

## 探討健康塑身 App 使用者回饋與評價之關係

賴維祥<sup>1</sup> 李晶<sup>2</sup> 李嘉文<sup>3</sup> 黃子峻<sup>4</sup>

### 摘要

近年健康塑身類 App 使用人數上升，平台上的健康塑身 App 數量呈成長趨勢，Google Play 等平台提及使用者回饋能幫助提升 App 評價，但未進一步說明。研究目的為透過關鍵字擷取與統計分析方式探討熱門健康塑身類 App 中使用者回饋與評價之關係。於 2020 年 9 月 1 日蒐集 Google Play 的「健康塑身」類別前 100 名熱門 App 的頁面資訊作為研究對象，後續至 2021 年 3 月 1 日共 6 個月為研究時間範圍。研究工具為關鍵字擷取系統與 SPSS 統計分析軟體。結果發現健康塑身 App 使用者回饋與評價之交互關係具顯著性，使用者回饋正負評數量愈多，能讓未來排名提升，使用者正負評中多在意資訊品質方面的問題。研究結果有助開發商能透過使用者回饋進行優化並改善產品／服務，提升 App 評價。

**關鍵詞：**關鍵字擷取、使用者評論、Google Play、平台

---

<sup>1</sup> 國立臺灣師範大學運動休閒與餐旅管理研究所碩士生

<sup>2</sup> 國立臺灣師範大學運動休閒與餐旅管理研究所教授

<sup>3</sup> 福州大學陽光學院文化產業管理學系副教授

<sup>4</sup> 國立臺灣師範大學運動休閒與餐旅管理研究所博士生

通訊作者：李晶，E-mail: t94002@ntnu.edu.tw

收稿日期：2021/10/21；接受刊登日期：2022/06/09

DOI:10.6284/NPUSTHSSR.202303\_17(1).3

## 壹、前言

資訊化時代來臨，手機應用軟體 (Application, App) 的使用比率呈成長趨勢。2016 年至 2018 年期間，關於健康塑身類 App 的使用者人數成長了三倍，使用者消費也成長 30% (App Annie, 2019)。在 2020 年創意臺灣政策白皮書 (國家發展委員會, 2016) 中，政府更將智慧化健康管理項目列入未來發展的重要政策，持續推動生活智慧化。

健康塑身 App 透過如 Google Play 這類平台提供使用者下載，平台成為資訊產出者與資訊接收者資訊交換的場域，資訊產出者的價值交換來自介紹與更新等資訊，而資訊接收者的價值交換來自評價 (Ciurumelea et al., 2017)。App 開發商能針對使用者回饋作為重要資訊去改善軟體品質與改正缺失的主題，說明了使用者回饋有助於改善 App (Pagano and Maalej, 2013)。使用者回饋中的評論內容能使需求分析人員可以立刻從使用者評論中得到回饋，不只從下載數與銷售數量上分析，並且可以評估功能重要性以利於開發 (Carreño and Winbladh, 2013)。

關於 App 如何提高在平台的評價，近年研究以文獻綜述方式解釋不同 App 平台中影響 App 評價之依據，主要為下載次數、評論和評分 (Karagkiozidou et al., 2019)。其中針對使用者回饋中評論的部分，App 平台提供的 Android Developer (2020) 提及回覆使用者評論有助於排名提升，但未說明應針對何種使用者回饋進行討論。過去研究發現 App 更新後會出現較多使用者回饋與評價變化 (Novak and Marchini, 2019)，如能針對更新後出現的使用者回饋進行探討，能釐清使用者對產品／服務之看法。

2020 年新冠肺炎 (COVID-19) 在三月疫情爆發期間，全球各地健身房健身人數下滑，轉而以居家健身為主，促使了健康塑身 App 在 Google Play 與 App Store 兩大平台上皆有不錯的成績。諸如 2020 年 12 月初 Google Play 公布年度最佳 App 榜單，健康塑身類的「Nuli-居家與重訓課表」與「HealthBuy-健康飲食消費指南」皆包含在其中。健康塑身 App 受重視的狀況下，如能針對使用者回饋作為改善軟體品質的重要資訊，將有助於提升 App 產品／服務。

基於上述研究背景與動機，本研究欲探討使用者回饋與評價之關係，並瞭解針對使用者回饋中何種類型之評論內容做改進，將有助於開發商能提升產品評價。以 Google Play 中熱門健康塑身 App 做為研究對象，將使用者回饋透過關鍵字擷取系統 (Tseng, 1998) 進行分類，並進一步討論不同特性的使用者回饋對評

價之影響。本研究目的為瞭解健康塑身 App 使用者回饋與 App 評價之關係，後續透過典型相關分析針對使用者回饋與 App 兩指標進行分析，討論兩指標間的交互關係。

## 貳、文獻回顧

本研究主要蒐集國內外之研究文獻資料並加以整理，探討熱門健康塑身 App 中使用者回饋與評價之相關研究，以下分別詳述：

### 一、使用者回饋

線上評論（Online Review）是消費者透過網路去評論產品或服務（Hennig-Thurau and Walsh, 2003），為電子口碑（electronic word-of-mouth）的一種形式。使用者在不熟悉產品情況下，能先查看產品相關評論以獲取資訊，線上評論成為衡量使用與否的依據。

