《心理科学进展》审稿意见与作者回应

题目：提取练习在临床人群中的应用研究：一项系统综述
作者：冉光明，李艺，陈迎

第一轮

审稿人 1 意见：

本文是一篇问题导向的系统综述，研究方向明确，结构完整，具有一定的创新性，具备发表潜力。但是在科学问题阐述以及核心内容阐述上建议作者根据以下意见进行重大修改。

回应：感谢您在百忙之中能审阅我们的稿件，并给予我们详尽而专业的意见。您提出的每一条建议都让我们受益匪浅，也为文章的修改完善提供了至关重要的方向。我们已依照您的指导逐一认真斟酌并进行了修改(修改后的文字在正文标记为了红色)，在此谨向您表达最诚挚的谢意。

意见 1：作者对本文科学问题的呈现不足，系统性综述的作用首先是将单个、分散、有时结论矛盾的研究，整合成一个清晰的“证据体”，其次才是清晰地展示哪些问题已有共识，哪些结论仍存争议，哪些领域根本缺乏研究。因此我认为本文的科学问题首选是首次系统论证了“提取练习在临床人群中普遍有效但具有特异性”这一整体性结论，其次才是创新性地提出了三个未来方向。目前来看，本文对科学问题的阐述在引言以及结果和讨论部分都是欠缺的。

回应：感谢您的意见。根据您的建议，我们对引言、结果和讨论的写作逻辑进行了重大修改，修改后的逻辑如下：

首先，引言的**第一段**为引出“提取练习”的相关基本概念，并说明在正常人群中的适用性，紧接着**第二段**我们将提取练习引入到了临床人群中，并逐步推导提取练习可能如何在该类人群中发挥作用，进一步说明已有研究的不足(如证据分散)，最后提出了本文拟讨论的问题“提取练习是否适用于临床人群？”以及“其效果的条件与边界如何？”。**第三段**中我们阐述了以往相关综述(de Lima et al., 2020)的不足，以及本研究的必要性。

其次，在结果部分，我们将每一类疾病分为了小节(3.1 神经系统疾病。3.2 精神障碍。3.3 其他)，每一种疾病的相关描述修改为独立一段，放弃了原有的“摘要的集合”。在其中我们首先对相关研究的结果进行了精炼总结，之后我们指出了在该种疾病研究中的不足和未来方向，最后我们在每一类疾病(每一小节)的最后一段进行了跨疾病的比较。指出了提取练习对该类疾病是否有效，以及有哪些特异性和研究不足。以期符合系统综述的写作逻辑。

最后，我们将原标题“4.1 提取练习在临床人群中的效度”修改为了“4.1 提取练习在临床人群中的适用性与特异性”。“4.1.1 跨群体适用性的普遍性”修改为“4.1.1 跨群体的适用性”，具体阐述部分做了细微修改。在 4.1.2 人群特异性与效应边界部分，我们提炼出来 3 点，分

别是(1)疾病的核心认知缺陷、(2)学习材料的复杂性和生态性、(3)多治疗手段的支持，并分别进行了阐述。

意见 2：首先，当前“结果”部分的主要问题是写法类似“叙述性摘要的集合”：对每项研究进行了近乎独立的、细节丰富的描述(如具体实验流程、被试数量、细微结果)，像是在为每篇论文写迷你摘要。但是，综合与提炼不足：读者需要自己从大段的文字描述中，费力地提炼出“对于某一类疾病(如 DLD)，提取练习的整体效果如何？关键影响因素是什么？”这样的核心信息。这样的写法淹没了综述的“主线”：即综合证据、发现模式、指出空白。

回应：感谢您的建议。另一位审稿专家也指出了该问题，结合两位专家的建议，我们对结果部分做了重大修改。

在结果部分，正如我们对意见 1 的回应。我们每一类疾病分为了小节(3.1 神经系统疾病。3.2 精神障碍。3.3 其他)，将一种疾病的相关描述修改为独立一段。在每一种疾病的阐述中，我们首先对相关研究进行了精炼总结，之后我们指出了在该种疾病研究中的不足和未来方向，最后我们在每一类疾病(每一小节)的最后一段进行了跨疾病的比较，指出了提取练习对该类疾病是否有效，以及有哪些特异性和研究不足。修改后符合“综合证据、发现模式、指出空白”的写作逻辑。

意见 3：其次，在 4.1.2 人群特异性与效应边界这个标题下，依然是叙述性摘要的写法，没有提炼出特异性即效应边界是什么？

回应：感谢您的意见，我们根据您的意见，做了如下修改。

在**“4.1.2 人群特异性与效应边界”中**，我们提炼出来 3 点，分别是(1)疾病的核心认知缺陷、(2)学习材料的复杂性和生态性、(3)多治疗手段的支持，并分别进行了阐述**(文章第 11-12 页)**。

意见 4：4.2.1 对共病的论述较充分，但 4.2.2 和 4.2.3 部分可适当增加临床案例或假设，增强说服力。

回应：感谢您对我们的肯定和意见，我们根据您的意见，做了如下修改。

在“4.2.1 提取练习在共病人群中的研究”中，我们优化了部分语言表述，保证表达精准、流畅。

在“4.2.2 提取练习的间接效应”和**“4.2.3 临床人群的个体差异变量在提取练习中的作用”**中，我们增加了对部分临床疾病的阐述，以增加说服力，详见 4.2.2 第 2 段的 4–10 行，4.2.3 第 2 段的 2–6 行**(文章第 13–15 页)**。

意见 5：在 4.2.3 “临床人群的个体差异变量在提取练习中的作用”这一小节，当前的逻辑确

实存在跳跃和焦点模糊的问题。主要问题在于，它混杂了两个不同层面的论证：基于健康人群研究的推论：在健康人群中，个体差异对提取练习的影响不显著，可能是因为这些变量的“变异性”不足。而临床人群在这些变量上变异更大，因此将他们纳入研究(与健康人一起)可以提高整体变异性，从而可能检测出效应。但是实际上，现有研究中不乏以健康被试做对照组，来探讨个体差异对提取练习效应的影响。针对临床人群研究的批评与建议：现有的临床人群提取练习研究，本身严重忽略了个体差异变量(仅将其作为控制变量或组间匹配依据)。鉴于这些变量在理论上可能很重要，未来应直接在临床人群内部检验它们的作用。也许临床群体本身的个体差异才是未来研究的重点，比如认知损伤的程度？是否存在提取练习发挥效应的基线水平？

回应：感谢您给我们意见，并且非常感谢您帮我们理清了该部分的写作思路，我们根据您的意见，做了如下修改。

在“4.2.3 临床人群的个体差异变量在提取练习中的作用”中，我们重新梳理了写作逻辑，决定**只对临床人群的个体差异变量对提取练习的影响这一点进行论述**。首先我们指出当前研究没有将个体差异变量作为核心变量，之后我们论述了个体差异变量可能在对提取练习的影响，最后我们在该节第二段举了相关例子和理论，用于推导这些变量如何影响提取练习，以增加说服力**(文章第 14–15 页)**。

意见 6：部分长句可拆分，提升可读性(如：“提取练习可促进精神分裂症患者对新信息的学习并减少前摄干扰……”)。

回应：感谢您的意见。**针对可读性**，我们邀请了 1 位非专业人士和 2 位心理学工作者分别进行了挑剔性阅读，并且我们进行了全文的出声朗读，随后进行了相应修改，以提升文章的可读性。

