

中小學教師通用思維教學技能測評工具的開發與驗證

Development and Validation of Teaching Thinking Assessment Tool

曹舒旻¹, 王寸尺^{2*}

¹ 珠海高新區永豐小學

² 北京師範大學教育技術學院

*roy.wang@live.com

【摘要】 培養通用思維技能是 21 世紀教育改革的關鍵，教師的通用思維教學技能水準對學生思維培養至關重要。然而目前評估教師通用思維教學技能的工具匱乏，限制了其進一步提升。本研究基於 TPACK 理論，開發教師通用思維教學技能評估工具。通過 Rasch 模型檢測顯示試題內部擬合度和外部擬合度 MNSQ 均接近理想值 1，信度檢驗分析結果符合相關統計學指標，表明使用此測評工具對教師通用思維教學技能知識進行評估，其評估結果具有可靠性，可為教師通用思維教學技能的專業發展提供有力支持。

【關鍵詞】 通用思維教學技能；教師思維教學測評；TPACK 理論

Abstract: This study develops an assessment tool for teachers' general thinking teaching skills based on TPACK theory. It extracts indicators from 24 demonstration videos and uses the Delphi method to create a theoretical framework and draft questions, which are revised by experts and teachers. The Rasch model tests 121 data sets, showing reliable assessment results.

Keywords: General Teaching Thinking Skill; Teaching Thinking Assessment; TPACK

1. 問題的提出

各國教育檔均強調學生需在特定領域掌握知識、技能和思維。相比學科思維，開展通用思維教學可減少資源消耗，加速教育進程。Fisher 提出教師應系統引導學生運用通用思維技能（Fisher, 2013）。通用思維教學技能的評估能幫助教師發現不足，明確改進方向。因此，開發高質量的通用思維教學測評工具，對教師專業成長和教育品質提升意義重大。

2. 文獻綜述

通用思維技能（General Thinking Skill）由 McPeck 最早提出，指適用於多種情境和學科領域的普遍思維技能（Moore, 2004）。Marzano 通過研究美國多份課程相關檔，發現多數檔對學生有推理、決策等相似要求，為通用思維的存在提供了有力證據（Marzano, 1998）。目前，通用思維技能被認為包括批判性思維、分析和綜合等廣泛適用的認知技能。通用思維教學技能則是指教師通過課程設計、學習活動安排和教學策略，培養學生這些技能的能力。教師的通用思維教學技能則是提升學生思維能力的關鍵，對其評估是提升其專業能力的必要途徑。目前，思維測評工具多以學生為對象，採用李克特自填式量表，主要測評學習者的傾向、態度和價值觀（Swartz & McGuinness, 2014）。然而，針對教師通用思維教學技能的測評工具多集中於教師的教學傾向和對學生思維的引導，而非直接測量教師自身的教學技能。因此，本研究旨在開發有效的測評工具，以評價教師的通用思維教學技能，具體研究問題包括：

1. 如何開發有效的教師通用思維教學技能測評工具？
2. 通用思維教學技能測評工具的信效度如何？

3. 教師通用思維教學技能工具開發

3.1. 通用思維教學技能測評的過程與方法

1986 年, Shulman 提出“學科教學知識 (PCK)”概念 (Gudmundsdottir & Shulman, 1987)。2001 年, Pierson 在此基礎上提出“整合技術的學科教學知識 (TPCK)” (Pierson, 2001), 將技術知識納入其中。基於 TPCK 模型, 眾多學者開發了教師測評工具 (楊絮, 2023; 陳紅君, 2022)。本研究以 TPACK 模型的核心要素搭建測評工具框架: 一級指標包括思維知識 (CK)、思維教學方法知識 (PK) 和技術支持教學的知識 (TK), 其中 TK 特指思維工具知識。二級指標通過分析 24 節思維教學示範視頻並經 7 位專家兩次修訂, 最終形成測評框架, 如圖 1 所示。

