

【附件三】教育部教學實踐研究計畫成果報告格式

教育部教學實踐研究計畫成果報告

Project Report for MOE Teaching Practice Research Program (Cover Page)

計畫編號/Project Number：PED109069

學門專案分類/Division：教育

執行期間/Funding Period：2020/08/01-2021/07/31

將遊戲式學習整合進入適性化鷹架型三層次學術英語測驗之研究

(「學術英語」及「論文寫作」課程)

計畫主持人(Principal Investigator)：陳攸華

執行機構及系所(Institution/Department/Program)：中央大學網路學習科技所

成果報告公開日期：立即公開 延後公開(統一於 2023 年 3 月 31 日公開)

繳交報告日期(Report Submission Date)：2021/09/05

將遊戲式學習整合進入適性化鷹架型三層次學術英語測驗之研究

一. 報告內文(Content)

1. 研究動機與目的(Research Motive and Purpose)

計畫主持人在去年的教學實踐研究計畫提出了「鷹架型三層次測驗」，「鷹架型三層次測驗」包含答案、原因及複習等三個部分，且利用鷹架輔助來減輕學習者的挫折，在分析學術英語期中考試成績(表三)之後，發現「鷹架型三層次測驗」可以改善學習成效($p < .05$)，並且學生之間成績的離散程度也縮小了，換言之，去年的計畫正在為教學現場帶來正面的影響。

但從面對面的對談發現，學生仍認為學習英語文法非常無趣，所以欠缺學習動機，並且高先備知識者覺得此測驗對他們的幫助有限。針對這些問題，在今年度所提出的教學實踐研究計畫裡，將「鷹架型三層次測驗」，優化為「適性化遊戲式鷹架型三層次測驗」，換言之，適性化和遊戲式學習乃是此項計畫的創新點，關於適性化，則是利用雙軌道的三層次測驗來滿足不同先備知識學習者的需求，關於遊戲式學習，乃是因為遊戲式學習可以提升學生的學習動機 (Erhel & Jamet, 2013)，因為它包含了不同的遊戲元素，這些遊戲元素吸引學生的好奇心，並讓他們產生動機完成更多的學習目標 (Spires, Rowe, Mott, & Lester, 2011)，此動機使他們在心理上收到需要改變學習態度的訊號，進而驅使出積極的學習行為。

在另一方面，計畫主持人在教學的過程中，發現先備知識確實會影響英語學習的成效，故此計畫也以實證研究來探討先備知識對使用「適性化遊戲式鷹架型三層次測驗」之影響，以建立能容納學習者個別差異性的個人化學習環境為終極目標。

2. 文獻探討(Literature Review)

線上測驗的效率與效益被許多研究認同 (Veenmana, Bavelaar, De Wolf, & van Haaren, 2014)，最普遍的線上測驗方式是選擇題 (Gikandi, Morrow, & Davis, 2011)，但是只用選擇題評量學習成效可能沒有辦法了解學生的問題 (Mora, Sancho-Bru, Iserte, & Sánchez, 2012)，因為學生可以猜測答案 (Schoenfeld, 1988) 而並不了解為什麼是這個答案(Sealey, 2014)，這個問題在高等教育尤其嚴重，所以有需要改善選擇題的設計(Andrade-Aréchiga, López, & López-Morteo,

2012)。針對這個需要，Haslam 與 Treagust (1987) 提出雙層次測驗的想法，第一層為選擇題或是非題，著重在評量學習者對學科內容了解的多寡。第二層為說明部分，要求學習者為他們的答案選擇適當的理由，以判斷學生是否真的了解 (Adodo, 2013)。這種評量方式受到研究者的關注 (Won, Yoon, and Treagust, 2014)，因為它可以分析學生的錯誤類型跟學習障礙 (Lin, Yang, & Li, 2016)。由於此項優勢，計畫主持人的研究團隊也用雙層式測驗來評量學生對微積分課程的理解，第一層是答案，第二層是原因，以診斷學生是否真的知道這個答案，研究結果顯示雙層式測驗確實可以幫助學生發展深層思考的技巧，而不是只依賴記憶來答題，所以雙層次測驗不僅是有效的評量工具 (Yang, Chen & Chen, 2016)，也是非常有用的教學策略 (Hsieh, Lin, & Hou, 2016)。