#### （一）使用者於平台中之定位

關於使用者於平台互動模式之研究，Norman（1983）提出使用者、設計師與系統間皆具備各自的心智模式（Mental Model），系統必須能正確引導，讓設計師與使用者的概念能盡量符合，此為早期平台互動模式的概念。而一般的使用者無從去認識真實系統運作模式和原理，只能透過介面所透露的各種暗示與訊息，去建立自認為合理的心智模式，藉此臆測與系統間的互動（葉謹睿，2010）。目前在平台模式中，由平台中的控制者制訂使用規範與管理，界面提供者則是建構產出者、接收者與平台的互動方式，產出者將所生產的產品或服務放到平台，而接收者則是至平台購買或使用這些產品／服務，彼此藉以交換價值觀或資訊（Van Alstyne et al., 2016）。平台為產出者與接收者資訊交換的場域，產出者的價值交換來自介紹與更新等資訊，而接收者的價值交換來自評價（Ciurumelea et al., 2017）。對應目前 App 平台的模式，可以將產出者視為開發商，而接收者為下載安裝 App 產品／服務的使用者，平台則提供管道讓使用者在評論欄位對 App 評分與評論。

#### （二）使用者回饋重要性

互聯網興起後，能透過平台建構的網絡建立事業、貢獻內容，以及購買產品和服務（Bonchek and Choudary, 2013），而平台上使用者留下評論的電子口碑方

式，將更快的提供有關產品／服務的相關訊息（Huete-Alcocer, 2017）。平台中諸如電商網站、旅遊網站與應用程式下載平台，皆提供使用者評論欄位讓使用者能對產品做出使用者評析。Google Play 平台中，使用者回饋有評論與評分機制，評論為留下文字訊息說明對 App 的使用者體驗；評分則是依照一至五顆星等，負面評論為評分為 1 至 2 顆星，正面評論則是 4 至 5 顆星（Liu, 2020）。

許多 App 開發商會透過各種形式去收集有關使用者滿意度的數據，例如：焦點團體、調查和錯誤報告。App 使用者也可以在下載應用程式的平台上發布評論中來表達對 App 的意見（Carreño and Winbladh, 2013）。針對使用者回饋的重要性，Pagano 與 Maalej（2013）發現能從使用者評論中獲取回饋，不只從下載數與銷售數量上分析，並能評估功能重要性以利於開發；而多數使用者回饋在新版本後不久提供，頻率數量會隨著時間遞減，因此分析每一版本推出後的使用者回饋，有助開發商發現產品／服務可改善之方向。

## 二、App 評價

Google Play 的評分從平均所有版本的評分改為參考最近版本的評分為主，意指開發商無法依賴舊版本評分吸引使用者。如能透過更好的方式來提昇平台的評價將使 App 被越來越多人看見。Gabor（2017）提出使用者流量並非主要考量，能提高存量才是重點，透過 RARRA 框架邏輯探討 App 上的使用者。RARRA 的框架邏輯來源於：強調留存（Retention）之重要性，並努力讓用戶體會產品／服務的核心價值（Activation）；這樣同時也能達到讓用戶自主分享（Referral），透過使用後的口碑邀請身邊好友一起使用該服務／產品；最後自然增加營收（Revenue）並且獲取到新用戶（Acquisition）。後續 Clark（2019）針對如何從使用者行為在後台之數據評價平台可分為流量、參與度、轉換與營收四種指標。FlipWeb 數位資產觀點（2017）提出針對檢測 App 表現可在平台上查看故障次數、整體評分、分類排名、下載次數與營收等五種指標。如何在平台上的資訊作為評價指標中，過去研究提到透過指標反映 App 表現常會出現虛榮指標（Vanity Metrics）的問題，像是平台上的註冊人數、頁面瀏覽人數、下載量都屬於虛榮指標的一種（Rogers, 2018），開發商並無法從單一虛榮指標看出使用者對於產品的反應，必須透過多項指標做比較。

## 三、使用者回饋影響評價之研究

關於如何萃取使用者評論的回饋，將有助於提升 App 評價，何淑君等人

(2019) 從線上評論找出使用者在意的行動應用軟體品質特徵進行分類，並從 ISO25010 針對企業軟體品質的模型中建構以使用者為中心的軟體品質特徵模型，主要分為系統品質、資訊品質與服務品質。系統品質主要是 App 本身提供給使用者的系統基本功能是否符合需求；資訊品質多評論該 App 上提供給使用者的資訊內容，與是否滿足使用者的目的；最後服務品質為使用者評論針對軟體服務的看法，包含了使用者情緒性字眼與滿意度。上述提出的行動軟體品質特徵，主要為提供開發商能以使用者經驗為中心檢視行動軟體品質，未進一步根據不同類別的 App 進一步探討。關於健康塑身 App 中使用者回饋的評論，Li 等人 (2020) 萃取 Google Play 中的健康塑身 App 使用者負面評論，依據負評中的關鍵詞頻分成功能性、互動性與批判性進行討論。功能性主要為使用者針對該健康塑身 App 在頁面介紹中提及的功能是否解決使用目的，部分使用者也會提及更新後的功能使用問題；互動性則是關於 App 的操作問題，諸如基本設定與登入登出之操作動作；最後則是有關於使用者的情緒性字眼皆歸類於此構面，透過三種構面瞭解使用者對於健康塑身 App 的使用多抱怨何種使用上的問題。

