

廓清腦機介面的倫理視野：  
變與不變

Ethical Horizons of  
Brain-Computer Interface  
Technology: Changes and  
Constants

王 珏

Wang Jue

Abstract

“The Ethics of Thinking with Machines” (Lyreskog et al. 2023, 11–34) focuses on the ethical issues and debates caused by the

---

王 珏，西安交通大學哲學系教授，中國西安，郵編：710049。  
Wang Jue, Professor, Department of Philosophy, Xi'an Jiaotong University, Xi'an, China, 710049.

基金專案：國家社科基金一般項目《人工智慧醫學應用的倫理框架與治理研究》  
(專案編號：20BZX127)。

《中外醫學哲學》XXI:2 (2023 年)：頁 47–51。  
*International Journal of Chinese & Comparative Philosophy of Medicine* 21:2 (2023), pp. 47–51.

© Copyright 2023 by Global Scholarly Publications.

development of brain-computer interface technology, and describes a changing moral landscape: some boundaries are blurred or broken, while some core values are reinvented. However, a constant concern underlies these changes, namely concern about “who I am.” This concern has become even more pressing in the current age of artificial intelligence.

本期的主旨論文《與機器一起思考：腦機介面技術》(Lyreskog et. al 2023, 11–34) 從技術發展不確定後果出發，聚焦腦機介面技術發展引發的倫理問題和倫理爭論，為讀者描繪了一幅正處於震變中的倫理視域：某些邊界被模糊或者突破，某些核心價值被重塑。如作者所呼籲的，如果生命倫理學家想要在腦機介面技術領域繼續發揮作用，他們應該準備好接受概念方面的劇烈變動，接受對醫療和保健領域中許多核心價值的概念重塑。全文思路開闊，同時又富有預見性，提供了許多值得進一步追問的思想線索。

雖然作者更關注的是變化的方面，然而如果細讀文章的話，不難發現在變動的視野中存在著一些貫通的、不變的東西。對繪製完整道德圖景的任務而言，如何去描摹、定位和把捉這些變化中不變的東西可能是同樣重要，甚至更為艱難的挑戰。作者因為篇幅所限，或者受限於分析的視角，沒有凸顯這些不變的因素，本文願意就此做一些補充探討。下文將圍繞著作者重點分析的三個核心價值——自主 (autonomy)、精神完整性 (mental integrity)、和統一性 (identity) ——展開論述。

腦機介方面臨最主要的倫理爭議是自主性問題。一方面，植入腦機介面的初衷是用來提升或促進病人的自主，但是人們很容易產生直觀疑慮，機器直接干預認知過程是否會影響病人的自主行動能力。這一疑慮又會因為不透明性而變得更加尖銳：腦機介面讓我的認知過程暴露在機器的影響下，但機器的運作方式對我而言不是完全透明的，這就會引起可解釋性問題，也即著名的“黑箱問題”。作者建議通過拓寬自主性概念的範圍來解決這個疑慮，將自主性概念轉變為一個外部化的、實踐上可解釋的範疇：

即，如果我們能夠有效合理化我們的選擇，並將之關聯於我們的意願，我們就可以把這些選擇看作是自主做出的，儘管機器介入認知的內部過程對我而言依然是黑箱狀態。

但是這一概念拓展也是有前提的，或者說是有邊界。如作者所說的，能否實現這一拓展很大程度上取決於我們如何理解真實或本真 (authenticity)。由此關於自主性的討論就很自然地關聯到文章重點探討的第二個價值：精神完整性。精神完整性指的是一個人精神生活的整全和融貫，與自主性、本真性、心理連續性等概念都有重疊之處。自主性概念更偏重是理性思考和自我規劃生活道路的能力，而精神完整性概念則偏重主觀的自我感。作者以一個假想的情緒調節的案例來說明兩者之間的差別。腦機介面可以以使用者無法察覺的方式來監控腦部活動，相應調節情緒和激素水準，從而減少使用者的焦慮、抑鬱症狀。這種情況並不會損害使用者的自主性，甚至隨著負面情緒的減少，她的理性思考和控制能力還會有明顯的提升。但是在這種狀況下，使用者卻很有可能遭受內在自我感的削弱，她可能會覺得現在這個情緒更加樂觀的自己不是自己，感覺這個新生的自我是難以接受的陌生人。對自主性和精神完整性的探討總是會自然延伸到對人格同一性問題的探究，並被表述為如下問題：腦機介面會否在某種更為基本層面上導致人格的解體？問題由此而被導向更為純粹的哲學討論。

由自主性，到精神完整性，再到人格同一性，我們不難看出有一條隱隱約約貫通的線索，即對“我是誰”的關切。這種關切在我們身處的人工智慧時代變成更為迫切的思想任務。如許煜指出的，“在資訊機器或控制論式機器時代，機器本身成為了資訊組織者，而人類就不再是中心了，即使他們仍把自己當成機器的指揮者和資訊的組織者來考慮。這就是人類受自己對機器的刻板類型的信念之苦的一刻：他們誤把自己指認為中心，而在這樣做的時候，就面臨著持續的挫敗感和對身份的恐慌的追尋。” (Hui

2023) 在人工智能時代追問“我是誰”的問題，就是在被技術急劇更改的現實中尋求人類得以安身的位置，以試圖保存一些支撐著我們成為我們的人類現實。《與機器一起思考》很好地描繪了變化的那一面，但不變的那一面可能同等重要，甚至更為重要。對“我是誰”的不變的關切構成了決定技術未來走向何方的重要指標，成為變化中規約變化的隱形邊界。

從這個視角看，該論文對人格同一性問題的回答就顯得還不夠徹底。該文主張腦機介面時代我們需要接受一種擴大了的人格同一性概念，自我和同一性的邊界不應該限定在直接的物理或心理的連續性上，而是允許將一些輔助思維的外在工具都看作是自我心智的某種延展。然而我們還需要進一步追問，在這種拓展中人機之間的關係的實質。人機關係可以有不同的走向，正如論文題目中的“與機器一起思考”(thinking with machines) 本身就可以有歧義的解釋。“與機器一起思考”的一種可能是機器服務於人，解放人的潛能；還有一種可能就是機器代替人思考，或者更準確的說，機器會悄悄改變人對思的定義和思的經驗。如果我們想要讓人工智能技術沿著前一條路徑發展，將技術現實整合到人類現實中，那麼我們就很有必要持續不斷追問“我是誰”的問題，並將之建立為變動的道德視野的不變的邊界。

## 參考文獻 References

- 萊瑞斯科、佐赫尼、辛格、薩烏萊斯庫：〈與機器一起思考：腦機介面技術〉，《中外醫學哲學》，2023年，第XXI卷，第2期，頁11-34。
- David M. Lyreskog, Hazem Zohny, Ilina Singh, Julian Savulescu. “The Ethics of Thinking with Machines: Brain-Computer Interfaces in the Era of Artificial Intelligence,” *International Journal of Chinese & Comparative Philosophy of Medicine* 21, no. 2 (2023): 11-34.
- Hui, Yuk. “ChatGPT, or the Eschatology of Machines,” e-flux, no.137, June 2023.

<https://www.e-flux.com/journal/137/544816/chatgpt-or-the-eschatology-of-machines/>