

# 聊天機器人性別形塑對使用者信任感受之影響

Gendering chatbots by design and its impacts on users trust towards shopping chatbots

<sup>1</sup>游曉貞 <sup>2</sup>陳繹介 <sup>3</sup>蕭方雯

<sup>1</sup>國立臺中科技大學多媒體設計系 | 助理教授 | hcyous@gmail.com

<sup>2</sup>國立臺中科技大學多媒體設計系 | 碩士 | skyway605@gmail.com

<sup>3</sup>國立臺中科技大學多媒體設計系 | 研究生 | sharon22332001@gmail.com

<sup>1</sup> Hsiao-Chen You <sup>2</sup> Yi-Chieh Chen <sup>3</sup> Fang-Wen Xiao

<sup>1</sup> Department of Multimedia Design, National Taichung University of Science and Technology, Assistant professor, hcyous@gmail.com

<sup>2</sup> Department of Multimedia Design, National Taichung University of Science and Technology, Master, skyway605@gmail.com

<sup>3</sup> Department of Multimedia Design, National Taichung University of Science and Technology, Graduate student, sharon22332001@gmail.com

## 引用本文：

游曉貞、陳繹介、蕭方雯 (2022)。聊天機器人性別形塑對使用者信任感受之影響。《感性學報》，10(1)，51-80。

**隨** 人工智慧 (AI) 科技的成熟，企業紛紛應用聊天機器人 (chatbot) 於其行銷服務當中。而大量 AI 服務的出現，讓人機互動領域學者開始重視 AI 的性別化議題，希望理解形塑 AI 性別的設計手法與 AI 性別化對使用者的影響。本研究透過兩階段實驗，分別探索不同聊天機器人的視覺設計作為性別提示的效果，以及購物情境中聊天機器人性別對於使用者信任感受之影響。前導實驗先檢視聊天機器人頭像設計中不同視覺元素對使用者性別感受之影響。再進一步藉由操作聊天機器人的性別提示組成四種條件的組間設計實驗，招募受測者進行模擬網路購物的互動實驗，了解不同性別購物聊天機器人給人的信任感受差異。結果發現聊天機器人頭像的髮長、顏色、與眉型都可以明顯影響人們對聊天機器人的性別感受；但聊天機器人的性別對於人們的信任感受僅有些微影響；反倒是受測者本身的性別與年齡對聊天機器人的信任感受影響較大。

關鍵詞：聊天機器人、性別提示、信任、感性工學、對話式使用者介面

**W**ith the rise of artificial intelligence technology, companies have integrated chatbots into e-commerce marketing services nowadays. With the emergence of a vast number of AI applications, the researchers in HRI have begun to pay attention to the gender issue of AI, hoping to understand the impact of AI gender on users. This study investigates the effects of chatbot gender on user trust in online shopping contexts using a two-stage experimental design. First, this study examined the impact of the visual elements of AI avatars on the user's gender perception in the pilot study. Then, this study conducted a between-subject design experiment to examine how gender factors affect user trust towards shopping chatbots. The results found that gender cues have apparent impacts on the gender perception of chatbots, but the gender of chatbots has very little influence on users' trust towards them; instead, the gender and age of participants have more impact on the user trust.

Keywords: Chatbot, Gender Cues, Trust, Kansei Engineering, Conversational User Interface

## 一 緒論

對話式使用者介面 (Conversational User Interface, CUI) 泛指能模擬真人對話的方式，讓使用者跟電腦進行溝通的使用介面。其目標是讓人們可以用日常的語言直接告訴電腦該怎麼做，無須透過制式的電腦指令語言或圖像化界面 (Graphical User Interface, GUI) 等傳統方式與系統互動 (Brownlee, 2016)。這種讓電腦能模擬人類與使用者對話的概念可以追溯到 1950 年代，Alan Turing 提出圖靈測試 (Turing Test) 讓受試者從人機對話中，來判斷對話的一方是電腦或真人。礙於當時電腦科技水準發展有限，直到今日人工智慧 (Artificial Intelligence, AI) 與自然語言處理 (Natural Language Processing, NLP) 技術提升到電腦能夠分析、理解、和創造人類語言時，CUI 才開始展現其實用價值 (Shevat, 2017)。近年來類似 Apple 的 Siri、Google 語音助理、Amazon 的 Alexa… 等的對話式助理服務陸續出現，使用者可以隨時透過交談方式來取得 AI 助理的協助，有效提升工作效率與人們生活品質，CUI 儼然成為形塑人類未來智慧化生活模式的重要指標 (Yonck, 2017)。

聊天機器人 (Chatbot) 是一種透過聲音或文字與使用者進行對話的電腦程式，換言之，它是透過 CUI 手法讓軟體化服務得以曝光的程式應用，也叫做對話式助理 (Conversational Agents)、聊天助理 (Chat Agents)、對話式介面 (Conversational Interfaces) (Shevat, 2017)，是當今科技下體現 CUI 的主流形式。近年聊天機器人逐漸成為各大企業不可或缺的發展項目，主要原因除了前述 CUI 技術成熟與對話式商務興起之外，即時通訊平台的使用人數遽增，更促使各大企業企圖藉由聊天機器人客服在 Line、Facebook、Slack 等社群平台，進行電子商務行銷與維護顧客關係，使得 CUI 的服務應用範圍更加擴大 (Nguyen, 2017; Casas et al., 2020)。

CUI 的興起，是電腦使用者介面發展中的一大創新，因為 CUI 消除了人們與機器溝通的技術障礙，讓人機之間的互動就像日常生活中人與人交談那樣地自然。CUI 的重心不在於使用者介面 (User Interface) 本身，而是人際對話 (Human-Human-Conversation, HHC)，這樣的對話已然不是電腦人機互動 (Human-Computer-Interaction, HCI)，對 HCI 領域產生了典範轉移 (Paradigm Shift)，許多 HCI 理論與準則已不再適用於 CUI，諸多問題有待釐清 (Chiriyankandath, 2018)。特別是 CUI 將如何影響人類與電腦系統的互動體驗？使用者是否會將人類社會的慣例與特徵投射到擬人化的電腦系統，進而影響人類對待這些虛擬助理的態度與行為？深入檢視這類議題將有助 CUI 領域的進一步發展。性別刻板印象對人類行為產生深

遠影響。過去許多研究指出性別偏見會影響人們對擬人化電腦系統的看法，使用者傾向根據電腦的性別設定來評定系統的表現，而非依據系統的性能 (McDonnell & Baxter, 2019)。但是針對聊天機器人此一新興科技應用，我國在聊天機器人性別因素對人機互動影響的討論較少。因此本研究透過實驗設計來討論聊天機器人的性別提示的效果，以及電商聊天機器人性別對於使用者信任感受之影響。

## 1.1 研究背景與動機

隨著線上聊天機器人客服的大量興起，人們對它的接受狀況卻不如預期樂觀，Følstad, Nordheim 與 Bjørkli (2018) 認為這主要關鍵在於聊天機器人尚未能給人們足夠的信任感。信任是人們在進行社會互動時不可或缺的部分，當一個人或是系統被認為是具備信任感時，人們才會有願意依賴他們的服務。Madsen 及 Gregor (2000) 認為人機之間的信任是促成使用者聽從系統建議、甚至願意採取行動的指標。為了提升聊天機器人互動品質，不少專家提出賦予聊天機器人適當的擬人化個性的建議 (Saffer, 2017; Dua, 2017)。而聊天機器人個性的展現可以透過：聊天機器人的命名、性別、背景設定…等角色設定來達成 (Hay, 2017)。

人類社會互動演化至今，「性別」常是社會化生活中人際互動參考的依據。隨著各式 AI 服務或機器人產品上市，人們對於 Siri、Pepper、甚至掃地機器人的性別討論也慢慢浮出檯面 (Knibbs, 2014; Newitz, 2009; Turk, 2014)。Schwartzman (1999) 認為機器人的性別能提供使用者一種與其互動的「錨點」；賦予它們性別標記，能讓冰冷的機器人更具人性化。Robertson (2010) 呼應 Schwartzman 的主張，認為機器人性別形塑其實是賦予機器人一種社交符號、一種身份的標誌，讓使用者更可以理解機器人在日常生活上所代表的意義。

## 1.2 研究目的

任何人都可以指派一個特定的性別給機器人，但這樣的性別化過程中，往往也會將人們的性別偏見帶進 AI 的人機互動之中，所以 Coren (2017) 指出 AI 性別化的挑戰在於理解 AI 服務的性別要如何塑造？進而藉此設計出對抗性別偏見、有助性別平等的 AI 服務。不少機器人的人機互動 (Human Robot Interaction, HRI) 學者陸續針對實體的機器人性別化與它對使用者的感受與行為的影響開始調查 (Wang, 2014; Eyssel & Hegel, 2012; Siegel, Breazeal, & Norton, 2009; You & Lin,

2019)。但是較少深入探討聊天機器人的性別議題，因此，本研究以網路購物情境中的購物聊天機器人為例，透過實驗設計來理解聊天機器人的性別化對我國使用者信任感受的影響。首先以感性工學方法探討聊天機器人頭像的設計元素對使用者性別感受之影響；再進一步藉由操作聊天機器人性別提示的互動模擬實驗，來釐清聊天機器人的人機互動中性別因素對使用者信任感受的影響，以提供未來聊天機器人設計師形塑其性別角色之參考。本研究主要探討問題有三：

- (1) 聊天機器人視覺設計元素可否作為「性別提示」影響人們對它的性別判斷？
- (2) 於網路購物中，購物聊天機器人的性別是否會影響人們對它的信任感受？
- (3) 使用者本身的性別是否也會對購物聊天機器人的信任感受產生影響？