意见 7：英文摘要语法基本正确，但建议请英语母语者或专业编辑再次润色，确保表达地道。

回应：感谢您的意见。**针对英文摘要**，我们邀请了一位具有英语学历背景的心理学副教授做了润色，确保表达地道，相关表述已用红色字体突出。

意见 8：参考文献格式基本规范，建议统一检查页码范围符号(如“1151-1159”是否应为“1151–1159”)。

回应：感谢您的仔细审阅。页码范围符号的确是我们之前没有发现的细小错误，我们已经做了全面检查和修改。

意见 9：检索词中“临床(Clinical)”等是否包含对应英文术语(如“clinical populations”)？建议在“方法”中说明中英文检索词的对应关系。

回应：感谢您的意见。**针对检索词的中英文对照关系**，我们改写了在“方法”中的部分表述(文章第 3 页)，并在附录中新增了表 3 中英文检索词对照表(文章第 33 页)，以说明中英文检索词的对照关系。

意见 10：质量评估工具基于 de Lima 等(2020)，建议简要说明其适用性及本研究的调整(如有)。

回应：感谢您的意见。**针对质量评估工具**，我们简要说明了在本研究中的适用性(文章第 4 页)，以下是增加的具体内容：

“这 7 方面从随机化、数据缺失、评分、结果报告等多个维度进行评价，符合心理学研究一般标准。其中方面 6 和方面 7 针对提取练习实验中的多次最终测试和初始学习进行评价，表明该方法适用于提取练习实验的相关研究”

意见 11：纳入文献数量较少(21 篇)，是否因检索词或数据库限制？建议在“局限性”中说明。

回应：感谢您的意见。纳入文献梳理较少的确是本文的局限之一，我们为此新增了“5 本研究的不足与局限”部分进行阐述(文章第 15 页)。以下是具体阐述内容：

“本研究主要存在以下不足与局限。第一，应更广泛的纳入参考文献。本研究纳入的文献只有 21 篇不够全面，其原因可能是以下两方面。(1)检索时间限制，2020 年前的相关文献已有类似综述进行了研究(de Lima et al., 2020)，因此本研究只检索了 2020 年至 2025 年的相关文献。(2)根据纳入和排除标准，在纳入的文献中，学习方式变量需要有两个及以上的不同水平的文章(如提取练习和重复学习)，因此只单纯研究提取练习的文章则被排除(如：Smith et al., 2024)。第二，应进行定量的元分析，以获取数据支持。本研究是对当前研究领域的系统综述与未来研究展望，未来可对相关研究进行提取练习效应的元分析，进一步从数据分析层面分析提取练习在临床人群中的效应量大小。”

意见 12：未来研究方向是否可结合“临床转化路径”或“干预方案设计”稍作展开，增强实践指导性。

回应：感谢您的意见。**针对未来研究方向的实践指导性**，我们新增了“4.2.4 干预方案的设计与实施”，阐述了三点建议。以下是具体内容：

“提取练习临床干预方案的设计与实施，需要系统性思考和多维度整合。应注意以下几方面。首先，明确干预目标。这需要系统评估患者的认知缺损模式及日常生活需求，可使用与认知缺陷特点或患者日常生活相关的材料。其次，加强指导和宣教。有研究认为人们可能不愿意使用提取练习这一策略(Karpicke et al., 2009)，该现象可能在临床人群中同样存在。因此要提高患者进行自我调节提取练习的意愿和能力，让患者在日常学习和生活中也更愿意使用提取练习。最后，采用渐进式实施策略。可先将提取练习作为现有治疗方案的补充环节，

后逐渐加大干预频率和强度，最终成为独立的治疗手段。”

审稿人 2 意见：

研究选题具有重要的理论价值和临床意义。

回应：诚挚感谢您在繁忙工作中抽出时间审阅本文，您提出的每一条意见都极具参考价值，为我们的稿件打磨提供了重要指导。我们已严格对照所有修改建议完成细致修订(修改后的文字在正文标记为了红色)，再次向您致以由衷的感谢！

意见 1：摘要部分对研究结果的表述过于概括，缺乏精确性。建议修改摘要。

回应：感谢您的建议，根据您的建议，我们修改了摘要，以增加摘要的精确性。以下是具体内容：

“但其效果受到疾病的认知缺陷特点、材料复杂性和生态性的调节，并且常需要作为辅助治疗手段。当前研究存在共病人群探索不足、间接效应未明确等问题。未来需关注共病人群干预、间接效应及个体差异影响。本研究针对提取练习干预方案的设计实施提出了明确干预目标、加强指导与宣教、采用渐进式实施策略三方面建议，以提升提取练习在临床场景的应用价值。”

意见 2：作者在自检报告中提到研究目标是提取练习“如何发挥作用”、“发挥怎样的作用”、“受哪些因素影响”，但略显宽泛，几乎是所有机制性研究都会提出的问题。建议精确化研究关注的重点，同时，建议在引言的最后一段，应将宽泛的目标转化为具体、可验证的研究问题。

回应：感谢您的建议。根据您的建议，我们将我们的研究目标细化为：“提取练习是否适用于临床人群？”以及“其效果的条件与边界如何？”并且修改了自检报告和引言，在引言最后一段对研究目标进行了阐述(文章第 1–2 页)。

意见 3：引言部分罗列了事实，但未构建围绕主题的论证。如第二、三段简单描述了不同疾病的症状和研究，但未能清楚表述“为什么提取练习应该对这些疾病有效”的理论框架。例如，提到精神分裂症的语音记忆、情景记忆缺陷，但未点明与提取练习之间的逻辑关联，如提取练习可通过强化编码和巩固来针对此缺陷的。

回应：感谢您的意见，我们修改了引言，在**引言第 2 段 5–10 行**中，阐述了提取练习在临床人群中起作用的可能原因(精细化编码、激活相关脑区)。以下是具体阐述：

“提取练习则是一种具有潜力的认知干预手段。首先，提取练习可能诱发精细化编码进而提高记忆表现(廖陈陈, 张锦坤, 2024)。其次，提取练习效应的产生与内侧、腹外侧前额叶皮层，以及额下回、腹侧纹状体和中脑等脑区相关(Van den Broek et al., 2013; Ye et al., 2020; Zhuang et al., 2022)。这些区域一方面与记忆表现相关，另一方面可能与部分临床疾病(如阿尔茨海默症、精神

分裂症)的记忆缺陷有关(Liu et al., 2021; Williams et al., 2023)。因此提取练习可能会通过精细化编码或激活相关神经回路来帮助临床人群改善认知缺陷。”

意见 4：文章在结果部分的整体逻辑结构较为松散，段落之间相对割裂。目前更像研究摘要的集合体，而非系统综述应有的需呈现的具有分析性与综合性的结果报告。同时，在结果部分罗列了不同疾病的研究发现，但未体现不同研究、不同人群之间的模式、异同点和特异性，且缺乏研究证据支持。建议在“结果”部分的最后，或为每个疾病种类(如神经系统疾病)设立一个总结段落，进行跨研究的比较。

回应：感谢您的意见，您和另一位审稿专家同时指出了该问题，说明该问题的确是本文的缺点，因此我们也结合两位专家的意见，对结果部分做了重大修改。

我们放弃了原有的“摘要的集合”，重新按照“综合证据、发现模式、指出空白”的逻辑进行写作。将每一类疾病分为了小节(3.1；3.2；3.3)，将一种疾病的相关描述修改为独立一段，在其中我们首先对相关研究进行了精炼总结，之后我们指出了在该种疾病研究中的不足和未来方向，最后我们在每一类疾病(每一小节)的最后一段进行了跨疾病的比较，指出了提取练习对该类疾病是否有效，以及有哪些特异性和研究不足。

意见 5：现存问题和未来展望部分篇幅较长，但需基于详细且具体的证据链条进行阐述。

回应：感谢您的意见。**针对“现存问题与未来展望”部分**，我们修改了相关表述，增加了对部分例子的阐述，以提高说服力。以下是增加的部分表述(详见文章第 12–15 页)：

