本研究欲重製Kerwer等人(2021)之研究，Kerwer等人之研究目的在於驗證白話摘要(plain language summaries)是否能夠將科學研究的發現傳遞給更廣泛的讀者。此外也觀察白話摘要所帶來的easiness effect。本研究的作者認為需要複製的原因是想了解上述結果是否能在不同語言、主題及文化中穩定被觀察到。本研究作者使用中華心理學刊所發行之學術摘要及編輯部改寫之白話摘要作為材料，以台灣大學生為主要研究參與者，試著重製Kerwer等人之研究。結果發現在可理解性上複製過去研究的發現，關於知識獲取則與過去研究的結果不同，簡易效應相關指標則部分複製過去研究發現。在閱讀過程中發現本文的文字流暢性尚嫌不足，會交替使用相似的概念詞，對文獻結果的陳述也不甚清晰，以致在閱讀時會有不太理解作者意思的情況，建議進行通篇調整並精確用詞，相關疑惑可見以下意見；除此之外，我有以下的疑慮與建議提供貴刊編輯與作者參考。

核心問題

1. 作者僅在摘要中提及複製Kerwer等人(2021)之研究的主要原因在於想要檢驗簡易效應是否跨語言、跨文化成立，但在本文中完全沒有提及複製的原因，也沒有說明為什麼簡易效應可能受到或不受到語言、文化的影響，因此難以展現複製此研究之必要及重要性。唯一相關的論述是緒論的第一段，也僅說明白話摘要有助於科學訊息的傳播，這段落僅是在說明此議題的重要性，並未直接對應到重製Kerwer等人(2021)之研究的必要性。建議作者針對此重點更加清楚地說明。最後針對結果討論時也沒有回應到跨文化是否成立的議題，建議作者增加說明。
2. 在本文的文獻回顧及研究範疇中並未討論白話摘要對於虛假訊息傳播的議題，均僅限於討論白話摘對讀者主觀感受、相關決策等影響，其研究內容均為真實訊息，僅在撰寫或呈現方式上有所差異，並未操弄訊息內容的真偽，因此建議作者將論述聚焦於此即可，第3頁第8-10行「也可以是傳播虛假訊息的工具」、「或匯合根據虛假資訊的極端立場」等較廣泛性的論述建議可以避免以降低讀者的困惑。「或匯合根據虛假資訊的極端立場」、「值得從心理學源頭探究」的語意也相當模糊，不太清楚作者想要表達甚麼意思。
3. 建議作者在說明白話效應相關指標時能保持全文一致，以減少讀者對應的困惑。倘若是用以說明該指標的評估或測量方式時則建議可以清楚地說明其對應於哪一個指標。例如第3頁第13行將指標稱為「信賴感、真實感、與專業感」，同頁第21-23行又改稱為「可信度、判讀資訊真實性的信心、個人專業判斷能力」，第4頁第6-7行則改稱為「資訊可信度、對資訊的判斷信心、以及獲取進一步資訊的意願」。而在假設的部分(第6頁第14-20行)則稱為「理解、獲取知識、可信度、真實性、參考更多資訊」。本研究的測量工具說明的部分也會交替使用不同的名詞，例如第9頁第4-5行為「內容科學性、來源可信度」第6行又改稱為「科學性、可信度」。其中可信度與來源可信度應指涉完全不同的東西，因此會不確定作者到底測了甚麼。在第11頁結果的部分所使用的指標似乎是「文本內容科學性、內容真實性、參考充分度」。另外一個需要請作者多加說明的部分是，因著這些名詞的交互使用，其定義似乎也不甚清楚，建議作者可以清楚說明這些概念的意涵及在文內說明給參與者看到題目的樣子，參與者是否能夠理解這些指標需要進行的判斷是甚麼。
4. 建議作者重新架構第4頁第2段關於Scharrer等人（2019）的研究相關內容。第一是因為其研究中所操弄的來源可信度與本研究所關注的主軸無關，三種被視作依變項的評價指標也不完全與本研究相關，建議可保留相關，建議可強調本研究相關的概念，清楚說明並一致用詞。第二是該段落的文句在閱讀上不甚清楚，「只有容易理解的文章內容提高評價的可靠性假設（Robustness hypothesis）；只有來源可信度高的文章提高評價的可覺察假設（Awareness hypothesis）；容易理解且來源可信度高的文章才會有效提高評價的調節假設（Moderation hypothesis）」以上文句應該不是三種影響途徑，而是對於自變項與依變項關係間的預期，若不在Scharrer等人（2019）的研究上下文的脈絡中，直接援引「假設」來稱呼且未說明其間關係，不甚容易理解。此外「評價」一詞也沒有被說明清楚，若其為包含對文章主張的同意度、訊息真實性的信心、應用知識的可行程度的綜合指標，則在本研究的脈絡下，建議作者更仔細地說明對與本研究相關的指標是否產生影響、產生甚麼影響。例如「同意文章主張的評價符合調節假設」是否是指容易理解且來源可信度高的文章才會提高文章主張的同意度？
