

結合 QAR 教學策略之英語學習聊天機器人

A ChatBot with QAR teaching strategy for English Learning

區國良，王綉嬭

國立清華大學 學習科學與科技研究所

klou@mx.nthu.edu.tw

【摘要】 本研究探討生成式人工智慧技術應用於高中文本閱讀與提問能力學習產生的影響，本研究設計並開發了名為 Qally 的對話機器人，輔助學生掌握問答關係（QAR）策略。研究以台灣北部某公立高中 36 名高一學生為對象，以《小王子》作為指定文本，透過 Discord 平台進行英語閱讀課程。學生在與 Qally 機器人互動中學習辨別 QAR 四種問題類型，並利用該策略提升閱讀理解與提問能力。實驗結果顯示，學生的 QAR 認知測驗前後測成績顯著提升，而閱讀理解測驗分數無顯著變化。在學習動機(ARCS)問卷中，注意、相關、信心及滿足四個構面證實 Qally 機器人對學生的學習有積極影響。Qally 能提供即時回應與針對性反問，幫助學生深化思考，同時在忠實度及回應相關性方面表現優於 ChatGPT。本研究證明基於大型語言模型的對話機器人在教育中具有潛力，不僅解決教師個別指導的時間限制，亦促進學生自主學習與認知成長，為未來應用於不同學科與教育環境奠定基礎。

【關鍵詞】 問答關係策略、對話機器人、大語言模型、學習動機

Abstract: This study investigates the impact of generative artificial intelligence technology on high school students' learning of text reading and questioning skills. The research involved the design and development of a chatbot named "Qally," which assists students in mastering the Question-Answer Relationship (QAR) strategy. The participants were 36 first-year high school students from a public school in northern Taiwan. Using *The Little Prince* as the designated text, English reading lessons were conducted via the Discord platform. During interactions with Qally, students learned to identify the four types of QAR questions and applied the strategy to enhance their reading comprehension and questioning skills. The experimental results revealed a significant improvement in students' pre- and post-test QAR cognitive assessment scores, though there was no notable change in reading comprehension test scores. In the ARCS motivation questionnaire, the mean scores for the four dimensions—Attention, Relevance, Confidence, and Satisfaction indicating a positive impact of the Qally chatbot on students' learning. Qally provided real-time responses and targeted follow-up questions, fostering deeper student thinking. Additionally, it outperformed ChatGPT in fidelity and response relevance. This study demonstrates the potential of dialogic chatbots based on large language models in education. They not only address the time constraints of individualized teacher guidance but also promote autonomous learning and cognitive development among students. The findings lay a foundation for future applications across different subjects and educational contexts.

Keywords: Question-Answer Relationship Strategy, Chatbot, Large Language Model, Learning Motivation

1. 研究背景與動機

提問與反思的能力對於提升閱讀理解能力扮演了關鍵的角色，Sinabariba (2020)認為要提升學生的閱讀能力，除了培養閱讀理解技巧外，更應提昇學生的提問技巧，並鼓勵學生運用高層次的思維技巧提出問題，以及積極與教師或同儕參與討論。根據 Hartman(2001)的研究，

自行提問有助於深化理解並促進高層次的學習，也就是透過示範和實踐，教師能引導學生自主思考，從中了解問題與答案之間的關係，進而加深反思與確認。

常見的閱讀提問策略包含 K-W-L、5W1H、SQ3R 以及 QAR(Question-Answer Relationship) 等；然而，對於初學者而言，往往會遇到不知道如何從閱讀文本中找到相關問題，或是如何更進一步理解內容的情況，尤其在東方社會中，由於文化背景和教育環境的特性，導致學生在課堂上較少主動提問。即使勇於發問，受制於課堂時間有限，學生的提問機會相對較少。對教師而言，在面對一對多的教室中使用提問策略時，面對不同程度的學生將難以提供充分的個別指導，因此，如何促進學生提問的能力，並在學習中實現個別化教育，是當前教育面臨的挑戰。

大型語言模型的發展，為個別化學習帶來了全新的可能，例如 Arun et al. (2024) 為學習解剖學課程的學生提供以檢索增強生成 (RAG) 技術為基礎之對話機器人，有效地提供學生必要的知識及輔助；Haindl and Weinberger (2024) 則引導學生使用 ChatGPT 作為 Java 程式語言學習輔助工具，發現對於程式語法、設計概念以及演算法等知識具有良好的輔助效果，但仍有適用性的問題，除此之外，使用 LLM 提供的對話界面（例如 ChatGPT）作為學習平台時，教師將難以掌握學生的學習進度，同儕之間也不易互動交流。

