

從面具到化身：媒介、展演與文化在 VTubing 媒介生態演化中的動態交織*

柯籙晏**

投稿日期：2024 年 7 月 17 日；通過日期：2025 年 1 月 7 日。

* 本文作者感謝三位匿名審查人與編輯委員的寶貴建議，特此致謝。

** 柯籙晏為國立政治大學傳播學院博士，ORCID: 0009-0004-6065-7749, e-mail: duress.ko@gmail.com

本文引述格式：

柯籙晏（2025）。〈從面具到化身：媒介、展演與文化在 VTubing 媒介生態演化中的動態交織〉，《新聞學研究》，163，107-168。https://doi.org/10.30386/MCR.202504.0008

《摘要》

本研究探討 VTubing（虛擬網紅直播）如何重拾蒙面展演（masked performance）並使其在主流媒介文化中煥發新生。研究採用綜合了展演理論與技術能供性理論的廣義媒介生態學方法論，結合文獻資料分析與線上田野觀察，從媒介技術、媒介化展演、文化與歷史脈絡三個層次，描述 VTubing 如何演出並探討其為何能（相對）成功全球化。

以 Hololive 所屬 VTuber 直播為例，本研究在分層描述 VTubing 現象後，首先指出即時動作捕捉所驅動的化身作為創新形式的面具，如何強化 VTubing 的展演功能，包括放大觀眾沉浸感、避免觀眾假戲真作、降低展演者面對觀眾的心理負擔等。這些功能使 VTubing 在串流直播生態中具有相對於非蒙面直播的優勢。其次本研究指出，VTubing 作為蒙面展演的創新形式，根植於日本技術泛靈論的媒介文化，該文化偏好技術的擬人化，從動畫、虛擬偶像再到 VTuber，皆可見其影響。

最後，本研究探討 VTubing 在當前媒介生態中的演化過程，指出其全球化的成功源自它能適應串流直播生態、搭上日本動畫全球擴張浪潮，以及掌握 COVID-19 疫情期間全球大封控的機遇。本文建議未來可探討 VTubing 的全球在地化變體，從而揭示它們在媒介生態中可能扮演的角色。

關鍵詞：化身、串流直播、面具、媒介化展演、媒介生態學、
虛擬網紅直播

壹、研究背景與問題

2020 年夏，我和一位久未謀面的同學重聚。我們兩個阿宅湊在一起，總是在聊些動畫、漫畫與遊戲（anime, comic and game, ACG）話題。他介紹我一種來自日本的新鮮玩物——VTubing（虛擬網紅直播）。這是一種新形式的媒介化展演（mediated performance），¹由展演者透過動作捕捉技術（motion capture）操控通常是日式動畫風格的虛擬化身（virtual avatar）來扮演虛擬網紅（virtual YouTuber, VTuber）的角色，並透過即時串流直播（live streaming）平臺演出。在那當下我不覺得這特別有趣，然而到了同年秋天，出國工作的我在防疫旅館長達十四天的隔離生活中迷上了看 VTubing。而像我這樣在 COVID-19 大流行的防疫隔離中百無聊賴的大量觀眾，推動了 VTubing 在 2020 年的全球爆發性成長。

一、研究動機

這個新媒介現象引起我的興趣，但我一時不確定它有什麼研究價值。直到我在一場研討會上發表對 VTubing 的初步觀察時，一位顯然覺得 VTubing 很奇怪的聽眾問我：「為什麼要用化身直播？這和普通直播有什麼不同？」這個問題讓我意識到，正如 McLuhan 所言「直到擱淺，魚才發現水的存在」（M. McLuhan, E. McLuhan, & Zingrone, 1996, p. 31），我「不覺得 VTubing 很奇怪」的知覺缺乏反思，而這種知覺或許

¹ “Mediated”一詞在本文中依脈絡需要翻譯為「媒介化」或「中介」。

與 VTubing 的媒介形式有關。於是我進一步聯想到，這個現象或許可以從媒介認識論（media epistemology），亦即不同媒介會以不同方式影響其使用者的知覺與行為（Postman & Postman, 2005）的角度來加以考量。²

儘管有了初步想法，考量到 VTubing 可能只是防疫隔離期間的短暫流行，我一直沒有進一步研究。然而，VTubing 在後疫情時代持續蓬勃發展。以全球頂尖四千多名 VTuber 直播主的總觀眾收視時數為例，從疫情結束的 2022 年第四季的 3.034 億小時增長至 2024 年第一季的 4.0321 億小時，增幅近 33%。在這些 VTuber 中，來自日本以外者占比已近半數（47.4%），其中來自美國（15.5%）與韓國（4.9%）者尤其顯著，來自臺灣者也不算少（2.2%，Dempsey, 2024, May 1）。這表明 VTubing 已突破原本只是日本在地現象的地域與文化限制，展現全球化的穩健增長。隨著 VTubing 至少在日本已經成為一門穩固的產業（詳見業界龍頭 Hololive 母公司 COVER 的 CEO 訪談，PIVOT 公式チャンネル，2023 年 12 月 2 日），我決定重拾這個題目。

二、VTubing 相關研究缺口

基於前述考量，我開始檢視相關研究，特別關注媒介認識論觀點的

² 「不同媒介以不同方式影響其使用者的知覺與行為」的概念，被許多不同理論家使用不同的術語加以指涉。目前比較流行的術語包括 Innis 提出並經由 McLuhan 普及的「媒介偏向」（media bias），以及 Postman 的「媒介認識論」（media epistemology）。本文借用後者。需要注意的是，這一概念並非 Postman 的首創。Tresch（2012）的研究將這個概念追溯到 19 世紀的法國哲學家 Maine de Biran，並將其稱作「物理唯靈論」（physio-spiritualism）。此外，Tresch 整理了基於此概念進行研究的學者，包括 Merleau-Ponty、Bergson、Deleuze、Latour、McLuhan 在內。

文獻，以進一步理解 VTubing 與普通直播如何不同，以及為何要用化身直播這兩個問題。如前所述，VTuber 與 VTubing 並非同一概念，前者意指展演者演出的虛擬網紅角色，後者指涉的是 VTuber 的串流直播。在需要與觀眾即時互動的直播中，展演者一般會使用即時動作捕捉技術操控 VTuber 化身；然而 VTuber 的演出形式不限於直播，例如在預製影片中，展演者通常不需操控 VTuber 角色的動作，只需要配音。由於 VTubing 只是 VTuber 展演形式的一種，一般較少特意使用這個詞。³這種使用習慣也反映在相關研究多以「VTuber」為關鍵詞，因此尋找 VTubing 的相關文獻需要同時搜尋這兩個詞。

搜尋結果顯示，多數論文與媒介認識論無關。例如 Lu et al. (2021) 從閱聽人研究角度探討觀眾如何與 VTuber 互動並感知其身分；Brett (2022) 聚焦性別與種族批判，從文化研究角度探討當前 VTuber 化身的造型設計為何更多呈現日式動畫美少女 (anime bishōjo) 風格；Yuan (2024) 聚焦 VTubing 的勞動條件，從政治經濟學角度探討 VTuber 透過數位平臺所中介的勞動關係；Chen & Hu (2024) 從身分建構角度出發，探討展演者如何演出超女性化的 VTuber 等。

除此之外，也有少數媒介認識論觀點的研究。Li (2023) 的研究主旨雖然是為了揭示準社會互動 (parasocial interactions) 是影響觀眾觀看 VTubing 意願的關鍵因素，但他同時指出 VTuber 化身的視覺設計是另一個關鍵因素，並且指出 VTuber 的角色與化身等因素，使得某些觀眾對 VTuber 與對真人主播的看法不同；Jiang et al. (2023) 發現，VTuber 觀眾更偏好 2D 而非更寫實的 3D 化身，並且假設了幾個可能的解釋，包括

³ 根據 Google 搜尋趨勢統計，“VTuber”一詞於 2018 年 1 月開始出現顯著搜尋量，而“VTubing”則在 2020 年 9 月後才開始被廣泛使用，且前者搜尋量約為後者百倍 (Google Trends, “VTuber, VTubing”, 2024/1/15)。

觀眾反而對更寫實的虛擬角色會感到疏離或不適的「恐怖谷效應」（uncanny valley effect），以及 2D 化身與觀眾熟悉且喜愛的日式動畫風格更契合等；Mou et al.（2023）透過分析網路評論，探討觀眾對 VTuber、真人網紅和傳統名人的認知差異，結果顯示，觀眾傾向把使用化身的 VTuber 視作虛構角色，並且傾向把真人網紅視作傳統名人；Chinchilla & Kim（2024）的實驗指出，與真人露臉或不露臉的直播主相較，VTuber 帶給觀眾的社會在場感（social presence）較弱；Wan & Lu（2024）指出使用化身確實影響直播主的自我呈現與性別。這些研究顯示，觀眾或直播主自身確實對於 VTuber 有不同於真人展演者的認知。然而這些主旨在探討觀眾或直播主對 VTuber 認知的研究，並未（也不需要）區分本文先前所提出「VTuber、VTuber 演出與 VTubing」的概念差異，當然也沒有探討我所關切的「有何必要在直播中使用化身演出」的問題。

於是我進一步調查「為何要在直播中使用化身演出」的相關研究。Bredikhina（2021）指出，使用即時動作捕捉驅動的化身讓 VTubing 能夠創造出虛擬角色自由活動和說話的錯覺，這種技術源自日本傳統蒙面戲（masked theater）與傀儡戲（puppetry），並且在直播主扮演虛擬角色時發揮了關鍵作用；隨後，Bredikhina & Giard（2022）進一步解釋，為什麼某些男性直播主會用日式動畫美少女風格的化身進行直播，受訪的直播主表示，這是因為他們把 VTubing 當作逃避社會壓力的展演（performance），在演出過程中可以不受現實生活約束，與粉絲共同創造出一種全新身分。這種透過詢問直播主動機所得到的解釋，雖然絕非錯誤，但用 Geertz（1973）的隱喻來說，可能還不夠「厚」（thick）——即未能透過更細緻的脈絡分析來理解特定文化的實踐意義。粉絲對次文化的癡迷背後可能有更複雜、尤其是非個人動機的原因；而且，顯然

並非所有觀眾都會產生 VTuber 栩栩如生的錯覺。不過，Bredikhina 的研究確實提供了一個理解 VTubing 的好起點，即可以將其視作蒙面與傀儡戲的創新形式。

在蒙面戲中，展演者透過面具彰顯所扮演角色的同時，也隱藏了本來面貌與身分；在傀儡戲中，展演者操控傀儡來演出角色；同樣地，VTubing 中，展演者在操控化身演出 VTuber 的同時，也隱藏了本來面貌與身份。由此可見，VTubing 確實是一種蒙面與傀儡戲的創新形式。從文化與歷史脈絡來看，蒙面展演可追溯到人類文化初始：文化的核心包括記憶與符號，記憶維繫傳統，符號強化記憶。面具作為無文字或文盲社會的文化符號，有助於保存社會記憶，充當重要溝通工具，提供一個由展演者所想像並在觀眾集體記憶中鞏固的宇宙觀（Edson, 2005）。非西方的傳統劇場普遍使用面具，如日本能劇（Noh）、印尼的哇揚翁戲（Wayang Wong）等等。西方則是在古希臘的悲劇與喜劇，以及在 16 世紀成形的義大利即興喜劇（Commedia dell'arte）中也使用面具，直到 18 世紀以後隨著義大利即興喜劇的衰弱，蒙面戲才退出西方主流媒介文化。即使如此，蒙面戲從未消失，只是邊緣化與小眾化（Devlin, 1989, pp. 53-81; Wingert, n.d.）。

三、研究問題

因此，VTubing 不是什麼全新現象，而是體現了 M. McLuhan & E. McLuhan（1988）所提出媒介定律（law of media）中的「重拾」現象——曾經被主流媒介文化淘汰（obsolesce）的蒙面展演，透過新媒介擴充（extend）為 VTubing 並且被重拾（retrieve）而再次流行起來。這就為「為何要用化身演出」這個問題增加了一個媒介生態學（media

ecology) 的維度：長期被主流媒介文化淘汰的蒙面展演為何會被重拾？這個問題於是乎比我當初想得更有趣。要更全面地回答這個問題，根據廣義媒介生態學方法論（詳見本文後段方法論架構的討論），不僅需要更詳細地描述 VTubing 如何實踐，還需要將其置於更大的、人們使用或不使用面具演出的媒介生態及其文化與歷史脈絡中，才能適當地加以解釋。

綜上所述，本文於是將原本的研究問題重構為：在當前媒介生態中，化身、直播主與觀眾「如何」以及「為何」共同參與 VTubing 這種蒙面展演的創新形式，從而使邊緣與小眾的蒙面展演能重回主流媒介文化？

貳、方法論架構

如前所述，本文從媒介生態學觀點重新審視了 VTubing 作為蒙面展演的重拾現象，並提出需要從更大的媒介生態脈絡加以理解。為回答所提出的研究問題，以下本文將說明廣義媒介生態學方法論及其相關理論。首先探討廣義媒介生態學，提出一個用以分析媒介化實踐的通用方法論架構；其次探討技術能供性理論，聚焦於如何具體分析媒介化實踐的媒介層次；最後討論展演與後設溝通理論，聚焦於如何具體分析作為一種特殊媒介化實踐——媒介化展演的 VTubing。這些理論將共同構成本研究的方法論架構。

一、廣義的媒介生態學方法論

「媒介生態學」一詞雖然由 Postman 提出，但其理論概念已經超越

依循 McLuhan 與 Postman 學派傳統的範疇（參見 Lum, 2006, pp. 9-11），形成更普遍的方法論取徑。以下本文整合多個相關理論，包括 Ihde（1990）的技術現象學（technological phenomenology）、Hutchins（1996）的分散認知理論（distributed cognition theory）、Latour 的行動者網絡理論（actor-network theory, ANT）與 McLuhan 的媒介理論。這些理論從不同角度探討技術如何形塑人類的知覺、實踐與文化，並指出人類的使用方式和文化脈絡會反過來影響技術的發展與創新。這些理論的共通點在於「非個人主義」的方法論，強調人類的知覺與實踐是人類、技術與環境互動的產物，而非僅僅源自個人意向與行為。