“例如，在提取过程本身及之后获得的成败经验，是影响个体后续提取或学习行为的重要因素(Carpenter et al., 2020)。而在部分临床患者(如焦虑抑郁或与其共病)的心理敏感性较高，表现为十分在意外界看法和负面经验(Fan et al., 2021)。因此利用提取练习可能让患者获得成功经验，以改善临床患者在元认知判断上的缺陷(Drueke et al., 2023; Kubik et al., 2022)，进而提升学习和生活的自信心和效能感”

意见 6：该论文中，现存问题和展望与该研究自身的发现脱节，如章节指出的共病、间接效应、个体差异等问题，均未能与本综述“结果”部分的具体发现相挂钩。

回应：感谢您的意见。**针对“现存问题和展望与该研究自身的发现脱节”的问题**，我们本次修改后，在结果部分中对每一类/种疾病的相关研究增加了对现有研究问题和未来展望的相关表述。因此我们在“4.2 现存问题与未来展望”部分中，则选择创新性的提出了未来研究的几个方面。

意见 7：此外，在论述“间接效应”时，大量引用健康人群的研究来论证其在临床人群中的潜力，存在逻辑上的问题。未拿出有力的论据证明这些机制在目标人群(即本综述所关注的各

类患者)中同样存在且是关键。

回应：感谢您的意见。**针对“4.2.2 提取练习的间接效应”中引言大量健康人群研究的问题，**该部分的确存在一定的逻辑问题，因此我们进行了修改。首先我们在该节**第一段**简要阐述了提取练习的间接效应，**第二段**阐述了在部分临床人群中，提取练习可能会如何发挥间接效应，逐步推导出间接效应在临床人群中同样存在并且是十分重要的。以下是修改后的内容：

“现有研究多聚焦于提取练习对陈述性记忆(如事件记忆、语义记忆)或语言学习的直接影响。但提取过程中的体验亦能强化元认知监控与调节能力，这种能力可帮助个体有效识别已知与未知信息，进而优化学习决策(Carpenter et al., 2020; Rivers, 2021; Endres et al., 2023)。这些影响被界定为提取练习的间接效应(Endres & Renkl, 2022)。

与健康群体相比，部分临床人群不仅存在记忆或语言能力的损伤，还常伴随多维度的认知缺陷。具体可能表现在元认知判断(Drueke et al., 2023)、元认知监控(Lehmann & Ettinger, 2023)、自我效能感(Olsson et al., 2024)及持续性注意(Lai et al., 2024)等方面。提取练习对这些因素表现出间接干预潜力。例如，在提取过程本身及之后获得的成败经验，是影响个体后续提取或学习行为的重要因素(Carpenter et al., 2020)。而在部分临床患者(如焦虑抑郁或与其共病)的心理敏感性较高，表现为十分在意外界看法和负面经验(Fan et al., 2021)。因此利用提取练习可能让患者获得成功经验，以改善临床患者在元认知判断上的缺陷(Drueke et al., 2023; Kubik et al., 2022)，进而提升学习和生活的自信心和效能感。此外，提取练习有助于减少前摄干扰(Li et al., 2024; Mulligan & Buchin, 2025; Mulligan et al., 2022)。这使得患者在后续学习或回忆时可能无需分配过多注意资源去抑制干扰信息，从而会间接缓解注意资源的竞争，最终改善注意集中能力。

综上所述，鉴于临床人群认知缺陷的复杂性，以及提取练习对认知功能存在的间接干预潜力，未来研究可进一步探索提取练习在临床人群中的间接效应，为临床认知干预提供新的思路。”

意见 8：文章整体较倾向于未来的研究展望，而不是一篇系统综述。引言、结果和分析似乎都在为研究展望铺垫，与标题“提取练习在临床人群中的应用研究: 一项系统综述”的契合度较弱。缺乏对现有证据进行综合、结构化呈现，较多的是单个研究的描述，无法从“结果”中清晰地看到提取练习在不同人群中的效应模式、效应量大小及争议点等。

回应：感谢审稿专家的意见。我们在结合专家意见修改后，已经对纳入的文献进行了跨研究的比较，修改了大部分内容的写作逻辑，将文章主线修改为了“综合证据、发现模式、指出空白”，以期符合系统综述的写作逻辑。但在未来展望方面，本研究的确较为详细。综合考虑后，我们决定将文章题目修改为“提取练习在临床人群中的应用研究”。

意见 9：语言表达不够规范，建议进一步优化语言表达。格式上也存在一些问题，如，符号

使用存在中英文格式混淆(文内存在英文格式的逗号和引号，应为中文格式的逗号和引号)等。

回应：感谢审稿专家的意见。**针对语言表达，**我们邀请了 1 位非专业人士和 2 位心理学工作者分别进行了挑剔性阅读，并且我们进行了全文的出声朗读，随后进行了相应修改，以提升文章的可读性。**针对文章格式，**我们已经做了全面检查和修改。

第二轮

审稿人 1 意见：

感谢作者认真、细致地回应了第一轮审稿意见，并对全文进行了系统性修改。从回复与修改稿来看，作者在科学问题凝练、结构重组、逻辑梳理等方面付出了努力，稿件质量有明显提升。为进一步提升稿件的论证深度与临床转化价值，提出以下几方面建议，供作者进一步修改参考。

回应：非常感谢您再次审阅了我们的稿件，并为文章的第二轮修改提供了宝贵建议。我们已遵照您的意见逐一完善，在此深表谢意(为和文中已有的蓝色和红色区分，我们将本次修改后的文字在正文标记为了绿色)。

意见 1：关于“人群特异性”论证的深化。作者在 4.1.2 中提炼出三个调节因素(疾病缺陷、材料复杂性、多治疗支持)，这是一个比较好的框架。但目前更多是观点陈述，部分维度在各类疾病中的证据支撑不够均衡。尤其是“材料复杂性”纬度和“多治疗支持”纬度，需要进一步补充研究案例。

回应：感谢您的意见，在第一轮修改中，我们对人群特异性的论述的确较为薄弱，结合两位专家对此部分的意见，我们在研究后决定，将原 4.1.2 人群特异性与效应边界，修改为：4.2 提取练习的人群特异性与效应边界。并将原有的三个调节因素分别修改为：4.2.1 疾病核心认知缺陷的类型、4.2.2 学习材料的复杂性和生态性、4.2.3 多治疗手段的支持。并分别进行了阐述。具体阐述如下**(正文第 12~15 页)：**

“4.2.1 疾病核心认知缺陷的类型

提取练习在临床人群中的效应并非完全一致，可能受到疾病核心认知缺陷类型的调节。Stamate 等(2020)指出，AD 患者的真正挑战在于需要更多次的学习才能将信息成功编码进入记忆系统，患者的记忆障碍主要表现为编码阶段的困难。Small 和 Cochrane（2020）的研究亦指出，提取练习的长期效果部分依赖线索的支持，当支持被撤除时，提取练习的长期维持效果可能会减弱。这说明，在以编码缺陷为特点的疾病中，提取练习的补偿作用可能存在一定的上限，或需借助其他支持性条件，才能发挥长久的作用(Stamate et al., 2020)。

相比之下，精神分裂症患者的认知缺损更多表现为干扰抑制功能下降与提取效率较低(Weisbrod et al., 2000)。Tan 等(2024)和 Wen 等(2025)的研究发现，提取练习能通过减少前摄

干扰、促进语义组织能力，进而改善记忆效果。这说明，对于精神分裂症患者，提取练习的有效性可能在于优化信息表征与抗干扰能力，从而有助于调动其尚存但效率较低的认知控制资源。