但是之後我們發現雙層式測驗無法讓學生有達到深入的學習，也無法使學生保留長期的記憶 (陳彥翔，王振漢，楊子奇，陳攸華，2018)，因此在去年的教學實踐研究計畫提出三層次測驗，三層次測驗除了包括雙層次測驗的答案及原因兩個部分之外，還包括複習的部分，主要目的是藉著複習來加深學習者所學到的概念，以便日後能夠使用到此所學到的概念。然而，三層次測驗還是屬於一種評量機制，在評量的過程中，學習者仍然會覺得沉悶和壓力，因而缺乏學習動機 (Chen & Tseng, 2019)。所以需要與別的數位學習工具一起整合，例如遊戲式學習就是一種能夠提升學習動機的數位學習工具 (Chen & Chang, 2020)。

Lin et al. (2018) 設計了一款遊戲，目的是為了促進學生學習在地文化，並幫助學生在結束學習活動之後，仍保有反覆學習機會，實驗結果證明遊戲式學習組的學習成績及學習動機都優於傳統學習組。但是，遊戲式學習提供了複雜且豐富的多媒體素材，包含文本、聲音、視覺元素等，這些素材造成過多的感官刺激，進而會增加學習者的認知負荷 (Paas, van Gog, and Sweller, 2010)。由此可知，遊戲式學習並非完美無缺，在另一方面，每一個學生都有不同的特質、技能和背景 (Chen & Macredie, 2010)。因此不是每一個人都可以克服這些缺點，換句話說，必須要更慎重地考慮學習者間的個別差異性 (Individual Differences)。先備知識乃是最常被探討的個別差異性，有研究指出先備知識的高低會對學習有影響，例如：Gao、Wang 和 Gao (2015) 指出學習者的先備知識會影響他們的學習行為，高先備知識者偏好索引字母方式，而低先備知識者偏好階層式地圖。Yang 和 Quadir (2018a) 應用角色扮演遊戲開發了一個遊戲式英語學習環境，他們發現低先備知識者較高先備知識者不僅有比較不好的的學習表現，而且經歷比較多的負面情緒。

上述文獻探討顯示出三層次測驗是一種有效的評量機制，但可能需要與遊戲式學習整合，因此在本計畫將發展一個「適性化遊戲式鷹架型三層次測驗」，此外先備知識扮演一個重要的角

色，會影響學習者如何使用數位學習工具，故本計畫亦探討先備知識對於學習者使用「適性化遊戲式鷹架型三層次測驗」的影響。

3. 研究問題(Research Question)

本計畫以實證研究來探討先備知識對於使用「鷹架型三層式測驗」的影響，將包括以下兩個研究問題：

- (1) 在使用「適性化遊戲式鷹架型三層次測驗」時，先備知識如何影響學習者的學習成效?
- (2) 在使用「適性化遊戲式鷹架型三層次測驗」時，先備知識如何影響學習者的學習行為?

4. 研究設計與方法(Research Methodology)

「適性化遊戲式鷹架型三層次測驗」的第一層次是答案部份，第二層次是原因部分(圖一)。詳述如下：

- 答案部份：一共有 10 種類型的題目，每一個問題皆由具有專業的英文知識背景的老師出題，而每一個問題之答案有四個選項，受測者從其中選擇一個做為答案，而他們有二次做答的機會，每答錯一次，扣除 5 分的基本分數。
- 原因部分：除了給予答案之外，受測者亦需從四個原因選項中，選擇正確答案的原因，但是他們只有一次做答的機會，若答錯的話會扣除 5 分的基本分數。

第三層次有兩個路徑，如果學習者對主要問題給出了錯誤的答案，則要求他們回答一個複習題。相反，如果學習者對主要問題給出了正確答案，則要求他們回答進階題(圖二)。

此外，「適性化遊戲式鷹架型三層次測驗」亦會運用鷹架輔助來幫助學習者進行測驗，也會運用扣分機制，來激勵學習者，詳述如下：

- 鷹架輔助：若是受測者不曉得該如何回答問題或是對自己的觀念有所疑惑的時候，可以使用鷹架輔助來獲得幫助，鷹架輔助之方式包括三種提示：「句意提示」、「圖片提示」、及「語音提示」。
- 扣分機制：在畫面的左下方，每一位受測者一開始都會有 200 分的基本分數，每答錯一次，皆會扣除 5 分基本分數，而每使用一次提示會扣除 2 分基本分數。

第1題：

Because of Individual differences exist between learners, no one instructional method is suitable for all learners.