在這些理論中，Ihde（1990）指出，技術不是中立的工具，使用技術改變了人類與環境的互動方式，因而重塑了人類對環境的知覺與經驗；技術與人的互動總是互相影響，兩者共同塑造了人對環境的理解。Hutchins（1996）透過對於軍艦導航的經驗研究，打破傳統認知理論，指出認知是一個分散式現象，不僅存在個人大腦之中，還分散在人與技術、環境、社會的互動之間。Latour（2005）指出，社會並非僅由個人或集體行動者所建構，而是由人與非人行動者（如物質、技術、社會組織等）之間的對等互動，動態地建構出具有共同實踐的網絡所構成，而這種網絡還可能隨著持續建構而重新定義。McLuhan 指出，媒介是人類能力的延伸，使用不同媒介塑造了人類理解與適應環境的不同方式，這不僅影響了個人的思考與行為，並潛移默化地改變了社會與文化。在此基礎上，他進一步提出媒介定律，強調媒介使用的四個相互作用的動態演化過程：強化（enhancement）、反轉（reversal）、淘汰（obsolescence）、重拾（retrieval）。首先，新媒介能放大（intensification）使用者的知覺或擴充（extension）實踐；然而隨著強化推向極端，該媒介可能反過來對使用者造成與該媒介原本使用目的相

反的效果，即反轉。其次，強化會淘汰（或更精確地說，邊緣化）某些舊的媒介；然而這些舊媒介不會完全消失，還可能在新脈絡中以擴充的新形式被重拾（McLuhan, Fiore, & Agel, 1996; M. McLuhan & E. McLuhan, 1988; M. McLuhan, E. McLuhan, & Zingrone, 1996）。

綜合來說，這些理論的共同特徵可以稱作方法論的「非」個人主義（methodological non-individualism），即：人類的認知與行動並非如「方法論個人主義」所預設僅由個人意向所決定，而是分散在人類、技術與環境之間，並在三者的互動過程中湧現。其特徵包括：（1）非人行動者：技術並非人類的被動工具，而是能主動改變人類與環境的行動者；（2）技術中介（mediated）的人類實踐：技術在人類認知、行動與經驗的形塑中具有關鍵作用；（3）媒介生態的演化觀點：將人類、技術與環境的關係視為動態、互相依賴的生態系統，在文化與歷史脈絡中發展與演化；（4）跨學科觀點：這些理論跨越傳統學科界限，結合哲學、心理學、人類學、社會學等領域的觀點，提供對人類、技術與環境關係的全面理解。本文將這種非個人主義的方法論概括為「廣義的媒介生態學方法論」。基於這種方法論，對於所有種類媒介化實踐的分析都需要包括三個層次：（1）媒介技術，對應非人行動者；（2）媒介化實踐，對應技術中介的人類經驗；（3）媒介化實踐的文化與歷史脈絡，對應媒介生態的演化過程。

廣義媒介生態學方法論為研究包括 VTubing 在內的媒介化實踐提供了一個通用分析架構，然而要如何落實到具體分析上？例如，在媒介層次，哪些技術該被挑出來分析？在實踐層次，哪些類型的行動該被挑出來分析？以下，本文將借用技術能供性（technological affordance）理論與展演理論來處理這個問題，為後續基於廣義媒介生態學方法論的經驗分析，提供更具操作性的理論工具。

二、技術能供性理論

能供性概念最早由 Gibson (1979) 提出，指的是環境供應與約束了棲息於其中生物的行動可能性，而這些生物能直接知覺到這種可能性。例如小水窪能讓青蛙而非大象直接知覺到它的游泳可能性。其後，Norman (1988) 將此概念引入設計領域，把能供性重新定義為「感知能供性」(perceived affordance)，即使用者如何理解物體使用方式的設計線索，例如設計師在門上設置的推拉標誌，讓使用者能夠直接理解這扇門的開閉方式。感知能供性概念在設計領域影響很大，但也引來批評。其中 Kaptelinin & Nardi (2006) 指出，Norman 過於強調設計如何影響使用者感知，忽略了 Gibson 的原始定義中，能供性的「行動可能性」是實際存在於使用者與環境互動過程的特性，而不僅僅是使用者透過設計所感知到的設計師意向。其後，Kaptelinin & Nardi (2012) 進一步融合能供性概念與 Vygotsky (1978) 關於媒介化行動的認知理論指出，對於使用技術而且有文化與歷史的人類而言，媒介化行動的能供性並非天生本能所能直接感知，而是後天學習而得的。他們因此提出媒介化人類行動 (mediated human action) 與技術能供性的概念，將技術能供性定義為媒介化人類行動的可能性。對於熟練的媒介使用者而言，技術能供性能夠被直接知覺，但對於非熟練使用者則需要進一步學習才能知覺到。

綜上所述，技術能供性理論符合廣義媒介生態學方法論，同樣強調媒介、實踐與文化三者互動，而技術能供性的分析就是對媒介化實踐中技術條件的分析。技術能供性理論提示研究者應聚焦於能夠供應與約束實踐的技術條件，例如，聚焦於使 VTubing 成為一門可能潛在職業的化

身控制系統與直播平臺的線上支付系統，並為本文後段所要說明的展演理論鋪墊，使後者能利用前者的分析結果，進一步探究這些技術條件所形塑的 VTuber 與觀眾之間的審美互動。

三、展演與後設溝通理論

前述廣義媒介生態學方法論的通用分析架構，適用於任何一種媒介化實踐，例如 Hutchins (1996) 的軍艦導航或是本文的 VTubing 研究。然而就哪些具體項目需要分析而言，軍艦導航與 VTubing 顯然非常不同。為了能夠更具體地分析作為一種特殊媒介化實踐的 VTubing，需要更清楚地定義其特殊之處何在。

本研究將 VTubing 定義為一種特殊媒介化實踐，是直播主透過操控化身扮演虛擬網紅角色進行直播的一種「媒介化展演」。基於技術能供性理論，所謂媒介化是指這種展演需要透過特定技術的中介才能實現。如化身使 VTubing 得以實現。

關於展演的定義，本文主要採用 Bauman (1975) 的觀點並參考了 Schechner (2020)：展演首先是一種在溝通中用於詮釋所溝通訊息的類型的後設訊息 (meta message)，即 Bateson (1987) 所謂的框架 (frame)。在展演框架中溝通的訊息，具有至少區別於表面意義 (literal) 的後設意義。例如，笑話框架指出所說的話並非其表面意思，VTubing 框架指出所演出的動作亦非其表面意思。其次，展演展現在展演者向觀眾呈現其溝通才能 (communication competence) 的責任，這種才能有賴展演者能以觀眾認可的審美型態 (aesthetic pattern) 進行演出。有鑑於演出的溝通才能是一種技術，包括物質的 (如操作化身)，與非物質的 (如語言表達)，這呼應了本文前述，技術透過改變

人類與環境（在此指展演者與觀眾）的互動方式而重塑人類經驗的觀點（Hutchins, 1996; Ihde, 1990; Kaptelinin & Nardi, 2006）。

另一方面，演出（performing）本身成為觀眾鑑賞的對象，觀眾會期待透過觀賞演出來強化（enhance）自己的經驗，因此演出會引起觀眾的高度關注，而觀眾一再讚賞的演出則會形成約定俗成的慣例（convention），未能獲得觀眾認可的演出則會邊緣化。這呼應了前文所述，人與技術所共建的實踐會不斷演化的觀點（Latour, 2005; M. McLuhan & E. McLuhan, 1988）。

綜上所述，基於廣義媒介生態學方法論、技術能供性理論與展演理論的綜合，VTubing 是一種媒介化展演。這種展演形式涉及展演者與觀眾的審美互動以及慣例的演化過程。據此，本文對於「VTubing 如何展演」的分析與描述將聚焦於三個層次，其中包括：（1）文化與歷史脈絡，說明 VTubing 從哪些既有的媒介化展演形式中演化而來；（2）媒介技術，說明 VTubing 依賴哪些媒介技術的中介才得以呈現在觀眾面前；（3）作為展演的媒介化實踐，說明 Vtuber 如何利用這些技術演出，而觀眾又是如何鑑賞這些演出，乃至於某些受到觀眾普遍喜愛的演出型態成為慣例。

參、研究方法

以下基於廣義媒介生態學方法論，說明本研究的研究設計以及案例選擇與資料蒐集程序。

一、研究設計

基於非個人主義的廣義媒介生態學方法論，本研究將從互相關連的三個層次來描述 VTubing 現象，分別是：（1）媒介層次，分析哪些媒介技術供應與約束了 VTubing 的展演；（2）實踐層次，描述 VTubing 的展演特徵；（3）文化層次，分析 VTubing 在其媒介文化的脈絡中的意義。為完成這些工作，本研究採用歷史資料分析與線上田野調查中的非介入式觀察兩種研究方法。

由於超出直接觀察的範圍之外，本研究對於 VTubing 文化與歷史脈絡的分析，不得不藉助二手文獻資料，主要包括相關產業報告與新聞報導。至於 VTubing 的媒介技術與實踐本身的分析，本研究採用非介入方式觀察 VTubing 田野，以採取一手資料。觀察重點首先聚焦於 VTubing 的媒介化展演，即 VTuber 的演出以及與觀眾的互動如何透過媒介技術（如化身）得以實現；其次聚焦於媒介技術本身（如即時動作捕捉驅動的化身）的特徵與其在展演中的作用。

以下說明非介入式線上田野觀察的具體操作方式。由於當前 VTubing 的演出慣例上包含大量「雜談」，亦即 VTuber 與觀眾彼此閒談生活與工作上的瑣事，而這樣的閒談往往涉及他們詳盡的自我揭露（如獅白ぼたん，2024 年 8 月 18 日，YouTube 直播），這使得 VTubing 本身就是極佳的田野場，讓研究者可以在不介入的情況下深入瞭解 VTubing 的各種面向，不僅包括公開展示的前臺活動，還能窺見幕後的運作狀況。透過觀察不同 VTuber 與不同場次的直播並比較其中的觀察結果，研究者能得到偏差較小的資料。這種方法比傳統問答式訪談有幾個優勢：（1）與主要採取受訪者個人認知的訪談法相比，它更符合本

研究的「非」個人主義方法論；（2）它提供了持續觀察的機會，能夠捕捉到更多現象細節與變化；（3）在這種自然情境中，VTubing 參與者將表現出不受研究者干擾的行為。與面對權力不對等的研究者的訪談法相較，後者可能導致受訪者給出偏向較大的資料；最後，透過交叉驗證對於不同場次直播的觀察結果，資料的可靠性得以提高。

二、案例選擇與資料蒐集程序

根據針對 2024 年第一季全球最受歡迎四千餘名 VTuber 的統計，VTubing 市場呈現幂律分佈（power-law distribution）：最大的經紀公司 Hololive Production 占 22.5% 的觀眾收視時數，第二大的 NIJISANJI 占 19%，第三大驟降至 5.6%，其後依序為 3.4%、2.6%、2.4%、2%、0.7%、0.5%，其餘小型公司合計占 5.4%，個人勢（獨立經營）合計占 36%。這種集中狀況也反映在地域（日本 VTuber 占全數 52.6%）、平臺（YouTube 平臺觀眾收視時數占全平臺 60.7%）和經營模式（企業勢，即企業經營的 VTuber 占全數 64%）等面向。其中，Hololive Production 旗下的女性 VTuber——Hololive members（一般簡稱 Holomem 或 Holo 成員）表現尤為突出：在全球最受歡迎的前十名 VTuber 中占據八席，並在這前十名 VTuber 所吸引的總觀看人次中得到 84.5% 的占比（Dempsey, 2024, May 1）。

基於上述 VTubing 的市場特徵，本研究選擇 Holo 成員直播作為主要觀察對象。這個選擇確實可能存在代表性偏向、商業模式特殊性與文化偏向等限制，但基於以下原因：（1）在全球前四千餘名 VTuber 中，Holo 成員吸引的觀眾收視時數占近四分之一，這種主導地位使其成為產

業典範；（2）相較於個人勢 VTuber 一般活動期間較短，⁴ 且演藝生涯結束後，先前直播的內容存檔往往散失；企業勢的 Holo 成員大多長期穩定地直播，更重要的是這些直播大多有完整公開存檔，能為研究提供連續且豐富的資料；（3）雖然 VTubing 包含諸多個人勢的小眾型態，但作為企業勢代表的 Holo 成員直播代表了 VTubing 被「重拾」，（相對）成功進入全球主流媒介文化的型態；（4）Hololive 全球化成果顯著。截至 2024 年 9 月共 66 名成員中，包括 9 位印尼語與 19 位英語分部成員；而根據 PIVOT 公式チャンネル（2023 年 12 月 2 日）訪談，其非日本觀眾比例已達 30%。這些資料顯示，Holo 成員直播可以作為本文探究「VTubing 如何以及為何能使蒙面展演重回主流媒介文化」的理想案例。

在資料蒐集方面，Holo 成員主要透過 YouTube 直播，大部分直播都有公開存檔。此外，公司策略也鼓勵粉絲將直播內容剪輯成精華片段（clip）以利推廣。本研究對 VTubing 的現象描述，主要基於對這些公開資料的觀察，研究範圍包括活動中以及停止活動但直播存檔仍公開，共 70 位成員的直播內容。時間跨度從第一位 Holo 成員開始直播到本文完成之間的 2017 年 9 月至 2024 年 9 月，共計七年。

由於資料量極大，單單 2023 年，Hololive 旗下頻道總直播時數就高達 52,838 小時（ジェネリック集計人（Holo データ分析），2024 年 2 月 6 日），研究者不可能看過所有直播。此外，主要是即興演出的直播資料不像報紙或傳統廣電節目資料那樣高度結構化。在一場連續 27 小時的直播中——這是某位 Holo 成員的紀錄——對於本研究有意義的內容可能只有十分鐘，且分散在不同時段；而另一場只有三十分鐘的直播

⁴ 根據日本 VTuber 圈的說法，VTuber 的平均壽命大約是八個月，但這個說法缺乏權威統計證實。

可能從頭到尾都有意義。這使得直播資料既不適合內容分析取徑的隨機抽樣，因為研究者無法預先知道有意義的資料可能會在那個時段發生，也不適合用紮根理論取徑的理論抽樣，因為要在巨量且非結構的直播內容中抽出足夠的理論飽和樣本，將耗費不敷成本的大量時間和精力。