## 二 文獻探討

### 2.1 對話式使用者介面 (CUI) 與聊天機器人 (Chatbot)

只要是允許人們使用自然語言，模擬人際對話方式與電腦系統進行溝通的互動形式，無論是透過文本或語音溝通，都屬 CUI 的一種 (Brownlee, 2016)。電腦語言和人類日常使用的語言不同。過去受限於當時科技水準，為了要能於電腦溝通，使用者必須學習指令語言或是透過圖像化使用介面 (GUI) 才能與電腦互動。隨著 AI 以及 NLP 技術的進展，今日的人機互動已發展到人們可以使用自然語言來與電腦對話的 CUI 模式，使得數位化服務的人機互動更直覺、自然和有效率 (Mielke, 2016)。Messina (2016) 指出人們未來會更趨向利用 CUI 與企業對話，並提出「對話式商務」(Conversational Commerce) 一詞，指的就是電腦系統可以透過 CUI 的互動模式，提供顧客即時且客製化的購物、客服、諮詢 .. 等商業服務。隨 CUI 應用的盛行，更衍生出可以混用文本、語音或 GUI 元件 (如：按鈕、選單、圖像、影片) 的複合介面，讓使用者與電腦系統的對話體驗更為豐富。但因主要對話媒介管道的不同，CUI 可分為語音使用者介面 (Voice User Interface, VUI) 與文本使用者介面 (Message User Interface, MUI) 兩大類。前者使用者主要是透過口語來完成溝通任務，如 Apple Siri 和 Amazon Echo 的互動模式；後者則多是透過網頁或行動裝置介面讓人們以文本輸入來提問和檢索信息的互動方法 (Pan, 2017)。

聊天機器人是指透過 AI 技術，以電腦程式模擬的對話內容與使用者互動，來幫使用者解決購物、訂餐、叫車、資料查詢... 等日常瑣事的軟體服務。因此，目前提

供使用者對話式介面的各式軟體程式與助理服務，都可以通稱為聊天機器人，是當今科技下體現 CUI 的主流形式 (Shevat, 2017)。聊天機器人一詞起源於九十年代，是一種旨在模擬人類智能與對話方式的軟體系統。因為它們使用自然語言與使用者輪流交換對話內容，聊天機器人也被稱為對話式助理。使用者可透過文本輸入或語音命令與聊天機器人進行對話 (Di Gaetano & Diliberto, 2018)。如同上述，由於對話內容傳遞渠道（聽覺與視覺）的差異 CUI 可區分為 MUI 與 VUI 兩類，聊天機器人也因對話媒介、所需設備與應用平台分成兩大類，一是透過 VUI 的語音對話的聊天機器人，也稱為語音助理 (Voice Assistant)；另一類則是常出現於社群平台、透過 MUI 以虛擬助理形式出現的電商客服或購物聊天機器人 (Pan, 2017)。

聊天機器人可藉由對話蒐集使用者資訊，逐步分析其喜好，進而適時提供商品的建議。對於企業而言，它不只為顧客提供更加個人化的服務，同時也可蒐集顧客的數據，遂成為推動電子商務服務的一個重要項目 (Nguyen, 2017)。聊天機器人所擁有的高效率即時性回覆能力，不僅可同時與大量客戶進行對話、節省客戶因客服人員忙線而等待的時間；同時也可作為即時監測資訊的幫手，協助客戶在短時間內完成服務 (優拓資訊, 2017)。

## 2.2 聊天機器人的性別化

人類的性別包括兩個層次：生理性別 (Sex) 與社會性別 (Gender)。「生理性別」是人類出生因生理結構上的明顯差異所產生的男、女性區分 (王振寰、瞿海源, 2009)。「社會性別」是一種社會與文化因素對生理性別差異所抱持的特有文化或態度，經常被區分成陽剛氣質與陰柔氣質兩類 (Calhoun, Light, & Keller, 2001)。由於聊天機器人與其它的擬人化電腦系統 (例如：實體機器人與虛擬助理) 都是人造物，本身不具備性別，它的性別化須倚賴設計者加諸於電腦系統的性別提示 (Gendered cues)，例如：名稱、聲音、外型等，來喚起使用者對於電腦系統的性別認知 (Wang, 2014)。

為了要開發出能提升客戶關係，有效協助使用者提升其工作與生活品質的聊天助理，不少機器人互動設計 (Human-Robot Interaction, HRI) 專家提出賦予聊天機器人適當的角色設定來達成穩定一致的個性的建議，包括設定聊天機器人的名字、性別、人格、背景設定…等 (Saffer, 2017; Dua, 2017; Hay, 2017)。各科技大廠也確實花費不少功夫為旗下的聊天機器人打造鮮明的角色，例如：Siri 是「風趣、狡獪的助理」，性別預設為女性，可自行變更；Google Assistant 是「中性、專業的文青

圖書館員」；微軟的 Cortana 是「酷酷的怪咖女性」，來符合企業品牌形象 (Vlahos, 2019)。Feine 等人 (2019) 針對在 chatbots.org 平台上的 1,375 個文本式聊天機器人的調查，發現應用於品牌形象、客服和銷售的聊天機器人 多被設定成女性，她們多數具有女性名字、女性頭像，並被描述為女性聊天機器人。設計師依循社會慣例的刻板印象來打造聊天機器人性別與角色設定的做法，雖然可以快速被企業或使用者接受，但卻潛藏了促成性別刻板印象、甚至性別偏見的可能。

## 2.3 機器人性別化對 HRI 之影響

Nass 與 Yen (2010) 指出使用者會下意識地將電腦系統視為「人類」來看，採用人際溝通的社交法則與其互動，並將社交禮儀與性別刻板印象等社會規則帶入人機互動中。Lee, Nass 與 Brave (2000) 利用聲音作為性別提示，並讓受測者去與不同配音的電腦系統互動，結果發現受測者會因為電腦的聲音給人的性別感受，而對電腦產生類似人類社會互動時的性別刻板反應。過去多項研究顯示即便在任務與互動內容完全一致的情況下，僅藉由性別提示賦予實體機器人或虛擬 AI 助理特定性別，就可以影響受測者對 HRI 的感受或行為 (Wang, 2014; Eyssel & Hegel, 2012; Siegel, Breazeal, & Norton, 2009; You & Lin, 2019; Niculescu et al., 2010; Zambaka et al., 2006; Crowelly et al., 2009)。由於我國目前較少針對 CUI 或聊天機器人的性別形塑與影響深入討論的研究，因此將借鑑上述的 HRI 性別化的研究成果，理解實體機器人與虛擬角色的性別化手法與它們對使用者互動感受的可能影響。

### 2.3.1 實體機器人性別化對 HRI 之影響

Eyssel 與 Hegel (2012) 讓受測者與功能完全相同，但是頭髮長度不一的機器人互動後進行評價，發現頭髮長度不但會影響受測者對機器人性別的認知，受測者也會將人類傳統社會中性別偏見移轉到他們對機器人的評價上。短頭髮的機器人被認為是更具備男性特質、更適合擔任傳統男性社會角色的任務；反之，長頭髮的機器人則被認為具備女性特質，更適合社會認定的女性角色任務 (圖 1)。Wang (2014) 僅用不同性別代名詞 (He、She) 來描述機器人，就能夠促使受測者對機器人產生符合人際互動中性別刻板印象的互動差異。Siegel, Breazeal 與 Norton (2009) 透過真人預錄的男女性聲音做為性別提示，讓一個中性外型的機器人透過募款實驗來探索機器人性別對其說服力的影響，研究發現僅透過聲音就能夠清楚提示機器人的性別，

且受測者對異性機器人遊說能力的評價與捐款金額都高於同性的機器人。You 與 Lin (2019) 對國人進行了類似的募款實驗，聲音之外、研究發現透過調整機器人頭髮長短與身體顏色兩項性別提示（圖 2）也能有效影響受測者對機器人性別的感受。同時根據不同性別機器人對受測者進行勸募的結果發現，我國受測者傾向給予同性別的機器人較高的信任評價；但在實際捐款行為上，男性受測者則較傾向於捐款給女性機器人。

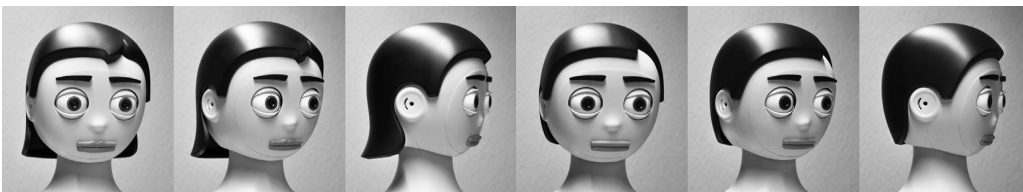


圖 1 | Eyssel 與 Hegel (2012) 實驗中的機器人性別提示，左：長髮；右：短髮。

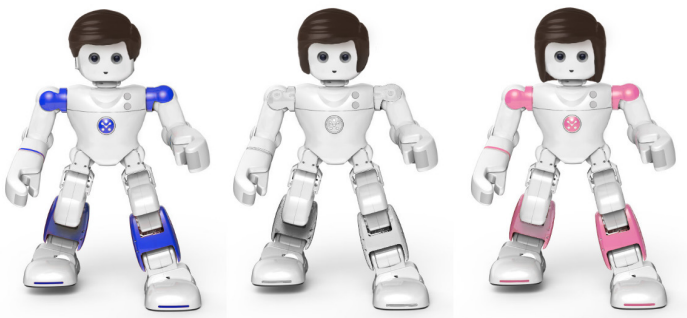


圖 2 | You 與 Lin (2019) 以頭長與顏色作為機器人的性別提示。

### 2.3.2 虛擬機器人與虛擬角色性別化對 HRI 之影響

Niculescu 等人 (2010) 以不同性別提示（男性、中性、女性）的虛擬人物頭像，透過螢幕與受測者進行問答，當使用具有女性性別提示的系統時，受測者普遍感覺到更舒適、更自信，且較不緊張。Zanbaka 等人 (2006) 探討演講者性別與視覺擬人程度高低對演講內容說服力的影響。實驗中以真人錄製的講稿搭配不同性別與不同視覺擬人程度的講者影片（真人錄影、虛擬人物影片、以及非人類虛擬角色影片）的組合做為實驗變項，如表 1 所示。觀看演講影片前，先讓受測者對特定議題進行主觀態度的前測；看過演講影片後，再進行一次相同議題態度的後測，以及對講者的評價。結果顯示，影片中演講者視覺擬人程度高低對其說服力並無明顯的影響，代表虛擬角色具有與真人同樣的說服力。但是性別因素確實會影響受測者對講者的評價。異性演講者的說服力明顯高於同性演講者，換言之，講者與聽眾的性別同時影