原因：

(A) One

(B) Two

(C) Three

(D) None

(A) 兩個子句要用連接詞來連接, 連接詞有and, but, ', ', so...等等

(B) Because of不能接子句,在三者或以上數量的人或物之間用 among

(C) ", "不能用來作為連接詞, 副詞用來修飾動詞

(D) difference不可+s, method為複數需要+s

送出答案 測驗答對:1 測驗答錯:0 原因答對:0 原因答錯:0
 任務分數:20 遊戲分數:0

送出原因

答題正確, 任務分數+20分, 請回答原因!

鷹架式提示 提示分數:2000 計分規則 下一題 下棋 01:59:13

圖一：適性化遊戲式鷹架型三層次測驗

進階題第1題：

Digital games provide joyful experience _____ they are popular with students.

(A) in that

(B) in which

(C) so that

(D) so which

送出答案 進階題答對:1 進階題答錯:0
 任務分數:35 遊戲分數:0

送出原因

答題正確, 任務分數+20分! 是否觀看文法說明?

鷹架式提示 提示分數:2000 計分規則 下一題 下棋 是 否 01:59:19

圖二：進階題的設計

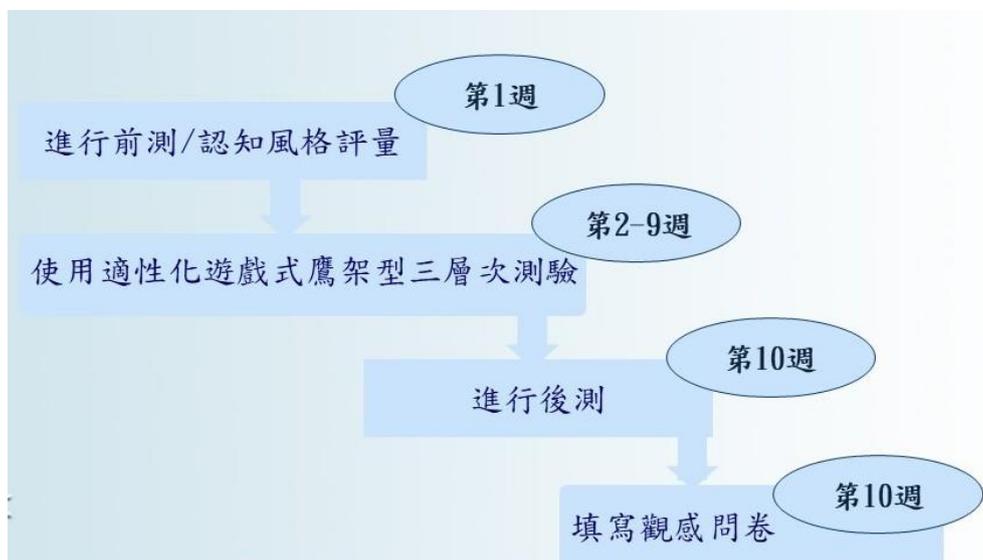
5. 教學暨研究成果(Teaching and Research Outcomes)

5.1 教學過程與成果

本計畫為探討高低先備知識對受測者在使用「適性化遊戲式鷹架型三層次測驗」的時候會有何影響，研究對象為41位來自台灣北部某國立大學且有受過專業的英文訓練的研究生，其

包括三個階段(圖三)：

- 在第 0 週的時候，教師使用兩小時講解英文學術論文使用到的文法知識，接著在第 1 週的時候，受測者進行認知風格的評量，再由前測評量他們的先備知識，測驗的時間約為 1 個小時，測驗內容為 40 題的專業英文的改錯題，每題至少有一個文法錯誤，例如：少了關係代名詞、連接詞使用錯誤或是詞性使用錯誤，受測者不僅需要找出錯誤的地方，還需要將錯誤的地方更改正確才可得到分數，測驗的基本分數為 100 分，每答錯一題扣 2.5 分基本分數。本研究將使用測驗的平均分數做為區分受測者的依據，平均分數大於或等於 60 分的受測者歸類為高先備組 (N=20)，小於 60 分的受測者則歸類為低先備組 (N=21)。
- 在第 2 週至第 9 週的時候受測者集中於同一個班級裡進行「適性化遊戲式鷹架型三層次測驗」實驗，在講解完「適性化遊戲式鷹架型三層次測驗」如何操作之後，讓受測者自行使用平板電腦操作鷹架型三層次測驗並回答問題，使用「適性化遊戲式鷹架型三層次測驗」之學習行為以及分數皆會被記錄在歷程檔。
- 在第 10 週的時候，受測者進行後測，後測之題目、內容與評分方式和前測相似，最後所得到的分數將作為後測成績。