因此，本研究借鑒田野調查法，除了在正式研究前已經是老觀眾的調查者本人，帶著詮釋學意義上的前見（prejudice）⁵——主要是身為粉絲的興趣以及廣義媒介生態學的方法論架構——進入直播田野四處閒逛之外，本研究主要依靠粉絲剪輯的精華，來快速鎖定值得觀察的直播內容。這種作法優點如下：（1）精華片段符合本研究（媒介化展演研究）的理論抽樣需求。根據展演理論（Bauman, 1975），觀眾高度關注的是更好的演出；（2）精華片段更易檢索。相較於長時間不分段的直播檔，這些帶有剪輯者自訂標題的精華容易透過關鍵詞搜尋（如「兔田、放屁」，詳見本文後段關於 VTubing 本真性的討論），方便更高效地篩選出目標樣本；（3）同一場直播通常有複數粉絲從各自面向進行剪輯，形成一種群眾智能（collective intelligence）的篩選機制。這種多元的剪輯視角，與傳統田野調查依賴少數報導人相比，能夠呈現更多觀察面向與可能被忽略的細節；（4）精華片段慣例上會提供原始直播檔的鏈接，研究者可回看原始內容，因此不會過度依賴精華片段的中介。

此外，在這些巨量資料中，研究者所觀察的同一主題通常會有多位 VTuber 在不同場合重複談論，研究者可以交叉比對這些不同來源，獲

⁵ 根據 Gadamer（1975）的詮釋學方法論，研究者無法在完全中立狀態下進行理解。相反，理解必然受研究者過去經驗和文化傳統的影響，這種「前見」（prejudice）構成了我們對世界的初步理解。雖然「前見」是理解的起點，但這並不意味著它是僵化的。當接觸到新證據時，研究者的前見可能會被挑戰，從而進一步修正其理解。

得更全面的資料（見本文後段關於多位 VTuber 的不同生理線索如何被觀眾視為 VTubing 本真性符號的討論）。而在分析這些資料的過程中，研究者可能識別出一些先前未考量過的關鍵詞（如 OBS，見本文後段關於 VTubing 技術能供性的分析）。利用這些關鍵詞可以進一步搜尋，從而發現更多相關資料。透過這種快速迭代的過程，研究者能夠逐步擴大資料蒐集範圍，在研究期限內不再有新發現時，即終止搜尋。

當然，這種作法也有限制。搜尋引擎的演算法是個黑盒子，研究者無法完全控制搜尋結果的代表性與可重複性。本研究使用 YouTube，即 Holo 成員直播的平臺內建的搜尋引擎，會根據使用者觀看記錄調整搜尋結果，使用者在不同時間、裝置上得到的結果總是有所差異，影響所能獲得的樣本。不過，透過迭代關鍵詞搜索、交叉比對不同來源資料以及回溯原始直播等作法，在一定程度上降低了搜尋演算法偏向對抽樣信效度的負面影響。

綜合而言，本研究的詮釋學循環（hermeneutic circle）式抽樣與編碼流程是——這個流程本身就是一種媒介化實踐，⁶ 體現了抽樣工具、研究實踐與學術文化的交織：（1）觀看直播，聚焦於方法論架構已經指出類目的內容，主類目如媒介、展演、文化，次類目如展演者演出型態、觀眾鑑賞型態等；（2）抽取關鍵詞；（3）使用關鍵詞透過 YouTube 搜尋快速鎖定相關精華片段，若有疑慮，則回看原始直播檔核實；（4）分析新資料，評估其研究意義，若發現新類目，則可據以擴

⁶ Gadamer（1975）指出，在詮釋學循環（hermeneutic circle）中，理解不斷在部分與整體之間迭代，即研究者透過對個別文本細節的詮釋，檢視整體文本的意義，而對整體的理解又會反過來影響對細節的詮釋。每一次的詮釋都是基於已有的前理解，但這些前理解並非固定不變，而是透過與新資料的對話而不斷被修正。理解作為一種對話過程，不僅是文本與詮釋者之間的互動，也受到歷史脈絡的深刻影響。

大關鍵詞庫；（5）持續迭代，直到在研究期限內不再發現新類目。雖然此程序有演算法偏向的限制，但相較於傳統紮根理論在如此龐大且非結構的直播資料中難以有效率地抽取足以達成理論飽和的樣本，透過搜尋過程的快速迭代以及多重資料來源的交叉比對與原始直播的核實，使本研究能在同樣的資料中更快速且有效地抽取樣本。

肆、VTubing 如何展演？以 Holo 成員為例

如前所述，以下本文對「VTubing 如何展演」的描述，需要包括它如何從既有的審美型態慣例中演化而來，並依賴哪些媒介技術得以展現；在此過程中，VTuber 如何利用這些技術來塑造其獨特的演出型態，觀眾又是如何鑑賞並推動某些演出型態成為慣例。從綜合了展演與技術能供性理論的廣義媒介生態學方法論出發，為了回答這些問題，以下本文將根據二手資料，描述 VTubing 文化層次的（1）文化與歷史脈絡；並且利用觀察 Holo 成員直播得到的一手資料，描述媒介層次的（2）技術能供性；以及展演層次的（3）審美型態。

一、VTubing 的文化與歷史脈絡

正如 McLuhan 所言「我們只能看著後照鏡駛向未來」（McLuhan, Fiore, & Agel, 1996, pp. 74-75），任何媒介化實踐都脫胎自既有形式。因此在描述「VTubing 如何展演」之前，先描述其前身的展演形式與這些形式所依賴的媒介生態。

有研究將 VTubing 遠溯到日本江戶時代（16-19 世紀）的大眾娛樂，包括蒙面戲（使用厚重的臉譜化妝）的歌舞伎（Kabuki）與傀儡戲

的人形淨琉璃（Ningyo Joruri）（Bredikhina, 2021）。若縮小範圍，VTubing 更直接的根源是 1970 年代成熟並在 1990 年代開始全球化的日本大眾娛樂——ACG 文化。在 1990 年代，業界開發出虛擬偶像（virtual idol），亦即「獨立於任何現實參照對象的媒介原生展演者」（Black, 2012, p. 209），推出以 3D 動畫呈現人物形象，並由配音員演出歌曲的虛擬偶像，如 1995 年的伊達杏子（Date Kyōko）。2005 年，模擬培養虛擬偶像的電玩遊戲【偶像大師】（THE IDOLM@STER）推出。2007 年，語音生成器 Vocaloid 2 與所搭配的虛擬角色初音ミク（Hatsune Miku）推出；2009 年起，初音以虛擬偶像的身分陸續舉辦多場虛實整合的擴增實境（augmented reality, AR）演唱會。沿著這條演化路徑，2016 年，被公認為第一位 VTuber 的絆愛（Kizuna AI）上線，主要活動是發布預製影片。2017 年絆愛開始直播；同年，第一位 Holo 成員時乃空（Tokino Sora）也開始直播。

VTubing 的另一條演化路徑，可追溯到面向日本國內 ACG 市場的影音共享平臺ニコニコ動画（Nico Nico Douga）。在 2016 年絆愛上線以前，一些小眾網紅在 Nico Nico 上已經形成扮演動畫風格造型的虛構角色，並用該角色的圖片代替本人入鏡的慣例。由於成本低廉，這種作法至今仍很常用。後來，圖片被升級為即時動作捕捉驅動的化身。雖然現在各大直播平臺都有使用化身的網紅在活動，但最早在商業成功的是選擇在 YouTube 上活動，並首先自稱為「VTuber」的絆愛，因此這類網紅被通稱為 VTuber。VTubing 是虛擬偶像與蒙面直播概念的綜合體。如前所述，VTuber 在 2016 年初出現時主要發布影片，到了 2017 才開始直播。事實上，根據 PIVOT 公式チャンネル（2023 年 12 月 2 日）的訪談，VTuber 在轉向直播後才開始爆紅。2024 年第一季，全球前四千名 VTuber 直播的總觀眾收視時數達到 4.0321 億小時（Dempsey, 2024, May

1)。2019 年，隨著日本動畫全球大擴張的浪潮，⁷ VTubing 也開始邁向國際，並在 2020 年由於防疫封控對於非接觸式娛樂需求的暴增所帶來的機遇，在全球範圍爆發式成長（PIVOT 公式チャンネル，2023 年 12 月 2 日）。2023 年，全球 VTuber 市場的估值為 13.5 億美元，收入來源包括直播、內容、商品銷售、授權與其他（Global Information Inc, 2024）。

二、VTubing 的技術能供性

如前所述，技術能供性是媒介化人類行動的可能性（Kaptelinin & Nardi, 2012）。VTubing 作為一種媒介化展演，由多項技術所共同供應與約束：直播平臺供應展演平臺，演出物（performing object）與舞臺供應展演形式，社交媒介平臺供應展演者的個人品牌經營，而網路基礎建設使所有這些媒介得以連結為整體。以下分析這些媒介如何供應與約束 VTubing 的實踐。

（一）直播平臺

以下聚焦於 Holo 成員使用的 YouTube，分析直播平臺如何供應與約束 VTubing，包括串流直播與影片代管、AI 內容審查、以及付費模式等三個面向。

⁷ 2013-2022 年間，日本動畫的國際市場產值從 0.3 兆日圓增長至 1.5 兆日圓，佔總市場份額的比例也從 19.1% 提升至 49.8%。這被認為主要得益於串流媒體如 Netflix 與 Disney+ 對動畫的龐大需求（Masuda et al., 2024）。

1. 串流直播與影片代管

YouTube 原本只有影片代管服務，2015 年為了因應 Twitch 的競爭開始提供串流直播，而這兩種技術分別對應預製影片與直播兩種展演形式。就 VTuber 而言，目前主流是直播，即 VTubing，但他們還是會發布預製影片，通常也會將直播存檔公開以供回看，而這些都需要影片代管服務。直播與影片的差別在於，前者讓直播主與觀眾能透過聊天室或其他方式即時（live）互動（見圖 1），並且就和其他型式的現場演出一樣，這種互動約束了直播的演出型態。例如，直播主與觀眾的互動，是直播演出的核心元素。這也使得 VTuber 需要在直播中使用即時動作捕捉驅動的化身（見下段說明），而非像在預製影片中那樣，通常只是為動畫中的化身角色後製配音。

圖 1：Holo 成員的 VTubing



資料來源：〈【新幹線0号】緊急で動画回してるんですけど…異変があります…〉【影音】，さくらみこ，2024 年 3 月 24 日，取自 <https://youtu.be/HVSq3sR9tBc?t=7638>

2. AI 內容審查

YouTube 會審查用戶上傳的內容，一旦發現含有不宜播放的成分，就會取消該內容的收益，嚴重者甚至會封鎖帳號。為了減輕人工審查的負擔，YouTube 從 2020 年開始使用 AI（artificial intelligence，人工智慧）輔助審查，但正確率大幅下降，被取消收益的內容在人工複審後有 50% 重新上架，是尚未引入 AI 審查之前的兩倍（Barker & Murphy, 2020, September 20）。這樣的技術問題約束了直播主的演出型態。為了避免因 AI 誤判而受到損失，許多直播主會刻意避開可能被誤判的內容，即使這些內容整體而言並無不妥。儘管如此，誤判還是經常發生。Holo 成員於是戲稱 YouTube 的 AI 審查系統是「YouTube 君」，形容它像一個受不了刺激內容的小男孩（如猫又おかゆ，2023 年 6 月 7 日，YouTube 直播）。這個隱喻具體化了 Latour（2005）將技術視為「非人行動者」的概念。

3. 付費模式

直播平臺的線上支付系統讓 VTubing 得以成為一門全球生意。除了廣告分潤外，Holo 成員使用的 YouTube 還提供二種主要付費模式讓粉絲直接支持創作者：（1）頻道會員：當 YouTuber 的訂戶數超過三萬時，他們可以建立收費頻道，提供定期定額付費的會員專屬服務，如專屬影片與直播、用於聊天室的徽章與表情符號貼圖等。當粉絲購買付費服務時，提示訊息會醒目地出現在留言版上。（2）超級留言（superchat）：閱聽人在收看直播時，可以透過超級留言隨時打賞，讓留言呈現在聊天室上，金額越高的留言越醒目（見圖 1）。2021 年在防疫封控下對非接觸娛樂的需求達到最高峰時，全球獲得超級留言打賞金額最高的前九名 YouTuber 全都是 Holo 成員，總金額超過九千萬美金

(Playboard, n.d.)。這種經濟上的成功吸引了更多創作者投身 VTubing 事業。

(二) 演出物與舞臺

以下分析作為演出物的化身及其舞臺對於 VTubing 的技術能供性。「演出物」是民俗符號學家 Proshan 的術語，意指「在敘事或戲劇展演中被創造、展示或操控的人類、動物或靈體 (spirit) 的物質形象」(1983, p. 4)。這個詞原本用來概括面具、傀儡等由展演者操縱演出人或擬人化角色的媒介，本文用這個詞來延伸概括與面具以及傀儡具有相同技術能供性的化身。

1. 即時動作捕捉驅動的化身

即時動作捕捉 (real time motion capture) 技術能將展演者的動作投射到虛擬化身上並加以呈現，使用這類演出物是使得 VTubing 不同於普通直播的關鍵特徵。即時動作捕捉驅動的化身在設備需求與視覺效果方面差異很大。包括 Holo 成員的大多數 VTuber 展演者，用的是在個人電腦上即可運作的低階系統。雖說技術不斷升級，這類低階化身 (見圖 1) 基本上只能呈現眨眼、開闔嘴、側身、前傾、後仰等基本動作，若要表現其他動作，則需要輔以非動作捕捉的操控方式，例如由展演者拖曳作為電腦桌面物件 (desktop object) 的化身，使其在螢幕上移動，用來呈現化身進出虛擬舞臺的視覺效果。另一方面，在專業攝影棚中運作的高階化身，可以把展演者大部分的肢體動作細節呈現出來。由於成本高昂，這類設備一般只有企業勢能夠負擔，並用於高收益的活動，像是以精彩舞蹈為賣點的 AR 演唱會中 (見 hololive_English, 2023, November 30, YouTube livestream)。

2. 虛擬攝影棚與虛擬舞臺

演出需要舞臺。對於以虛擬化身為演出物的 VTubing 來說，虛擬攝影棚是呈現虛擬舞臺的必要工具。虛擬攝影棚是一種軟體工具，用於設定和控制直播中的多模態元素，如影像、物件、圖層與聲音等。目前主流的是 Open Broadcaster Software (OBS)，在個人電腦上即可運作。在虛擬舞臺方面，Holo 成員最常用的舞臺設計由化身、聊天室、背景與其他物件所組成（見圖 1）。此外他們還經常使用一些具有讓玩家能操控電玩化身，在虛擬世界中自由探索、創造和改變環境的沙盒（sandbox）遊戲作為舞臺，例如【當個創世神】（Minecraft）就很受青睞，他們會在其中舉辦夜市、運動會等展演活動（如 Mumei, 2023, November 4, YouTube livestream）。