本研究整理行動應用軟體品質特徵與健康塑身 App 負評類型後，分成系統品質、資訊品質、更新品質與服務品質做後續探討，系統品質以 App 提供的基本設定功能為主；Google Play 在頁面資訊欄除了有該 App 的功能介紹之外，也包含更新內容資訊，因此將健康塑身 App 提供給使用者的資訊功能，分成資訊品質與更新品質，資訊品質為 App 介紹中提到的功能是否符合使用目的，而更新品質為使用者針對每次版本更新所作出關於更新好壞的評論；最後服務品質以使用完 App 後的情緒性字眼為主，討論使用者的認知價值與愉悅，與負評類型中的批判性相同。

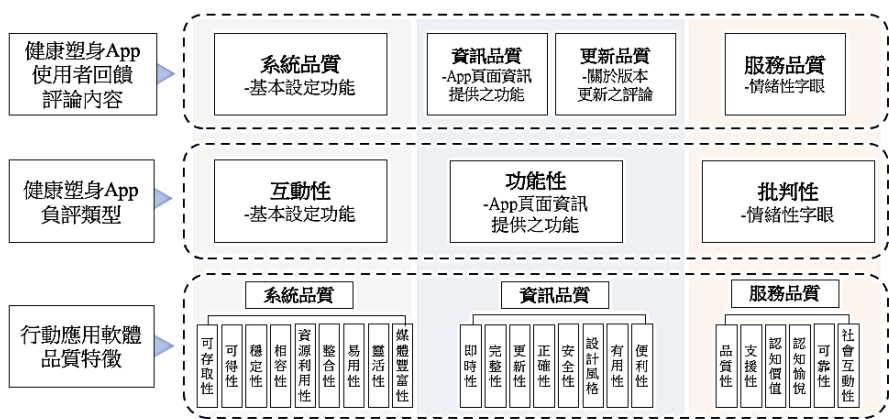


圖 1 健康塑身 App 使用者回饋評論內容

## 四、健康塑身 App

目前在應用程式開發上，常使用的 App 型態為原生型應用程式（Native App）、混合型應用程式（Hybrid App）與網頁應用程式（Web App），依照開發成本與技術而有所不同（趙子榮，2020）。本研究欲討論的 App 型態為原生型應用程式，於本研究中統稱為 App，此種 App 為使用者常見於 App Store（iOS）和 Google Play（Android）兩大平台上提供產品／服務，兩大平台提供原生型 App 的下載途徑，分別在各自的「健康與健身」與「健康塑身」類別皆有相關 App 提供使用者下載。

伴隨智慧手機的快速傳播，利用智慧手機進行運動消費的數量正在快速增加（Sama et al., 2014）。2020 年 12 月初 Google play 公布年度最佳 App 榜單，因疫情的影響，商業服務類的遠距協作 App 與健康塑身類的「Nuli 居家與重訓課表」與「HealthBuy 健康飲食消費指南」皆包含在其中。健康塑身 App 除每年度表現上升外，因應疫情年的影響也使此類 App 在市場中具備一定的使用人數與需求。

綜合上述，本研究欲探討目前熱門的健康塑身 App 使用者回饋與評價之關係，並以 2020 年 12 月前後三個月作為研究範圍，探討使用者回饋如何影響半年間評價變化，提供相關產品／服務能針對使用者回饋方面進行優化。

## 參、研究方法

### 一、研究架構

本研究為探討健康塑身 App 中使用者回饋與評價兩者之交互關係。欲釐清使用者回饋構面中的評論數、評論內容、正負評數是否影響 App 評價。使用者回饋為 2020/9/1 至 2020/12/1 前三個月累積之評論數量與評論內容，欲預測三個月後 2021/3/1 的評價包含排名和評分與半年間評價變化。

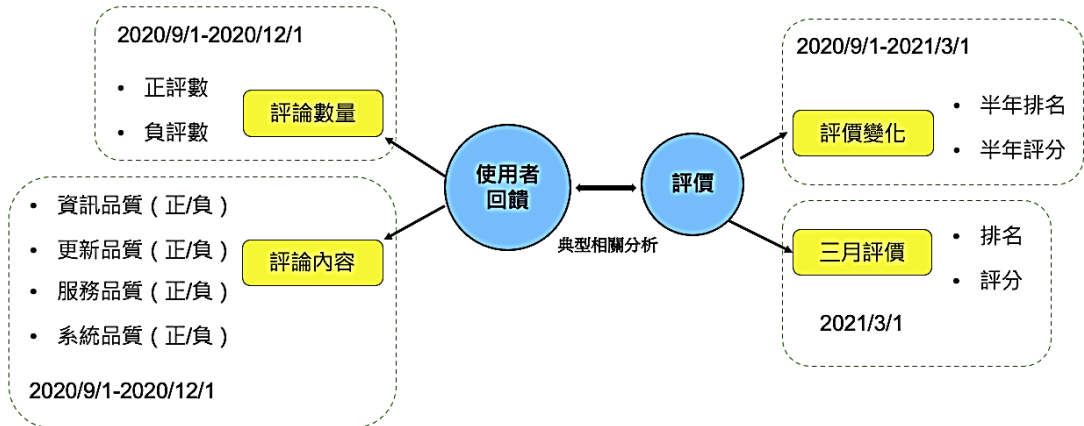


圖 2 研究架構圖

## 二、研究對象與資料蒐集

研究對象為 2020 年 9 月 1 日 Google Play 上臺灣地區「健康塑身」類別中排名前 100 名之免費類別 App。研究資料範圍為 2020 年 9 月 1 日至 2021 年 3 月 1 日共計半年期間。關於使用者回饋指標，為蒐集 2020/9/1 至 2020/12/1 前三個月間使用者回饋中的正面評論和負面評論，而評價部分，則是 2021/3/1 評價的排名／評分與半年評價變化，以此透過前三個月的使用者回饋討論後續評價變化。