圖三：實驗程序

5.2 教師教學反思

5.2.1 測驗成效 vs. 任務成效

二因子混合變異數分析被用來分析高先備學習者與低先備學習者在前測分數、學習任務分數以及後測分數上的差異(表 1)。結果表明，不同先備知識的學習者與前測分數、學習任務分數以及後測分數的交互處理下是有顯著差異($F = 25.996, p < 0.05$)。

因此，我們將進行單純主要效果檢定，來進一步分析不同先備學習者在前測分數、學習任務分數以及後測分數中的成效差異(表 2)。結果表明，有五項結果都具有顯著差異，第一項是高先備學習者的後測分數高於學習任務分數以及前測分數($F = 12.202, p < .05$)。第二項是低先備學習者的學習任務分數高於後測分數以及前測分數($F=55.169, p<.05$)。綜合上述的結果，我們針對單純主要效果的結果進行說明，首先是高先備學習者的後測分數(Mean=65.684, SD=13.153)是高於學習任務分數(Mean=55.289, SD=7.728)以及前測分數(Mean=49.842, SD=13.525)，更具體地說，高先備學習者在學習任務中藉由鷹架輔助提示獲得學習，因此高先備學習者在過程中是有持續進步。再者，低先備學習者的學習任務分數(Mean=45.432, SD=9.598)是高於後測分數(Mean=26.909, SD=17.811)以及前測分數(Mean=8.727, SD=8.154)。這可能是因為，低先備學習者在進行學習任務時，可以使用鷹架輔助提示來提升學習表現，但是當低先備學習者在沒有鷹架輔助的後測，便導致學習表現下降。但是，比較都是沒有鷹架輔助的前測與後測的學習表現，低先備學習者在學習任務中有學習到英文文法知識，因此低先備學習者的後測分數高於前測分數。

接著，第三項是高先備學習者的前測分數顯著高於低先備學習者($F = 143.345, p < .05$)。第四項是高先備學習者的學習任務分數顯著高於低先備學習者($F = 12.836, p < .05$)。第五項是高先備學習者的後測分數顯著高於低先備學習者($F = 61.151, p < .05$)。綜合上述的結果有一個有趣的發現，不論是前測分數、學習任務分數還是後測分數，都是高先備學習者的分數顯著高於低先備學習者。這可能是因為高先備學習者擁有較高的英文文法先備知識外，他們還會運用自己的先備知識針對問題進行更詳細的分析，以至於他們會擁有較高的問題解決能力(Gao, Wang & Gao, 2015)，因此高先備學習者在前測、學習任務以及後測中會有較好的成效，反之亦然。

5.2.2 複習題 vs. 進階題

卡方檢定被用來分析高先備學習者與低先備學習者在進入複習題與進階題次數上的差異(表 1)。結果表明，高先備學習者與低先備學習者進入複習題次數與進階題次數有顯著差異($\chi^2=20.101, df=1, p=.000$)。更具體地說，低先備學習者進入複習題的次數是多於高先備學習者。這可能是因為低先備學習者在測驗題的錯誤較多，導致有較多人進入到複習題。然而，高先備學習者因為擁有較好的英文能力以及問題解決能力，所以就減少了進入複習題的次數。另一方面，系統有提供鷹架輔助幫助學習者作答，因此造成低先備學習者也有相當的人數進入到進階題。

表 1: 複習題 vs. 進階題

	進階題	複習題	X^2	df	p
高先備學習者	515(476.9)	150(188.1)	20.101	1	.000***
低先備學習者	514(552.1)	256(217.9)			