（三）社交媒介平臺

任何網紅都需要經營社交媒介，作為虛擬網紅的 VTuber 也不例外。與一般網紅無異，VTuber 展演者的個人品牌經營也高度依賴即時語音通訊與社交媒介的技術能供性。在直播中，即時語音通訊讓身處異地的 VTuber 展演者能在同一個聲音空間中共同演出，藉以拉抬彼此的曝光度和影響力（如常闇卜ワ，2022 年 4 月 16 日，YouTube 直播）。社交媒介則讓展演者能即時表達形象、宣傳活動、回應粉絲，藉此保持與粉絲的持續互動，這種互動不僅強化了展演者與粉絲之間的親近感，還使他們與粉絲社群的緊密連結得以無縫延伸到直播之外，進一步鞏固了他們的品牌影響力。

（四）數位網路基礎建設

數位網路基礎建設使前述各種媒介得以連結在一起，共同構成整體

VTubing 生態的技術供應性。首先，數位網路基礎建設的全球連結使 VTubing 的全球化，包括跨國直播合作、粉絲互動、收益全球化等成為可能，讓 VTubing 得以抓住全球防疫封控下對於非接觸式娛樂需求的暴增所帶來的發展機遇。此外，穩定的網路才能保證流暢的直播，這對於以直播主與觀眾之間的即時互動為核心演出型態的串流直播，當然也包括 VTubing 而言尤為重要。在直播中，網路卡頓總會成為焦點事件（見角巻わため，2020 年 1 月 2 日，YouTube 直播片段）。由於不同地區的網路服務品質差異很大，因此直播主，包括 Holo 成員，甚至會為了更好的網路服務而搬家（見大神ミ才，2022 年 1 月 24 日，YouTube 直播）。

三、VTubing 的審美型態

根據本文對於展演的定義，展演者需要展現閱聽人認可的溝通才能；展演者的演出會成為閱聽人鑑賞的對象，被普遍認可的演出則會慣例化（Bauman, 1975）。據此，以下 VTubing 審美型態的分析將分成三部分：（1）展演者的特徵；（2）展演者的演出型態；（3）觀眾的鑑賞型態。

（一）展演者的特徵

作為被扮演的角色，VTuber 有角色設定，即創作者為其設計的外型、背景與性格。以下以 Holo 成員為例，就 VTuber 的化身與人格面具，即其角色的外型與性格、基於化身與人格面具的演出動作，以及超出人物設定範圍外的本真性，描述 VTuber 作為展演者的特徵。

1. 虛擬化身

為了迎合男性導向的市場，VTuber 大多是女性角色。2024 年第一季全球最受歡迎四千名 VTuber 中，女性占比 69.6% (Dempsey, 2024, May 1)。女性 VTuber 的化身大多採用日式動畫美少女的造型設計，美少女加上幻想與怪誕 (grotesquery) 元素的混合風格也很常見，用來營造天馬行空的直播情境 (見圖 1 與圖 3 右)。這類造型是用來假裝扮演 (make-believe play) 的。例如一位 Holo 成員的化身造型是半人半羊的美少女，他在直播中喜歡拿自己被做成羊肉料理來說笑，粉絲當然不會當真，也不會不識趣地把眼戳破，而是享受這個玩笑 (角巻わため，2020 年 9 月 10 日，YouTube 直播)。

VTuber 的化身俗稱「皮」(gawa)，扮演 VTuber 的演員則稱作「中之人」(nakanohido)，這些用語源自日本特攝 (tokusatsu) 劇的皮套演員傳統。目前 VTubing 的慣例是參與者要假裝相信 (make-believe) 「VTuber 就是 VTuber」，沒有中之人，例如某位 Holo 成員的自嘲與觀眾的回應：

慘了，[電玩世界中的玻璃窗]在反射[出男人的倒影]……
會被人說「中之人是男的」啊！哪些不大瞭解 VTuber，動不動就在說中之人是誰誰誰的討厭傢伙，會被那些人說我是男人啊……[對瞭解 VTuber 的人來說]VTuber 裡面什麼都沒有啊！……[唸出觀眾留言]「是男的也沒關係喔」，開什麼玩笑，我才不是男的！（さくらみこ，2024 年 3 月 24 日，YouTube 直播）

而就 VTubing 而言，這個慣例還延伸到演出之外。與一般動畫角色的配

音員不同，中之人的實際身份不會正式公開。⁸

2. 人格面具 (persona)

網紅通常會刻意設定其所演出的人格面具，VTuber 亦然，差別只在於他們由化身中介的人格面具，在慣例上比一般網紅更天馬行空。由於自我是與他人在社會互動過程中形成的 (Mead & Morris, 1962)，與一般網紅一樣，VTuber 的人格面具也會隨著在直播中與觀眾的長期互動而變化。這種變化有時幅度很大，例如某位 Holo 成員因此被粉絲戲稱有雙重人格，而他也順勢把雙重人格的人格面具用在演出上 (見赤井はあと，2021 年 2 月 7 日，YouTube 直播)。另一種常見的情況是，由於慣例上 VTuber 會在直播中大量閒聊自己的日常生活故事，這樣的自我揭露使得他們演出的人格面具，往往會調整成更貼近中之人的現實生活。但無論再怎麼貼近，演出仍然不是現實。

3. 演出動作

VTubing 目前通用的低階化身只能演出一些簡單的動作，看起來和紙娃娃一樣僵硬 (見圖 1)。雖然有些 VTuber 會積極探索如何讓低階化身呈現更戲劇化的動作，但他們的演出主要是透過聲音表情 (voice expression) 而非化身。Holo 成員亦然，他們全都有表情豐富，演技高超的女聲。

需要說明的是，作為虛構角色的 VTuber 可以與中之人的實際身分

⁸ VTuber 也可以真的沒有中之人。例如 2022 年開始在 Twitch 直播的 Neuro-sama 就是由大型語言模型 (large language model, LLM) 強化的聊天機器人 (chatbot)、語音生成器加上化身自動程式所驅動的 VTuber (見 Neuro-sama, 2023, January 3, YouTube livestream)。雖然目前只有實驗意義，但隨著 LLM 的能力日趨強化，這類 AI 驅動的 VTuber 前途無可限量。

無關，例如男性可以透過女聲變聲器演出女性 VTuber。正如指出「演出動作建構了性別」的 Butler 所言：「性別絕非各種動作所依據的穩定身分或能動性的所在；正相反，它是在時間中透過風格化的反覆動作脆弱地構成的身分」（1988, p. 519）。無論中之人實際身分如何，Holo 成員主要透過聲音演出女性人格面具。作為大眾娛樂，這些人格面具符合對於女性的各種社會刻板印象，例如性感的笑聲、小女孩的撒嬌等。當然，不同展演者風格各異。而這種聽覺導向的人格面具也會隨著與粉絲的互動而調整。例如在被許多粉絲戲稱為「嬰兒」後，一位 Holo 成員將自己的聲線逐漸調整成鼻音更重、更幼稚的聲線（見さくらみこ，2018 年 8 月 28 日，YouTube 直播與さくらみこ，2024 年 3 月 24 日，YouTube 直播的比較）。

另一方面，與低階化身中介的聽覺導向演出相對，高階化身中介的演出則是視覺導向。在使用高階化身的演出中，肢體動作細節所建構的人格面具會成為觀眾關注的焦點。例如，有粉絲從一位 Holo 成員的舞蹈風格正確猜出他是個美國妞：「我沒見過有哪個日本女人能像這樣扭腰，除非他苦練過！這不是日本人的屁股！不可能！」（Koefficient, 2020, June 29, YouTube livestream）

4. 本真性（authenticity）

本真性最初是用來分析現場（live）展演與媒介化（如錄音）搖滾樂之間關係的理論概念。Auslander 將本真性定義為「*展演者的特徵和他們對於音樂的呈現*」（2022, pp. 94-95）。例如，一個並不自己創作音樂的搖滾樂團會被聽眾認為是非本真的，而他們演出的音樂也因此是非本真的。

本文在這裡借用此概念，指的是 VTubing 的本真性主要體現在

VTuber 與中之人的綁定。這種慣例是在 VTubing 的歷史發展中演化出來的。2019 年，公認第一位 VTuber 絆愛被粉絲發現，他的經紀公司正在用其他中之人來逐步取代原本的中之人。就一般動畫角色而言，更換配音員的情況並非罕見，然而這個事件卻引發粉絲強烈抵制。粉絲認為，絆愛的成功源自原本中之人在直播中與粉絲互動所形成、無可取代的演出方式，公司卻對他過河拆橋。儘管公司在壓力下調整了經營方針，絆愛仍在 2022 年無限期中止活動。以此事件為鑑，包括 Hololive 在內的業界形成慣例，若 VTuber 原本的中之人離開，則該 IP（intellectual property，知識產權）就終止營運（如桐生ココ，無日期）。而這種作法又進一步強化了「VTuber 就是 VTuber，沒有中之人」的審美型態。

（二）演出型態

以下同樣以 Holo 成員為例，描述 VTubing 幾種較明顯的演出型態。

1. 中之人的肉體線索

如前所述，VTubing 的觀眾會假裝相信 VTuber 沒有中之人，實際上卻將 VTuber 與中之人綁定。這兩條相悖的審美型態要如何協調一致？其中一種方式，是透過中之人的肉體線索來暗示中之人的本真存在。

雖然 VTubing 在慣例上允許中之人呈現臉以外的肢體，中之人一般而言在視覺上缺席（absent）。⁹ 那麼，觀眾如何知道中之人沒有被調包，或者驅動化身的不是 AI？基於就算無意識甚至違反自身意願，溝通參與者的任何行為，都會被其他參與者當作訊息的「無法不溝通」（the impossibility of not communicating）的溝通公理（Watzlawick et al.,

⁹ 在英語圈的 VTubing，中之人不露臉的慣例並不像在日語圈中那樣嚴格。見本文圖二。

1967, p. 48) , VTubing 的本真性可以透過指涉中之人肉體狀態的線索呈現出來。如前所述, 視覺導向的高階化身能精準映射中之人的肢體動作, 這讓觀眾容易透過視覺確認中之人的本真存在, 例如透過舞蹈動作猜出某位 Holo 成員是本真存在的美國妞。然而大多數 VTubing 是低階化身中介的聽覺導向演出。這使得觀眾主要透過指涉中之人肉體狀態, 尤其是無意識肉體狀態的聽覺線索, 如噴嚏聲、肚子咕嚕聲, 來確認中之人的本真存在。這使得這類在一般專業播音中必須隱藏起來的噪音, 居然被某些 VTubing 觀眾視為一種值得打賞的演出型態:

糟糕, 噴嚏要……[噴嚏聲]……咦, 該不會沒靜音吧? ……[觀眾 A, 超級留言 250 日圓「噴嚏幫大忙了」]我要再問問怎麼靜音……[觀眾 B, 超級留言 5,000 日圓「噴嚏幫大忙了」]……[觀眾 C, 超級留言 500 日圓「噴嚏幫大忙了～」]……噴嚏沒幫大忙啊, 哈哈(癒月ちよこ, 2021 年 1 月 13 日, YouTube 直播)。

還有其他更間接的線索。如前所述, VTubing 通用的低階系統是個人電腦驅動的。若展演者離開電腦, 他就無法透過聊天室與觀眾互動, 系統也無法捕捉他的動作, 使得化身靜止不動。在這種情況下, 與觀眾互動的中斷以及化身動作的凍結, 就成為中之人本真存在的間接線索。這種情況在長時間的直播中更常見, 因為中之人總得要離席上廁所, 而無論是離席還是需要上廁所, 都是中之人具有肉身的間接線索。這種情況激發出一種更怪異的鑑賞型態——總是勸觀眾不要打賞的一位 Holo 成員的粉絲們, 居然形成在這位成員離席上廁所時給他打賞的慣例, 因為這時他就無法阻止粉絲們這麼做了(見犬山たまき/佃煮のりお, 2020 年 9 月 4 日, YouTube 直播)。

2. 表情勞動與超級留言時間

表情勞動（emotional labor）是管理感覺與表情以滿足工作所要求表情的過程（Hochschild, 1983）。與直播主直接向粉絲討要打賞的「掠奪式」演出型態（Zhang et al., 2019）不同，有些女性直播主發展出一套適應 YouTube 的表情勞動，其原則是直播主與粉絲的互動不只是一般意義上的演出，還包括彼此的生活經驗分享。YouTube 的超級留言機制創造了個人化的注意力經濟（attention economy），直播主會感謝打賞的觀眾，根據對話需要而非金額高低唸出超級留言，有時還會精心搭配一些生活故事（Guarriello, 2019）。許多 VTuber，包括 Holo 成員，都採用了這種演出型態：

[唸出超級留言]「……工作一直不順利，沒有朋友也沒有家人，想從這世上結束自己的生命，但這時你出現了……我的日語水準大概只能聽懂一半，但你的聲音就像你的手在抱著我一樣，消除了我的無力感[哽咽]……我賺得不多，不是……[金額最高的]紅色超級留言真的很抱歉[啜泣]……真的很感謝你。最喜歡你了喔」[啜泣]我什麼都做不到，但是可以像這樣聊天、玩耍、唱歌，成為某些人的力量的話，真的很開心[啜泣]……你說「抱歉，不是紅色超級留言」，不要在意，倒不如說，要給我的超級留言，不如拿去買些好吃的……為了給我打氣，用了這筆錢，謝謝你[啜泣]（角卷わため，2021 年 1 月 4 日，YouTube 直播）。

直播中的表情勞動並非 VTuber 專利，但與一般直播主相較，VTuber 的表情勞動有其獨特之處。表情勞動需要不斷控制自己的表

情，長期且高強度的表情管理往往會使勞動者感到倦怠，甚至導致表情耗竭（emotional exhaustion），無法再保持應有的工作表情（Hochschild, 1983）。VTuber 也面臨同樣的問題，但化身為他們提供了相對優勢——遮蔽了展演者的實際表情。例如某位 Holo 成員在直播中哭泣時，觀眾只能透過聲音線索與對話察覺，甚至可能沒注意到他哭了（常闇トワ，2022 年 4 月 16 日，YouTube 直播）。這使得 VTuber 不必隨時注意自己的表情變化，相對容易管理自己的表情勞動。