響了演講內容說服力。除了人類性別帶來的干擾外，Crowelly 等人 (2009) 曾探討機器人平台（實體機器人 vs. 虛擬 AI 系統）與機器人性別（女性 vs. 男性）兩個因素對使用者感受的影響。研究結果發現在虛擬 AI 系統中，受測者普遍覺得女性聲音的機器人給人較可靠的感受；但當機器人是實體的情況下，受測者反而覺得男性機器人較為可靠。因此機器人平台形式也會左右人們對不同性別機器人的感受。

表 1 | Zanbaka 等人 (2006) 實驗中使用的講者影像圖示

	人類	虛擬人物	虛擬角色
男性	 Jonathan	 Jack	 Tom
女性	 Amy	 Diana	 Ginger

## 2.4 電腦人機互動中的信任

信任 (Trust) 對於理解人際關係和群體行為，管理有效性，經濟交流以及社會或政治穩定性至關重要，但是信任一詞因其使用程度廣泛，各有其脈絡，很難被精確定義 (Hosmer, 1995)。Botsman (2017) 比較上百個信任的定義後，得到的結論是：「信任是對預期結果的評價」，她認為一個對象是否值得信任，主要取決於能幹、可靠性、正直與關懷這四大特質。但是即便是同一個對象，人們會因為委派的任務不同而給予不同的信任評價。本研究是聚焦於電商購物情境中的人機互動信任感，但無論是傳統購物或是電子商務，信任都是商業行為的關鍵要素。而與傳統購物方式不同的是，電子商務主要是仰賴數位科技和服務規劃來建立與維護顧客的信任 (Cofta, 2006)。Gefen (2002) 從多方面審視有關電子商務中的信任定義，最後把信任的概念聚焦於三個量測構面：能力 (Ability)、正直 (Integrity) 與關懷 (Benevolence)。能力是指對於網路商家提供產品的品質以及服務技術和能力的評價指標；正直則是商家遵守對於消費者的規則與承諾的程度；關懷則是商家在合理獲利行為之外，對於消費者仍抱有好意，期望消費者有正向感受。下面就虛擬性別與人機互動信任感受關聯性的相關研究成果做一介紹，以理解不同機器人服務角色設計對使用者信任感受的可能影響。

### 2.4.1 網路媒體信任感的性別偏見

在日常社會的人際互動中，不免存在著對於性別的偏見。這種偏見無意中也影響到人們對網路媒體的信任感。Armstrong 與 McAdams (2009) 曾針對部落格作者 (Blogger) 性別與所發表的文章可信度的關聯性進行調查。透過實驗設計，從三個具性別暗示的署名（男性名字 James Fitzgerald，女性名字 Ann Fitzgerald，中性筆名 Urbanite）任意指派一個到部落格文章，讓受測者閱讀後再對該文進行信任度的評估。結果發現受測者普遍給予男性作家的文章較高的信任評價。Flanagin 與 Metzger (2003) 調查作者性別對個人網頁內容的可信度之影響。研究人員分別以 Jeff Newmar（男子名）和 Julie Newmar（女子名）作為性別提示建立了個人網站，兩個網站中都發佈了輻射對於定期飛行的孕婦造成的潛在傷害訊息，讓受測者瀏覽後再做信任度的調查。實驗結果顯示不只是作者性別，讀者性別也會對網站可信度產生影響。男性讀者對於網頁可信度的評分顯著高於女性讀者；此外，相較於同性的網站作者，讀者傾向給予異性的網站作者較高的可信度評比。

### 2.4.2 實體機器人之信任研究

Ghazali, Ham, Barakova 與 Markopoulos (2018) 利用一個可以展示不同臉部特徵的實體機器人進行實驗，來解析機器人性別和臉部特徵對人們的信任感與說服力的影響。實驗中機器人擔任遊戲關主，跟受測者遊戲的過程中，會不時傳達訊息企圖去影響受測者。根據實驗中受測者的行為記錄以及受測者自我陳述的信任感受與心理抗拒量表的統計分析，研究發現眉毛和嘴唇上揚的機器人顧問（這也是人際互動中更受信任的人類特徵）更具說服力，喚起更多信任，產生的心理抗拒反應也更小。機器人的性別不影響信任感受，但是受測者在與異性機器人互動時會感受到較高的心理抗拒反應。因此 Ghazali 等人建議機器人的面部設計應該參考較具信任感的真人臉部特徵，並且可以客製化成與目標使用者相同的性別。

Siegel 等人 (2009) 透過一個募款實驗來探討受測者性別與機器人性別對 HRI 中的信任感受與捐款行為的影響。受測者對於機器人的信任感受是透過：信任、信譽與參與度三項構面的量表來取得。實驗同時記錄機器人獲得捐款次數與總金額做為信任行為的展現。結果發現受測者在量表上，普遍呈現異性偏好 (Cross-sex Preference)，對於異性機器人有較可信、較可靠，較具吸引力的評價；實際觀察受測者的捐款行為，同樣發現男性受測者傾向捐款給異性機器人，但機器人性別對女

性受測者的捐款行為則無顯著影響。該研究建議 HRI 的性別議題較為複雜，不應只就機器人性別做考量，也同時應納入使用者的性別於互動研究之中。

You 與 Lin (2019) 效法上述 Siegel 等人的機器人信任感實驗方法，以我國大學校園內研究室導覽機器人為例，進行了類似的性別化機器人的募款實驗。我國研究結果同樣發現僅考量機器人性別的情況下，機器人的性別並不會對受測者的信任感產生顯著影響；反而是機器人與受測者性別的相同與否才會對人機信任產生顯著影響。同樣地，男性受測者傾向捐款給女機器人，女性受測者則無顯著差異。但有別於 Siegel 等人的發現，我國受測者傾向給同性機器人較高的信任評價，與 Ghazali 等人 (2018) 的研究結果一致，發現實體機器人信任感的同性偏好 (Same-sex Preference)。

#### 2.4.3 虛擬機器人之信任研究

人們傾向把科技擬人化，且較願意相信具擬人特性的電腦系統 (Nass & Yen, 2010)。Waytz, Heafner 與 Epley (2014) 透過實驗設計來檢驗讓自駕車的電腦系統具備擬人特徵與否，是否會影響乘客對自駕車的信任。將 100 位受試者分散於三種實驗條件：(1) 一般汽車（控制組）、(2) 沒有擬人特徵的自駕車、(3) 被稱作 Iris（女性名稱）的自駕車，功能與沒有擬人特徵的自駕車完全相同，但是駕駛途中會撥放預錄的女性聲音。實驗結果發現受測者對具備擬人特性的自駕車系統最有信任感；同時研究人員在實驗中故意安排了車禍事故，搭乘自駕車的組別相較於控制組都傾向將事故歸罪於車子，但是搭乘 Iris 的受測者對自駕車的究責較少。這個實驗顯示僅僅賦予電腦系統名字，與撥放女性聲音就能影響人們的信任。

McDonnell 與 Baxter (2019) 利用具性別提示的名稱（男性、中性、女性）來賦予聊天機器人性別角色，並搭配銀行（性別角色中立）或機械（偏向男性角色）兩個不同專業領域的對話腳本組成六種條件，讓受測者與其互動並填寫滿意度和性別刻板印象的量表。研究發現聊天機器人性別確實會影響受測者的滿意度與處理特定專業領域議題的偏好。當聊天機器人提供的服務是屬於具性別投射的專業領域，且聊天機器人的性別設定不符合社會預期時，受測者較傾向採用性別刻板印象來貶低聊天機器人的能力，對聊天機器人的滿意度也降低。所以除了聊天機器人性別之外，賦予聊天機器人的任務角色也會影響人們對它信任與滿意。這一結論也呼應前述 Botsman (2017) 的主張，信任的評定與所委派的任務有關。

### 三 研究方法

本研究主要參考 Zanbaka 等人 (2006) 與 Siegel 等人 (2009) 針對機器人的性別提示在人機互動中對使用者信任感之影響的研究架構與實驗設計，規劃出兩階段實驗分別探索：(1) 設計師能否透過頭像視覺設計的提示來塑造聊天機器人性別？與 (2) 購物情境中聊天機器人的虛擬性別是否會影響使用者對它的信任感受？首先藉由前導研究來檢視聊天機器人的視覺元素對使用者性別感受之影響。根據其結果再進行正式實驗，在相同的對話腳本下，僅操作聊天機器人的性別提示以產生不同性別的聊天機器人，招募受測者透過組間設計的網路購物模擬實驗，來理解人們是否會因為購物聊天機器人的性別不同而影響網路購物時的信任感受。

#### 3.1 前導研究

本階段實驗是透過感性工學方法探討聊天機器人頭像的設計元素聊天機器人給人的性別感受之影響？並藉此為後續正式實驗找出有效塑造聊天機器人性別的頭像性別提示。根據國內外相關文獻的中曾運用過的性別提示，本研究挑選出：頭髮長度 (Eyssel & Hegel, 2012)、眉毛粗細 (Ghazali et al., 2018)、顏色 (You & Lin, 2019) 三個頭像設計項目作為聊天機器人頭像性別化設計的自變項，受測者對聊天機器人頭像的性別評價為依變項，整體實驗之研究架構如圖 3 所示。