5.3 學生學習回饋

高先備學習者和低先備學習者在使用「適性化遊戲式鷹架型三層次測驗」有相似的行為，就是當高先備學習者和低先備學習進入測驗時，他們會先使用三項不同的提示，分別是句意提示、語音提示以及圖像提示後，才送出答案。而且，學習者會在使用句意提示後，使用語音提示或是圖像提示，接著再送出答案。另外，當學習者答對後會使用句意提示後，再送出原因。上述的行為顯示出不論是高先備學習者還是低先備學習者都會使用提示來協助他們答題，尤其是句意提示，這可能是因為句意提示提供與題目相關的英文文法概念。另一方面，他們之間也存在行為差異，附件中說明了此類行為差異。

此外，我們使用開放式問卷來確定他們的學習觀感。從問卷的結果得知，高先備學習者和低先備學習者都認為「適性化遊戲式鷹架型三層次測驗」可以幫助他們獲得許多英語文法知識，黑白棋遊戲增加了樂趣，可激勵自己認真答題。但是高先備學習者仍然不知道如何建立論文的邏輯架構，而低先備學習者仍然有閱讀論文的問題。

6. 建議與省思(Recommendations and Reflections)

針對高先備學習者和低先備學習者的成效做分析與比較之後，發現「適性化遊戲式鷹架型三層次測驗」可以改善碩博士生對英語文法的了解，然而從觀感問卷中可以得知撰寫學術論文不能只有正確的英語文法，也需要有清楚的邏輯架構和合理的論證辯述，碩博士生要具備

這些能力，必須仰賴大量的英文學術論文閱讀，可是碩博士生通常缺少閱讀英文學術論文的動機和能力。另外個別差異性是一個值得考慮的因素，因為他們所遇到的英文學術論文閱讀的問題，不完全相同。

呈上所述，為提升學習者學術論文寫作的品質，高等教育急需要發展一個工具不僅能夠提高他們閱讀學術論文的動機，也可以兼顧他們的個別差異性，針對這個需要，計畫主持人在今年的教育部教學實踐研究計畫將結合客製化及適性化的機制來開發遊戲式電子書，希望這個計畫也能展現豐碩的成果。

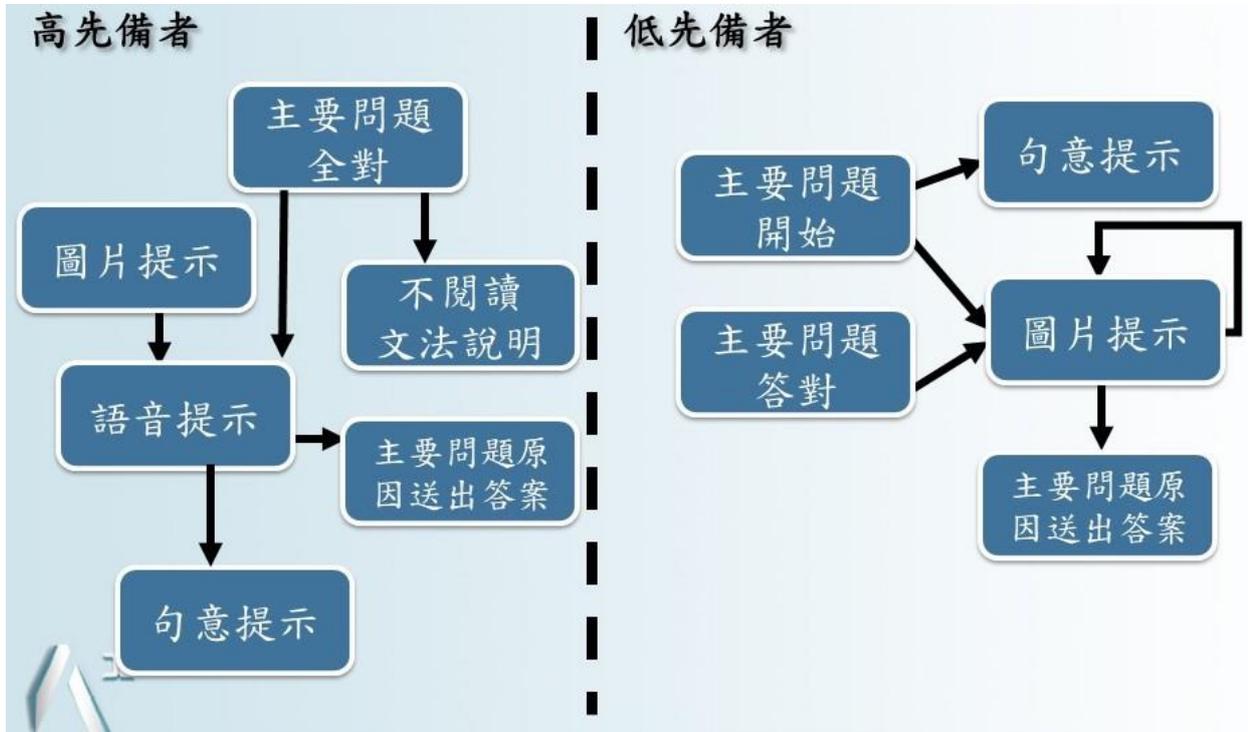
7. 參考文獻(References)