本研究將利用生成式人工智慧技術，運用 OpenAI GPT 4 模型建構英語自主學習問答機器人，以 Discord 社群軟體為學習討論互動環境，讓學生可以透過提問與問答機器人互動學習閱讀策略，並促進學生的認知層次和思考能力。以下為本研究之主要研究目的：

以生成式人工智慧建構輔助學習 QAR 策略之對話機器人

本研究將運用生成式人工智慧技術，建立限定特定文本的問答對話機器人，以社群軟體作為互動介面，採用 QAR 策略針對教師指定的文本內容與學生進行的深度討論，提供即時解答及反問，以促進學生理解與掌握 QAR 策略的相關知識及問答技巧。

探討使用對話機器人對於 QAR 問答關係學習成效與動機之影響

本研究探討使用問答策略的對話機器人對學生學習成效與動機的影響，包括活動前後學生對策略的理解及其在自主學習中提問的情況，並分析提問內容與生成回應。

2. 文獻探討

2.1. QAR 教學策略

Taffy E. Raphael 提出的問答關係策略(Question-Answer Relationship, QAR)可協助學生運用四種類型的問題更有效地識別和回答閱讀中的問題，這些類型包括：Right There、Think and Search、Author and Me 以及 On My Own，其中前兩種類型問題的答案在文本中(In the Text)，而後兩種類型問題的答案則是在讀者大腦中(In the Head)。Right There 要求學生從文本中特定句子找出答案；Think and Search 則需要整合文本不同部分的資訊；Author and Me 題要求學生結合個人知識推測作者意圖；On My Own 則根據個人知識回答問題，無需依賴文本。這些策略不僅促進了深度閱讀和理解能力，也提升了學生的批判性思維和自主學習能力，對他們的認知和閱讀理解技能有著深遠的影響(Raphael, 1986; Caldwell and Leslie, 2005)。

問答關係策略被證實是提升閱讀理解能力的有效方法(Afriani et al., 2020)。這策略鼓勵學生發展關鍵性問題提出技巧，並學會依據問題類型尋找答案，深化他們對閱讀材料的理解，建立問題與答案間的連結。這種方法不僅培養學生的深度思考能力，還能作為後設認知工具，讓他們意識到自身學習的過程和方法，有效運用問答關係策略來確認答案(Green, 2019)。

有關運用問答關係策略(QAR)提升學生閱讀能力的相關研究表明，這一策略在數學和科學教學中均有積極效果。McIntosh 和 Draper(1995, 1996)的研究指出，教師教授 QAR 策略能有

效提升學生對數學文本的閱讀理解能力。Kinniburgh 和 Shaw(2009)的研究則表明，在科學教學中引入 QAR 策略能夠提高學生對科學文本的理解能力。研究顯示，教授 QAR 策略並讓學生理解其四種問題類型，有助於學生辨別問題類型並加深對文本的理解，從而顯著提升閱讀理解能力(Afriani et al., 2020; Baqi, 2019; Nurhayati et al., 2019; Wila et al., 2020)。這些研究強調，QAR 策略不僅幫助學生更有效地找到答案，還能增強其閱讀動機和參與度，並在標準化閱讀測試中展現正面影響(Green, 2019)。

2.2. 人工智慧及對話機器人在教育的應用

人工智慧(AI)是指機器具備人類般智慧的能力，能進行學習、判斷和解決問題(Shabbir & Anwer, 2018)。在教育領域，AI 技術可支援個人化學習體驗的線上平台，能根據個別學生的學習情況調整學習內容，提供適性化學習環境。近年來因生成式人工智慧 (GAI) 的發展，已有許多研究提出了基於 GAI 技術的對話機器人作為輔助教學系統(Chen et al., 2020)，研究顯示，研究指出，ChatGPT 在醫學教育中展示出一致性、準確度及獨特見解，具有提供專業知識支援的潛力(Kung et al., 2023)。因此，許多研究開始以 GAI 作為對話機器人的對話產生來源，融入教學環境當中，例如 Neumann et al. (2024)以 ChatBot 結合 GAI 設計了一套名為 MoodleBot 的對話機器人，融入大學生的資料庫課程中，該研究除了討論學生對於生成式人工智慧對話機器人的接受程度外，也證明能有效支持學生的自主學習 (SRL) 活動，Liu et al. (2024)則同樣以 GAI 所建構的對話機器人這些系統作為小學生自主學習英文寫作的輔助工具，其效果也顯示其學習成效以及使用 SRL 策略皆優於未使用對話機器人的對照組。