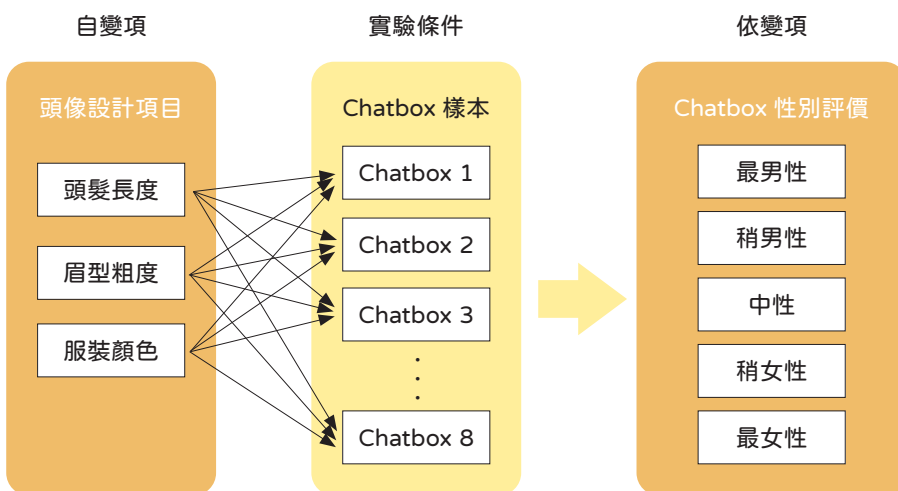





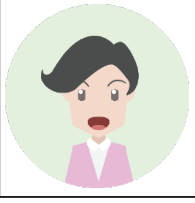




圖 3 | 前導研究之實驗架構

自變項的設計項目下各有兩個類目：目頭髮長度（長／短）、眉毛粗細（粗／細）、服裝顏色（粉紅／天藍），可組出如表 2 的八個聊天機器人頭像。依變項的性別感受則是讓受測者透過五點李克特量表 (Likert scale) 進行評估，1 點為最男性，5 點為最女性，3 點則為中性或無意見。整體實驗是利用線上問卷形式發放於社群平台。透過滾雪球抽樣方式匿名招募受測者。每位受測者都被要求逐一觀看表 2 所示的八個聊天機器人頭像，並立即用五點李克特量表對每一個聊天機器人進行性別感受評量。根據回收的問卷數據的統計分析，可以確認每一頭像設計元素能否做為聊天機器人的性別提示，並得到最具男性及女性感受極端值的聊天機器人頭像設計，做為下一步聊天機器人性別對使用者信任感受影響的研究（正式實驗）中有效操控聊天機器人性別之提示手段。

表 2 | 作為前導研究中做為實驗樣本的 8 個聊天機器人頭像

頭像設計元素	短髮		長髮	
	天藍衣服	粉紅衣服	天藍衣服	粉紅衣服
粗眉				
細眉				

### 3.2 正式實驗

正式實驗將透過聊天機器人性別提示的操作，檢視在網路購物情境下的聊天機器人性別對使用者信任與互動感受的影響。本次研究所設定情境為「流行服飾線上購物情境」，以 Facebook Messenger 作為實驗平台，實驗所使用之網路購物聊天機器人是透過 Chatisfy 聊天機器人協作平台的開發工具來建置，可根據預設腳本自動回覆受測者，來模擬人們使用 Messenger 與聊天機器人進行線上購物的互動情境。實驗任務則是要求受測者透過聊天機器人的服務，完成一項商品購物，並填寫量表。

實驗自變項為聊天機器人的性別，分「男性聊天機器人」與「女性聊天機器人」兩類。實驗中聊天機器人性別是透過其常用的兩種性別提示，「頭像」與「命名」，來進行塑造。頭像是取自前導實驗中，被評定最男性與最女性的頭像設計組合。性別提示中增加命名的主因為符合實驗設定的線上購物情境，聊天機器人需要先自我介紹，故參考 Armstrong 與 McAdams (2009) 之作法，從國人常見名字中，挑選男性化的「聖博」、中性化的「雨帆」、與女性化的「敏慧」。實驗依變項是受測者在完成聊天機器人購物體驗後，對聊天機器人的信任與互動感受的評價，本研究參考 Gefen (2002) 有關電子商務中信任的定義與 Crowelly 等人 (2009) 機器人性別對受測者感受度的研究方法，選定「能力」、「正直」與「關懷」與「互動信賴度」四個構面問項並製作信任感受評價量表，並首先邀請 30 名受測者進行量表的前測，以瞭解量表之信效度。正式實驗之研究架構如圖 4 所示。

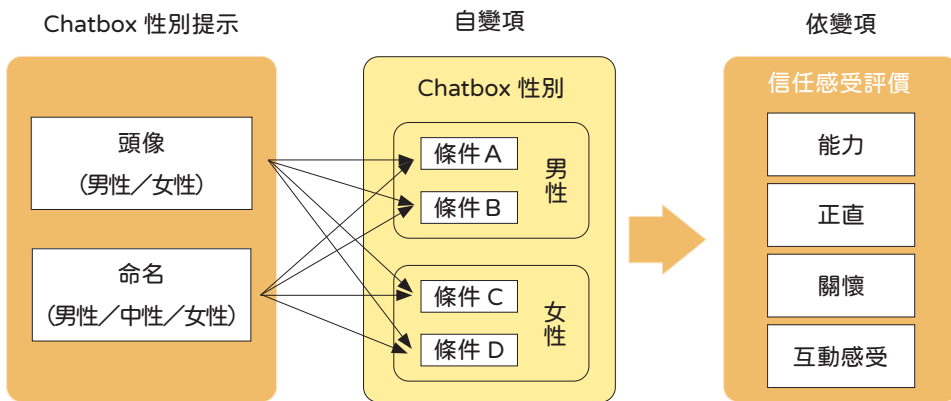


圖 4 | 正式實驗之研究架構

表 3 | 正式實驗中不同聊天機器人性別提示組合下的四種實驗條件

性別提示	男性頭像	女性頭像
男性名字 (聖博)	條件 A (男性聊天機器人)	
中性名字 (雨帆)	條件 B (男性聊天機器人)	條件 C (女性聊天機器人)
女性名字 (敏惠)		條件 D (女性聊天機器人)

### 3.2.1 實驗聊天機器人設計

正式實驗中控制聊天機器人性別的性別提示有頭像（男性頭像、女性頭像）與命名（聖博、兩帆、敏慧），共可以組合出六種聊天機器人條件。但是男性頭像搭配女性名字「敏惠」，與女性頭像搭配男性名字「聖博」的兩種組合，容易造成聊天機器人判別的困難，予以排除。因此產生如表 3 所示的四種實驗條件：(A) 男性頭像 / 男性名字；(B) 男性頭像 / 中性名字；(C) 女性頭像 / 女性名字；(D) 女性頭像 / 中性名字。條件 A 與條件 B 屬於男性聊天機器人，條件 C 與條件 D 屬於女性聊天機器人。

四種實驗條件除了性別提示不同之外，其他功能與對話內容完全一致。本研究參考市面上現有購物型聊天機器人對話內容與常用流程，規劃一套通用的購物腳本。腳本內容涵蓋：歡迎訊息、通用回應（8 個選項）、禮貌問候、無法準確回應（5 個選項）、互動提示、功能說明、推薦商品（3 個選項）、回應商品問題（6 個選項）、協助提示（2 個選項）、詢問滿意度（2 個選項）、服務內容（10 個選項）、友善提醒、確認購物完成、結束聊天機器人體驗…等互動情境。對於受測者的實驗設定上，則要求受測者需有使用 Messenger 的經驗，並對於網路購物具有一定熟悉程度（包含：購物流程、操作使用），且曾獨立完成網路購物者。本實驗之條件 C 之網路購物聊天機器人與受測者互動過程螢幕擷圖如圖 5 所示。

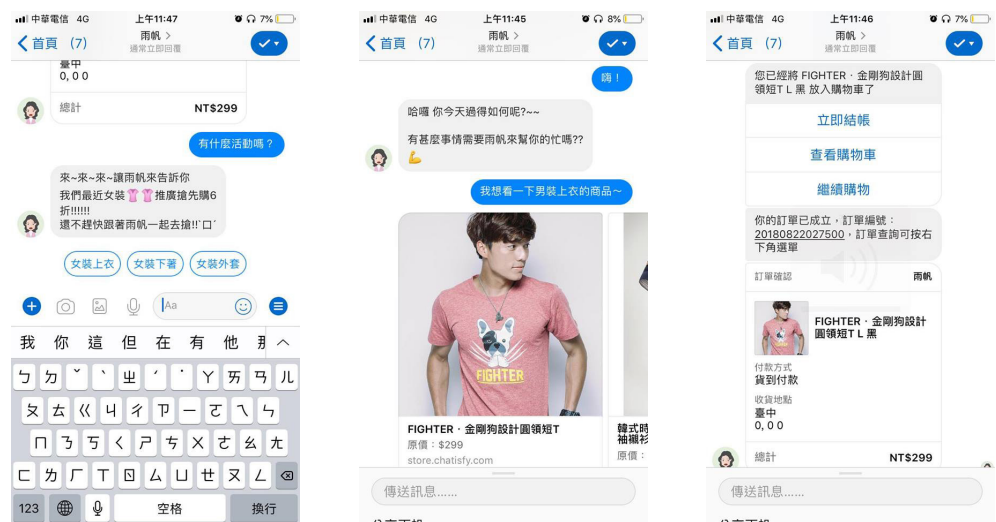


圖 5 | 正式實驗之聊天機器人（條件 C）與受測者互動過程螢幕擷圖

### 3.2.2 信任感受量表設計

本實驗的依變項是受測者在與具性別提示的聊天機器人互動後，對聊天機器人的信任與互動感受的評價，參考 Gefen (2002) 與 Crowelly 等人 (2009) 評量機器人給人的互動感受的量表構面與題項，根據本次實驗的情境任務與目的，設計了一個包含「能力」（10 題）、「正直」（8 題）、「關懷」（9 題），與「互動感受」（8 題）四個構面的信任感受量表，讓受測者透過五點李克特量表進行評估（1 點為非常不同意，5 點為非常同意）。因為問項較多，表 4 僅摘錄四個構面量表的部分問項。