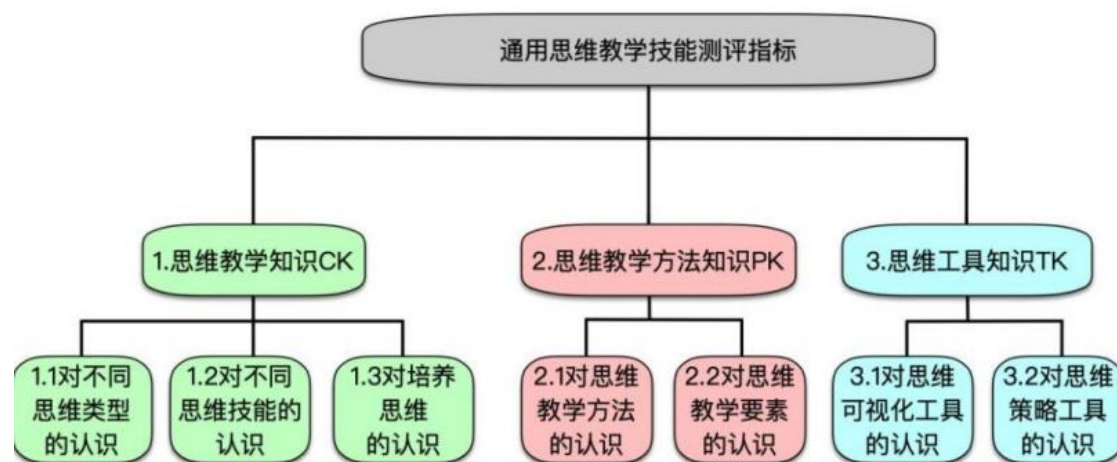


圖 1 基於 TPACK 模型的通用思維教學技能測評指標框架

3.2. 通用思維教學技能測評的題目來源與類型

由於思維技能測評工具主要以客觀題為主 (Facione, 2000), 本研究採用客觀題進行測評, 題目基於思維教學論文和示範課研討內容編寫, 經 7 位專家和 17 名優秀思維教學教師審議, 最終確定 23 道題。以下為示例及說明:

思維導圖和概念圖用途類似, 只是不同的名稱。

正確 B. 錯誤

該題對應“思維工具的知識 (TK)”維度, 考察教師對思維可視化工具的理解。思維導圖和概念圖是新手教師常見的迷思概念, 該題目的在於評估教師能否正確區分二者。

4. 教師通用思維教學技能工具驗證與分析

使用 Rasch 模型檢測測評工具有效性已被眾多研究採用 (Alonzo & Steedle, 2009; 李月明 et al., 2024)。在對 108 名教師 (涵蓋學科語文, 數學, 英語, 物理, 化學) 的通用思維教學技能知識進行測試後, 將數據導入至 Winsteps 軟體, 進行數據分析, 本研究的被試表現如下表 1 所示。

表 1 Rasch 模型整體品質分析

	Measure	S.D.	Infit		Outfit		Seperation	Reliability
			MNSQ	ZSTD	MNSQ	ZSTD		
被試	0.45	0.95	1	0	1.02	0	1.63	0.73
專案	0.00	1.49	0.98	0	1.02	0.3	5.92	0.97

測評結果顯示, 內部擬合度和外部擬合度的 MNSQ 值接近 1, 表明數據與 Rasch 模型擬合良好, ZSTD 值接近 0, 符合理想模型。分離度顯示, 被試能力集中 (1.63), 專案難度區分度大 (5.92)。信度方面, 被試信度為 0.73, 專案信度為 0.97, 均高於理想值 0.5, 表明信度良好。單維性指所有專案是否考察的是同一內容。單維性分析顯示, 大部分題項的因數載

荷在理想範圍 (-0.4~0.4) 內，標準殘差的未解釋率為 10.8%，總解釋率為 60.5%，超過理想值 50%，說明測評工具對通用思維教學技能知識具有較強的解釋能力，如圖 2 所示。

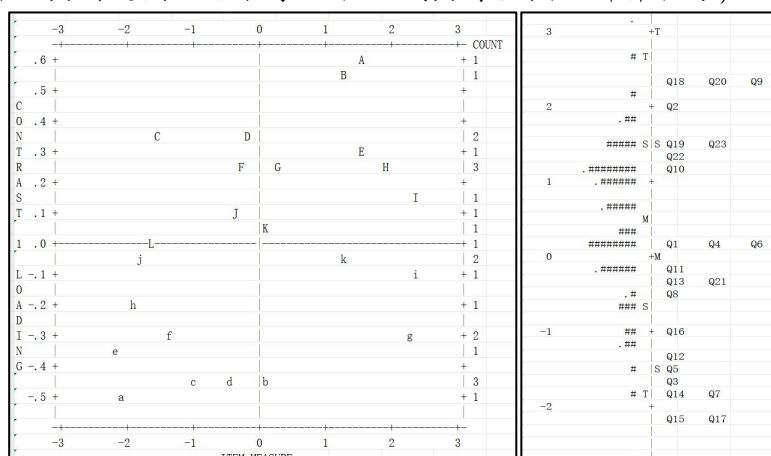


圖 2 通用思維教學技能標準殘差對比

圖 3 通用思維教學技能知識懷特圖

在 Rasch 模型的應用中，懷特圖 (Item-Person Map, 亦稱 Wright Map) 是一種直觀展示測評工具整體水準合理性的工具。本研究的懷特圖如圖 3 所示，我們可以發現教師的通用思維教學技能能力值的分佈呈現出明顯的正態分佈特徵。23 個题目的難度跨度廣泛且分佈均勻，有效地避免了難度集中於某一特定區域的問題。綜上所述，該測評工具的題目設計能夠與教師的能力水準形成良好的對應關係。並證明了該測評工具在評估教師通用思維教學技能知識方面的適用性和準確性。