對話機器人是人工智慧的典型應用，透過自然語言處理技術，能在教育中提供個性化學習和即時回饋、批判性思維和學習滿意度(Chang et al., 2022; Belda-Medina & Calvo-Ferrer, 2022)。對話機器人根據其操作方式可分為三種類型 (Hien et al., 2018)：規則式模型、檢索式模型和生成式模型。規則式模型基於預設規則生成回應，具高精確度但缺乏靈活性，無法處理未知問題(Binkis et al., 2021)。檢索式模型依賴於知識庫快速匹配答案(Qian et al., 2021)，限於知識庫範圍內的回應。生成式模型利用大型語言模型生成多樣和自然的回應，雖靈活但可能產生不準確信息或「幻覺」，在導入教學使用前，必須施以微調(Fine-Tuning) 或 RAG (Arun et al., 2024)處理，下節將說明本研究如何針對指定文本進行微調後，作為對話機器人融入教學使用。

3. 研究方法

本研究以台灣北部某公立高中的 36 名高一學生為研究對象，男生 20 人，女生 16 人，由教師將學生每 3 人至 4 人為單位分成 10 個小組進行實驗。研究期間為 2024 年 4 月，融入該班英語閱讀課程進行，其中兩堂 50 分鐘的課程內容指定閱讀文本為 Antoine de Saint-Exupéry 的著作：小 Le Petit Prince 《小王子》，由於其內容淺顯易懂且充滿深刻寓意，且故事情節生動有趣，適合作為 QAR 策略使用。

研究首先進行特定文本的分析和處理，並開發名為 Qally 的對話機器人，其名稱結合了問答(Question-Answer)和夥伴(Ally)的意涵，除了可利用 OpenAI GPT 語言模型回答學生提出的問題，還能辨別問題類型並提供反問，從而幫助學生理解 QAR 策略。Qally 機器人運行於 Discord 伺服器中，學生可以在各種設置的頻道中進行小組或個人討論，這些討論和互動可在社群軟體中保留，便於後續回顧和學習歷程的追溯。Qally 機器人採用 Python 語言與 LangChain 框架開發，該框架專為支援大型語言模型應用而設計，提供便捷的功能建置和整合能力。LangChain 的基本元件包括載入、分割、嵌入和相似性搜索，這些元件能有效連接文本數據到語言模型進行後續處理。為了進行文本嵌入，使用 OpenAI 的 text-embedding-ada-002 模型，將文本轉換為向量並建立向量資料庫。使用者需設置 OpenAI

API key 並引入相關套件，以實現文本向量化。由於模型在處理文字數量上的限制，文本需事先分割成小段，以避免超出 Tokens 數量限制並節省成本。接著對學生進行 QAR 策略教學，介紹 QAR 四種問題類型後，進行「QAR 策略認知與閱讀理解測驗」之前測，接著讓學生與 Qally 機器人進行互動和提問練習，之後再進行後測及填寫學習動機問卷。其中測驗參考 Pathways to English Reading Literacy (張碧珠, 2021) 中的 QAR 問題類型。學習動機問卷則採用 John. M. Keller 的 ARCS 學習動機理論衍生的教材動機量表(IMMS)，評估學生對於使用 Qally 機器人的動機與看法，包括注意、關聯、信心和滿足四個構面。本研究依照原問卷進行修改編制，問卷共 36 題，用以分析學生對 Qally 機器人的看法，採用李克特氏五點量表，選項從 1(非常不同意)到 5(非常同意)，以表達對各題的看法之態度。