综上，提取练习的临床效应受核心认知缺陷的调节：其调节作用的核心可能在于，提取练习并非简单地“补偿”编码过程中的缺陷，而更多是在于“激活”个体相对低效的提取或抑制功能。当缺损集中于编码层面时，提取练习可能需要借助外部支持才能发挥较为持久的作用。当缺损集中于提取效率与干扰控制层面时，提取练习本身或可作为一种相对有效的认知康复手段。

4.2.2 学习材料的复杂性和生态性

提取练习在临床人群中是否有效，还受材料复杂性与生态性的影响。在复杂性方面，当 DLD 儿童学习动词时，提取练习能促进其学习效果，但相较于名词学习时的效应量更小，且提取练习未能改善其语法形态转换能力(Leonard et al., 2023)。这可能与学习动词和形态转换时需要的认知负荷有关。在学习复杂材料时，任务本身已对工作记忆构成高负荷，提取练习所带来的额外认知成本可能超过了其记忆巩固效果(Zheng et al., 2023)。这表明，将提取练习应用于临床实践并非简单地“使用”即可，而是一个精细的调控过程。这就要求研究者不断探索材料的“最优难度区间”，找到既能激活巩固机制、又不至于造成认知超载的难度平衡点，从而真正发挥其干预作用。

在生态性方面，Stern 和 Halamish(2023)的研究中并未发现提取练习在 ADHD 人群中的优势。一方面，前文提到这可能与被试的服药状态有关，另一方面，实验所采用的高生态性材料(文章全文或段落)可能对注意力调控能力本就有限的 ADHD 人群构成了额外认知负荷，反而削弱了提取练习的潜在效益。相比之下，Souto 等(2025)将提取练习嵌入绘本阅读这一自然互动情境中，结果显示出良好的生态效度，提示提取练习在符合日常学习方式的情境中更易发挥其促进作用。

综上所述，材料的复杂性与生态性共同调节提取练习的干预效果。在临床应用中，需根据患者的认知负荷承受能力与日常习惯，审慎选择与之匹配的学习材料，以期最大化提取练习的干预效益。

4.2.3 多治疗手段的支持

首先，药物治疗对提取练习效果的调节作用，揭示了该策略对个体基础认知状态的依赖。Stern 和 Halamish(2023)要求 ADHD 被试在实验前停药，这可能是导致提取练习未表现出干预效果的原因之一。反之，Minear 等(2023)的研究证明，药物能通过改善 ADHD 患者在学习阶段的注意、动机和努力水平(Addicott et al., 2019; Rubinson et al., 2019)，从而影响患者的记忆编码能力，间接促进提取练习的干预效果。由此可见，药物治疗所提供的认知支持，是提取练习这一干预方法生效的重要前提。

其次，物理治疗的介入则暗示，提取练习的干预效果在一定程度上依赖于特定神经环路

功能的有效性。Pan 等(2025)采用 tDCS 对精神分裂症患者的 L-DLPFC 皮层进行刺激，显著提升了提取练习的干预效果。这一促进作用可能来源于两方面机制：一方面，tDCS 通过影响 γ -氨基丁酸的抑制性功能并促进长时程增强(Long-term Potentiation)，为提取练习提供了更有利的神经环境。另一方面，tDCS 刺激并激活的 L-DLPFC 脑区，可能与提取练习所涉及的神经机制相吻合(Zhuang et al., 2022)。通过外部物理手段优化这一脑区的功能活性，相当于为提取练习的有效运作提供了必要的“硬件支持”，使其在信息编码与提取过程中的作用得以充分发挥。

综上，无论是药物改善认知准备状态，还是物理手段优化神经环境，其与提取练习的协同作用均表明：在临床群体中，提取练习是一种条件性的认知补偿策略，其有效性一定程度上依赖于由其他治疗手段所提供的认知与神经层面的基础支持。”

意见 2：关于“干预方案”的可操作性与临床转化。第四部分 4.2.4“干预方案的设计与实施”的内容偏向原则性建议，可以简短补充已有干预研究进行说明。

回应：感谢您的意见。针对干预方案的可操作性和临床转化，我们在结合两位专家的意见后决定，将原 4.2 现存问题与未来展望部分，独立为第 5 部分“未来研究方向与临床转化”其中 5.1、5.2 和 5.3 部分分别阐述三个未来研究方向，5.4 部分单独对临床转化进行详细阐述。具体修改内容如下：

5.4 干预方案的临床转化与实施

“提取练习干预方案的临床转化与实施，需要系统性思考和多维度整合。应注意以下几方面。

首先，明确干预目标并据此选择恰当的干预材料。例如，Middleton 等(2022)的实验目标直接靶向失语症患者的词汇-语义理解障碍，并因此选用了最小语义对的图片-词汇作为材料，每组由两个语义相近但不同的概念及其对应的图片组成(例如，蜘蛛和蝎子)。该研究通过将抽象的干预目标具体化为基于产出的提取练习和基于理解的提取练习，不仅精准匹配了临床需求，也最终验证了提取练习在改善特定认知缺陷方面的干预效果。

其次，将策略宣教融入干预流程以提升依从性。有研究认为人们可能不愿意使用提取练习这一策略(Karpicke et al., 2009)，但通过直接指导，可以有效提高个体使用提取练习的策略的意愿和能力(Ariel & Karpicke, 2018)。上述现象与机制可能在临床人群中同样存在。为此，提取练习的干预设计需超越单纯的任务训练，构建一个包含原理阐释(知识传递)、亲身体验(技能训练)和正向激励(动机调节)在内的综合指导体系。这一设计旨在帮助患者内化学习策略，从而提高其在日常生活中主动应用提取练习的意愿，最终实现干预效果的泛化。

最后，采用渐进式实施策略。可将提取练习的训练分为以下三阶段。第一，引入阶段。将提取练习作为患者原有认知训练或心理教育环节的补充。例如，在每次治疗结束前，安排 5 至 10 分钟的提取练习，要求患者回忆核心知识点。该阶段不追求训练强度，重在建立初

步的练习习惯。第二，扩展阶段。随着患者适应程度提高，可逐步提升干预剂量与任务挑战性。一方面延长提取练习时长，另一方面采用更具功能性的个性化材料，如记忆新姓名或购物路线等生活信息。至此，提取练习已成为干预方案的核心环节。第三，泛化阶段。该阶段的重点在于策略的迁移。患者开始学习如何将提取练习自主应用到日常生活中，如睡前回忆当日发生的事件。通过这一过程，帮助患者将提取练习内化为一种自主的、自我调节的学习工具，实现临床干预的目标。”

意见 3：关于检索策略与标题的微调。建议在“方法”部分补充说明是否进行了“引文追溯”或“手工检索”以补充遗漏文献，增强检索全面性的说明。

回应：感谢您的建议。在修改前，我们提到了手动检索，但的确不够精确和详细，因此，我们根据您的建议，对此部分进行了补充阐述，具体阐述内容如下：

“...之后还进行了补充性的手工检索，主要包括两个方面。第一，对纳入研究的参考文献列表进行回溯，以查找可能遗漏的相关研究。第二，手动搜索该研究领域内相关研究者或团队的研究成果。通过手工检索，补充获得了 14 篇文献。...”