- 陳彥翔，王振漢，楊子奇，陳攸華(2018). 從先備知識的角度探討英語字彙學習系統的使用. 第22屆全球華人電腦教育應用大會
- Adodo, S. O. (2013). Effects of Two-Tier Multiple Choice Diagnostic Assessment Items on Students' Learning Outcome in Basic Science Technology (BST). *Academic Journal of Interdisciplinary Studies*, 2 (2), 201.
- Andrade-Aréchiga, M., López, G., & López-Morteo, G. (2012). Assessing effectiveness of learning units under the teaching unit model in an undergraduate mathematics course. *Computers & Education*, 59(2), 594-606.
- Chen, S. Y. & Chang, Y. M. (2020) The Impacts of Real Competition and Virtual Competition in Digital Game-based Learning. *Computers in Human Behavior*. 65, 359-368
- Chen, S. Y. & Tseng, Y. F. (2019) The Impacts of Scaffolding E-Assessment English Learning: A Cognitive Style Perspective. *Computer Assisted Language Learning*.
- Chen, S. Y., & Macredie, R. (2010). Web-based interaction: A review of three important human factors. *International Journal of Information Management*, 30(5), 379-387.
- Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education*, 67, 156-167.
- Gao, Q., Wang, D., & Gao, F. (2015). Impact of knowledge representations on problem-oriented learning in online environments. *International Journal of Human-Computer Interaction*, 31(12), 922-938.
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, 57(4), 2333-2351.
- Haslam, F., & Treagust, D. F. (1987). Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument. *Journal of Biological Education*, 21 (3), 203-211.
- Hsieh, Y. H., Lin, Y. C., & Hou, H. T. (2016). Exploring the role of flow experience, learning performance and potential behavior clusters in elementary students' game-based learning. *Interactive Learning Environments*, 24 (1), 178-193.
- Lin, Y. C., Yang, D. C., & Li, M. N. (2016). Diagnosing Students' Misconceptions in Number Sense via a Web-Based Two-Tier Test. *Eurasia Journal of Mathematics, Science & Technology Education*, 12 (1), 41-55.
- Lin, Y. T., Tseng, Y. M., Lee, Y. S., Wang, T. C., Tsai, S. I., & Yi, Y. J. (2018). Development of a SoLoMo Game-Based Application for Supporting Local Cultural Learning in Taiwan. *Educational Technology & Society*, 21(4), 115-128.