## 三、研究工具

本研究主要探討健身塑身 App 中使用者回饋與 App 評價間之交互關係，針對使用者回饋與評價兩構面各研究工具內容分述如下：

### (一) 使用者回饋分類

使用者回饋中分為評論數量與評論內容。評論數量依據 Google 評論機制可分成正負評，1 至 2 顆星為負評，4 至 5 顆星為正評；評論內容為透過關鍵字擷取系統 (Tseng, 1998) 將健康塑身 App 的評論內容進行關鍵詞頻萃取，並將各關鍵詞頻數量依照行動軟體品質特徵中提到的系統、資訊、服務、更新四種品質之構面進行分類，依品質定義對該評論內容之詞頻數計次。

表 1 使用者回饋之評論內容分類

構面	定義		操作
使用者回饋 (評論內容)	系統品質	與App系統提供給使用者的系統基本功能有關，系統基本功能為各個App皆具備。	將評論內容做詞頻萃取，依品質定義對該評論之詞頻數計次。
	資訊品質	資訊品質主要軟體提供資訊是否符合使用者最初的使用目的。	
	更新品質	App 頁面中關於每次更新皆提供相關資訊，針對每次更新後有關更新好壞的評論皆屬於此構面。	
	服務品質	服務品質為使用者與軟體間的互動是否獲得處理，包含使用者對於產品／服務滿意度與相關情緒性字眼。	

## (二) 評價分類

針對本研究中的 App 評價指標分為評分與排名，兩者數據於 App Annie 平台取得，此平台透過彙整 Google Play 各類別 App 資訊，提供行動應用產業相關數據。2017 年 Google Play 取消 App 各類別排名，但部分合作廠商諸如 App Annie 仍可透過 API (Application Programming Interface) 調用方式取得評價相關資料。以此計算 2020 年 9 月 1 日至 12 月 1 日與 2020 年 12 月 1 日至 2021 年 3 月 1 日間的評價變化。評價中的評分為 App Annie 取用 Google Play 頁面資訊的整體評分欄位，而使用者回饋的正負評論則是平台用於提供使用者篩選與查看不同回饋內容。

## 四、資料處理與分析方法

本研究將健康塑身 App 評論內容透過關鍵字擷取系統做詞頻萃取，依照行動軟體四種品質特徵進行整理。關鍵字擷取系統為用以擷取文件中有意義且有代表性的字串代表性字串 (string)、片語 (key phrases)、詞彙 (keywords) 或內容片段 (key segments) 的自動化技術。該系統特色為能夠自動擷取新生詞彙、各領域的專有名詞，故適用於快速對欲研究文檔進行關鍵詞頻分析。本研究將使用者評論作為文本，以關鍵字擷取系統獲取詞頻與詞頻數量。後續透過 IBM SPSS Statistics 24 版中文套裝軟體對健康塑身 App 使用者回饋與評價進行敘述統計與典型相關分析。



## 肆、研究結果

本研究分析 Google Play 平台上前 100 大熱門健康塑身類 App 使用者回饋與評價間交互關係，透過典型相關分析討論兩構面間彼此之影響，釐清使用者回饋是否影響未來 App 評價。

### 一、使用者回饋評論內容之關鍵字萃取結果

透過關鍵字擷取系統將 100 個健康塑身 App 於 2020 年 9 月 1 日至 2020 年 12 月 1 日之使用者正負評進行萃取後，根據行動軟體特徵模型的四種品質特徵：系統、資訊、更新與服務品質進行分類，萃取較常被使用者提及的關鍵詞，統整如下圖。系統品質中，較多人提及的關鍵詞為訂閱、閃退、斷線、介面、通知與設定，屬於 App 系統基本功能，正評較多人提及的為「介面」基本功能，在負評方面則是出現比較多關鍵詞「閃退」的討論，共有 61 個詞頻；資訊品質為比對各 App 頁面資訊內容的介紹欄位，關鍵詞內容主要有紀錄、睡眠、跑步、手錶、訓練與計畫。正評資訊品質多為「紀錄」與「睡眠」，負評的資訊品質則以「睡眠」與「手錶」最多；更新品質相對詞頻內容較少，詞頻內容為更新、更換，使用者多會提及 App 更新後是否比前一版本好用，負評提及「更新」的關鍵詞頻比正評多 163 次；服務品質在此次關鍵字萃取後多為情緒性字眼，為客服、垃圾、差勁、爛、方便、好用，多是有關正負評論使用者對於使用後的滿意度／心得。