意见 4：作者将标题从“一项系统综述”改为“应用研究”，可能弱化了文章的方法学属性。若稿件仍遵循系统综述的撰写规范(如 PRISMA 流程图、系统检索、质量评估)，建议保留“系统综述”，以更准确反映文章类型。

回应：感谢您对标题提出的修改意见。我们在第一轮修改时认为，文章主线已经修改为了“综合证据、发现模式、指出空白”，符合了系统综述的写作逻辑，但我们在未来展望方面阐述得较为详细。因此在第一轮修改中，我们从文章内容的角度考虑，删除了“系统综述”的表述。但此轮修改中，您给我们明确指出了“系统综述”是方法学属性，因此我们认为第一轮修改中的上述考虑有失偏颇，应从文章的方法学角度出发考虑该问题。我们的研究在两轮修改后仍然是遵循系统综述的写作规范，因此我们决定将题目改回原始题目，即“提取练习在临床人群中的应用研究：一项系统综述”

意见 5：文章当中部分语句有多字(第 7 页第 2 段第 6 行)，重复(第 10 页第 3 段第 6 行)的情况，以及标点符号使用不规范的情况，请通读全文进行校对。

回应：感谢您的仔细审阅，发现了细节错误，我们已经进行了修改，除此之外，我们还对以下内容进行了修改：

第一，文字措辞上：

1. 将“...在时常面临认知挑战的临床人群中，提取练习的应用前景如何...”**(引言部分第 2 段第 1 行)**，修改为：对于存在认知功能缺陷的临床人群，提取练习的应用前景如何...。
2. 将“...因此，研究补偿或改善此类患者的认知/记忆缺陷的有效策略至关重要...”**(引**

言部分第 2 段第 4 行)。修改为：因此，探索能改善此类患者认知或记忆缺陷的有效策略，具有重要的临床意义。

3. 将“...也未能整合近年来快速增长的、涵盖更广泛临床人群的新证据...”**(引言第 3 段第 1 行)**修改为：也未能梳理近年来涉及更丰富病种的新证据

4. 将“...其效果表现存在怎样的特异性和效应边界...”**(引言第 3 段第 4~5 行)**修改为：其效果存在怎样的特异性和效应边界

5. 删除“为临床康复实践提供基于证据的指引，并最终为未来研究指明创新方向”**(引言部分第 3 段最后 1 行)**。并将其上一句“并进一步指出现有研究存在的薄弱或空白领域。”修改为：并进一步指出现有研究存在的薄弱和未来研究方向，最后为临床干预方案的实施提出指导性意见。

6. 将“...从 2020 年 1 月到 2025 年 7 月”修改为“2020 年 1 月至 2025 年 7 月”**(在 2.1 文献检索部分第 1 段第 1 行，及其他相关部分)**

7. 将“...项目成功提取后停止重学, 且持续提取练习...”**(第 6 页第 2 段第 6 行)**，修改为：...初步掌握后继续提取，不再重学...

8. 将“...Wang 等(2020)的研究综合证据表明，间隔提取练习的效果显著优于集中提取练习，且单词对学习效果优于图片对...”**(在 3.1.4 失语症中第 1 段第 2 行，正文第 8 页)**。修改为：Wang 等(2020)的证据表明，间隔提取练习的效果显著优于集中提取练习，且词语学习效果优于图片。

9. 将“...尽管这些发现富有洞察力，但 PWA 患者的亚型颇多，患者之间的特异性较大，前人研究中的被试量均偏少。此外，提取练习在 PWA 患者跨领域联想学习效果和深层认知机制，以及提取练习在更复杂任务中的具体效果仍是未来研究需要填补的空白...”**(在 3.1.4 失语症中第 7~9 行，正文第 8 页)**修改为：“...尽管这些发现富有洞察力，但 PWA 患者亚型多样、个体间差异显著，这可能限制了结论的适用范围。同时，提取练习对该群体在复杂任务及跨领域联想学习中的干预效果及其深层认知机制，仍是当前研究中的薄弱环节....”

10. 将“...Lázaro(2024)的实验发现即时和间隔提取均能显著优于无提取条件，有效促进 DHH 儿童的 9.单词学习与记忆表现”**(正文第 9 页第 1 段第 2 行)**。修改为：...de Diego-Lázaro(2025)的实验发现 DHH 儿童即时和间隔提取的记忆效果显著优于无提取条件。(我们投稿时该文章为在线提前出版，当时文献信息不全，且我们本轮修改发现了作者姓氏错误的问题，遂进行了更正)

11. 将“...可促进语义信息的精细化组织，并增强提取练习的效果，与提取练习联合提升记忆维持能力(Pan et al., 2025)**(正文第 10 页第 2 段第 8~9 行)**。修改为...可促进语义信息的精细化组织，进而增强提取练习的干预效果(Pan et al., 2025)...

12. 将“...提取练习的补偿作用可能存在天花板效应或必须依赖其他因素，才能发挥长

久的作用...”。修改为：...提取练习的补偿作用可能存在一定的上限，或需借助其他支持性条件，才能发挥长久的作用...**(在 4.2.1 疾病核心认知缺陷的类型部分，正文第 13 页第 1 段第 5~6 行)**

13. 将“...为提升认知治疗在复杂病例中的实效提供科学支撑与实践指引...”。修改为：...为提取练习在复杂病例中的应用提供科学支撑与实践指引...**(在 5.1 提取练习在共病人群中的研究部分的第 3 段第 6 行)**

第二，标点符号校对上：

首先，关于摘要的冒号，参考“进展”的以往文章，中文摘要和关键词后没有冒号，因此我们的稿件中此处的标点符号和字体使用不规范，现已做出修改，具体修改如下：“**摘　要** 提取练习(Retrieval Practice)是通过测试....”。“**关键词**　提取练习...”

其次，关于全文的引号和逗号，我们根据第一轮审稿时一位审稿专家明确指出的意见(文内存在英文格式的逗号和引号，应为中文格式的逗号和引号)，再次检查了全文的标点符号。做了如下修改：

1. 全文的引号均修改为了中文格式

2. 在英文全称与缩写之间的逗号均修改为中文格式，例如(Transcranial Direct Current Stimulation，tDCS)、(Developmental Language Disorder，DLD)。

3. 将句子“...尽管已有综述关注了该领域, 但其范围多局限于特定类型的获得性障碍”中部的英文逗号修改为了中文格式。**(正文第 2 页第 2 段第 1 行)**

4. 将句子“...项目成功提取后停止重学, 且持续提取练习...”修改为：...初步掌握后继续提取，不再重学...。本句我们修改了表述，并将句子中的英文逗号修改为中文。**(第 6 页第 2 段第 6 行)**

5. “...早期精神病(EP)患者常存在情景记忆损伤, 尤其在自发提取信息时表现更差。该句中部的英文逗号和后半段已被删除。**(3.2.2 早期精神病部分第 1 段第 1 行)**

第三，参考文献上，我们更换了原 Lázaro, B. D. D. (2024). Retrieval practice and word learning in children who are hard of hearing. *International Journal of Speech-Language Pathology*, 11–14. Advance online publication.。因为该版本为提前在线出版的版本，没有页码，且我们发现作者的姓氏出现了错误，已经进行了修改，更正为：de Diego-Lázaro, B. (2025). Retrieval practice and word learning in children who are hard of hearing. *International Journal of Speech-Language Pathology, 27*(5), 735–748.