Discord 社群軟體原被用於遊戲社群，提供玩家即時互動交流的頻道，在疫情期間被許多研究利用 Discord 作為教學與學習媒體，獲得大多數學生的接受，並認為其介面為具吸引力的介面，功能完整且容易使用 (Arifianto & Izzudin, 2021)。此外，在 Discord 伺服器中能加入各式機器人，使用者能透過觸發命令或是關鍵字等與機器人互動，而機器人能連結相關資訊與數據，或透過串接 API 取得所需結果，適合作為研究使用。在 Discord 的學習頻道中（如圖 1），學生可以透過提及 @Qally 機器人並輸入問題，啟動自主學習流程。機器人首先判斷問題類型，如「Right There」、「Think and Search」等，然後生成相應回答並提出一個反問，鼓勵學生深入思考。例如，學生可能問「小王子為何來到地球？」，機器人將解釋這是哪一類型的 QAR 問題，提供答案，然後以另一個問題引導學生更深入地探索文本內容。這種方法不僅支持即時問答，還促進了學生對問題類型和文本理解的反思和學習。

使用者提問後，系統利用向量資料庫進行相似性搜索，提取相關文本區塊，結合問題和提示詞傳遞給語言模型，生成回應後傳回給使用者。本研究使用 LangChain 框架的 ConversationalRetrievalChain 實現這一過程，支援上下文和關聯記憶的傳遞。

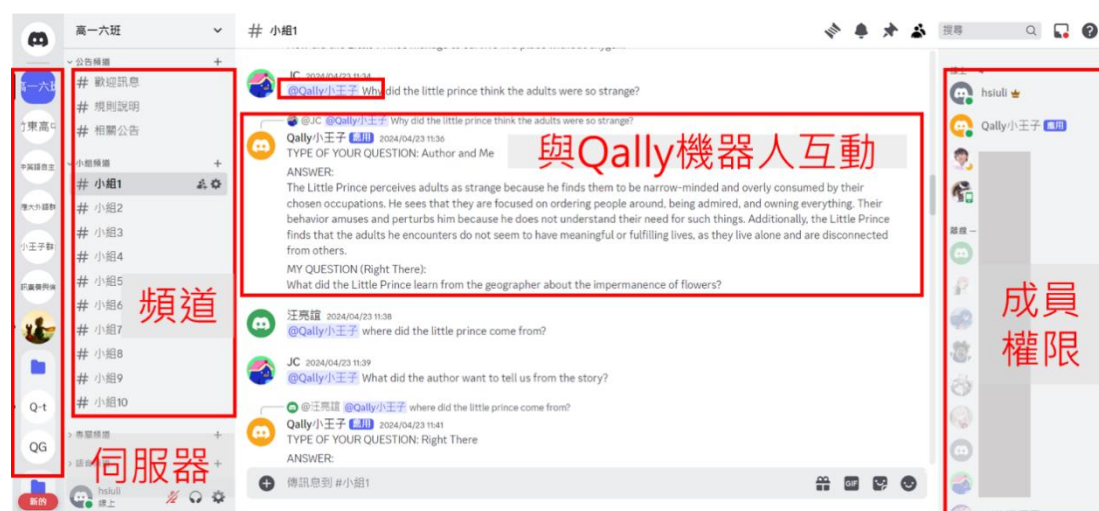


圖 1. 系統畫面說明圖

4. 研究結果

本研究將參與實驗之 36 位高中生分為 10 組，每組學生在閱讀完教師指定的文字內容後，以小組為單位利用平板登入 Discord 平台及加入教師指定的討論頻道後，與 Qally 對話機器人進行對話，對話的內容經教師分析其對應 QAR 策略後統計如表 2，其中共 6 組成功提出四種問題類型（第一、二、四、五、七、九組）；而第三、六、十組只提出三種問題類型（缺少 On My Own），第八組僅提出兩個問題（缺少 Think and Search 以及 On My Own）。

表 1.
各組提問狀況統計表

Group ID	1	2	3	4	5	6	7	8	9	10
Right There	2	1	3	1	1	2	1	1	1	2
Think and Search	3	3	3	4	4	2	3	0	2	1
Author and Me	3	3	2	1	2	5	1	1	3	3
On My Own	2	2	0	2	2	0	1	0	1	0
Total	10	9	8	8	9	9	6	2	7	6

本研究在兩次教學實驗活動中，設計 QAR 策略與閱讀理解測驗，包含小王子文本之閱讀理解及對應 QAR 問題類型辨別各 6 題，共 12 題，每題 10 分。

針對閱讀理解前後測成績做成對樣本 t 檢定，信賴區間設為 95%，即設置顯著性水準為 5%。依據閱讀理解測驗六題一題 10 分，滿分 60 分。將其轉換成滿分 100 分，前測平均值為 90.28，後測平均值為 89.35， t 值為 .320 得出 p 值大於 .05，表示前後測之間沒有明顯差異，並且前後測平均值均接近滿分，分析其可能原因。學生對文本有足夠認知並正確選出答案，所以前後測均相當高。

表 2.