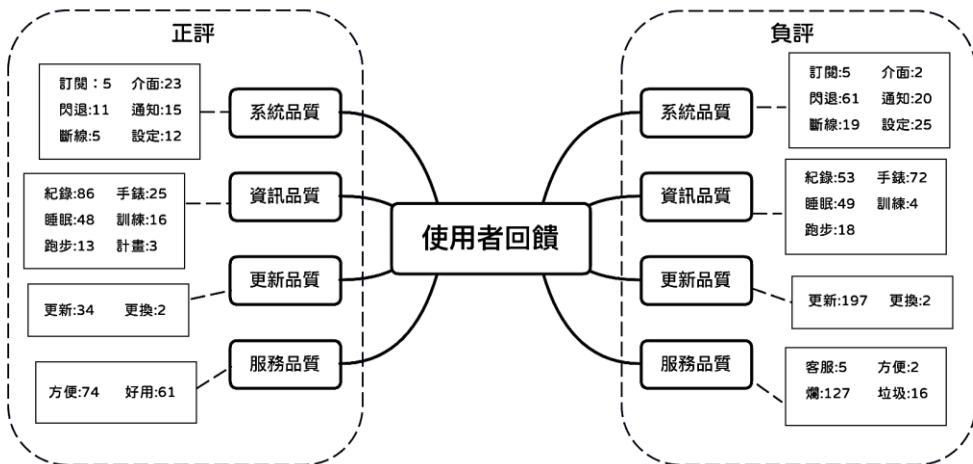


圖 3 使用者回饋評論內容萃取結果

## 二、前 100 大健康塑身 App 之使用者回饋與評價

將使用者回饋與評價兩變項資料進行敘述統計。使用者回饋變項中，負評內容特性的詞頻數平均值以資訊與系統品質較多，分別是 4.82 與 4.71，正評則以資訊品質為 5.87 最多，健康塑身 App 中較多使用者回饋以討論資訊品質為主。

使用者回饋資料最大值與最小值顯示前 100 大健康塑身 App 中，單一 App 討論的使用者回饋數量，表中可見有些 App 於三個月期間會有無評論的情況。蒐集評論數量中，此 100 個 App 正評數共有 5191 則，負評數則是 1011 則，於 2020 年 9 月至 12 月期間負評數比正評數少。評價變項中，三月排名與評分平均值分別為 80.37 與 4.00，半年排名變化平均為退後 29.87 名，而評分變化平均增加 0.05。

表 2 使用者回饋、評價與更新資料

構面	變項	最小值	最大值	平均值	標準差
使用者回饋	負評（資訊）	0	95	4.82	12.26
	負評（更新）	0	54	2.22	7.14
	負評（系統）	0	56	4.71	10.63
	負評（服務）	0	90	2.04	9.81
	正評（資訊）	0	124	5.87	13.74
	正評（更新）	0	10	.34	1.32
	正評（系統）	0	40	2.14	4.69
	正評（服務）	0	33	2.42	5.38
	正評數	0	147	51.91	107.1
	負評數	0	840	10.11	21.11
評價	3 月排名	1.00	500.00	80.37	90.56
	3 月評分	1.80	4.90	4.00	.88
	半年排名變化	-464	44	-29.87	83.6
	半年評分變化	-0.1	2	0.05	0.36

## 三、使用者回饋評論數量與評價之關係

先針對健康塑身 App 使用者回饋數量與評價關係進行典型相關分析，此兩組典型相關變量有三組特徵值與典型相關係數。維度遞減分析中 1 TO 3 與 2 TO 3 的維度遞減分析之  $F$  值皆具有顯著水準，需考慮兩組典型相關係數。第一組典型相關係數為 0.84，解釋變異量 71.92%；第二組典型相關係數為 0.69，解釋變

異量 26.98%，皆為高度相關。此兩組典型相關函數累積變異量為 98.91%，能解釋此模型大部分變異量。

表 3 使用者回饋數量與評價典型相關檢驗結果

典型相關	特徵值	解釋變異量%	累積變異量%	典型相關係數	典型相關係數平方	Wilk's 值	F 值
1 To 3	2.36	71.92	71.92	.84	.70	.15	21.27*
2 To 3	.89	26.98	98.91	.69	.47	.51	12.45*
3 To 3	.04	1.08	100.00	.19	.03	.97	1.70

\* $p < .05$

典型因素中的結構係數（典型負荷量）可用來解釋原始變項與典型因素之關係。典型因素之結構係數絕對值  $\geq 0.3$  即可視為有意義，代表此一變數對於所對應之線性組合具有顯著之解釋能力；若典型因素之結構係數絕對值  $\geq 0.5$ ，即為高相關（王正華、陳寬裕，2011）。從下表 4 可發現，第一組典型相關係數為 .84，在使用者回饋數量此原始變項中，正評數量的結構係數為 -.73，負評數量的結構係數為 .51；評價中，三月評分的結構係數為 -.93，三月排名的結構係數為 .56。當正評數越多時，會影響半年後的三月排名往前，而評分表現會變佳；反之，當負評增加時，排名與評分表現則會變差。在第二組典型函數為 .69，其中負評的結構係數為 -.83，評價構面的三月排名為 .80，當前三個月的正負評數量變少時，會影響三個月後的排名增加且評分上升。整體而言，當正負評論數量越多時，將影響 App 評價表現。