第四，我们还更换了图 1 文献筛选流程图，使其清晰度更高**(正文第 5 页)**。

意见 6：结果部分的“3.1 神经系统疾病”这个标题可以改为“提取练习在神经系统疾病人群中的作用”，然后下面的六类疾病可以都加上小标题，例如“3.1.1 发展性语言障碍”。3.2 部分也是同理。

回应：感谢您的建议，原版的二级标题的确存在表达不够精准的地方，我们已经按照您的建议进行了修改。分别是：3.1 提取练习在神经系统疾病人群中的作用、3.2 提取练习在精神障碍人群中的作用、3.3 提取练习在 HIV 感染者和唐氏综合征儿童中的作用。此外，我们也添加了 3 级标题，在 3.1 部分中分别是：3.1.1 发展性语言障碍、3.1.2 注意缺陷/多动障碍、3.1.3 阿尔茨海默症、3.1.4 失语症、3.1.5 听力障碍、3.1.6 脑卒中和 3.1.7 不同神经系统疾病个体的提取练习差异。在 3.2 部分中分别是：3.2.1 精神分裂症、3.2.2 早期精神病和 3.2.3 不同精神障碍个体的提取练习差异。

审稿人 2 意见：

论文做了大篇幅修改。如下意见供参考。

回应：非常感激您再度审阅我们的修改稿，您为第二轮优化所提出的意见，对我们至关重要。我们已经根据您的建议，逐一细致考量并修订了稿件。特此向您致以最诚挚的感谢。(为和文中已有的蓝色和红色区分，我们将本次修改后的文字在正文标记为了绿色)。

意见 1：摘要部分相较于已有综述的独特价值或新发现(如首次系统梳理特定临床人群、发现了特定的调节模式)未体现。

回应：感谢您的意见，我们已做出相应修改。此外，我们还精简了摘要的部分表述，以期符合“进展”对摘要字数的要求(修改后摘要为 272 个字符)。英文摘要做了相应修改。具体修改如下：

摘　要　提取练习(Retrieval Practice)是通过测试、自由回忆等方式从记忆系统中提取信息以提高记忆效率的策略。部分临床患者可能会面临记忆或认知能力的缺陷。本文首次系统梳理了 2020 年 1 月至 2025 年 7 月提取练习在临床人群中的研究，最终纳入了 21 篇文献。结果显示，提取练习在临床人群中表现出良好的适用性，可改善记忆、语言障碍等认知损伤。提取练习的效果受疾病的核心认知缺陷、学习材料的复杂性和生态性，以及是否有其他治疗手段的调节。未来可关注共病人群、提取练习的间接效应及个体差异等问题。在临床转化方面可通过明确干预目标、加强指导与宣教、采用渐进式实施策略三方面措施，提升提取练习的临床应用性。

关键词　提取练习，临床，记忆，认知康复

Abstract: Retrieval practice is a strategy that enhances memory efficiency by extracting information from the memory system through methods such as testing or free recall. Some clinical patients may face memory or cognitive impairments. This paper systematically reviews, for the first time, research on retrieval practice in clinical populations from January 2020 to July 2025, resulting in the inclusion of 21 studies. The results demonstrated that retrieval practice has good

applicability in clinical populations, leading to improvements in cognitive impairments such as memory and language disorders. However, its effectiveness of retrieval practice was moderated by the core cognitive deficits associated with the condition, the complexity and ecological validity of the learning materials, as well as the presence of other therapeutic interventions. Future studies should focus on interventions for comorbid conditions, indirect benefits, and the role of individual differences. In terms of clinical translation, measures such as clarifying intervention goals, strengthening guidance and education, and adopting a gradual implementation strategy should be taken to enhance the clinical applicability of retrieval practice.

Keywords: Retrieval Practice，Clinical，Memory，Cognitive Rehabilitation

意见 2：引言中在神经机制的论证逻辑链条可更严谨。如在引用脑区研究后，直接推论“因此提取练习可能会通过…激活相关神经回路来帮助临床人群改善认知缺陷”，这个推论中间缺了一环逻辑，略显突然。

回应：感谢您的细致审阅。我们对此做了修改，修改后相关表述如下**(正文第 2 页第 1 段第 3 至 10 行)**：

“.....提取练习则是一种具有应用潜力的认知干预手段。首先，提取练习可能诱发精细化编码进而提高记忆能力(廖陈陈, 张锦坤, 2024)。其次，研究发现提取练习效应与内侧前额叶、腹外侧前额叶、额下回、腹侧纹状体及中脑等脑区密切相关(Van den Broek et al., 2013; Ye et al., 2020; Zhuang et al., 2022)。值得注意的是，这些脑区不仅在记忆巩固与提取中扮演关键角色，其功能异常还与阿尔茨海默症、精神分裂症等疾病的记忆缺陷病理机制密切相关(Liu et al., 2021; Williams et al., 2023)。因此，基于上述共同的神经基础，提取练习不仅可能通过精细化编码来提升记忆表现，更有可能调节临床人群功能异常的神经回路，从而在神经机制层面为改善其认知缺陷提供干预途径。....”

意见 3：方法中质量评估部分表述采用了 7 项检查表，但结论表述为“除了一项研究以外，其他研究在七个方面的四个方面都被评估为低风险”。这一表述欠透明，“四个方面”是哪四个方面？有多少研究在关键项目(如随机化、结果测量)上存在高风险？这些高风险偏倚会如何影响研究结论的可靠性吗？

回应：感谢您对这一部分的意见，原稿对该部分确实阐述不够透明，因此我们做了修改。**首先，**为保证质量评估的准确性，我们三位作者对所有文献进行了再次评估(之前是两位作者)，并进行了讨论，最终形成了最新的质量评估结果。**其次，**我们将 2.3 部分拆分为了 2.3.1 评估工具和 2.3.2 评估结果。2.3.1 部分介绍了评估所用的工具，主要包括随机化等 7 方面。我们新增了 2.3.2 部分用于报告在关键方面(随机化、结果测量和选择性报告)的评估结果。以及在其余 4 方面中的偏倚情况，并解释了评估为风险不明或高风险的可能原因，并阐述了可

能带来的影响，并且进行了总结。**最后，**我们在讨论后认为，纳入的研究虽然在部分方面存在偏倚，但在关键方面，如随机化、结果测量和选择性报告方面，纳入的文献质量较好，因此，尽管部分研究的存在高风险或风险不明，但这并不影响我们对于“提取练习可应用于临床人群”这一初步结论的判断。具体修改如下**(正文第 4 页)**：

2.3 质量评估

2.3.1 评估方法

为了评估纳入研究的方法学质量，本研究进行了偏倚风险评估。根据 de Lima 等人(2020)依据 Higgins 等(2019)和 Ding 等(2015)的模板开发的 7 方面检查表进行了评估。该清单共包含 7 个方面，分别评估：随机化过程(方面 1)、缺失结果数据(方面 2)、结果测量(方面 3)、计分(方面 4)、选择性报告(方面 5)、多次最终测试(方面 6)和初始学习(方面 7)。该工具从随机化、数据缺失、评分、报告等维度进行评价，符合心理学研究的一般标准。其中方面 6 和方面 7 专门针对提取练习实验设计，适用于本领域研究。各方面的操作定义和评分标准详见补充材料。评估由三位作者共同完成。

2.3.2 评估结果

评估结果显示(见附录表 4)，纳入的 21 项研究在核心方法学维度上总体质量较好。随机化过程(方面 1)中，18 项研究为低风险，3 项研究因报告不足被判定为风险不明，无高风险研究。结果测量(方面 3)同样无高风险，但 8 项研究因细节报告不足而风险不明。选择性报告(方面 5)全部为低风险。

证据的不确定性主要体现在：第一，多次最终测试(方面 6)是主要偏倚来源，共有 12 项高风险研究，因测试本身构成额外练习，这可能高估提取练习的长期效果。第二，计分偏倚(方面 4)方面，10 项研究风险不明，这可能多因电子自动计分而未阐述是否盲评或报告评分者信度。第三，缺失数据(方面 2)方面，1 项高风险研究、6 项风险不明，部分研究未报告退出者与完成者的系统差异。第四，初始学习信息(方面 7)方面，3 项高风险研究、2 项风险不明，学习阶段报告不足可能影响对提取练习效应的归因。

总体而言，上述结果表明，关于提取练习对临床人群具有积极作用的结论是相对稳健的。纳入研究在随机化、结果测量和选择性报告等核心方法学维度上普遍质量较好，增强了证据的可信度。但效应量精确估计和长期效果的确定性受多次测试、缺失数据处理及报告透明度的制约。未来研究应控制多次测试干扰，并详细报告计分流程、缺失数据及学习阶段表现。