閱讀理解前後測分數成對樣本 t 檢定($N=36$)

	平均數	標準差	t	p
前測	90.28	10.06	.320	.751
後測	89.35	15.51		

QAR 策略認知的前、後測成績做成對樣本 t 檢定，信賴區間設為 95%，即設置顯著性水準為 5%。依據 QAR 問題類型辨別勾選六題一題 10 分，滿分 60 分。將其轉換成滿分 100 分，前測平均值為 61.11，後測平均值為 91.20， t 值為 -9.932 得出 p 值小於 .001，表示前後測之間有明顯差異，即學生針對 QAR 認知理解與辨別有明顯進步，經過教學與 Qally 機器人自主學習活動後，大部分同學能明確分辨 QAR 問題類型。

表 3.

QAR 認知測驗前後測分數成對樣本 t 檢定($N=36$)

	平均數	標準差	t	p
前測	61.11	11.95	-9.932	.000***
後測	91.20	12.27		

*** $p < .001$

本研究問卷依據 ARCS 學習動機量表之 IMMS 之四個構面：注意、相關、信心、滿意四大構面分量表的 Cronbach's α 信度係數分別為 .886、.795、.785 和 .779。

表 4.

學習動機四大構面量表之平均分數

	注意(A)	相關(R)	信心(C)	滿足(S)
平均	3.49	3.76	3.55	3.75

四個面向的平均分數分別為 3.49、3.76、3.55、3.75，在數據顯示個構面均有超過 3 分，其中較高的為相關構面與滿足構面，表示 Qally 機器人的回應確實與其提問有關連，而在滿足構面也達 3.75 分，整體來看確實有讓學生滿意。

5. 結論

本研究應用大型語言模型建構了 Qally 機器人，支援高中生進行自主學習活動。機器人專注於特定文本內容，能理解並回應學生提問，提供即時反饋和支持。這解決了老師時間有限的問題，增強了學生的閱讀理解和反思能力，同時促進了個性化教育和自主學習。

研究也發現在教導學生掌握 QAR 策略，並透過與 Qally 機器人的互動來強化他們的閱讀理解能力。這種策略幫助學生有效分析問題與答案之間的關係，並提升他們的提問技巧。研究顯示，基於 QAR 策略的學習環境能夠增進學生對 QAR 四種問題類型的識別能力，同時提升他們的思辨能力。Qally 機器人不僅回答學生的問題，還能引導他們深入探討，這些互動能在 Discord 平台上被記錄下來，為個性化教學提供了豐富的數據來源。此外，我們將學生在互動中產生的問題輸入 ChatGPT 進行比較分析，結果顯示，Qally 機器人在忠實度和內容回應相關性方面均優於 ChatGPT，這證明了我們建置的學習環境比一般的 ChatGPT 更有效地支持學生的學習需求。

透過高中生自主學習課程實際應用後的學習動機調查，結果顯示學生對使用 Qally 機器人有正面的學習動機反應，IMMS 量表中各構面平均分超過 3.49 分。此外，QAR 認知理解前後測統計分析顯示，Qally 機器人對學生學習有正向影響，證明了其在教學和學習效果上的有效性及可行性。本研究展示了 ChatGPT 及基於 GPT 模型的對話機器人在教育中的應用潛力和實際效益，可作為未來研究進一步探索這些技術在不同教育環境下的應用，推動教育技術的創新發展。