3. ACG 與日式偶像演出型態的借用與挪用

如前所述，VTubing 源自日本 ACG 文化，因此 VTubing 有其次文化風格的演出型態。而要瞭解這些演出型態的獨特之處，需要瞭解這種次文化的脈絡。在日本，「阿宅」（otaku）通常指 ACG 迷，但也涵蓋其他狂熱愛好者，如鐵道、攝影、日式偶像粉絲等。不同阿宅圈有各自的審美和參與方式，例如 ACG 宅的二次創作，或偶像宅的「御宅藝」（ヲタ芸）和「打 call」（コール）。同屬日本次文化的這些圈子雖然不同，但彼此間相對容易溝通。目前的 VTubing 大致沿用了 ACG 宅圈既有的演出型態，包含演唱 ACG 歌曲，談論阿宅話題和長時間的電玩直播。Holo 成員的獨特之處在於，他們率先以 VTubing 形式融合虛擬偶像與日式偶像的演出型態。Hololive 的母公司 Cover 原本是一家資訊公司，一開始經營 VTuber 的目的是要行銷自家的化身控制系統。因此，初期的 Holo 成員與其他 VTuber 並無明顯差異。2019 年，Cover 參考經典日式偶像團體 AKB48 的經營模式，¹⁰ 成立將自家 VTuber 當作偶像來經營的經紀公司 Hololive，並利用自身的技術優勢，以虛實整合的方式

¹⁰ AKB48 是 2005 年以來最成功的日本女子偶像團體，其商業模式不只在日本成功，也輸出印尼、中國等地。參見森重崇文（2019）。

舉辦偶像演出中必不可少的「現場」(live) AR 演唱會——除了「livestreaming」的另一個「live」。然而，Hololive 並未完全移植經典偶像的經營模式，而是採取一種混合模式。根據 PIVOT 公式チャンネル (2023 年 12 月 2 日；2023 年 12 月 3 日) 的訪談，Holo 成員除了原本的直播，也可以自主決定是否參與公司提供的歌舞訓練，在演唱會上演出。

這種混合模式激發出一些有趣的演出型態。例如，慣例上傳統偶像必須維持純潔形象，「偶像不會大便」的粉絲口號正是這一理想的體現。¹¹ 而這恰好與前述粉絲透過中之人肉體線索追求 VTubing 本真性的怪異鑑賞型態相悖。在此脈絡中，雖然也有 Holo 成員選擇依循傳統偶像的自我形象管理，把上廁所委婉地稱作「摘花」，刻意不在直播中上廁所，但也有一些 Holo 成員創造出反諷偶像形象的後設演出，像是將上廁所戲稱為「偶像會議」(idol meeting，見 Baelz, 2022, May 16)，甚至以「偶像」的身分大開屎尿屁玩笑：

[唸出觀眾留言]「就算[你]會放屁，我也一樣喜歡。」
好，偶像兔田ぺこら [Usada Pekora]是不會「噗噗」放屁的。
放出來的只有棉花糖。麻煩大家記住，只有棉花糖和軟糖。以上兩種在[這位 Holo 成員的虛構家鄉]Peko Land 都買得到。哈哈，有夠糟糕，你們不要因為現在是半夜，就講這種骯髒話題啊。說不定有人現在才吃飯耶。真抱歉，我們的直播本來很

¹¹ 2013 年日本電視節目【FNS 27 時間テレビ】中，受邀的偶像在直播中表示要上廁所，這一幕引起了部分網友質疑，表示不相信偶像也需要上廁所，甚至熱烈討論偶像會不會大便(播磨谷拓巳，2016 年 4 月 28 日)。這些粉絲當然不是真的相信偶像不會大便，但他們確實認為在觀眾面前隱藏生理線索是偶像與經紀公司的責任。

乾淨的（兔田ぺこら，2021 年 5 月 25 日，YouTube 直播）。

這種拉伯雷式（Rabelaisian）玩笑與 Bakhtin（1965/Iswolsky, Trans., 1984）評論俗民如何在狂歡節中透過這類象徵無窮生命力的肉體玩笑顛覆日常生活規範不謀而合。粉絲普遍認可這類演出型態——如前所述，觀眾甚至透過包括排泄在內的肉體線索確認 Holo 成員中之人的本真存在——取笑這樣玩的 Holo 成員是諧星而非偶像。而這也並不妨礙同一群粉絲享受 Holo 成員如「真偶像」般載歌載舞的演出。

（三）鑑賞型態

與一般串流直播一樣，觀眾會透過聊天室與 VTuber 即時互動，並以訂閱或打賞等方式肯定 VTuber 的演出。除此之外，VTubing 的觀眾還有自阿宅文化中演化出來的其他鑑賞型態。

1. 二次創作

粉絲參與喜愛作品的二次創作，是粉絲圈（fandom）的基本鑑賞型態（Jenkins, 2013），VTuber 的粉絲們也不例外。而 Hololive 允許粉絲在特定條件下營利化二次創作，進一步激發了 Holo 成員粉絲的創作力。常見的二次創作形式包括：（1）精華片段剪輯與後製，尤其是翻譯字幕。由於大多數 VTuber 使用日語，這對於 VTubing 在國際市場上的推廣尤其重要。（2）同人圖（fan art）與限圖（internet meme）。延續 ACG 圈慣例，粉絲會利用 VTuber 的造型與直播內容進行二次創作。其形式除了插圖與 ACG 等常見形式外，尤其是但不限於英語圈的粉絲會創作只為了讓圈內人看懂，因此特別能強化圈內共識的限圖。這些創作不僅娛樂粉絲，也容易被 VTuber 看見並用於直播，VTuber 也經常會委託粉絲創作（如天音かなた，2021 年 4 月 6 日，YouTube 直播）。這

種 VTuber 與粉絲之間的正向互動，讓粉絲更願意支持 VTuber。

2. AR 現場活動中的鑑賞型態

在防疫隔離解封後，雖然 VTubing 主要仍是線上活動，但像 Holo 成員這樣的企業勢還增加了 AR 實體現場的線下與線上混合活動，讓觀眾不只能觀看線上直播，還能線下親臨實體活動現場。若是後者，粉絲會沿用 ACG 聚會的慣例，使用 IP 商品或自製道具打扮自己或是角色扮演（cosplay）VTuber 以示支持。如果參加的是演唱會，粉絲在 VTuber 演出時也會相應地打 call（見 hololive_English, 2023, November 30, YouTube livestream）。雖然這些鑑賞型態與一般現場活動相同，但對粉絲來說，還具有確認 VTubing 本真性的額外意義，這種本真性並非體現在 VTuber 與中之人的綁定，而是帳號觀眾與具身（embodied）觀眾的綁定。某位粉絲在參加 Hololive 現場活動後分享了他的感受：

平常看直播同接[同時在線觀看人數]都是幾千、幾萬……我們卻容易忘記每一個數字的背後都是像你我一樣有血有肉真實的人，而演唱會將這些人帶來現場……希望 V[Tuber]們可以親眼見見現場，讓他們知道他們的世界不是虛擬的，而是由我們這幾千幾萬甚至幾百萬真實的人們在為他們歡呼和支持著（AU TW TG，2023 年 7 月 4 日的置頂留言）。

伍、為何要用化身演出？

以下處理「為何要用化身演出」的解釋性問題。本文前段描述「VTubing 如何展演」的研究結果顯示，除了化身的技術能供性所中介的一些特殊審美型態，例如對於 VTubing 本真性的悖論追求，VTubing

與普通直播並無區別。然而，這些特徵並不足以解釋，對 VTubing 外行人來說，有何必要選擇這種在外行人眼中顯得怪異的形式。

為了避免用內行人才能理解的特徵，來解釋為何內行人選擇這種形式的循環認證，我們需要從更根本之處探究「為何要用化身演出」，也就是「在當前媒介生態中，化身、直播主與觀眾為何共同參與 VTubing 這種蒙面展演的創新形式，從而使得邊緣與小眾的蒙面展演能重回主流媒介文化」的問題。為了處理這個問題，以下本文將借鑒 M. McLuhan & E. McLuhan (1988) 與 Latour (2005)，透過媒介化實踐的生態演化觀點進行分析。分析將聚焦於三個面向：（1）作為強化面具的化身，其技術能供性如何使 VTubing 在直播生態中具有優勢？（2）傳統蒙面戲在什麼特殊的媒介文化中，擴充為 VTubing 的創新形式？（以上兩點對應媒介定律的「強化」）；（3）曾經被主流媒介文化淘汰／邊緣化的蒙面戲，如何能以 VTubing 的創新形式，在當前媒介生態中被重拾？（對應媒介定律的「重拾」）。

一、化身在當前直播生態中的展演功能

首先說明化身在直播中的技術能供性——本文把這種特殊能供性稱作「展演功能」——如何使 VTubing 在當前直播生態中具有優勢。如前所述，展演是一種相對於表面意義框架的後設框架。在演出過程中，這個過程本身對其觀眾表明「要用特殊而不是表面意義來詮釋我的演出」（Bauman, 1975）。然而觀眾如何確認眼前所發生的事不是表面上的意義？基本的作法是，把某個在展演中的元素當作符號，用來指出展演的後設框架，即「框架設定訊息」（frame-setting message）。例如，舞臺提示了舞臺上的殺人是演出，而不是表面意義的殺人。一般而言，演出

中最引人注目的元素會被視作主要框架設定訊息。在蒙面戲中是面具，傀儡戲中是傀儡，在 VTubing 中就是化身了。相較之下，非演出物展演的框架設定訊息就沒那麼明顯。因此，化身在直播中的首要功能是，它可以放大觀眾對於「VTubing 具有超越表面意義的後設意義，而非現實意義的單純再現」的知覺。要進入這種超越的審美經驗中，觀眾先得假裝相信「VTuber 就是 VTuber，沒有中之人」才能入戲（參見 Meineck, 2018; Piris, 2014, pp. 40-41; Pollock, 1995, pp. 592-593）。

另一方面，化身作為強化的框架設定訊息，也能防止 VTubing 「玩過火」、「假戲真作」。與真人露臉的直播主相較，日式動畫風格的化身有著誇張五官比例（見圖 2），讓觀眾很容易明確辨識出 VTuber 作為化身所中介的角色是虛構的（參見 Suan, 2013, p. 202），因此更能有效地放大觀眾對於「VTubing 非現實」的知覺，從而避免演出與現實的混淆而引發問題。

圖 2：同一位直播主的真人與化身形象比較



資料來源：圖 2 左〔実写📷〕おわかれです〕【影音】，kson ONAIR，2024 年 3 月 18 日，取自 https://youtu.be/PuA8oRf_RD0?t=2619；圖 2 右〔おかえりしろ〕TORA-CON ニューヨークから今日帰国ッ！〕【影音】，kson ONAIR，2024 年 3 月 18 日，取自 <https://youtu.be/uFbao1VYWdQ?t=333>

缺乏明確框架設定訊息會有什麼問題？展演這類後設活動的麻煩之處在於，它看起來就和它的表面意義一樣，容易使參與者感到混淆。演出的殺人像殺人，難道不是真殺人？類似的情況還有，嬉鬧像打架，難道不是真打架？說笑像言語攻擊，難道不是真的言語攻擊（參見 Bateson, 1987）？更糟糕是，人們會用嬉鬧和說笑為藉口，來掩飾真正的打架和言語攻擊（參見 Goffman, 1974, pp. 195-197）。如果缺乏明確框架設定訊息加以區分，在演出／玩耍／開玩笑（playing）這類後設活動中，很容易假戲真作玩過火。傳統劇場有物理空間的約束，只有少數觀眾能近距離接觸展演者，相較之下，串流直播放大了直播主與觀眾的接觸，讓直播主更「近」地暴露在數目無上限的觀眾面前。再加上網路匿名的特性，使得雙方關係更加失衡。直播主的應對只要稍有不當，就可能引發觀眾的炎上（framing，網路用語，透過網路發布攻擊式訊息的行為）。擔心觀眾流失的直播主，在觀眾「開玩笑」時通常選擇隱忍。例如有身心弱勢的直播主表示，當遇見粉絲的殘酷留言時，除了告誡自己要明白能夠承受的極限在哪裡，還得有「厚臉皮」（thick skin, Johnson, 2018）。有女性直播主表示自己曾經被觀眾當作情色女主播（cam girl），要求進行「私人演出」，有類似處境的性少數直播主也表示，要應對這類狀況得要有「厚臉皮」（Freeman & Wohn, 2020）。

若說前述直播主的所謂「厚臉皮」只是隱喻，那麼 VTuber 化身就是能夠保護直播主的真正「厚臉皮」了。一方面，化身非現實的外觀能夠放大觀眾對其所中介的角色與演出並非現實的知覺；另一方面，化身透過遮蔽直播主的實際面貌，賦予直播主與觀眾平等的匿名性，這種匿名性不僅保護了直播主的身分隱私，還讓直播主能更好地管理表情勞動。這使得化身在觀眾玩過火時能起到良好的緩衝作用。像是本文先前所舉的例子，那位 Holo 成員在面對不無性騷擾意味的觀眾留言「就算

[你]會放屁，我也一樣喜歡」時——這種「性騷擾」與透過中之人肉體線索確認 VTubing 本真性是一體兩面——以拉伯雷式幽默加以化解。在化身中介角色的保護下，這種以 VTuber 而非中之人身分的自嘲，賦予這位 Holo 成員更好地應對觀眾「開玩笑」的能力。萬一還是有觀眾玩過火，根據「VTuber 就是 VTuber」的慣例——如前所述，這種慣例延伸到直播外，中之人的實際身分通常不會正式公開——化身也能起到替身的作用，讓攻擊集中在 VTuber 而不是中之人身上。例如某位 Holo 成員向觀眾這樣解釋他無將限期休假的理由與方式。注意，這段話是以 VTuber 而非中之人身分說出來的：

現在說的都是シオン[Sion]我個人的想法……[要無限期休假的原因是]對網路上……和自己相關……的負面消息……透過社交媒介流傳，看到這些覺得很難受。我很喜歡這份工作……絕對不想變得討厭……想說先拉開點距離，這是主要原因……這需要滿長一段時間……因為是以積極的心態去休息的，不會沒有通知就消失，大家不必擔心（紫咲シオン，2023 年 12 月 22 日，YouTube 直播）。

正如 Bakhtin (1965/Iswolsky, Trans., 1984) 描述的狂歡節並非為了顛覆社會，而是為了調節被規範化的日常生活所壓抑的旺盛肉體生命力。作為狂歡節面具 (carnival mask) 的化身，讓展演者在得以擺脫日常身分狂歡演出的同時，也能順利重回日常生活。