意见 4：结果呈现方式可更丰富，建议在 3.1 节之前，增加一小段“纳入研究总体特征”的文字概述，并提炼 1-2 个关键总体特征。

回应：感谢您的意见，我们在 3.1 节前增加了描述总体特征的表述。我们首先提炼了两方面特征：第一，在疾病分布方面。第二在干预与评估的内容方面。其次我们提出了总体性的结论，即在大多数研究中，提取练习均显示出优于重复学习或其他对照条件的效果。尽管具体

效应存在差异，但这一发现初步验证了该干预手段在临床人群中的应用潜力。具体阐述如下**(正文第 6 页)**：

3 结果

本综述共纳入 21 项研究，主要考察提取练习作为一种认知干预策略在不同临床人群中的应用效果(见附录表 1)。这些研究总体上呈现以下特征：首先，在疾病分布方面，研究高度集中于神经系统疾病(共 15 篇)，其中以发展性语言障碍(Developmental Language Disorder，DLD)相关研究最多(5 篇)。其次，注意缺陷/多动障碍(Attention Deficit/Hyperactivity Disorder，ADHD)3 篇，阿尔茨海默症(Alzheimer Disease，AD)、失语症(People with Aphasia，PWA)、听力障碍(Deaf and Hard of Hearing，DHH)各 2 篇，脑卒中幸存者(Stroke Survivors，SS)1 篇。在精神障碍领域中，共纳入 4 篇研究，其中 3 篇关注精神分裂症(Schizophrenia，SP)，1 篇关注早期精神病(Early Psychosis，EP)。此外，还包括 HIV 感染者 1 篇，以及唐氏综合征(Down Syndrome，DS)1 篇。其次，在干预与评估内容上，多数研究聚焦于提取练习对陈述性记忆(词汇与事件记忆)的促进作用，仅少数研究探讨了其在复杂认知任务(如动词形态转换)中的迁移效果。在大多数研究中，提取练习均显示出优于重复学习或其他对照条件的效果。尽管具体效应存在差异，但这一发现初步验证了该干预手段在临床人群中的应用潜力。下文将按疾病类别对研究结果进行详细阐述。

意见 5：对于同一疾病内不同研究结论的异质性(如 ADHD 研究中，不同提取形式的比较结果)，主要是描述差异，而未初步分析可能导致异质性的原因(如样本差异、干预剂量不同)。

回应：感谢您对异质性结果阐述的建议，我们做了如下修改：

第一，在 ADHD 的研究中，我们做了如下修改和补充，将不同研究的异质性原因归纳为药物的作用，具体阐述如下**(正文第 7 页第 2 段)**：

“研究表明，ADHD 患者在提取练习条件下的表现比无练习条件更优，这提示提取练习对 ADHD 患者具有积极作用(Knouse et al., 2020)。该研究报告了患者的服药情况，但并未探究药物治疗情况对提取练习的影响。进一步地，Stern 和 Halamish(2023)要求 ADHD 患者在实验前停药，该研究虽证明了全文自由回忆比段落自由回忆具有更优的记忆效果，但该研究未观察到提取练习效应。这可能与被试在实验前暂停服药有关。Minear 等(2023)则进一步考察了患者服药状态对提取练习的影响。研究显示，在服药状态下，提取练习会比在未服药状态下更有效。这表明，提取练习虽有效，但可能无法弥补基础编码问题，其效果可能受到患者的编码能力的调节。对于未服药患者，其编码能力较弱，提取练习可能不足以完全补偿其缺陷(Stern & Halamish, 2023)。而对于经药物改善编码功能的患者，提取练习则能充分发挥其促进提取和巩固的优势。”

第二，针对 DLD 儿童，我们在神经系统疾病的总结性讨论时，分析了提取练习对 DLD 儿童的动词形态转换任务作用不显著的原因，具体阐述如下**(正文第 10 页第 1 段)**：

“其次，效应的迁移性存在局限。在 DLD 儿童的动词形态转换中，提取练习的“近迁移”效果显著，但“远迁移”至新情境或复杂规则应用的优势有限(Leonard et al., 2023)。这或许与实验的难度有关，动词形态转换对 DLD 儿童可能存在地板效应，最终导致提取练习的效果有限。”

意见 6：讨论开篇表述提取练习在“高度异质的临床人群中展现出稳定适用性”，并推测其机制具有“跨群体普适性”。但本研究仅纳入 21 篇文献，从有限的、可能存在发表偏倚的证据中直接推导出“普适”的结论，削弱了系统综述结论应有的谨慎性。

回应：感谢您对我们的批评指正，跨群体的适用性表述的确过于绝对化，让研究结论不严谨。因此我们修改了相关表述**(正文第 12 页)**。具体修改如下：

将原小标题“4.1.1 跨群体的适用性”修改为“提取练习在临床人群中的适用性”。

将正文中“跨群体普适性”修改为“可能具有一定程度的跨疾病稳定性，能够在异质性的临床群体中展现出一定的应用潜力。”。

研究结论与贡献部分也做了相应修改。

意见 7：“讨论”的核心应是解释和综合本综述的发现，但此部分超过三分之二的篇幅(4.2.1 至 4.2.4)用于阐述“现存问题与未来展望”。4.2.1(共病人群)和 4.2.3(个体差异)所讨论的问题，恰恰是本次综述因其纳入/排除标准而未能充分涵盖的领域。大篇幅讨论本综述“未研究什么”，而非深入分析“研究了什么及其意义”，讨论重心有所偏离。

回应：感谢您的批评指正。针对这一部分，我们讨论重心的确有所偏离。在综合考虑后，我们决定对讨论部分的框架和内容进行调整，区分讨论和未来方向两个部分，并扩展讨论部分。具体修改如下：

首先，将原 4.2 现存问题与未来展望，独立为第 5 部分，即 5 未来研究方向与临床转化，原 4.2.1、4.2.2、4.2.3 和 4.2.4 做了相应调整，原 5 本研究的不足与局限和原 6 研究结论与贡献顺位为第 6 和第 7 部分。

其次，我们结合两位专家的意见，删除原标题 4.1 提取练习在临床人群中的适用性与特异性。将原 4.1.1 跨群体的适用性和 4.1.2 人群特异性与效应边界，分别修改为：4.1 提取练习在临床人群中的适用性、4.2 提取练习的人群特异性与效应边界。

在修改后的 4.1 中，我们精简了部分疾病名称，在此之后，我们新增了一段内容，来进行讨论，添加的内容如下**(正文第 12 页)**：

“...这一结果提示，提取练习对记忆的增强作用可能具有一定程度的跨疾病稳定性，能够在异质性的临床群体中展现出一定的应用潜力。其潜在机制或许涉及对海马等记忆核心脑区功能的调节(Wiklund-Hörnqvist et al., 2021)，但这一推断仍需在未来针对不同临床群体的研究中进一步验证。总体而言，现有证据初步提示，提取练习作为一种认知干预策略，其具

备一定的认知干预效果和临床应用潜力，这为其后续迈向更广泛的临床标准化应用提供了依据。”

在修改后的 4.2 中，我们将三方面的调节因素提升为 3 级标题，并在部分维度添加了支持证据。具体修改如下**(正文第 12~15 页)**：

4.2 提取练习的人群特异性与效应边界

尽管总体结果支持提取练习的有效性，但提取练习的效应强度与作用路径仍存在人群特异性。在临床人群中，提取练习的本质是一种受到疾病类型调节，且有条件性的认知补偿策略，其特异性和效应边界主要体现在以下三方面。