參考文獻

- Afriani, Z. L., Anggraini, M., & Riswanto, R. (2020). The Effect of Question Answer Relationship (QAR) Strategy in Enhancing Students' Reading Comprehension. *Journal of English Education and Teaching*, 4(4), 548-558.
- Arifianto, M., & Izzudin, I. (2021). Students' acceptance of discord as an alternative online learning media. *International Journal of Emerging Technologies in Learning (iJET)*, 16(20), 179-195.
- Arun, G., Perumal, V., Urias, F. P. J. B., Ler, Y. E., Tan, B. W. T., Vallabhajosyula, R., Tan, E., Ng, O., Ng, K. B., & Mogali, S. R. (2024). ChatGPT versus a customized AI chatbot (Anatbuddy) for anatomy education: A comparative pilot study. *Anatomical sciences education*, 17(7), 1396-1405.
- Baqi, F. A. (2019). implementing question-answer relationships (qar) strategy to improve students' reading comprehension skill. *English Journal*, 13(2), 88-98.
- Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using Chatbots as AI Conversational Partners in Language Learning. *Applied sciences*, 12(17), 8427.
- Binkis, M., Kubiliūnas, R., Sturienė, R., Dulinskienė, T., Blažauskas, T., & Jakštienė, V. (2021). Rule-based chatbot integration into software engineering course. *Information and Software Technologies: 27th International Conference, ICIST 2021, Kaunas, Lithuania, October 14–16, 2021, Proceedings 27* (pp. 367-377). Springer International Publishing.
- Caldwell, J. S., & L. Leslie. 2005. *Intervention strategies to follow informal reading inventory assessment: So what do I do now?* Boston: Pearson.

- Chang, C.-Y., Kuo, S.-Y., & Hwang, G.-H. (2022). Chatbot-facilitated Nursing Education: Incorporating a Knowledge-Based Chatbot System into a Nursing Training Program. *Educational Technology & Society*, 25(1), 15-27.
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *Ieee Access*, 8, 75264-75278.
- Green, S. (2019). The Effects of Question-Answer Relationship (QAR) Instruction on Standardized Reading Comprehension Test Scores of Third Grade Students Thesis (Ph.D.)--The University of North Carolina at Charlotte, 2019.
- Haindl, P., & Weinberger, G. (2024). Students' Experiences of Using ChatGPT in an Undergraduate Programming Course. *IEEE Access*, 12, 43519-43529.
- Hartman, H. J. (2001). Teaching metacognitively. In *Metacognition in learning and instruction: Theory, research and practice* (pp. 149-172). Springer.
- Kinniburgh, L. H., & Shaw, E. L. (2009). Using Question-Answer Relationships to Build: Reading Comprehension in Science. *Science activities*, 45(4), 19-28.
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., ... & Tseng, V. (2023). Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS digital health*, 2(2), e0000198.
- Liu, Z.-M., Hwang, G.-J., Chen, C.-Q., Chen, X.-D., & Ye, X.-D. (2024). Integrating large language models into EFL writing instruction: effects on performance, self-regulated learning strategies, and motivation. *Computer Assisted Language Learning*, 1-25.
- McIntosh, M. E., & Draper, R. J. (1995). Applying the Question-Answer Relationship Strategy in Mathematics. *Journal of adolescent & adult literacy*, 39(2), 120-131.
- McIntosh, M. E., & Draper, R. J. (1996). Using the Question-Answer Relationship Strategy to Improve Students' Reading of Mathematics Texts. *The Clearing house*, 69(3), 154-162.
- Neumann, A. T., Yin, Y., Sowe, S., Decker, S., & Jarke, M. (2024). An LLM-Driven Chatbot in Higher Education for Databases and Information Systems. *IEEE Transactions on Education*, 1-14.
- Qian, H., Dou, Z., Zhu, Y., Ma, Y., & Wen, J.-R. (2021). Learning implicit user profile for personalized retrieval-based chatbot. *proceedings of the 30th ACM international conference on Information & Knowledge Management*,
- Raphael, T. E. (1986). Teaching Question Answer Relationships, Revisited. *The Reading teacher*, 39(6), 516-522.
- Shabbir, J., & Anwer, T. (2018). Artificial intelligence and its role in near future. *arXiv preprint arXiv:1804.01396*.
- Sinabariba, Y. E. (2020). the improving students' reading comprehension by using questioning techniques in narrative text. *Jurnal Suluh Pendidikan*, 8(1), 15-19.
- Wila, S., Rina, H., & Didih, F. (2020). question answer relationship (qar) strategy in critical reading on narrative text. *jall (Journal of Applied Linguistics and Literacy)*, 4(1), 95-102.