在直播中，化身作為強化框架設定訊息的優勢，還可以透過它與美顏濾鏡 (beauty filter) 的差別突顯出來。單就技術而言，化身與濾鏡同屬 AR，都能用來改變使用者原本外貌。但是兩者分屬不同實踐與文化脈絡，化身隱藏使用者的真實身分；而美顏濾鏡則修飾並強化真實身

分。因此，使用者對它們也會有不同的審美期待。即使所呈現的都是直播演出的角色，VTubing 的觀眾透過化身期待看到的是幻想角色（參見 Jiang et al., 2023），而真人直播的觀眾透過美顏濾鏡希望看到「真實」的直播主。用 Bolter & Grusin（2000）的術語來說，化身所強化的是媒介的「超中介」（hypermediacy）邏輯，讓觀眾清楚意識到他們所看到的是經過媒介處理的內容。而濾鏡所強化的是媒介的「去中介」（immediacy）邏輯，透過隱藏媒介的存在，讓觀眾產生他們所看到的就是現實的錯覺。因此，雖然就技術而言，濾鏡也能呈現一眼就可看出是虛構的誇張形象，真正受使用者青睞還是「美顏」濾鏡（Ryan-Mosley, 2021）。

由於使用者——無論展演者或觀眾——的不同期待，美顏濾鏡中介的直播也展現出不同於化身中介直播的審美型態。在中國的直播平臺上，容貌不出色的女直播主會被要求開美顏濾鏡，需要隨時注意「自己的」容貌不能有被一眼識別的「瑕疵」。在這種演出型態中，使用美顏濾鏡往往被視作直播主願意爭取男性觀眾性注意的框架設定訊息。而在這種被男性觀眾期待或可建立「私人關係」的情況下，直播主之間仍然不得不競爭這些觀眾的喜愛（王怡霖、許瑩琪、董晨宇，2023）。相較美顏濾鏡的問題，使用化身展現出這樣的優勢：在同樣對女性直播主不友善的直播環境中，化身能幫助直播主擺脫容貌焦慮，其所提供的匿名性能夠降低性騷擾和性別歧視的發生，還能降低討論敏感話題的風險，從而促進直播主與觀眾之間更坦誠的對話（Wan & Lu, 2024）。

綜上所述，化身作為面具的強化，能促使入戲，防止假戲真作，有效減輕直播主應對觀眾時的心理負擔。這些都是在串流直播生態中，使用化身的 VTubing 相對於非蒙面直播所具有的優勢。

圖 3：前現代的面具與從未現代過的面具



資料來源：圖 3 左：“*Masks and masking: Faces of tradition and belief worldwide* (p. 59)” by G. Edson, 2005, Sage；圖 3 右：“百鬼あやめ”，上網日期：2025 年 3 月 18 日，取自 <https://hololive.hololivepro.com/talents/nakiri-ayame/>

二、VTubing 在當前媒介生態中的文化意義：技術泛靈論的視角

雖然可以從其展演功能在直播生態中的優勢來解釋「為何要用化身演出」，但這個問題之所以是問題，更可能源自提問者在審美知覺上感到化身很怪異。通常是日式動畫風格的化身作為一種強化框架設定訊息，固然能讓習慣這種風格的觀眾迅速入戲，但同時也能讓不習慣者難以入戲。這種現象不只存在 VTubing 中。在前現代無文字或低識字率的文化中，面具是用來吸引注意力、操控知覺與表情的重要媒介。面具代表了現實世界中的非現實存在，或作為強化框架設定訊息將觀眾帶入非現實世界。因此，面具的外觀如同魔法，既現實又非現實，其造型取決

於觀眾的視角，並隨不同文化而異。例如西非 Igbo 文化用來戲劇化生者和在世亡靈之間的互動，混合了人與動物特徵的 Mgbedike 面具（圖 3 左）。若觀眾由於文化中心主義而不理解面具背後的意涵，就難以欣賞這類與自己習俗不同的人們所創造的面具（Edson, 2005; Meineck, 2018, pp. 66-74）。

隨著工業化與技術創新所帶來的現代化進程，對於超凡世界的信念逐漸遭到除魅（disenchantment）。在這樣的文化與歷史脈絡中，儘管蒙面戲與其姊妹形式的傀儡戲曾經屬於劇場主流，但自 18 世紀以來，這類演出物的展演形式在西方逐步被所淘汰／邊緣化（參見 Bell, 2014, 2001）。不僅在西方，隨著現代化的全球擴散，許多非西方的蒙面與傀儡戲傳統也逐漸沒落。如前所述，技術能供性理論指出，對於媒介化實踐可能性的直接知覺來自學習（Kaptelinin & Nardi, 2012），因此，對於已經習慣西方現代主流媒介文化的觀眾而言，VTubing 與其他前現代蒙面戲一樣會在審美知覺上引發他們的怪異感也合乎情理。

因此要從審美知覺的角度來解釋「為何要用化身演出」，需要把這個問題放在更大的文化與歷史脈絡中來考察。VTubing 不僅是 ACG 文化的產物，更與日本傳統劇場文化密切相關。與其他非西方國家相比，日本更好地保留了包括劇場在內的傳統文化。源自江戶時代（17 至 19 世紀）的日本傳統蒙面戲「歌舞伎」與傀儡戲「人形淨琉璃」時至今日仍然是大眾娛樂，並且被認為是 ACG 文化的主要根源之一（Okuno, 2014; Suan, 2013）。Bredikhina 則直接主張，作為 ACG 文化產物的 VTuber，其反映理想而非現實的演出型態植基於歌舞伎，而 VTubing 是「人形淨琉璃風格的現場[live]娛樂的虛擬或數位形式……是日本傳統藝術的新形式」（2021, p. 22）。對於 VTubing 的審美知覺就是在這樣的媒介文化中養成的。

這種觀點雖有道理，但無法解釋其他文化的蒙面戲傳統，為什麼沒有演化出這類甚至能全球化的大眾化創新形式。本文在這裡提出的替代解釋方案是，這是因為 VTubing 不單只是新媒介形式的蒙面或傀儡戲，還反映了日本技術泛靈論 (techno-animism) 的媒介文化。Jensen & Blok (2013) 指出，作為日本本土知識系統的神道 (Shinto) 認為，自然界與人間遍布著各種被稱作神 (kami) 的非人存在 (non-human beings)，它們具有與人類似的情感與行動。這種與其他前現代文化類似的泛靈論信念，在日本的現代化過程中並未消失，而是轉化為被 Allison (2006) 稱作「技術泛靈論」的新信念。技術泛靈論模糊了人與非人、文化與自然的界線，認為作為非人的技術也具有與人相同的能动性 (agency)，這與現代西方認為文化與自然二元對立的信念大為不同。¹²

技術泛靈論在日本媒介文化中的體現之一，就是偏好讓非人能夠像人一樣活起來各種演出物的創新形式。例如對人形淨琉璃的愛好 (Boyd, 2020)、對人形機器人的積極接受 (Haring et al., 2015)，甚至為無法修復的機器人舉辦慰靈儀式 (Jensen & Blok, 2013) 等。對動畫角色的迷戀當然也是，更甚者，日本動畫作品中有各式各樣人與非人的混合體在活動 (參見 Allison, 2006; Brienza, 2013)，這類主題或可稱作「幻想蒙面與傀儡戲」，像是小女孩使用魔法道具變身成幻想角色，或是主角駕駛巨大機器人等。而作為這個主題的延伸，正如本文先前描述 VTubing 前身時就已經指出的，自 ACG 文化中衍生出來的虛擬偶像，

¹² 讀者可能注意到，技術泛靈論與 Latour (1991/1993; 2005) 的行動者網絡理論「社會是由彼此對等的人與非人行動者相互建構而成」之間的相似之處。確實如此。Jensen & Blok (2013) 認為技術泛靈論提供了一個有價值的視角，可以挑戰和擴展行動者網絡理論。

以及作為其後繼形式的 VTuber 於是被創新出來。這也可以解釋為什麼 VTuber 的化身造型充滿各種人與非人混合體。例如，一位 Holo 成員的造型融合了美少女與動物特徵，其設計元素來自日本傳說妖怪「鬼」（oni）。他戴著一個名為「般若」（hannya）的鬼面具，這原本是日本傳統能劇中用來表現女性超自然化的嫉妒與憤怒情感的面具。這類怪誕風格反映了本文先前所述「面具是通往超凡世界的媒介」這種前現代信念的技術泛靈論詮釋（見圖 3 右）。

綜上所述，本文主張，於蒙面展演在現代主流媒介文化中屬於邊緣小眾的脈絡中，技術泛靈論能夠解釋為何不同文化群體對於化身的審美知覺迥異，並連帶解釋了 VTubing 為什麼是在日本興起。

三、VTubing 全球化的媒介生態演化解釋

技術泛靈論只解釋了 VTubing 作為一種創新演出物的展演形式，為什麼會在日本出現以及被日本觀眾的審美知覺所接受，卻不能解釋它為什麼能突破地域與文化限制而（相對）被全球主流媒介文化接受，如果技術泛靈論是日本特有文化的話。要解決這個循環論證的問題，需要回到本文先前討論的，化身的展演功能在直播生態——這個以 YouTube 與 Twitch 為主要平臺的生態——不是技術泛靈論的產物中具有優勢來解釋。同樣使用化身，原本以製播影片為主的 VTuber，是在轉向 VTubing 之後才開始爆紅（PIVOT 公式チャンネル，2023 年 12 月 2 日）。這在事實上支持，使用化身在直播生態中確實有其優勢。

因此，我們可以綜合技術能供性與文化偏好的解釋，來更全面地理解 VTubing 的全球化過程，以回答本文的核心問題「為何要用化身演出」。本文以下基於廣義媒介生態學方法論，借用生態學隱喻來描述

VTubing 的整個發展過程。這個隱喻有助於理解 VTubing 如何在特定媒介文化中產生，並在直播生態中成功占有一席之地，甚至進一步為全球主流媒介文化所重拾。

首先，在全球主流媒介文化中，蒙面展演屬於邊緣，但在日本技術泛靈論中仍屬主流。這種媒介文化特別偏好擬人化技術，如傀儡、動畫角色、虛擬偶像等各種類型的演出物。其次，作為受到在地歡迎的演出物的創新形式，使用化身演出的 VTuber 在日本產生。其三，VTuber 演化出 VTubing 的展演形式，迅速地適應以提供全球服務的 YouTube 為主要平臺的串流直播生態。這個生態本身不是技術泛靈論的產物。其四，Covid-19 的防疫封控使得全球對非接觸式娛樂的需求暴增，擴大了直播生態的環境空間。藉此機遇，VTubing 在乘著日本動畫全球大擴張浪潮的同時，挾其相對於非蒙面直播的優勢，快速搶占更多全球化直播生態的環境空間，成功突破原本的地域與文化限制。其五，擴大了種群的 VTubing，在從業者努力下，繼續開發出一些新型的審美型態變體（variant），像是 Hololive 所開發 VTuber 與日式偶像的綜合，並獲得進一步成功。終於，作為一種創新的蒙面與傀儡戲，VTubing（相對地）被全球主流媒介文化所接受。

綜上所述，使用化身的 VTubing，其全球化的成功並非僅取決於所具有的直播展演功能或參與者的媒介文化偏好，而是技術、實踐、文化以及偶然因素，即 Covid-19 的防疫封控，動態地互相依賴且共同作用的結果。

陸、結論

以下從研究結果以及研究限制與未來展望二方面，總結本文。

一、研究結果

從探詢 VTubing 在當前媒介生態中的意義出發，本研究基於廣義媒介生態學方法論、技術能供性與展演理論，以 Holo 成員的直播為例，透過描述 VTubing 作為媒介化展演的三個層次，包括：（1）媒介技術，如即時動作捕捉驅動的化身；（2）媒介化展演，如 VTubing 本真性的審美型態；（3）文化與歷史脈絡，如 VTubing 的技術泛靈論根源，回答以下問題：在當前媒介生態中，化身、直播主與觀眾如何以及為何共同參與 VTubing 這種蒙面展演的創新形式，從而使邊緣與小眾的蒙面展演能重回主流媒介文化。

本研究首先指出，除了使用化身以及伴隨而來的特殊審美型態，例如對於 VTubing 本真性的悖論追求之外，VTubing 與普通直播並無明顯差異。這並不足以說明，為什麼 VTubing 是直播參與者的一個好選擇。為了深入解釋，本研究進一步從媒介生態演化的觀點進行分析。

首先，本研究從技術能供性的角度出發，探討了 VTubing 在當前直播生態中，相對於非蒙面直播的優勢。化身作為強化的框架設定訊息，具有多重展演功能：它不僅能強化觀眾的入戲感，還能讓觀眾更明確地感知「VTubing 是虛構的」，避免觀眾混淆了演出與現實。例如，VTubing 常用的日式風格化身，其誇張的五官比例能讓觀眾清楚辨識出其虛構本質。相較於非蒙面直播，這讓 VTubing 能夠有效降低直播中「玩過火」的情況。另一方面，化身透過遮蔽直播主的實際面貌，在觀眾「玩過火」時能起到良好的緩衝作用。這種保護不僅體現在身分隱私的保護，還包括表情勞動的管理上，能有效降低直播主在面對觀眾時的心理負擔。這一優勢在處理性騷擾等敏感情況時尤為明顯，化身讓直播

主能以更幽默、更從容的方式化解可能的衝突。

其次，本研究從媒介文化的角度出發，闡明了 VTuber 源自日本特殊的技術泛靈論媒介文化。在現代化過程中，許多國家的蒙面展演已經邊緣化與小眾化，但在日本則否。在日本，前現代的泛靈論信念在現代化過程中轉化為「技術泛靈論」，體現在對於演出物的偏好上。基於這種媒介文化，日本不斷透過新媒介技術創新演出物：從傳統的人形淨琉璃，到現代的動畫角色、虛擬偶像，最終演化為 VTuber。這解釋了為何 VTuber 在日本興起，以及為何不同文化背景的觀眾對 VTuber 的審美知覺會有差異。VTuber 不僅重拾了面具作為通往超凡世界媒介的傳統功能，更擴充了這種可能性，創造出一種能夠適應全球化的創新展演形式，即 VTubing。