5. 第4頁第21-23行「Lisa Scharrer 團隊的研究顯示，儘管簡易效應的量化與質性測量所顯示的結果分別支持不同假設，在嚴格控制條件的實驗，同類指標的量測結果有高度一致性」，不清楚作者想表達的意思，建議作者重新調整文句。
6. 第4頁第23行至第5頁第22行「相較於Scharrer 團隊的研究， Kerwer 等人（2021） 取用來自第一手學術論文及論文作者撰寫的白話摘要，嚴格控制來源可信度，只操作文本可理解程度，能檢驗簡易效應各項指標的一致性。」此部分論述我有幾個疑惑：在Scharrer等人的研究中將來源可信度視為一個變項，並沒有不嚴格控制。操作，應為操弄。簡易效應各項指標間的一致性是甚麼意思？為什麼需要進行這個檢驗？在Scharrer等人的研究中可能不是使用相同的指標，若真的要檢驗簡易效應指標的一致性也並不是不能檢驗。
7. 第5頁第9行，我不是很理解為什麼「根據第一手學術論文準備的摘要，也具備跨語言比較的優點。」
8. 第5頁第15行，建議作者直接說明四種可讀性指標為何。前面(第5頁第13行)指的是哪個指標？可讀性分數是否就是指標之一？因這是作者欲複製的關鍵研究，建議可將其發現更清楚地說明。且後續結果部分亦都有提及此些指標的分析，因此清楚說明這些指標的意涵是有必要的。包含在中文可讀性評估的指標及量尺(SVM1.0、2.0)上都建議作者可以清楚說明之。後續關於研究材料的說明中也有提及SVM1及SVM2的分數，且此二指標的發現不一致，但亦未說明此指標的意涵，建議作者補充說明。
9. 在產生研究問題與假設時作者直接將Kerwer等人所列的五項假設條列出來，但完全沒有說明為什麼要產生這五項假設。雖然本研究是重製研究，但簡要地敘述假設產生的脈絡與預期仍是需要的。建議作者可以稍加敘述。此外，在Kerwer等人的假設中對於三種不同版本的摘要形式間有更仔細的差異性比較，但作者只取了其中關於白話摘要優於學術摘要的部分，並忽略有無副標題版本間的比較，此處調整亦須請作者說明。
10. 本研究為複製研究，因此應該在方法、程序及資料分析上盡可能貼近原始研究，我可以理解在材料取得及可讀性指標計算上會有其限制，除此之外，以下是我對於方法、程序及分析上差異的疑惑。
    * 為什麼要多測量「科學活動經驗的調查」？
    * 我想確認一下本研究在知識理解測驗的題目編撰。首先是題數的部分在第7頁第3行的部份說一共有24題，但在第8頁第2行的部份說一共有48題。第二是在原始據Kerwer等人的研究中只有24題，且每位參與者均是填答此24題，並沒有隨機抽選24題的程序，不確定研究者為何要改變此程序。第三，我覺得最關鍵的部分是「知識理解測驗題由12篇學術摘要與白話摘要各取一題」，但在Kerwer等人的研究中並沒有區分題目是屬於學術摘要或白話摘要，其強調無論閱讀何種形式的摘要都可以正確答題。因此作者這個敘述會令我有些擔心是否某些題目需要搭配特定形式的摘要方能正確回答。這個程序的改變也須請作者說明原因。
    * 在Kerwer等人原始的研究程序中是讓參與者至特定場所參與研究，因此可以控制參與者在作答期間的活動。然在本研究中是透過台灣心理科學研究協作群來招募參與者，因此想確認實驗場域是否會至固定的實驗室中並由主試者協助並控制作答過程的活動呢？倘若沒有進行這樣的控制，參與者是否可能在回答問題時搜尋其他資料或將摘要拍攝下來待答題時再次觀看呢？
    * 資料清理的部分，本研究刪除了作答時間超過200分鐘或理解題正確率未超過50%的參與者，而Kerwer等人的篩選標準是100分鐘，且並未以正確率作為門檻刪除資料。正確率過低某種程度也反映了參與者無法理解呈現的摘要，我不太確定改變刪除標準的原因，想請問為什麼作者設立了不同的資料篩選標準呢？
11. 根據Kerwer等人（2021）的發現，其結論應該是有副標題的白話摘要的索取全文意願會高於學術摘要，並非作者所述之「白話摘要似乎能提昇索取完整論文的意願」第7頁第9行，建議作者更精確地陳述其發現。
12. 本文內並沒有提及表二，僅插入表二於第10頁處，建議作者補充說明表二相關敘述。
