並沒有絕對的關係，因此建議未來相關研究可以以內容與圖像的抽象、特定、一般程度作為內容不同種類的依據，而不是直接使用新聞的類別作為內容種類的依據。

除了以上建議之外，由於視障者與明眼人對於視覺經驗的不同，因此聽取圖像替代文字時的感受會有所差異，並且引發的想像也可能有所不同。因此除了先對明眼人進行本研究推論結論的實驗，建議未來研究可以以視障者與明眼人為兩組比較組別，以了解視障者與明眼人對於圖像替代文字感受的差異。在 Elean Di Giovanni 對於改善電影口述影像的相關研究中，研究者以明眼人觀看電影時眼動儀的視覺注意力分配數據作為依據修改口述影像並且讓盲人受測聽取修改前後的口述影像，研究最終證實根據明眼人眼動儀數據修改的口述影像比起傳統的口述影像更能讓視障者理解電影 (Di Giovanni, 2014)。故建議未來相關研究可以本研究樣本對盲人做實驗，比較明眼人認為適合的圖像替代文字是否能給盲人帶來更好的體驗。

雖然經由文獻探討了解到，經由已經致力於口述影像多年的專業口述影像製作人員表示，在對畫面製作口述影像時，畫面元素概念的報讀順序也會對於視障者聽取口述影像時造成影響。當在報讀時，推薦的報讀方式會是先報讀較大範圍的概念，例如說可以先報讀出畫面的場景概念，再繼續報讀較小的物件概念，這樣會更有助於視障者對畫面完形，然而這個報讀順序，與目前臉書的報讀順序恰巧相反，推測會有這種認為適合的方式恰巧完全相反可能來自於文化差異，類似的現象可以由許多不同文化在書寫會包含很多層級的資訊習慣看出，舉例來說台灣在書寫地址時與西方寫地址的順序是完全相反的，台灣是由最大層級開始寫或者報讀，但是西方在呈現地址時是由最小的層級開始。所以臉書會提出應該先報讀較小層級也是有其中的道理。故本研究也建議未來的相關研究可以探究在圖像替代文字中是否對於不同文化會有不同的適合方向。

五 參考文獻

- Armitage, L. H., & Enser, P. G. (1997). Analysis of user need in image archives. *Journal of information science*, 23(4), 287-299.
- Barthes, R. (1993). *Rhetoric of the Image*: na.
- BBC. (2020). Retrieved from <https://www.bbc.com/news>

- Caro, M. R. (2016). Testing audio narration: the emotional impact of language in audio description. *Perspectives*, 24(4), 606-634. doi:10.1080/0907676X.2015.1120760
- Di Giovanni, E. (2014). Visual and narrative priorities of the blind and non-blind: Eye tracking and audio description. *Perspectives*, 22(1), 136-153.
- Facebook. (2016a). Learning to Segment. Retrieved from <https://research.fb.com/learning-to-segment/>
- Facebook. (2016b). Under the hood: Building accessibility tools for the visually impaired on Facebook. Retrieved from <https://code.fb.com/ios/under-the-hood-building-accessibility-tools-for-the-visually-impaired-on-facebook/>
- Facebook (2017). Accessibility Research: Developing automatic-alt text for Facebook screen reader users. Retrieved from <https://research.fb.com/blog/2017/02/accessibility-research-developing-automatic-alt-text-for-facebook-screen-reader-users/>
- Facebook. (2018). Using AI to help people with visual impairments share images on Facebook. Retrieved from <https://research.fb.com/using-ai-to-help-people-with-visual-impairments-share-images-on-facebook/>
- Facebook. (2019). 自動替代文字如何運作? . Retrieved from https://www.facebook.com/help/216219865403298?__tn__=-UK-R
- Facebook (Producer). (2021). How Facebook is using AI to improve photo descriptions for people who are blind or visually impaired. Retrieved from <https://ai.facebook.com/blog/how-facebook-is-using-ai-to-improve-photo-descriptions-for-people-who-are-blind-or-visually-impaired/>
- Hollink, L., Schreiber, A. T., Wielinga, B. J., & Worrying, M. (2004). Classification of user image descriptions. *International Journal of Human-Computer Studies*, 61(5), 601-626. doi:<https://doi.org/10.1016/j.ijhcs.2004.03.002>
- Jaimes, A., & Chang, S.-F. (1999). Conceptual framework for indexing visual information at multiple levels. Paper presented at the Internet Imaging.
- Jørgensen, C., Jaimes, A., Benitez, A. B., & Chang, S.-F. (2001). A conceptual framework and empirical research for classifying visual descriptors. *Journal of the American Society for Information Science and Technology*, 52(11), 938-947. doi:10.1002/asi.1161

- Kemp, S. (2019). DIGITAL 2019: GLOBAL DIGITAL OVERVIEW. Retrieved from <https://datareportal.com/reports/digital-2019-global-digital-overview>
- Kress, G. R., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*: Psychology Press.
- Kruger, J.-L. (2010). Audio narration: re-narrativising film. *Perspectives*, 18(3), 231-249. doi:10.1080/0907676X.2010.485686
- Lazar, J., Allen, A., Kleinman, J., & Malarkey, C. (2007). What Frustrates Screen Reader Users on the Web: A Study of 100 Blind Users. *International Journal of Human-Computer Interaction*, 22(3), 247-269. doi:10.1080/10447310709336964
- Morris, M. R., Zolyomi, A., Yao, C., Bahram, S., Bigham, J. P., & Kane, S. K. (2016). With most of it being pictures now, I rarely use it: Understanding Twitter's Evolving Accessibility to Blind Users. Paper presented at the Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.
- Shatford, S. (1986). Analyzing the subject of a picture: a theoretical approach. *Cataloging & classification quarterly*, 6(3), 39-62.
- Times, T. N. Y. (2020). Retrieved from <https://www.nytimes.com/>
- Voykinska, V., Azenkot, S., Wu, S., & Leshed, G. (2016). How blind people interact with visual content on social networking services. Paper presented at the Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing.
- Walczak, A., & Fryer, L. (2017). Creative description: The impact of audio description style on presence in visually impaired audiences. *British Journal of Visual Impairment*, 35(1), 6-17.
- Wu, S., & Adamic, L. A. (2014). Visually impaired users on an online social network. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Wu, S., Wieland, J., Farivar, O., & Schiller, J. (2017). Automatic alt-text: Computer-generated image descriptions for blind users on a social network service. Paper presented at the Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing.

口述影像服務網 (2019)。什麼是口述影像 (Audio Description) ?。取自：

<https://www.ad.org.tw/%e4%bb%80%e9%ba%bc%e6%98%af%e5%8f%a3%e8%bf%b0%e5%bd%b1%e5%83%8f-audio-description%ef%bc%89%ef%bc%9f/>

翁煌德 (2019)。口述影像：給視障者、也給所有人的全人可能性。取自：<https://hef.org.tw/journal365-3/>

臧國仁、蔡琰 (2017)。敘事傳播：故事 / 人文觀點。台北：五南。

趙雅麗 (2002)。言語世界中的流動光影：口述影像的理論建構。台北：五南。