最後，本研究整合了技術能供性與媒介文化的觀點，說明 VTubing 在當前媒介生態中的演化過程，以解釋其全球化的（相對）成功。研究指出，VTubing 的發展經歷了幾個關鍵階段：首先，VTuber 源於日本特殊的技術泛靈論媒介文化。隨後，藉其展演功能上的優勢，VTubing 成功適應以提供全球服務的 YouTube 為主要平臺的直播生態。最後，COVID-19 防疫封控帶來的暴增需求，為 VTubing 提供了全球化的機遇，讓它能在搭上日本動畫全球擴張浪潮的同時，還能利用前述展演功能的優勢，快速搶占全球直播市場份額、擴大影響力，使得作為蒙面展演創新形式的 VTubing，突破了原本的地域與文化限制，成功全球化。

綜上所述，本研究揭示了使用化身的 VTubing 不單只是技術創新的媒介化實踐或文化產物，還是人與非人對等互動所建構的複雜現象。其成功取決於媒介、實踐、文化、加上防疫封控的偶然因素，動態交織的結果。

另一方面在理論貢獻上，本研究所建立的分析架構，揭示了媒介、

展演與文化在媒介化展演中的動態交織關係。這個架構不僅有助於理解 VTubing 現象，更為研究其他創新形式的媒介化展演提供了方法論參照。

總結而言，本研究不僅對於 VTubing 研究做出貢獻，其分析架構與研究發現也對其他類型媒介化展演的研究具有啟發意義。

二、研究限制與未來展望

如前所述，由於研究問題與資料取得的需求，本研究主要以 Holo 成員的直播為例，聚焦於目前在全球市場上最成功的 VTubing 審美型態。然而這樣的選擇也意味著本研究忽略了諸多相對邊緣的其他型態，這限制了對 VTubing 更全面的理解。而在那些相對邊緣的型態中特別值得注意的，是因應全球在地化而產生的變體。以下簡要討論它們。

全球在地化（glocalization），亦即在全球化過程中，根據不同地區的文化、語言與市場需求，調整產品與服務，是 VTubing 產生變體的主要因素。以 Hololive 為例，其母公司 COVER 的 CEO 即表示，拓展海外市場最大的挑戰之一，在於如何在本土業者競爭下適應在地市場（PIVOT 公式チャンネル，2023 年 12 月 2 日）。而在這種全球化與在地化拉扯的市場環境中，本土 VTuber 業者一方面借鑑日式 VTubing 的成功模式，另一方面也積極將其演出融入在地媒介生態。這種嘗試從企業勢到個人勢都有所體現。

在企業勢方面，例如已經相當成功——截至 2024 年 9 月近 80 萬訂戶——的韓國 VTuber 團體「Plave」（플레이브），就是基於與日本流行文化在國際市場上分庭抗禮的韓式流行音樂（K-pop）與網路漫畫（Webtoon，웹툰），形成迥異於日式 VTubing 的韓式變體。他們的化

身是韓式漫畫風格，演出型態是經典韓式偶像的擴充，主要吸引 K-pop 粉絲更勝於 ACG 阿宅（參見플레이브，2024 年 6 月 27 日，YouTube 直播）。在個人勢方面，例如臺灣 VTuber「高雄林小姐」——截至 2024 年 9 月有 10 萬訂戶——的臺式變體 VTubing，化身不同於日式動畫風格，演出型態也幾乎與日式 VTubing 無涉，而是臺灣在地廣播與電視政治評論節目的擴充，所吸引的觀眾更不是 ACG 阿宅，而是臺灣政治關注者。對於熟悉臺灣老派政論節目的人來說，「高雄林小姐」這個名字能讓他們會心一笑，因為在這類節目中，主持人慣用「地名+姓氏+先生/小姐」的形式來稱呼 call-in 進節目的觀眾（參見高雄林小姐，2024 年 9 月 5 日，YouTube 直播）。

以上二例代表全球在地化下 VTubing 變體的不同發展路徑。Plave 展示如何在既有的強大在地媒介文化與產業基礎上發展變體。這種方式能夠快速整合資源，創造出符合主流市場需求的高品質內容；與此相對，高雄林小姐則代表了如何在更邊緣的在地媒介文化基礎上發展變體，這種方式雖然缺乏大規模資源支持，但能創造出更草根的內容，並在特定小眾市場中取得成功。這些多元演化的在地實踐不僅顯示 VTubing 的全球適應能力，更預示隨著其與在地媒介生態的持續交織，可能繼續塑造 VTubing 的新格局。正如 2019 年 VTubing 進入國際市場初期，彭博社的評論指出：VTuber 至今幾乎完全是日本現象，但其背後的核心技術以及透過增加互動性提高通俗文化可信度的公式，具有普世性（Alpeyev & Furukawa, 2019, September 18）。基於此，未來研究或可聚焦於 VTubing 全球在地化過程中產生的變體，分析這些變體在不同媒介文化中的演化過程。這不僅有助於更全面理解 VTubing 現象，也有助於揭示它們在未來媒介生態中可能扮演的重要角色。

參考書目

- AU TW TG (2023 年 7 月 4 日)。〈【Hololive】現場體驗洛杉磯 EN 演唱會！海外粉絲熱情無限，Yagoo 還驚喜現身合照！〉【影音】。
https://www.youtube.com/watch?v=knjwIQOy6Kg&list=FLy06XQNkctcKu0xV3NSCN2jA&index=4&t=465s&ab_channel=AU_TW_TG
- 王怡霖、許瑩琪、董晨宇 (2023)。〈操縱“美麗”：中國秀場直播平臺中的美顏工具及其關係性使用〉，《婦女研究論叢》，4，116-128。
- 高雄林小姐 (2024 年 9 月 5 日)。〈柯文哲被收押禁見！王世堅不用跳了 | 姚立明任司法副院？法界：著作乏人引用〉【影音】。
<https://www.youtube.com/watch?v=vIjDOYQffvU>
- 森平崇文 (2019)。〈女團偶像商業模式的傳播——從「秋葉原」到「上海」〉，《戲劇學刊》，29，87-99。
- kson ONAIR (2024 年 3 月 26 日)。〈【おかえりしろ】TORA-CON ニューヨークから今日帰国ッ！〉【影音】。
<https://youtu.be/uFbao1VYWdQ?t=333>
- kson ONAIR (2024 年 3 月 20 日)。〈[実写]〉【影音】。
https://youtu.be/PuA8oRf_RD0?t=2619
- PIVOT 公式チャンネル (2023 年 12 月 2 日)。〈【VTuber ビジネス、爆速成長の秘密】〉【影音】。
<https://youtu.be/u3OESVB0thY?si=LUU4SGWD4rMV0rk>
- PIVOT 公式チャンネル (2023 年 12 月 3 日)。〈【VTuber の次はメタバース・ゲームを攻める】〉【影音】。
<https://youtu.be/EP4W65RuaJA?si=umruiLmMsey65bWM>
- PLAVE_플레이브 (2024 年 6 月 27 日)。〈[4 PD]〉【影音】。
<https://youtu.be/eJ65biGZAOW?t=89>
- VSTATS・ジェネリック集計人 (Holo データ分析) (2024 年 2 月 6 日)。〈Vtuber 業界地図 2023 年総集編〉，《VSTATS blog》。
<https://note.com/vstats/n/ndd38f56e99d2>
- さくらみこ (2018 年 8 月 28 日)。〈【神回】初生放送★さくらみこ 1000 人突破記念！夏最後の花火大会【#01】〉【影音】。
<https://www.youtube.com/watch?v=iaih7CGee8A&t=63s>
- さくらみこ (2024 年 3 月 24 日)。〈【新幹線 0 号】緊急で動画回してるんですけど…異変があります…〉【影音】。
<https://www.youtube.com/watch?v=HVSq3sR9tBc&t=483s>

- 大神ミオ（2022年1月24日）。〈【#朝ミオ】ちょっとお久しぶりのおはみおーん!!!風邪なおった〜〉【影音】。https://youtu.be/Yuck9sBGjCU?t=2448
- 天音かなた（2021年4月6日）。〈#フリー素材かなたそ で作品募集した結果、とんでもないクオリティのものがきた〉【影音】。https://www.youtube.com/watch?v=lGgk6RpbGFE&t=131s
- 犬山たまき／佃煮のりお（2020年9月4日）。〈【桐生ココ】任侠ドラゴン襲来!?世界一の女になったって本当?たま×ココ地獄対談!!〉【影音】。https://youtu.be/gosliffuKQw?t=2475
- 兎田ぺこら（2021年5月25日）。〈【ドラクエ35周年記念】モンスター35体仲間にする!!!ぺこ!DAY8〉【影音】。https://youtu.be/ulhc-rUJ94?t=1758
- 角巻わため（2020年1月2日）。〈【配信できなくても】謝罪動画 full.ver【投稿ならできる】〉【影音】。https://www.youtube.com/watch?v=h6IwhTuKSRk
- 角巻わため（2021年1月4日）。〈【マリオカート8DX】大会に向けて練習だ!!!〉【影音】。https://www.youtube.com/watch?v=miqsFPzwaew&t=17227s
- 角巻わため（2020年9月10日）。〈【朝活】一緒にごはん食べて、つのまきじゅんけんしよう!〉【影音】。https://www.youtube.com/watch?v=KOilp0IaamE&t=15865s
- 赤井はあと（2021年2月7日）。〈#FREEHAACHAMA〉【影音】。https://www.youtube.com/watch?v=8JjjscSWLi0&t=1625s
- 桐生ココ（無日期）。【Kiryu Coco YouTube Channel】。上網日期：2024年9月1日，取自 https://www.youtube.com/@KiryuCoco
- 常闇トワ（2022年4月16日）。〈【APEX】Startend、明日は本番。今日が最後の練習と前夜祭!〉【影音】。https://www.youtube.com/watch?v=SBvnsWCGwT8&t=6570s
- 猫又おかゆ（2023年6月7日）。〈【げりら歌枠】深夜に歌いたくなるとき、あるよね〉【影音】。https://www.youtube.com/live/byoEKpVP0wo?t=4117s
- 紫咲シオン（2023年12月22日）。〈お休みについてのお話。〉【影音】。https://youtu.be/zOTxwiOsO50?t=187
- 獅白ぼたん（2024年8月18日）。〈【雑談】最近あった事やってた事とかファンミと4周年のお礼とかもういろいろ今日は喋ろう!〉【影音】。https://www.youtube.com/live/ceNHUKD0N4U?t=3358s
- 播磨谷拓巳（2016年4月28日）。〈アイドルのうんこ問題〉をドルオタに聞く〉，《Buzzfeed》。https://www.buzzfeed.com/jp/takumiharimaya/idol
- 癒月ちよこ（2021年1月13日）。〈【#ちよこつとクッキング】火を使わない白菜のミルフィーユ鍋風作る!〉【影音】。https://www.youtube.com/watch?v

=WbQbJK7u1ys&t=3553s

- Allison, A. (2006). *Millennial monsters: Japanese toys and the global imagination*. University of California Press.
- Alpeyev, P., & Furukawa, Y. (2019, September 18). How virtual streamers became Japan's biggest YouTube attraction. *Bloomberg*. <https://www.bloomberg.com/news/articles/2019-09-17/how-virtual-streamers-became-japan-s-biggest-youtube-attraction?embedded-checkout=true>
- Auslander, P. (2022). *Liveness: Performance in a mediated culture* (3rd ed.). Routledge.
- Baelz, H. (2022, May 16). <APEX> please carry ft. Towa-sama and Mumei-shi. [Video]. <https://www.youtube.com/watch?v=h8GrEGsrSKI&t=5876s>
- Bakhtin, M. M. (1965/1984). *Rabelais and his world* (H. Iswolsky, Trans.). Indiana University Press.
- Barker, A., & Murphy, H. (2020, September 20). YouTube reverts to human moderators in fight against misinformation. *Financial Times*. <https://www.ft.com/content/e54737c5-8488-4e66-b087-d1ad426ac9fa>
- Bateson, G. (1987). A Theory of Play and Fantasy. In G. Bateson (Ed.), *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology* (pp. 138-148). Jason Aronson.
- Bauman, R. (1975). Verbal art as performance. *American Anthropologist*, 77(2), 290-311. <https://doi.org/10.1525/aa.1975.77.2.02a00030>
- Bell, J. (2014). Playing with the Eternal Uncanny: : The persistent life of lifeless objects. In D. N. Posner, C. Orenstein, & J. Bell (Eds.), *The Routledge companion to puppetry and material performance* (pp. 43-52). Routledge.
- Bell, J. (2001). *Puppets, masks, and performing objects*. MIT Press.
- Black, D. (2012). The virtual idol: Producing and consuming digital femininity. In P. W. Galbraith & J. G. Karlin (Eds.), *Idols and celebrity in Japanese media culture* (pp. 209-228). Palgrave Macmillan UK.
- Bolter, J. D., & Grusin, R. (2000). *Remediation: Understanding new media*. MIT Press.
- Boyd, M. (2020). *Japanese contemporary objects, manipulators, and actors in performance* (人形演劇の現在 モノ、モノ遣い、アクター). Sophia University Press.
- Bredikhina, L. (2021). Virtual theatrics and the ideal VTuber Bishōjo. *REPLAYING JAPAN*, 3, 21-32. <https://doi.org/10.34382/00014532>
- Bredikhina, L., & Giard, A. (2022). Becoming a virtual cutie: Digital cross-dressing in Japan. *Convergence*, 28(6), 1643-1661. <https://doi.org/10.1177/13548565221074812>
- Brett, N. (2022). Why do we only get anime girl avatars? Collective white heteronormative avatar design in live streams. *Television & New Media*, 23(5), 451-461. <https://doi.org/10.1177/15274764221080956>
- Brienza, C. (2013). Objects of otaku affection: Animism, anime fandom, and the gods of ... consumerism? In G. Harvey (Ed.), *The handbook of contemporary animism* (pp.