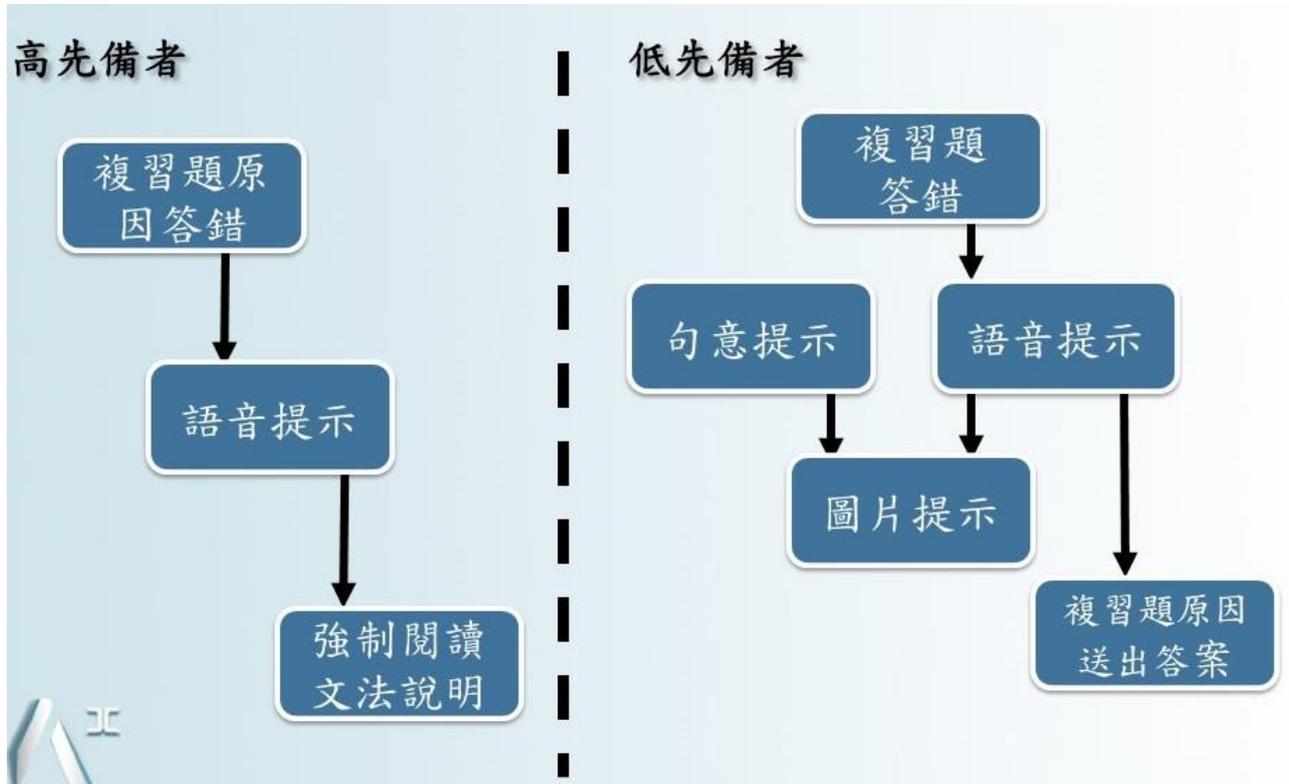
- Mora, M. C., Sancho-Bru, J. L., Iserte, J. L., & Sánchez, F. T. (2012). An e-assessment approach for evaluation in engineering overcrowded groups. *Computers & Education*, 59(2), 732-740.
- Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review*, 22(2), 115-121
- Schoenfeld, A. H. (1988). When good teaching leads to bad results: The disasters of well-taught mathematics courses. *Educational Psychologist*, 23(2), 145-166.
- Sealey, V. (2014). A framework for characterizing student understanding of Riemann sums and definite integrals. *The Journal of Mathematical Behavior*, 33, 230-245.
- Spires, H. A., Rowe, J. P., Mott, B. W., & Lester, J. C. (2011). Problem solving and game-based learning: Effects of middle grade students' hypothesis testing strategies on learning outcomes. *Journal of Educational Computing Research*, 44(4), 453-472.
- Veenman, M. V., Bavelaar, L., De Wolf, L., & Van Haaren, M. G. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and Individual Differences*, 29, 123-130.
- Won, M., Yoon, H., & Treagust, D. F. (2014). Students' learning strategies with multiple representations: Explanations of the human breathing mechanism. *Science Education*, 98(5), 840-866.
- Yang, J. C., & Quadir, B. (2018a). Effects of Prior Knowledge on Learning Performance and Anxiety in an English Learning Online Role-Playing Game. *Educational Technology & Society*, 21 (3), 174-185.
- Yang, T. C., Chen, S. Y., Chen, M. C. (2016). An Investigation of a Two-Tier Test Strategy in a University Calculus Course: Causes vs. Consequences. *IEEE Transactions on Learning Technologies*. 9(2), 146-156.

8. 附件(Appendix)

主要問題



複習題



學術論文:

- Chen, S. Y. & Tseng, Y. F. (accepted) The Impacts of Scaffolding E-Assessment English Learning: A Cognitive Style Perspective. *Computer Assisted Language Learning*. (SSCI, IF=2.018, Cited=0)
- Tseng, J. & Chen, S. Y. (2020) Incorporating a Digital Game into a Three-Tier Test: A Prior Knowledge Perspective. *Proceedings of Taiwan Academic Network Conference*.