表 4 使用者回饋數量與評價之典型相關摘要表

使用者回饋數量	典型因素		評價	典型因素	
	$\chi^1$	$\chi^2$		$\eta^1$	$\eta^2$
負評數量	.51	-.83	半年排名變化	-.27	-.31
正評數量	-.73	-.60	半年評分變化	.15	.32
			2021/3 月排名	.56	.80
			2021/3 月評分	-.93	.34
抽出變異量%	42.04	41.68	抽出變異量%	22.54	11.28
重疊量數	29.53	19.57	重疊量數	32.10	24.04
$\rho^2$	.70	.47			
典型相關 ( $\rho$ )	.84	.69			
$p$	.00	.00			

#### 四、使用者回饋內容與評價之關係

後續針對使用者回饋內容與評價間之關係進行典型相關分析，以此討論正負評論內容特性是否與評價表現會出現交互關係，此兩組典型相關變量有四組特徵值與典型相關係數，如下表 4。維度遞減分析中 1 TO 4 與 2 TO 4 的維度遞減分析之  $F$  值皆具有顯著水準 ( $p$  值  $< .05$ )，需考慮兩組典型相關係數，其餘兩組典型相關函數則不具顯著水準。第一組典型相關係數為 0.80，解釋變異量為 57.20%；第二組典型相關係數為 0.73，解釋變異量為 35.95%，皆為高度相關。此兩組典型相關函數累積變異量為 93.15%，能解釋此模型的大部分變異量。

表 5 使用者回饋內容與評價典型相關檢驗結果

典型 相關	特徵值	解釋變 異量%	累積變 異量%	典型相 關係數	典型相關 係數平方	Wilk's 值	F 值
1 To 4	1.80	57.20	57.20	.80	.64	.14	6.37*
2 To 4	1.13	35.95	93.15	.73	.53	.38	4.17*
3 To 4	.19	5.89	99.05	.40	.16	.82	1.34
4 To 4	.03	.94	100.00	.17	.03	.97	.45

\* $p < .05$

表 6 中第一組典型相關係數為 .80，在使用者回饋內容此變項中，研究發現正評中的四種品質特徵相較於負評的品質特徵的結構係數絕對值較大，表示使用者回饋中正評對於評價影響較明顯，且反映在 App 排名上。當正評評論內容越多，App 排名將會靠前。正評內容中又以資訊品質的結構係數較大為 -.80，其次是服務品質、系統品質與更新品質，使用者在健康塑身 App 上提及的資訊介紹是否符合使用目的會影響該 App 的評價表現。使用回饋中評論內容與評價間的交互關係中第二組典型相關係數具顯著性，為 .73。其中負評的四種品質特徵比正評的品質特徵之結構係數較大，表示使用者回饋中負評對於評價影響較具有解釋能力。在負面評論中的四種品質特性中，資訊品質、系統品質與服務品質的結構係數絕對值皆高於 .8；評價變項中，三月排名與評分的結構係數為 -.33 與 -.79。在此典型相關係數中，當資訊品質、系統品質與服務品質等負評內容越多時，評價中的評分變化較大會影響評分表現下降，但在排名上會稍微靠前。

表 6 使用者回饋內容與評價之典型相關摘要表

使用者回饋	典型因素		評價	典型因素	
	$\chi^1$	$\chi^2$		$\eta^1$	$\eta^2$
負評（資訊）	-.17	.87	半年排名變化	-.23	-.16
負評（更新）	-.10	.65	半年評分變化	.36	-.34
負評（系統）	-.04	.86	2021/3 月排名	.89	-.33
負評（服務）	-.21	.82	2021/3 月評分	-.60	-.79
正評（資訊）	-.80	.25			
正評（更新）	-.20	.37			
正評（系統）	-.42	.30			
正評（服務）	-.61	.24			
抽出變異量%	32.90	21.43	抽出變異量%	21.39	11.66
重疊量數	17.48	13.78	重疊量數	33.26	21.95
$\rho^2$	.64	.53			
典型相關（ $\rho$ ）	.80	.73			
$p$	.00	.00			

綜上所述，研究發現在 2020 年 9 月 1 日至 2021 年 12 月 1 日此三個月期間，健康塑身 App 使用者回饋整體正評數量是高於負評的。負評內容中，需要注意的是資訊品質與系統品質，平均較高，使用者在意健康塑身 App 本身頁面資訊上提及的功能是否達到使用目的，且像是基本設定或出現閃退等系統品質問題都造成負評變多。透過典型相關分析，使用者回饋數量中正評變多能解釋後三個月評價表現變好，負評數量則為影響評價變差，當正負評整體數量變多，代表有越多聲量，在排名上會變好。使用者回饋內容方面，當出現正評資訊品質變多時能夠影響評價表現提升，最後負評的資訊、服務與系統品質變多，則會造成評分變差。總體而言，App 開發商如欲提升自身產品的評價，能多針對使用者回饋做改善，將有助於釐清產品改善的方向並且提升產品／服務在該類別未來排名與整體評分上的表現。

## 伍、結論與建議

針對健康塑身 App 使用者回饋與 App 評價之交互關係進行統整，並依據研究目的、研究結果與討論提出結論與建議。

## 一、結論

透過 2020 年 9 月 1 日前 100 名健康塑身 App 前三個月的使用者回饋與半年後的評價，能釐清健康塑身 App 中兩者間的關係，對於本研究之結果提出結論如下：