表 4 | 正式實驗信任感受量表問項摘錄

構面	問項	來源
能力	我認為聊天助理提供的資訊是可靠的	Gefen (2002)
	我認為聊天助理提供的資訊是我所需要的	
	我認為聊天助理提供的建議反而讓我找不到我想要的商品	
	我認為聊天助理可以幫助我更快找到我要的商品	
正直	聊天助理給我的承諾是值得依賴的	
	當我做重要決定時，聊天助理會顧及我的權益	
	對於聊天助理的誠信我不會去懷疑	
	我會期待聊天助理給的建議是最公正的判斷	
關懷	我會期望聊天助理準備好並願意協助和支持我的	
	聊天助理會真誠地關心我的交易內容是否成功	
	我會期望聊天助理對我有正向的意圖	
互動感受	我覺得聊天助理是可靠的	Crowelly 等人 (2009)
	我覺得聊天助理是具有能力的	
	我覺得聊天助理是值得信任的	

正式實驗也是利用線上問卷形式發放於社群平台。透過滾雪球抽樣方式匿名招募受測者。因為正式實驗採受測者組間設計，將依受測者參與實驗的順序，分配於四種實驗條件中。每位受測者僅與表 3 其中一個實驗條件之機器人互動，進行網路購物模擬。當受測者瀏覽完機器人介紹的商品或是覺得已經體驗完成，決定中止與聊天機器人的互動，聊天機器人會引導受測者到一個線上問卷形式的「信任感受量表」，要求針對剛剛的互動體驗進行評量。

## 3.3 資料分析

在性別提示的前導實驗中，透過三項性別提示的自變項中所得出的八個聊天機器人頭像，首先透過單因子變異數分析 (One-way ANOVA)，了解三項性別提示是否可使受測者對聊天機器人的性別感知上具有顯著差異。並在性別提示所交互組合產生的八組頭像中透過 Duncan 事後檢定的方式，選出兩項分別最具代表的男性及女性的頭像。在正式實驗分析中，首先邀請 30 名受測者進行前測問卷，透過 Kaiser-



Meyer-Olkin (KMO) 及 Bartlett 球形檢定對該量表的四個構面內問項可否使用因素分析進行判斷，針對此次實驗，對於不同性別聊天機器人與不同性別之受測者，是使用獨立 T 檢定的方式進行分析，另外在受測者的「年齡」上，則分為 5 項組間進行變異數分析，另外比較其他因素對於聊天機器人性別之研究分析。

## 四 研究結果

### 4.1 前導實驗結果分析

透過回收的 134 份前導實驗問卷單因子變異數分析中發現，顯著性  $p$  值均小於 .001，代表它們對性別感知是具有顯著影響的。分析結果三者的 F 檢定值依序為 562.005，69.366，84.488，顯示這三個因子對性別感知造成的差異：頭髮長度 > 服裝顏色 > 眉毛粗細。再根據問卷數據的 Duncan 事後檢定的結果（表 5），八個實驗條件分佈在六個子集。可知短髮、粗眉、天藍色服裝的聊天機器人頭像組合，最給人男性的感受；而長髮、細眉、粉紅色服裝的頭像組合則最為女性的感受。前導實驗的結果，確認了髮長、顏色、眉型這些符合人類社會對性別典型的視覺特徵，對聊天機器人而言仍是有效的性別提示。接著，本研究選用前導實驗中性別感受差距最大的兩個頭像樣本作為後續正式實驗之機器人的性別視覺提示。

表 5 | 八個實驗條件 Duncan 事後檢定表

實驗條件	個數	子集					
		1	2	3	4	5	6
短髮_粗眉_天藍	134	1.19					
短髮_細眉_天藍	134		1.48				
短髮_粗眉_粉紅	134			2.55			
短髮_細眉_粉紅	134			2.63			
長髮_粗眉_天藍	134				3.02		
長髮_粗眉_粉紅	134				3.22		
長髮_細眉_天藍	134					4.15	
長髮_細眉_粉紅	134						4.83
顯著性		1.000	1.000	0.532	0.125	1.000	1.000

### 4.2 正式實驗結果分析

正式實驗採受測者組間設計，共招募 120 位受測者，平均分配於四種實驗條件中。120 位受測者男女各半，年齡分佈為「20-29 歲」57 人 (47.5%)，「30-39 歲」24 人 (20.0%)，「40-49 歲」10 人 (8.3%)，「50 歲以上」29 人 (24.2%)。最高學歷

以「大學」63人最多(52.5%)，「高中」15人(12.5%)，「研究所」42人(35.0%)。曾經使用過聊天機器人購物的有31人(25.8%)，沒有經驗的有89人(74.2%)。

#### 4.2.1 量表信度檢定

在問卷信效度檢測中「能力」構面的 Kaiser-Meyer-Olkin (KMO) 檢定為 0.828、Bartlett 球形檢定顯著性小於 .001；「正直」構面的 KMO 檢定為 0.752、Bartlett 球形檢定顯著性小於 .001；「關懷」構面的 KMO 檢定為 0.606、Bartlett 球形檢定顯著性為 .003；最後一個「互動感受」構面的 KMO 檢定為 0.878、Bartlett 球形檢定顯著性小於 .001，均符合可以進行因素分析之水準。再透過因素分析將各構面的問項進行成分分類，並剔除僅有一個問項之分類。留下的量表構面與問項為：「能力」構面有 9 個問項、「正直」構面有 2 個成分 7 個問項、「關懷」構面有 3 個成分 7 個問項、「互動信賴」構面 5 個問項。

再就這 28 個問項以五點李克特量表收集的數據進行量表的信度分析，得到量表構面與成分因素之 Cronbach's  $\alpha$  值 (表 6)。除了「關懷」構面下的「正向意圖」與「關切服務」成份的 Cronbach's  $\alpha$  值未達尚可接受之 0.65 水準，予以排除；其餘構面與成分之 Cronbach's  $\alpha$  值問項皆大於 0.7，甚至「能力」構面  $\alpha$  值還高達 0.957，表示剩下的量表構面均具有高信度，即可檢視性別因素對聊天機器人信任感受的影響。

表 6 | 量表構面成分因素之 Cronbach's  $\alpha$  值

構面	成分因素	$\alpha$ 值	問項個數
能力 (Ability)		0.957 *	9
正直 (Integrity)	誠信度	0.855 *	4
	公正性	0.789 *	3
關懷 (Benevolence)	友善	0.783 *	3
	正向意圖	0.380	2
	關切服務	0.278	2
互動感受 (Engagement)		0.878 *	5

(\* 表示 Cronbach's  $\alpha$  值 > 0.7)

#### 4.2.2 聊天機器人性別對於受測者信任感受分析

有關聊天機器人性別的信任感受在獨立 t 檢定結果如表 7 所示，在信任感受的 28 個問項中，多數的問項的顯著性都高於 .05，代表網路購物聊天機器人的性別對

受測者信任感受普遍沒有顯著影響。其中顯著性  $p$  值小於 .05 的僅有兩題，顯示聊天機器人性別確實有顯著影響存在的是「能力」構面中有關資訊提供的問項「Ab02. 我認為聊天助理提供的資訊是我所需要的」，男性聊天機器人得分 (3.87) 高於女性聊天機器人平均數 (3.55)；以及「關懷」構面中有關友善用意的問項「Be05. 我會期望聊天助理的用意是友善的」，女性聊天機器人得分 (4.37) 高於男性聊天機器人平均數 (4.10)。

表 7 | 聊天機器性別對於受測者信任感受之獨立 t 檢定結果

構面	成分因素	問項	男性聊天機器人		女性聊天機器人		t 值	p 值
			平均數	標準差	平均數	標準差		
能力		Ab01	3.65	0.755	3.55	0.769	-0.719	0.474
		Ab02	3.87	0.791	3.55	0.946	-1.988	0.049 *
		Ab04	3.73	0.800	3.72	0.922	-0.106	0.916
		Ab05	3.37	0.823	3.37	0.938	0.000	1.000
		Ab06	3.43	0.745	3.52	0.854	0.570	0.570
		Ab07	3.87	0.812	3.80	0.898	-0.426	0.671
		Ab08	3.67	0.816	3.72	0.739	0.352	0.726
		Ab09	3.72	0.761	3.62	0.940	-0.640	0.523
		Ab10	3.53	0.929	3.68	0.873	0.911	0.364
		正直	誠信度	In06	2.93	0.936	2.80	1.086
In07	3.07			0.918	2.92	0.889	-0.909	0.365
In05	3.95			0.811	3.97	0.736	0.118	0.906
公正性	In03		3.22	0.825	3.03	0.882	0.242	0.242
	In02		3.57	0.789	3.35	0.777	-1.515	0.132
	In01		3.53	0.747	3.45	0.746	-0.611	0.542
關懷		In04	3.73	0.918	3.75	0.914	0.100	0.921
		Be05	4.10	0.730	4.37	0.637	2.133	0.035 *
		Be09	3.70	0.788	3.73	0.710	0.244	0.808
互動感受		Be06	4.10	0.775	3.97	0.843	-0.902	0.369
		En01	3.57	0.722	3.53	0.833	-0.234	0.815
		En07	3.67	0.774	3.60	0.785	-0.468	0.640
		En04	3.47	0.833	3.47	0.650	0.000	1.000
		En05	3.80	0.798	3.63	0.802	-1.141	0.256
		En03	3.82	0.748	3.77	0.810	-0.351	0.726