- 479-490). Acumen Publishing.
- Butler, J. (1988). Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre Journal*, 40(4), 519-531. <https://doi.org/10.2307/3207893>
- Chen, M., & Hu, R. J. (2024). The host in the shell: Constructing female identity in the embodied animated media. *New Media & Society*. <https://doi.org/10.1177/14614448231226213>
- Chinchilla, P., & Kim, J. (2024). VTuber for streamers: Exploring the role of social presence in the visual representation of streamers. *Communication Studies*, 75(6), 844-860. <https://doi.org/10.1080/10510974.2024.2337955>
- Dempsey, I. (2024, May 1). VTubers in Q1 2024 report — A new champion topped charts for watch time. *Streams Charts*. <https://streamscharts.com/news/vtubers-q1-2024-report>
- Devlin, D. (1989). *Mask and scene: An introduction to a world view of theatre*. Palgrave Macmillan.
- Edson, G. (2005). *Masks and masking: Faces of tradition and belief worldwide*. McFarland & Company.
- Freeman, G., & Wohn, D. Y. (2020). Streaming your identity: Navigating the presentation of gender and sexuality through live streaming. *Computer Supported Cooperative Work*, 29, 795-825. <https://doi.org/10.1007/s10606-020-09386-w>
- Gadamer, H. G. (1975). *Truth and method*. Seabury Press.
- Geertz, C. (1973). Thick description: Towards an interpretive theory of culture. In C. Geertz (Ed.), *The interpretation of cultures* (pp. 3-30). Basic Books.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Global Information Inc. (2024). *Vtuber (Virtual YouTuber) - Global market share and ranking, Overall Sales and Demand Forecast 2024-2030*. QYResearch. <https://www.gii.tw/report/qyr1559192-vtuber-virtual-youtuber-global-market-share.html>
- Goffman, E. (1974). *Frame analysis : An essay on the organization of experience* Harvard University Press.
- Guarriello, N.-B. (2019). Never give up, never surrender: Game live streaming, neoliberal work, and personalized media economies. *New Media & Society*, 21(8), 1750-1769. <https://doi.org/10.1177/1461444819831653>
- Haring, K. S., Silvera-Tawil, D., Takahashi, T., Velonaki, M., & Watanabe, K. (2015). Perception of a humanoid robot: A cross-cultural comparison. *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 821-826. <https://doi.org/10.1109/ROMAN.2015.7333613>
- Hochschild, A. R. (1983). *The managed heart: Commercialization of human feeling*. University of California Press.
- Holive English. (2023, November 30). **【Concert Performance】Connect The World**

- 【hololive English Original Song】**. [Video]. https://www.youtube.com/watch?v=C_AHQ2Rwx3s&ab_channel=hololiveEnglish
- Hutchins, E. (1996). *Cognition in the Wild*. MIT Press.
- Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Indiana University Press.
- Jenkins, H. (2013). *Textual poachers: Television fans and participatory culture* (2nd ed.). Routledge.
- Jensen, C. B., & Blok, A. (2013). Techno-animism in Japan: Shinto cosmograms, actor-network theory, and the enabling powers of non-human agencies. *Theory, Culture & Society*, 30(2), 84-115. <https://doi.org/10.1177/0263276412456564>
- Jiang, H., Pan, D., & Lu, J. (2023). Better technology, but less realism: The perplexing development and application of VTuber technology. *2023 International Conference on Culture-Oriented Science and Technology*, 168-173. <https://doi.org/10.1109/CoST60524.2023.00042>
- Johnson, M. R. (2018). Inclusion and exclusion in the digital economy: Disability and mental health as a live streamer on Twitch.tv. *Information, Communication & Society*, 22(4), 506-520. <https://doi.org/10.1080/1369118X.2018.1476575>
- Kaptelinin, V., & Nardi, B. (2012). Affordances in HCI: Toward a mediated action perspective. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 967-976). Association for Computing Machinery. <https://doi.org/10.1145/2207676.2208541>
- Kaptelinin, V., & Nardi, B. A. (2006). *Acting with technology: Activity theory and interaction design*. MIT Press.
- Koefficient. (2020, June 29). *My Favorite V-Tuber Just Made Her 3D Debut! Diary Of A Simpefficient Ep. 8*. [Video]. <https://www.youtube.com/watch?v=uMjFEzucpVA&t=345s>
- Latour, B. (1991/1993). *We have never been modern* (C. Porter, Trans.). Harvard University Press.
- Li, Y. (2023). Why does Gen Z watch virtual streaming VTube anime videos with avatars on Twitch? *Online Media and Global Communication*, 2(3), 379-403. <https://doi.org/10.1515/omgc-2023-0030>
- Lu, Z., Shen, C., Li, J., Shen, H., & Wigdor, D. (2021). More Kawaii than a real-person live streamer: Understanding how the Otaku community engages with and perceives virtual YouTubers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445660>
- Lum, C. M. K. (Ed.). (2006). *Perspectives on Culture, Technology and Communication: The Media Ecology Tradition*. Hampton Press.
- Masuda, H., TSudo, T., Koudate, T., Matsumoto, A., Rikukawa, K., Ishida, T., ... Hasegawa,

- M. (2024). *Anime industry report 2023 summary*. The Association of Japanese Animations.
- McLuhan, M., Fiore, Q., & Agel, J. (1996). *The medium is the massage: An inventory of effects*. HardWired.
- McLuhan, M., & McLuhan, E. (1988). *Laws of media: The new science*. University of Toronto Press.
- McLuhan, M., McLuhan, E., & Zingrone, F. (1996). *The essential McLuhan*. Basic Books.
- Mead, G. H., & Morris, C. W. (1962). *Mind, self, and society : From the standpoint of a social behaviorist*. University of Chicago Press.
- Meineck, P. (2018). The remains of ancient action: Understanding affect and empathy in Greek drama. In R. Kemp & B. McConachie (Eds.), *The Routledge companion to theatre, performance and cognitive science* (pp. 66-74). Routledge.
- Mou, Y., Lan, J., & Huang, Y. (2023). Good night versus goodbye? Comparing the mourning remarks of virtual and human uploaders through a data-mining approach. *New Media & Society*. <https://doi.org/10.1177/14614448231212822>
- Mumei, N. (2023, November 4). **【MINECRAFT】 MOOM ON !!! Team White ! #ホロライブ大運動会 2023**. [Video]. <https://www.youtube.com/watch?v=XwN95bpEaX0&t=9799s>
- Neuro-sama. (2023, January 3). *Neuro-sama the AI Vtuber learns to playosu! | Days 1 & 2*. [Video]. https://www.youtube.com/watch?v=q3RiFz1yu5s&ab_channel=Neuro-sama
- Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.
- Okuno, T. (2014). Roots of Cool Japan: From the Japanese Traditional Edo Culture to Anime and Manga. *Kwansei Gakuin University Social Sciences Review*, 19, 1-7.
- Piris, P. (2014). The co-presence and ontological ambiguity of the puppet. In D. N. Posner, C. Orenstein, & J. Bell (Eds.), *The Routledge companion to puppetry and material performance* (pp. 30-42). Routledge.
- Playboard. (n.d.). *Most superchatted YouTube channels worldwide in 2021*. Retrieved November 16, 2024, from <https://playboard.co/en/youtube-ranking/most-superchatted-all-channels-in-worldwide-yearly?period=1609459200>
- Pollock, D. (1995). Masks and the semiotics of identity. *The Journal of the Royal Anthropological Institute*, 1(3), 581-597. <https://doi.org/10.2307/3034576>
- Postman, N., & Postman, A. (2005). *Amusing ourselves to death: Public discourse in the age of show business*. Penguin Publishing Group.
- Proschan, F. (1983). The semiotic study of puppets, masks, and performing objects. *Semiotica*, 47(1-4), 3-44.
- Ryan-Mosley, T. (2021). Beauty filters are changing the way young girls see themselves. In J. Pooley (Ed.), *Social media & the self: An open reader (1st ed.)*. mediastudies.press. <https://doi.org/10.32376/3f8575cb.d4c8ee48>
- Schechner, R. (2020). *Performance studies: An introduction (4th Ed)*. Routledge.

- Suan, S. (2013). *The anime paradox: Patterns and practices through the lens of traditional Japanese theater*. Global Oriental.
- Tresch, J. (2012). *The romantic machine: Utopian science and technology after Napoleon*. University of Chicago Press.
- Vygotsky, L. S. (1978). *Mind in society: development of higher psychological processes* (M. Cole, Trans.; M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
- Wan, Q., & Lu, Z. (2024). Investigating VTubing as a reconstruction of streamer self-presentation: Identity, performance, and gender. *Proceedings of the ACM on Human-Computer Interaction*, 8, 1-22. <https://doi.org/10.1145/3637357>
- Watzlawick, P., Bavelas, J. B., & Jackson, D. D. (1967). *Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes*. W. W. Norton & Company.
- Wingert, P. S. (2024, February 16). Mask. In *Encyclopedia Britannica*. Retrieved February 16, 2024, from <https://www.britannica.com/art/mask-face-covering>
- Yuan, J. E. (2024). Making virtual celebrity: Platformization and intermediation in digital cultural production. *International Journal of Cultural Studies*, 28(1), 316-331. <https://doi.org/10.1177/13678779241230564>
- Zhang, X., Xiang, Y., & Hao, L. (2019). Virtual gifting on China's live streaming platforms: Hijacking the online gift economy. *Chinese Journal of Communication*, 12(3), 340-355. <https://doi.org/10.1080/17544750.2019.1583260>

From Masks to Avatars: The Dynamic Interweaving of Media, Performance, and Culture in the Evolution of VTubing's Media Ecosystem

Lu-Yen Ko*

ABSTRACT

Research Background and Problem Statement

The COVID-19 pandemic fueled demand for contactless entertainment, spurring the rise of VTubing originating in Japan. By 2024, VTubing had secured a significant share of the global livestreaming market.

This research defines VTubing as a mediated performance where performers employ motion capture technology to control anime-style avatars for livestreaming. While the main distinction between VTubing and traditional livestreaming lies in the use of virtual avatars, this also raises several key questions: Why use avatars for livestreaming? How does the use of avatars affect the performers and audiences of the livestreaming?

These questions carry different meanings for different livestreaming participants. For those familiar with Japanese anime culture, VTubing is a natural extension of their previous media practices; for those unfamiliar, it appears difficult to comprehend. This contrast reveals the close relationship between media and media practices and guides this research to reexamine VTubing from a media ecology perspective.

From the historical and cultural perspectives of media practices, an avatar

* Lu-Yen Ko holds a Ph. D. from the College of Communication at National Chengchi University. E-mail: duress.ko@gmail.com. ORCID: 0009-0004-6065-7749

that conceals the performer's true appearance and identity can be viewed as a new type of mask, and VTubing can be seen as a new form of masked performance. Therefore, VTubing is not an entirely new phenomenon, but rather embodies the retrieval phenomenon proposed in McLuhan and McLuhan's law of media - masked performance, once obsolesced by mainstream media culture, has been extended through new media into VTubing and gained popularity once again.

This research further poses media ecological research questions: Why has masked performance, long obsolesced by mainstream media culture, been retrieved? In the current media ecosystem, how and why do avatars, streamers, and audiences collectively participate in VTubing as an innovative form of masked

Methodological Framework

To address the aforementioned research questions, this research first explores the broader media ecology, proposing a general methodological framework for analyzing mediated practices. This framework integrates multiple related theories, including Ihde's technological phenomenology, Hutchins' distributed cognition theory, Latour's actor-network theory (ANT), and McLuhan's media theory. These theories examine from different angles how technology shapes human perception, practice, and culture, while pointing out how human usage patterns and cultural contexts in turn influence technological development and innovation. These theories share a non-individualistic methodology, emphasizing that human perception and practice are products of interactions between humans, technology, and environment, rather than solely originating from individual intentions and behaviors.

Second, this research explores Kaptelinin & Nardi's theory of technological affordances. It focuses on how to concretely analyze the media level of mediated practices.

Finally, this research synthesizes Bateson, Bauman, and Schechner, discussing performance and meta-communication theories. It also concretely evaluates VTubing as a special form of mediated practice - mediated performance.

These theories collectively form this research's methodological framework, in order to analyze three levels of VTubing: (1) Cultural and historical contexts, describing the existing forms of mediated performance from which VTubing has evolved; (2) Media technology, describing which media technologies VTubing relies on for mediation; and (3) Mediated practice as performance, describing how VTubers utilize these technologies to perform and how audiences appreciate these performances.

Research Methods

As it extends beyond the scope of direct observation, this research's analysis of VTubing's cultural and historical contexts necessarily relies on secondary literature, primarily including relevant industry reports and news coverage. For the analysis of VTubing's media technology and practices themselves, this research employs a non-interventional observation approach in online fieldwork, looking at the livestreams of VTubers under Hololive, a global industry leader. The observation covers 70 Hololive members' streaming content from September 2017 to September 2024. Given the massive volume of streaming content (52,838 hours in 2023 alone), this research adopts an iterative sampling strategy guided by fan-curated highlights, which are then cross-verified with original broadcasts to ensure data reliability.

Research Findings and Discussion

After describing VTubing's three levels through the case of Hololive livestreaming, this study finds that VTubing differs from unmasked livestreaming primarily in its use of avatars and the resulting aesthetic patterns.

To explain why mainstream media culture has retrieved masked performance through VTubing, this research conducts further analysis from a media ecosystem evolution perspective.

First, from a technological affordances perspective, this research explores VTubing's advantages in the mainstream livestreaming ecosystem. Avatars, as enhanced frame-setting messages, serve dual functions: enhancing audience immersion while clearly signaling fictional performance, thus preventing confusion between performance and reality and potential harassment. By concealing streamers' actual appearance, avatars provide both identity privacy and emotional labor protection. These features give VTubing distinct advantages over unmasked livestreaming.

Second, from a media culture perspective, this research clarifies that VTubers originate from Japan's techno-animistic culture. This media culture is reflected in preferences for performance objects. Based on this culture, Japan continuously innovates performance objects through new media technologies, from traditional Ningyo Joruri puppets to modern anime characters, virtual idols, and finally evolving into VTubers. This explains why VTubers arose in Japan and why audiences from different cultural backgrounds have varying aesthetic perceptions of VTubers. VTubers not only retrieve the traditional function of masks as media for accessing transcendent worlds, but also expand these possibilities, creating an innovative performance form adapted to globalization - namely, VTubing.

Finally, by integrating technological affordances and media culture perspectives and employing an ecological evolution metaphor, this research traces how VTubing has evolved into a globally successful phenomenon through three key stages: VTubers emerged from Japan's techno-animistic culture, adapted to YouTube-centered livestreaming ecology through advantages in performance functions, and expanded globally alongside

Japanese animation during a period when the pandemic drove demand for contactless entertainment.

In summary, this research reveals that VTubing is not merely a technologically innovative mediated practice or cultural product, but rather a complex phenomenon constructed through equal interaction between human and non-human actors. Its success depends on the dynamic interweaving of media, practices, culture, and the contingent factor of quarantine measures. This research also establishes an analytical framework of mediated performance, providing methodological reference for studying other innovative forms of mediated performance.

Research Limitations and Future Prospects

While focusing on Hololive's globally successful VTubing aesthetic patterns, this research overlooks other marginal patterns, particularly variants from global localization. Future research can examine these variants to enhance understanding of VTubing and to reveal its potential roles in media ecology.

Keywords: avatar, livestreaming, mask, mediated performance, media ecology, VTubing