### （一）健康塑身 App 使用者回饋多討論資訊品質

本研究以 2020 年 9 月 1 日 Google Play 上前 100 個健康塑身 App 為研究對象，並以探討使用者回饋與評價為主要內容。在討論兩者關係前，先整理目前健康塑身 App 使用者回饋與評價資料。前三個月的使用者回饋構面中，健康塑身 App 負評數量較正評少，其中正負評的資訊品質較多，使用者多在意 App 是否符合頁面介紹之功能；評價方面，整體前 100 名健康塑身 App 半年期間排名下降但評分有小幅提升。

### （二）使用者回饋正負評數量會影響下一季排名

本研究將使用者回饋構面分成正負評數量與內容兩者，分別對於評價的半年變化與三月表現進行典型相關分析，以此討論彼此交互關係。在正負評數量上，正面評論增多與負面評論下降皆有助於未來的評價表現，驗證 Google Play 平台提及針對使用者評論進行反饋對 App 的幫助。值得注意的是，當健康塑身 App 正負評數量變多時，意味討論聲量變高，有助於評價中的排名上升。但在 App 評分上會出現下降的情形，因此當開發商欲提高此類別中的 App 排名，確實能透過提高使用者回饋的評論量，製造越多使用者討論，來幫助產品／服務在更好的類別排名中。

### （三）處理正負評的資訊品質可提升未來評價

Google Play 的 Developer 指南說明回覆使用者回饋有助於排名提升，但未進一步說明能夠處理使用者回饋的方向，本研究蒐集健康塑身 App 使用者回饋內容以平台機制區分成正負評，經關鍵字擷取將關鍵詞以品質特徵模型進行分類。使用者回饋內容與評價的交互關係中，正面評論有關於資訊品質與服務品質的留言，會反映到三個月後的排名往前；在負面評論中，只要資訊、服務與系統品質類型的留言變多，將影響到平台上的評分表現變差，使用者透過網路提供的線上評論，會影響 App 產品／服務口碑，因此有相關評論出現需及時處理並進行優化。當大量評論出現時，開發商能夠先篩選出關於資訊品質的評論進行檢視，對於 App 在平台頁面資訊提及的功能如無法符合預期，使用者會有較直接的感受，也有利於提升產品／服務。

## 二、建議

### （一）實務應用建議

研究發現使用者回饋與評價之交互關係具有顯著性，透過增加產品的正負評論數量，能夠提升排名，因此開發商欲提高 App 在該類別的排名，能先增加該產品討論聲量。Pagano 與 Maalej（2013）發現使用者回饋數量會在更新版本出現後一時間遞減，須即時將使用者回饋擷取，因此能夠在版本釋出後針對使用者回饋內容進行分類。使用者回饋內容上，開發商需注意頁面資訊說明是否符合，使用者負評如出現有關資訊、系統與服務品質的內容會影響評價表現，需優先處理使用者回饋的負面評論。

### （二）未來研究建議

未來能針對不同平台如 App Store 與 Google Play 間進行討論，除了能釐清不同平台間同一類別的內容差異，也能瞭解使用者回饋與評價兩者關係在不同平台的變化。而 App 評價上因只取得頁面資訊中的評分與 App Annie 上的排名，如能取得參與度、流量與營收等後台數據指標，能進一步探討 App 評價上使用行為的多項指標。本研究以前 100 大健康塑身 App 資料為主，研究發現有些 App 會出現研究時間範圍內無評論之情形，後續研究能將期間內無評論與有評論的 App 分開討論，釐清兩者差異。

因研究對象主要是前 100 名健康塑身熱門 App，並透過典型相關釐清此 App 類別使用者評論與評價之關係。研究發現擷取關鍵詞頻後，會出現使用者雖然給該 App 五星評分，但因本研究依據正負評分類，而被歸類在正評，後續仍有建議內容可讓開發商提取做改進。故當開發商或未來研究針對使用者評論進行討論時，除提取負面評論的意見外，亦須將正評中的建議納入資料，進行優化方向之考量。

## 參考文獻

### 一、中文書目

- 王正華、陳寬裕 (2011)。《論文統計分析實務：SPSS 與 AMOS 的運用》。臺北市：五南。
- 何淑君、林勢敏、李珮琪 (2019)。〈從使用者觀點分析線上評論及建構行動應用軟體之品質特徵〉，《資訊管理學》，26, 1: 99-119。
- 國家發展委員會 (2016)。《ide@ Taiwan 2020 (創意臺灣) 政策白皮書》。臺北市：國家發展委員會。
- 葉謹睿 (2010)。《互動設計概論——後數位時代的網站、介面、產品及軟體設計的原則》。臺北市：藝術家出版社。
- 趙子榮 (2020)。《Progressive Web App: 探討未來 Web App 在 App 市場的趨勢及可行性》。臺北：國立臺北教育大學數位科技設計學系（含玩具與遊戲設計碩士班）碩士論文。

### 二、英文書目

- Bonchek, Mark, and Sangeet Paul Choudary (2013). Three elements of a successful platform strategy. *Harvard Business Review*, 92:1-2.
- Carreño, Laura V. Galvis, and Kristina Winbladh (2013). Analysis of User Comments: An Approach for Software Requirements Evolution. Paper presented at International Conference on Software Engineering (ICSE), San Diego, CA.
- Ciurumelea, Adelina, Andreas Schaufelbühl, Sebastiano Panichella, and Harald C. Gall (2017). Analyzing Reviews and Code of Mobile Apps for Better Release Planning. Paper presented at International Conference on Software Analysis, Evolution and Reengineering (SANER), Klagenfurt, AU.
- Hennig-Thurau, Thorsten, and Gianfranco Walsh (2003). Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the Internet. *International journal of electronic commerce*, 8, 2: 51-74.
- Huete-Alcocer, Nuria. (2017). A literature review of word of mouth and electronic word of mouth: Implications for consumer behavior. *Frontiers in psychology*, 8: 1256.