(\* 表示獨立 t 檢定  $p < .05$ )

#### 4.2.3 受測者性別對聊天機器人給人的信任感受分析

在不同性別受測者對聊天機器人的獨立 t 檢定的結果如表 8 所示，在四個信任感受構面中，僅有「正直」構面下的四個問項之  $p$  值低於 .05，代表受測者性別對聊天機器人信任感受僅在「正直」構面有顯著的影響，對「能力」、「關懷」和「互動感受」三個構面的影響則未達顯著性水準。

表 8 | 受測者性別對於聊天機器人信任感受之獨立 t 檢定結果

構面	成分因素	問項	男性聊天機器人		女性聊天機器人		t 值	p 值
			平均數	標準差	平均數	標準差		
能力		Ab01	3.68	0.748	3.52	0.770	1.203	0.231
		Ab02	3.82	0.833	3.60	0.924	1.349	0.217
		Ab04	3.80	0.819	3.65	0.899	0.955	0.341
		Ab05	3.47	0.929	3.27	0.821	1.250	0.200
		Ab06	3.55	0.811	3.40	0.785	1.029	0.150
		Ab07	3.92	0.907	3.75	0.795	1.070	0.287
		Ab08	3.75	0.836	3.63	0.712	0.823	0.412
		Ab09	3.73	0.880	3.60	0.827	0.855	0.394
		Ab10	3.77	0.927	3.45	0.852	1.948	0.054
		正直	誠信度	In06	3.10	0.969	2.63	1.008
In07	3.20			0.860	2.78	0.904	2.588	0.011 *
In05	4.08			0.636	3.83	0.827	1.792	0.076
公正性	In03		3.28	0.739	2.97	0.938	2.054	0.042 *
	In02		3.60	0.785	3.32	0.770	1.995	0.048 *
	In01		3.60	0.694	3.38	0.783	1.604	0.111
關懷		In04	3.78	0.825	3.70	0.997	0.499	0.619
		Be05	4.22	0.715	4.25	0.680	-0.262	0.794
		Be09	3.65	0.709	3.78	0.783	-0.978	0.330
互動感受		Be06	3.97	0.823	4.10	0.796	-0.902	0.369
		En01	3.68	0.770	3.42	0.766	1.902	0.060
		En07	3.77	0.851	3.50	0.676	1.900	0.060
		En04	3.48	0.676	3.45	0.811	0.244	0.807
		En05	3.82	0.770	3.62	0.825	1.373	0.173
		En03	3.85	0.820	3.73	0.733	0.822	0.413

(\* 表示獨立 t 檢定  $p < .05$ )

受測者性別會產生顯著差異的四個問項，有「誠信度」成分下的「In06 我會很放心地將我的重要資料提供給聊天助理（例：電話、住址）」、「In07 我相信聊天助理不會將重要資訊外洩給其他人」、「In03 對於聊天助理的誠信我不會去懷疑」，以及「公正性」成分的「In02 當我做重要決定時，聊天助理會顧及我的權益」。男性受測者在這四個問項的評價依序是 3.10、3.20、3.28 與 3.60，都持正面的肯定態度，與女性受測者偏向負面的評價（依序 2.63、2.78、2.97 與 3.32）相較，顯示不同性別受測者對聊天機器人的信任程度有差異，男性受測者對聊天機器人的「正直」感受程度明顯高於女性受測者，較願意將資料提供給聊天機器人並聽從其意見。

#### 4.2.4 受測者性別與機器人性別的交叉比較

本研究進一步就受測者與機器人的性別進行交叉比較，探究兩者性別對聊天機器人信任感受的交互作用。同樣採用獨立 t 檢定，分別就：(1) 聊天機器人性別對女性受測者信任感受的影響；(2) 聊天機器人性別對男性受測者信任感受的影響；(3) 受測者性別對女性聊天機器人信任感受的影響；(4) 受測者性別對男性聊天機器人信任感受的影響四種情況進行檢視，再對各個結果進行比較。

## (1) 聊天機器人性別對女性受測者信任感受的影響

聊天機器人性別對女性受測者信任感受的 t 檢定分析結果節錄如表 9。28 個問項中除了「能力」構面中「Ab02. 我認為聊天助理提供的資訊是我所需要的」的  $p$  值等於 .05，勉強達顯著水準；其餘構面問項  $p$  值均高於 .05 無顯著差異。唯一有顯著差異的問項「Ab02. 我認為聊天助理提供的資訊是我所需要的」，女性受測者給男機器人的評價 (3.83) 略高於女性機器人 (3.37)，代表女性受測者則認為異性機器人在提供資訊的能力上是略優於同性機器人的；其他問項則無顯著差異。

表 9 | 聊天機器性別對於女性受測者信任感受之獨立 t 檢定結果 (節錄能力構面分析)

構面	成分因素	問項	男性聊天機器人		女性聊天機器人		t 值	p 值
			平均數	標準差	平均數	標準差		
能力		Ab01	3.50	0.861	3.53	0.681	0.166	0.869
		Ab02	3.83	0.913	3.37	0.890	-2.005	0.050 *
		Ab04	3.73	0.785	3.57	1.006	-0.715	0.477
		Ab05	3.33	0.802	3.20	0.847	-0.626	0.534
		Ab06	3.43	0.774	3.37	0.809	-0.326	0.745
		Ab07	3.83	0.791	3.67	0.802	-0.810	0.421
		Ab08	3.60	0.814	3.67	0.606	0.360	0.720
		Ab09	3.73	0.828	3.47	0.819	0.804	0.245
		Ab10	3.43	1.040	3.47	0.629	0.150	0.881

(\* 表示獨立 t 檢定  $p \leq .05$ )

## (2) 聊天機器人性別對男性受測者信任感受的影響

聊天機器人性別對於男性受測者信任感受的 t 檢定分析結果節錄如表 10。除了「關懷」構面中 Be05 的  $p$  值小於 .05 達顯著水準；其餘構面問項  $p$  值均高於 .05。代表機器人性別對男性受測者唯一有顯著差異的是「Be05. 我會期望聊天助理的用意是友善的」，男性受測者較認為異性機器人的友善程度 (4.40) 是略優於同性機器人 (4.03)；其他問項則無顯著差異。

表 10 | 聊天機器性別對於男性受測者信任感受之獨立 t 檢定結果 (節錄關懷構面分析)

構面	成分因素	問項	男性聊天機器人		女性聊天機器人		t 值	p 值
			平均數	標準差	平均數	標準差		
關懷		Be05	4.03	0.718	4.40	0.675	2.038	0.046 *
		Be09	3.53	0.776	3.77	0.626	1.282	0.205
		Be06	4.03	0.765	3.90	0.885	-0.624	0.535

(\* 表示獨立 t 檢定  $p < .05$ )

### (3) 受測者性別對女性聊天機器人信任感受的影響

受測者性別對女性聊天機器人信任感受的 t 檢定分析結果節錄如表 11。僅有「正直」構面下的三個問項之  $p$  值低於 .05，代表受測者性別對女性機器人信任感受僅在「正直」構面有顯著的影響，對其餘構面則無顯著差異。三個有顯著差異的問項均屬「正直」下的「誠信度」成分，「In06 我會很放心地將我的重要資料提供給聊天助理（ex：電話、住址）」、「In07 我相信聊天助理不會將重要資訊外洩給其他人」、「In03 對於聊天助理的誠信我不會去懷疑」。男性受測者在這三個問項的評價依序是 3.10、3.17 與 3.27，對女性機器人都持正面態度，與女性受測者持負面態度的評價（依序 2.50、2.67 與 2.8）相較，顯示不同性別受測者對女性機器人的信任程度有差異，男性受測者對女性機器人的「正直」感受程度明顯高於女性受測者。

表 11 | 受測者性別對於女性聊天機器人信任感受之獨立 t 檢定結果（節錄正直構面分析）

構面	成分因素	問項	男性聊天機器人		女性聊天機器人		t 值	p 值
			平均數	標準差	平均數	標準差		
正直	誠信度	In06	3.10	1.094	2.50	1.009	2.209	0.031 *
		In07	3.17	0.913	2.67	0.802	2.253	0.028 *
		In05	4.13	0.730	3.80	0.714	1.787	0.079
	公正性	In03	3.27	0.828	2.80	0.887	2.107	0.039 *
		In02	3.47	0.900	3.23	0.626	1.166	0.248
		In01	3.57	0.774	3.33	0.711	1.216	0.229
		In04	3.90	0.803	3.60	1.003	1.279	0.206

(\* 表示獨立 t 檢定  $p < .05$ )

### (4) 受測者性別對男性聊天機器人信任感受的影響

受測者性別對男性聊天機器人信任感受的 t 檢定分析結果顯示，所有構面問項之  $p$  值均大於 .05，代表受測者性別對男性聊天機器人所有構面的信任感受均無顯著差異（構面問項均未達顯著水準，固略過 t 檢定結果）。

根據上述四種受測者與機器人性別因素交叉比對的分析，結果歸納如下：

1. 聊天機器人的性別對受測者（無論男性或女性）的信任感受影響不大，僅有少數問項（1/28，約 3.6%）具統計上的差異，受測者對異性機器人有稍高的評價。
2. 受測者的性別對女性機器人的信任感受有明顯的差異，男性受測者對女性機器人的「正直」感受抱持正面態度，但是女性受測者則持負面態度。但受測者的性別對男性機器人的信任評價沒有任何影響。