5. 討論與展望

測評結果顯示，教師對通用思維教學的認識得分最高，但在思維策略工具的應用上得分最低，說明教師雖瞭解通用思維教學的要素，但在實際應用中存在困難。高虹認為這源於教師缺乏相關理論指導 (高虹, 2023)。因此，建議教育決策者為教師提供思維教學應用培訓，以支持其設計和實施高階思維教學活動。此外，研究發現教師的教齡與通用思維技能教學能力無顯著差異，這與其他研究結論一致 (Long et al., 2022)。這可能是由於不同教齡教師的知識結構複雜多樣，導致評估技能時存在較大差異 (Loewenberg Ball et al., 2008)。本研究存在以下限制：樣本量較小 (121 名教師)，可能無法全面代表所有教師的通用思維教學技能水準。另外測評指標並未涉及交叉維度，未來研究可擴大樣本數量，繼續深化測評指標體系，並探索基於測評結果實施教師發展策略的路徑。

總體而言，本研究以 TPACK 理論為基礎，結合思維教學論文、教學視頻分析、專家意見及一線教師回饋，兼顧科學性與實用性。同時，該研究開發的教師通用思維教學技能測評工具展現了良好的信效度，並進一步揭示了教師通用思維教學技能在不同維度上的具體表現，為後續研究提供了更細緻的視角。該工具能夠幫助教育工作者更準確地評估和提升教學能力，為教育管理者制定更具針對性的教師培訓和支持策略提供有力依據。

中小學教師通用思維教學技能測評工具：<https://ffrq.bmwwx.cn/xmThii>

參考文獻

- Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education*, 93(3), 389–421. <https://doi.org/10.1002/sci.20303>
- Facione, P. A. (2000). The Disposition Toward Critical Thinking: Its Character, Measurement, and Relationship to Critical Thinking Skill. *Informal Logic*, 20(1), Article 1. <https://doi.org/10.22329/il.v20i1.2254>
- Fisher, R. (2013). *Teaching Thinking: Philosophical Enquiry in the Classroom*. A&C Black.

- Gudmundsdottir, S., & Shulman, L. (1987). Pedagogical Content Knowledge in Social Studies. *Scandinavian Journal of Educational Research*. <https://doi.org/10.1080/0031383870310201>
- Loewenberg Ball, D., Thames, M. H., & Phelps, G. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, 59(5), 389–407.
<https://doi.org/10.1177/0022487108324554>
- Long, T., Zhao, G., Li, X., Zhao, R., Xie, K., & Duan, Y. (2022). Exploring Chinese in-service primary teachers' Technological Pedagogical Content Knowledge (TPACK) for the use of thinking tools. *Asia Pacific Journal of Education*, 42(2), 350–370.
<https://doi.org/10.1080/02188791.2020.1812514>
- Marzano, R. J. (1998). What are the general skills of thinking and reasoning and how do you teach them?. *The Clearing House*, 71(5), 268-273.<https://doi.org/10.1080/00098659809602721>
- Moore T. (2004). The critical thinking debate: How general are general thinking skills? *Higher Education Research & Development*, 23(1), 3–18.
<https://doi.org/10.1080/0729436032000168469>
- Pierson, M. E. (2001). Technology Integration Practice as a Function of Pedagogical Expertise. *Journal of Research on Computing in Education*, 33(4), 413–430.
<https://doi.org/10.1080/08886504.2001.10782325>
- Swartz, R., & McGuinness, C. (2014). Developing and Assessing Thinking Skills: Final Report Part 2, Evaluating Current IB Programmes. <https://doi.org/10.13140/RG.2.1.3082.6080>
- 李月明, 王志航, 李繼友, 龐豔輝, & 周丹丹. (2024). 基於 Rasch 模型的高中地理學科核心素養測評工具的設計與研究——以“人地協調觀”為例. *地理教學*, 15, 42–47.
- 高虹. (2023). 思維可視化在小學英語閱讀教學中的應用研究 [碩士學位論文, 上海師範大學].
- 楊絮. (2023). 整合智能技術的學科教學知識I-TPACK 研究 [博士, 東北師範大學].
- 陳紅君. (2022). 初中物理教師 ST-PCK 模型建構與現狀測查研究 [博士, 東北師範大學].