- Karagkiozidou, Makrina, Christos Ziakis, Maro Vlachopoulou, and Theodosios Kyrkoudis (2019). App Store Optimization Factors for Effective Mobile App Ranking. In Kavoura, Androni, Efstathios Kefallonitis, and Apostolos Giovanis (Eds.), *Strategic Innovative Marketing and Tourism* (pp. 479-486). Northern Aegean, GR: Springer.
- Li, Ching, Wei-Shiang Lai, and Tzu-Chun Huang (2020). The Classification of Google's Top 50 Health and Fitness Apps from Health Promotion Approaches. Paper presented at the 2020 International Conference on the Development and Application of Big Data and Enterprise Resource Management, Taoyuan.
- Liu, Bing. (2020). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge: Cambridge university press.
- Norman, Donald A. (1983). Some observations on mental models. *Mental models*, 7, 112: 7-14.
- Novak, Ed, and Chris Marchini (2019). Android App Update Timing: A Measurement Study. Paper presented at the 20th IEEE International Conference on Mobile Data Management (MDM), Hong Kong, CN.
- Pagano, Dennis, and Walid Maalej (2013). User feedback in the appstore: An empirical study. Paper presented at 21st IEEE International Requirements Engineering Conference (RE), Rio de Janeiro, BR.
- Rogers, Richard (2018). Digital traces in context Otherwise engaged: Social media from vanity metrics to critical analytics. *International Journal of Communication*, 12: 23.
- Sama, Preethi R., Zubin J Eapen, Kevin P Weinfurt, Bimal R Shah, and Kevin A Schulman (2014). An evaluation of mobile health application tools. *JMIR mHealth and uHealth*, 2, 2: e3088.
- Tseng, Yuen-Hsien (1998). Multilingual keyword extraction for term suggestion. Paper presented at the 21st ACM SIGIR conference on Research and development in information retrieval, New York, NY.
- Van Alstyne, Marshall W., Geoffrey G. Parker, and Sangeet Paul Choudary (2016). Pipelines, platforms, and the new rules of strategy. *Harvard business review*, 94, 4: 54-62.

### 三、電子資料

FlipWeb 數位資產觀點 (2017)。〈五個檢測 APP 應用程式表現與指標〉。FlipWeb 數位資產觀點，網址：<https://medium.com/@flipweb/>。點閱日期：2021 年 6 月 1 日。

App Annie (2019). The state of mobile in 2019. Retrieved June 3, 2021, from App Annie. Website: <https://www.appannie.com/en/insights/market-data/the-state-of-mobile-2019/>

Android Developer (2020). Analyze user reviews to understand opinions about your app. Retrieved June 3, 2021, from Android Developer. Website: <https://developer.android.com/distribute/best-practices/grow/user-reviews>

Clark, Smith (2019). The Digital Marketing Metrics That Matter in 2020. Retrieved June 3, 2021, from Webmastered. Website: <https://www.webmastered.com/digital-marketing-metrics-that-matter-in-2020/>

Gabor, Papp (2017). Why Focusing Too Much on Acquisition Will Kill Your Mobile Startup. Retrieved June 7, 2021, from Phiture. Website: <https://phiture.com/mobile-growthstack/why-focusing-on-acquisition-will-kill-your-mobile-startup-e8b5fbd81724/>

# The Relationship between User Feedback and Evaluation of Health and Fitness Apps

Wei-Shiang Lai<sup>1</sup> Ching Li<sup>2</sup> Chia-Wen Lee<sup>3</sup> Tzu-Chun Huang<sup>4</sup>

## Abstract

The user number of health and fitness apps has increased showing the growth of health and fitness apps is a trend in recent years. Google Play had mentioned that user feedback was a way to enhance the evaluation of apps but without further explanation. Therefore, this study aims to explore the relationships between user feedback and app evaluation. The study applied content exploration and statistical analysis to analyze the information of the top 100 health and fitness apps on Google Play. The data collection time was from September 1st, 2020 to March 1st, 2021, and the research tool was the Keyword Extraction System and SPSS. The results showed that user feedback and app evaluation had a significant interactive relationship. The number of positive and negative reviews had an effective influence on the apps' future ranking. In addition, the content of positive and negative reviews showed that users care about information quality. The finding provides app developers with methods to evaluate their products and services in response to the user feedback and enhance app evaluation.

**Keywords:** Keyword Extraction, User Review, Google Play, Platform

---

<sup>1</sup> Master, Graduate Institute of Sport, Leisure and Hospitality Management, National Taiwan Normal University

<sup>2</sup> Professor, Graduate Institute of Sport, Leisure and Hospitality Management, National Taiwan Normal University

<sup>3</sup> Associate Professor, College of Modern Management, Yango University

<sup>4</sup> Ph.D. student, Graduate Institute of Sport, Leisure and Hospitality Management, National Taiwan Normal University

Corresponding Author: Ching Li, E-mail: t94002@ntnu.edu.tw

Received: 2021/10/21; Accepted: 2022/06/09