### 4.3 檢驗其他因素 (受測者年齡) 對機器人信任感受的影響

由於 4.2 的各項分析結果顯示受測者本身的性別對聊天機器人信任感受的影響大於聊天機器人性別，本研究不禁思考受測者其他特性是否也會對聊天機器人信任感受產生影響。因此就本次實驗的受測者年齡族群分類：20 歲族群（20-29 歲），30 歲族群（30-39 歲），40 歲族群（40-49 歲）與 50 歲族群（50 歲以上），共四類進行單因子變異數分析 (One-way ANOVA)，來檢驗不同年齡受測者對於聊天機器人信任感受是否存在顯著差異，來做為聊天機器人性別對信任感受影響程度的參考。

受測者年齡族群對於聊天機器人信任感受之單因子變異數分析結果如表 12 所示。顯著性  $p$  值低於 0.05 的問項有 Ab01、Ab02、In07、In05、In03、In01、Be05、Be09、En01 共九個，分布於四個信任感受構面。代表受測者年齡對聊天機器人信任感受的影響存在顯著差異（圖 5）。由於受測者年齡對其機器人信任感受之影響並非本研究探究的議題，故不深入檢視各信賴構面中不同年齡族群之間的差異變化。

表 12 | 受測者年齡族群對於聊天機器人信任感受之變異數分析

構面	成分因素	問項	20-29 歲	30-39 歲	40-49 歲	50 歲 ~	$p$ 值
能力		Ab01	3.67	3.96	3.10	3.34	0.003 *
		Ab02	3.84	3.92	3.40	3.38	0.046 *
		Ab04	3.68	4.04	3.40	3.66	0.171
		Ab05	3.37	3.38	3.30	3.38	0.996
		Ab06	3.39	3.71	3.20	3.55	0.241
		Ab07	3.88	4.08	3.30	3.72	0.084
		Ab08	3.70	3.88	3.40	3.41	0.481
		Ab09	3.75	3.88	3.40	3.41	0.135
		Ab10	3.63	3.83	3.40	3.45	0.398
		正直	誠信度	In06	2.98	3.00	2.50
In07	3.12			3.25	2.50	2.69	0.023 *
In05	4.19			3.96	4.00	3.48	0.001 *
公正性	In03		3.32	3.04	2.50	3.03	0.031 *
	In02		3.58	3.54	3.10	3.28	0.155
	In01		3.67	3.71	2.90	3.17	0.000 *
關懷		In04	3.70	4.08	3.80	3.52	0.152
		Be05	4.33	4.63	4.00	3.79	0.000 *
		Be09	3.70	4.00	3.20	3.69	0.038 *
互動感受		Be06	4.02	4.29	3.90	3.90	0.314
		En01	3.65	3.83	3.20	3.24	0.012 *
		En07	3.72	3.79	3.40	3.41	0.179
		En04	3.58	3.50	3.20	3.31	0.273
		En05	3.75	3.88	3.60	3.55	0.479
		En03	3.81	3.96	3.70	3.66	0.545

(\* 表示顯著性  $p < .05$ )

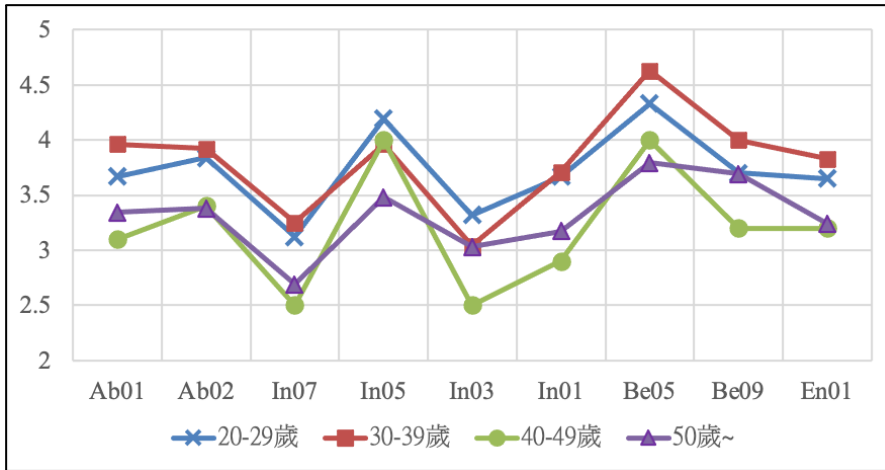


圖 5 | 不同年齡族群受測者對聊天機器人信任構面問項評價折線圖

相較於前面聊天機器人性別對於受測者信任感受分析（4.2.2 節）與受測者性別對聊天機器人信任感受分析（4.2.3 節），受測者年齡族群對聊天機器人信任感受的影響層面遍及能力、正直、關懷、互動感受四個構面（九個問項），顯示出年齡因素相較與性別因素對聊天機器人的信任感受影響更大，範疇也更廣。

## 五 結論

隨著 AI 技術日益成熟與普及，使用者對 AI 的需求從自動化與資訊計算功能，漸漸擴展到擬人化與情感連結，進入「情感運算」的階段。AI 科技在人類日常生活中同時扮演著工具、媒體、社會受測者的角色，使用者也逐漸傾向「人際互動」般地來與 AI 互動期望達到溝通、娛樂、情感支持等目的 (Yonck, 2017)。而透過賦予性別來為它增添一個鮮明形象或個性，不但讓機器人服務更貼近人性，也讓使用經驗更豐富有趣 (Shevat, 2017)。但是不少學者提出人類社會中的性別刻板印象也會隨著機器人的性別化擴散到人類與機器人的互動中 (Wang & Young, 2014; Lee et al., 2000)，促使 AI 性別隱含的偏見問題受到關注，並提升大眾對 AI 性別化設計的敏感度。本研究以近年盛行的購物 Chatobot 為例，檢視出現於社群平台中的 Chatobot 性別化手法的作用，與國人是否會因購物 Chatobot 的性別而產生信任感受的變化。最後歸納本研究的結論與對未來聊天機器人性別化設計的建議如下。



## 5.1 研究結果與討論

### (1) 頭像設計作為「性別提示」確實可影響聊天機器人給人的性別感受

前導實驗結果顯示，即便聊天機器人頭像在社群平台對話視窗中尺寸僅有 100 像素 x 100 像素，受測者確實可根據頭像中的性別提示感知其性別。頭像設計元素對聊天機器人性別感受的影響為：頭髮長度 > 服裝顏色 > 眉毛粗細。短髮、粗眉、天藍色服裝的頭像組合最能給人男性的感受；而長髮、細眉、粉紅色服裝則最為女性。這與國內外針對機器人的性別化研究結論相符 (Nass & Yen, 2010; Eyssel & Hegel, 2012; Wang & Young, 2014; Ghazali et al., 2018; You & Lin, 2019)，代表這種充分符合人類社會對性別特徵典型的期待的性別化手法可以精準傳達出聊天機器人的性別，且這種性別感受是全球化的。本研究前導實驗僅就頭像設計元素進行討論，但是對話互動介面中能做為聊天機器人性別提示的元素眾多，例如：姓名、指稱的代名詞、對話內容、修辭…等，期待未來研究人員的繼續探索。

### (2) 日常購物聊天機器人性別對使用者的信任感影響不大

本次針對線上服飾購物聊天機器人的信任感調查，實驗結果顯示受測者對不同性別聊天機器人的信任感受並無明顯差異，僅在資訊提供能力上男性聊天機器人被認為是略優於女機器人，與意圖友善上女性機器人略優於男性機器人，但是差異極小；合理的推斷是，國人對於電子商務中行銷日常用品的聊天機器人信任感沒有明顯的性別偏見存在。這樣的結果與國外的擬人化電腦系統信任調查，發現明顯存在性別偏見有別。例如：Armstrong 與 McAdams (2009) 與 Flanagin 與 Metzger (2003) 指出設定為男性作者執筆之資訊給人較具權威與信賴的感受。Crowelly 等人 (2009) 認為就實體機器人而言，男性機器人給人較可靠的感受；但虛擬 AI 系統反而是女性 AI 給人較可靠的感受。或許就本研究選定的電商情境購物中，因為聊天機器人協助受測者的僅是日常服飾的採購，這樣的任務或專業沒有明確的性別優勢，因此使用者對聊天機器人性別並不會有明顯的偏見或信任感受差異，未來可針對不同情境之聊天機器人探討是否對信任感或產生相同之結果。

### (3) 使用者性別比聊天機器人性別更能影響人們對機器人的信任感

研究顯示，受測者性別會對機器人信任感受產生影響。男性對聊天機器人的誠信與公正明顯抱持肯定態度，也較願意提供個人資料給聊天機器人，與參考其意見；但女性對聊天機器人誠信度則持保留態度。因此在討論性別因素對聊天機器人信任感受的影響時，機器人性別並無顯著差異，反倒是不同性別的使用者之間會有較大

的意見分歧。這種分歧在女性購物聊天機器人的情況下最為明顯。這種使用者性別與機器人性別的交互作用與國內外針對機器人的性別化研究結論相符 (Siegel et al., 2009; You & Lin, 2019)。因次想要提升電商聊天機器人給人的信任感，設計師必須同時思考聊天機器人預期服務的對象為何，針對潛在市場的顧客性別進行聊天機器人的角色設定，更有助於人機互動中的信任感提升。

#### (4) 使用者年齡比使用者性別對信任感受的影響更大

由於受測者自身因素對機器人信任感受之作用較為明顯，本研究以受測者的另一項因素「年齡」來進行分析，作為對聊天機器人信任感受影響的參考基準。研究結果顯示，年齡不同的受測者對聊天機器人信任感受有明顯差異，影響涉及的信任構面也更廣。由於年齡對機器人信任感受之影響並非本研究的主题，因此不再深受受測者年齡造成信任感差異的背後意義；但藉由這樣的比較，本研究可以確認電子商務中聊天機器人的性別對使用者的信任感受的影響非常細微，不如使用者本身的因素（年齡、性別）所造成的差異。