13. 我對於第10頁假設二結果的部分的疑慮可以參考第10點建議中的第四小點。因作者預先刪除了50%以下的正確率者，此作法與原始研究不同，且也沒有如原始研究一樣呈現難度區間，因此建議作者秉持複製研究的原則盡可能複製原作之作法以利結果的比較。另外，建議作者將檢定值一併在文內敘明。
14. 第10頁關於假設三至五的結果陳述部分令人有些困惑，以下是須請作者再次釐清的部分：
    * 此部分提及三種簡應效應的評估指標，包含可信度、真實性、參考更多資訊，但在行文時會僅以「指標」來稱呼，因此會不太確定是在說明哪個部分的結果。
    * 簡易效應評估指標中應為可信度(credibility)，但作者使用來源可信度(Source credibility)(第10頁第13行、第16行)，這應該是完全不同的概念，在目前複製的研究中應該沒有操弄來源，我不認為可以彼此交互取代使用，因此此部分的結果需要再次確認是否合宜。
    * 「此外，文本來源可信度評分與另一項有高度相關的文本內容科學性評分也顯示有顯著正相關，*r* = 0.80, *p* < .001。」這句話中有兩個須作者釐清的部分，第一是有高度相關的文本內容科學性評分是甚麼意思？第二是這個結果與假設三-五的關聯性為何？
    * 第17-19行處說明三種版本的兩種指標間分數比較時，所對應的表格應是表3，此處見表4的目的不甚清楚，建議作者再多做說明。
15. 第11頁關於知識背景共變作用的部分，作者將科學活動經驗的自我評估納入作為共變，因此可能會導致結果與原始研究無法直接對應，就複製研究而言可能需要盡可能依循原始文章進行再製為佳。若想延伸既有發現，也建議先呈現與原始研究一致之分析再行呈現延伸結果。
16. 在討論的部分作者認為「加入副標題也許讓參與者認為更好懂，卻可能不利正確理解。」第13頁第10-11行，這句話的意思並不太清楚，若僅是再次說明可理解性有文本類型間差異，但知識獲取度評估沒有差異，其實並沒有解釋這個結果的可能原因。後續關於文本設計的差異也沒有直接對應到本研究發現與原始研究差異的原因。關於作者提到以母語呈現能使參與者以「以習慣的思考模式理解摘要及題意」來做為解釋差異結果的原因我亦認為有待商榷，因為就結果而言，本研究的知識確認度的表現不及原始研究(且作者已刪除了一些正確率較低的參與者)，因此作者所提出之母語優勢顯然並不存在。
17. 關於索取全文意願的差異比較上，比較適切的討論應是比較文本間獲取全文的比例差異，此與本研究議題較為一致，並非比較兩個研究的獲取全文比例差異。一方面因為選材上有所差異，直接推論此差異是由參與者興趣所致並不適切，另一方面興趣與排版都不是本研究主要關注的議題，建議作者針對本研究及原始研究所關注的議題進行結果的比較與討論。最後白話摘要促進深度學習可行性的討論中(第14頁最後一段)作者亦以此無法進行統計分析的結果進行應用性延伸，建議在結果延伸應用上較為保守一些。
18. 在研究結論的論述上也建議作者能夠貼近研究發現進行論述，本研究的主要目的並非討論目前學術摘要寫作的現況，也並未討論虛假訊息的議題，關於正確傳播訊息之責任應不屬本文之結論。

其他細節(格式與繕打錯誤)

1. 目前文內交互使用了兩種中英文字型，建議作者應依據期刊搞規範統一全文同一種語言內之字型，例如第1頁第2行、第1頁第12行、第1頁第22行、第1頁第23行。此狀況通篇皆有不再一一列出，須請作者自行檢查確認。
2. 文內多處有位具意義的標示性底色，建議刪除，例如：第1頁第11行、第5頁第19行、第5頁第20行。
3. 文內應不需特別寫出作者名，例如第3頁第19行、第4頁第1行。
4. 因本文係以中文書寫，因此括號需統一為全形，例如第1頁第3行、第1頁第12行。此狀況通篇皆有不再一一列出，須請作者自行檢查確認。
5. 第9頁第22行，人後面多了一個空格。
6. 第9頁第22行，p值應斜體，此狀況通篇皆有不再一一列出，須請作者自行檢查確認。
7. 專有名詞的英文僅需在全文中第一次出現時呈現即可，須請作者自行檢查確認。
8. p值的呈現方式建議依據期刊要求全文統一，目前有多處p值有呈現整數位、未呈現實際值等問題，都須請作者自行檢查確認。r值呈現時亦須包含自由度，建議作者依據期刊格式規範重新調整之。