## 5.2 聊天機器人性別化的設計建議

綜上所述，電商購物情境中協助日常用品銷售的聊天機器人的性別對使用者信任感影響不大，換言之，國人對一般電商聊天機器人性別並無明顯的信任偏好。所以設計師在開發電商的購物聊天機器人時，設計師可以根據企業的品牌形象設定聊天機器人的角色。藉由符合人類社會中性別典型的特徵或符號，就能有效喚起特定的性別感受。無須過於擔心性別設定對顧客信任感的影響。僅考量信任感的情況下，電商環境中的聊天機器人是性別平等的。但是如果聊天機器人行銷的產品類別或代言的品牌專業是有明顯性別優勢（或性別刻板印象）的領域時，根據 Botsman (2017) 與 McDonnell 與 Baxter (2019) 的主張，聊天機器人信任的評定與所委派的任務有關，本研究的結論與設計建議恐不適用。

但值得注意的是，男女使用者對聊天機器人的信任態度是明顯有別的。比起女性，男性使用者更相信機器人的誠信與公正性，也較願意提供資訊與接受機器人的建議。同時異性的聊天機器人對於使用者具備些微的優勢（異性效應），所以根據目標客群的性別提供客製化的機器人服務也是一個可行的做法。AI 的人機互動是一種社會互動，對設計師或開發者而言，聊天機器人的性別化設計不應僅考量機器人的性別塑造，兼顧使用者性別的交互影響，將能更有效運用性別化創新概念於 AI 產品或服務的開發。

## 六 參考文獻

- Armstrong, C. L., & McAdams, M. J. (2009). Blogs of information: How gender cues and individual motivations influence perceptions of credibility. *Journal of Computer-Mediated Communication*, 14(3), 435-456.
- Botsman, R. (2017). *Who can you trust?: how technology brought us together—and why it could drive us apart*. Penguin UK.
- Brownlee, J. (2016). *Conversational interfaces, explained*. Fast Company. Retrieved from <https://www.fastcompany.com/3058546/conversational-interfaces-explained>.
- Calhoun, C., Light, D., & Keller, S. (2001). *Understanding Sociology*. New York: Glencoe-McGraw-Hill.
- Casas, J., Tricot, M. O., Abou Khaled, O., Mugellini, E., & Cudré-Mauroux, P. (2020). Trends & Methods in Chatbot Evaluation. In *Companion Publication of the 2020 International Conference on Multimodal Interaction* (pp. 280-286).
- Chiriyankandath, C. J. (2018) Moving from Human-Computer-Interaction (HCI) to Human-Human-Conversation (HHC), Retrieved Jan. 18, 2022, from <https://www.linkedin.com/pulse/moving-from-human-computer-interactionhci-cipson-jose-chiriyankandath>
- Cofta, P. (2006). Convergence and trust in eCommerce. *BT Technology Journal*, 24(2), 214-218.
- Coren, M. J. (2017). It took (only) six years for bots to start ditching outdated gender stereotypes. Retrieved August 15, 2021, from <https://qz.com/1033587/it-took-only-six-years-for-bots-to-start-ditching-outdated-gender-stereotypes>
- Crowelly, C. R., Villanoy, M., Scheutzz, M., & Schermerhornz, P. (2009). Gendered voice and robot entities: perceptions and reactions of male and female subjects. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 3735-3741). IEEE.
- Di Gaetano, S., & Diliberto, P. (2018). Chatbots and conversational interfaces: Three domains of use. In *Fifth International Workshop on Cultures of Participation in the Digital Age, Castiglione della Pescaia, Italy* (Vol. 2101, pp. 62-70).

- Dua, A. (2017). How to design the perfect chatbot, LinkedIn, December 4, 2017, Retrieved August 15, 2021, from <https://www.linkedin.com/pulse/how-design-perfect-chatbot-amit-dua/>
- Eyssel, F., & Hegel, F. (2012). (S) he' s got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology*, 42(9), 2213-2230.
- Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019, November). Gender bias in chatbot design. In *International Workshop on Chatbot Research and Design* (pp. 79-93). Springer, Cham.
- Flanagin, A. J., & Metzger, M. J. (2003). The perceived credibility of personal Web page information as influenced by the sex of the source. *Computers in human behavior*, 19(6), 683-701.
- Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What makes users trust a chatbot for customer service? An exploratory interview study. In *International conference on internet science* (pp. 194-208). Springer, Cham.
- Gefen, D. (2002). Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 33(3), 38-53.
- Ghazali, A. S., Ham, J., Barakova, E. I., & Markopoulos, P. (2018). Effects of robot facial characteristics and gender in persuasive human-robot interaction. *Frontiers in Robotics and AI*, 5, 73.
- Hay, S. (2017). Even bots need to build character. *Venture Beat*. Retrieved August 15, 2021, from <https://venturebeat.com/2017/09/22/even-bots-need-to-build-character/>
- Hosmer, L. T. (1995). Trust: The connecting link between organizational theory and philosophical ethics. *Academy of management Review*, 20(2), 379-403.
- Knibbs, K. (2014). What Gender Is Your Robot? Retrieved August 15, 2021, from <http://gizmodo.com/would-you-rather-own-a-boy-robot-or-a-girl-robot-1654656508>
- Lee, E. J., Nass, C., & Brave, S. (2000, April). Can computer-generated speech have gender?: an experimental test of gender stereotype. In *CHI'00 extended abstracts on Human factors in computing systems* (pp. 289-290). ACM.

- Madsen, M., & Gregor, S. (2000). Measuring human computer trust. In 11th australasian conference on information systems (Vol. 53, pp. 6-8).
- McDonnell, M., & Baxter, D. (2019). Chatbots and gender stereotyping. *Interacting with Computers*, 31(2), 116-121.
- Messina, C. (2016). 2016 will be the year of conversational commerce. Retrieved August 15, 2021, from <https://medium.com/chris-messina/2016-will-be-the-year-of-conversational-commerce-1586e85e3991>
- Mielke, C. (2016). Conversational Interfaces: Where Are We Today? Where Are We Heading? *Smashing Magazine*. Retrieved Jan. 20, 2022, from <https://www.smashingmagazine.com/2016/07/conversational-interfaces-where-are-we-today-where-are-we-heading/>
- Nass, C. & Yen, C. (2010). *The man who lied to his laptop: What machines teach us about human relationships*. New York, NY: Current.
- Newitz, A. (2009). What Gender Is Your Roomba? Retrieved August 15, 2021, from <http://io9.gizmodo.com/5268217/what-gender-is-your-roomba>.
- Nguyen, M. (2017). Chatbots are making online shopping and payments a breeze – see why retailers are jumping on this growing trend, Retrieved August 15, 2021, from <https://www.businessinsider.com/ecommerce-trends-conversational-commerce-chatbots-2017-10>.
- Niculescu, A., Hofs, D., Van Dijk, B., & Nijholt, A. (2010). How the agent's gender influence users' evaluation of a QA system. In *User Science and Engineering (i-USEr)*, 2010 International Conference on (pp. 16-20). IEEE.
- Pan, J. (2017). Conversational Interfaces: The Future of Chatbots. *Chatbot Magazine*. Retrieved August 15, 2021, from <https://chatbotmagazine.com/conversational-interfaces-the-future-of-chatbots-18975a91fe5a>.
- Robertson, J. (2010). Gendering Humanoid Robots: Robo-Sexism in Japan. *Body and Society*. 16.2 (2010): 1-36.
- Saffer, D. (2017). Personality, Not Specs: Designing Social Robots, Retrieved August 15, 2021, from <https://medium.com/@odannyboy/personality-not-specs-designing-social-robots-743f10bded91>
- Schwartzman, R. (1999). Engenderneered Machines in Science Fiction Film, *Studies in Popular Culture*, Vol. 22, No. 1 (OCTOBER 1999), pp. 75-87

- Shevat, A. (2017). *Designing bots: Creating conversational experiences*. O'Reilly Media, Inc.
- Siegel, M., Breazeal, C., & Norton, M.I. (2009, October). Persuasive robotics: The influence of robot gender on human behavior. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJInternational Conference on* (pp. 2563-2568). IEEE.
- Turk, V. (2014). *We're Sexist Toward Robots*, Motherboard. Retrieved August 15, 2021, from <https://www.vice.com/en/article/539j5x/were-sexist-toward-robots>.
- Vlahos, J. (2019). *Talk to Me: How Voice Computing Will Transform the Way We Live, Work, and Think*. Houghton Mifflin Harcourt.
- Wang, Y. (2014). *Gendering Human-Robot Interaction: exploring how a person's gender impacts attitudes toward and interaction with robots*. Retrieved August 15, 2021, from [http://mspace.lib.umanitoba.ca/bitstream/handle/1993/24446/wang\\_yan.pdf](http://mspace.lib.umanitoba.ca/bitstream/handle/1993/24446/wang_yan.pdf)
- Wang, Y. & Young, J. E. (2014). Beyond "Pink" and "Blue": Gendered Attitudes Towards Robots in Society. *Proceedings of Gender and IT Appropriation (GenderIT'14)*, pp 49 -59
- Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113-117.
- Yonck, R. (2017). *Heart of the Machine: Our Future in a World of Artificial Emotional Intelligence*, Arcade Publishing.
- You, H. C., & Lin, K. W. (2019, April). Gendered Tour-guide Robots and Their Influence on User Attitude and Behavior. In *Proceedings of the 5th International ACM In-Cooperation HCI and UX Conference* (pp. 32-35).
- Zanbaka, C., Goolkasian, P., & Hodges, L. (2006). Can a virtual cat persuade you?: the role of gender and realism in speaker persuasiveness. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1153-1162). ACM.
- 王振寰、瞿海源編 (2009)。社會學與台灣社會，臺北：巨流。
- 優拓資訊 (2017)。三分鐘讓你了解聊天機器人在紅什麼？。取自：<https://blog.yoctol.com/e21forum2017-2e2c0cb1d340/>