4.2.1 疾病核心认知缺陷的类型

提取练习在临床人群中的效应并非完全一致，可能受到疾病核心认知缺陷类型的调节。Stamate 等(2020)指出，AD 患者的真正挑战在于需要更多次的学习才能将信息成功编码进入记忆系统，患者的记忆障碍主要表现为编码阶段的困难。Small 和 Cochrane（2020）的研究亦指出，提取练习的长期效果部分依赖线索的支持，当支持被撤除时，提取练习的长期维持效果可能会减弱。这说明，在以编码缺陷为特点的疾病中，提取练习的补偿作用可能存在一定的上限，或需借助其他支持性条件，才能发挥长久的作用(Stamate et al., 2020)。

相比之下，精神分裂症患者的认知缺损更多表现为干扰抑制功能下降与提取效率较低(Weisbrod et al., 2000)。Tan 等(2024)和 Wen 等(2025)的研究发现，提取练习能通过减少前摄干扰、促进语义组织能力，进而改善记忆效果。这说明，对于精神分裂症患者，提取练习的有效性可能在于优化信息表征与抗干扰能力，从而有助于调动其尚存但效率较低的认知控制资源。

综上，提取练习的临床效应受核心认知缺陷的调节：其调节作用的核心可能在于，提取练习并非简单地“补偿”编码过程中的缺陷，而更多是在于“激活”个体相对低效的提取或抑制功能。当缺损集中于编码层面时，提取练习可能需要借助外部支持才能发挥较为持久的作用。当缺损集中于提取效率与干扰控制层面时，提取练习本身或可作为一种相对有效的认知康复手段。

4.2.2 学习材料的复杂性和生态性

提取练习在临床人群中是否有效，还受材料复杂性与生态性的影响。在复杂性方面，当 DLD 儿童学习动词时，提取练习能促进其学习效果，但相较于名词学习时的效应量更小，且提取练习未能改善其语法形态转换能力(Leonard et al., 2023)。这可能与学习动词和形态转换时需要的认知负荷有关。在学习复杂材料时，任务本身已对工作记忆构成高负荷，提取练习所带来的额外认知成本可能超过了其记忆巩固效果(Zheng et al., 2023)。这表明，将提取练习应用于临床实践并非简单地“使用”即可，而是一个精细的调控过程。这就要求研究者不断探索材料的“最优难度区间”，找到既能激活巩固机制、又不至于造成认知超载的难度平衡点，从而真正发挥其干预作用。

在生态性方面，Stern 和 Halamish(2023)的研究中并未发现提取练习在 ADHD 人群中的优势。一方面，前文提到这可能与被试的服药状态有关，另一方面，实验所采用的高生态性材料(文章全文或段落)可能对注意力调控能力本就有限的 ADHD 人群构成了额外认知负荷，反而削弱了提取练习的潜在效益。相比之下，Souto 等(2025)将提取练习嵌入绘本阅读这一自然互动情境中，结果显示出良好的生态效度，提示提取练习在符合日常学习方式的情境中更易发挥其促进作用。

综上所述，材料的复杂性与生态性共同调节提取练习的干预效果。在临床应用中，需根据患者的认知负荷承受能力与日常习惯，审慎选择与之匹配的学习材料，以期最大化提取练习的干预效益。

4.2.3 多治疗手段的支持

首先，药物治疗对提取练习效果的调节作用，揭示了该策略对个体基础认知状态的依赖。Stern 和 Halamish(2023)要求 ADHD 被试在实验前停药，这可能是导致提取练习未表现出干预效果的原因之一。反之，Minear 等(2023)的研究证明，药物能通过改善 ADHD 患者在学习阶段的注意、动机和努力水平(Addicott et al., 2019; Rubinson et al., 2019)，从而影响患者的记忆编码能力，间接促进提取练习的干预效果。由此可见，药物治疗所提供的认知支持，是提取练习这一干预方法生效的重要前提。

其次，物理治疗的介入则暗示，提取练习的干预效果在一定程度上依赖于特定神经环路功能的有效性。Pan 等(2025)采用 tDCS 对精神分裂症患者的 L-DLPFC 皮层进行刺激，显著提升了提取练习的干预效果。这一促进作用可能来源于两方面机制：一方面，tDCS 通过影响 γ -氨基丁酸的抑制性功能并促进长时程增强(Long-term Potentiation)，为提取练习提供了更有利的神经环境。另一方面，tDCS 刺激并激活的 L-DLPFC 脑区，可能与提取练习所涉及的神经机制相吻合(Zhuang et al., 2022)。通过外部物理手段优化这一脑区的功能活性，相当于为提取练习的有效运作提供了必要的“硬件支持”，使其在信息编码与提取过程中的作用得以充分发挥。

综上，无论是药物改善认知准备状态，还是物理手段优化神经环境，其与提取练习的协同作用均表明：在临床群体中，提取练习是一种条件性的认知补偿策略，其有效性一定程度上依赖于由其他治疗手段所提供的认知与神经层面的基础支持。

意见 8：文内出现多处标点符号使用不当或格式错误的问题，摘要的冒号等，中文引号的使用等。

回应：感谢您的仔细审阅，并帮助我们发现了细节错误。

首先，关于摘要的冒号，参考“进展”的以往文章，中文摘要和关键词后没有冒号，因此我们的稿件中此处的标点符号和字体使用不当，现已做出修改，具体修改如下：“摘　要 提取练习(Retrieval Practice)是通过测试....”。“关键词　提取练习...”

其次，关于全文的引号和逗号，我们根据第一轮审稿时专家明确指出的意见(文内存在英文格式的逗号和引号，应为中文格式的逗号和引号)，再次检查了全文的标点符号。做了如下修改：

1. 全文的引号均修改为了中文格式。

2. 在英文全称与缩写之间的逗号均修改为中文格式，例如(Transcranial Direct Current Stimulation，tDCS)、(Developmental Language Disorder，DLD)。

3. 将句子“...尽管已有综述关注了该领域, 但其范围多局限于特定类型的获得性障碍”中部的英文逗号修改为了中文格式。**(正文第 2 页第 2 段第 1 行)**

4. 将句子“...项目成功提取后停止重学, 且持续提取练习...”修改为：...初步掌握后继续提取，不再重学...。本句我们修改了表述，并将句子中的英文逗号修改为中文。**(第 6 页第 2 段第 6 行)**

5. “...早期精神病(EP)患者常存在情景记忆损伤, 尤其在自发提取信息时表现更差。该句中部的英文逗号和后半段已被删除。**(3.2.2 早期精神病部分第 1 段第 1 行)**

6. 删除了“图 1: 文献筛选流程图”的冒号，**修改为**图 1 文献筛选流程图

此外，我们还更换了图 1 文献筛选流程图，使其清晰度更高**(正文第 5 页)**。

最后，关于参考文献，我们更换了原 Lázaro, B. D. D. (2024). Retrieval practice and word learning in children who are hard of hearing. *International Journal of Speech-Language Pathology*, 11–14. Advance online publication.。因为该版本为提前在线出版的版本，没有页码，且我们发现作者的姓氏出现了错误，已经进行了修改。具体更正为：de Diego-Lázaro, B. (2025). Retrieval practice and word learning in children who are hard of hearing. *International Journal of Speech-Language Pathology, 27*(5), 735–748.

意见 9：文内多处句子在阅读过程中欠通顺。

回应：感谢您的意见，我们再次逐句阅读了文稿，主要做了如下修改：

1. 将“...在时常面临认知挑战的临床人群中，提取练习的应用前景如何...”**(引言部分第 2 段第 1 行)**，修改为：然而，对于存在认知功能缺陷的临床人群，提取练习的应用前景如何...。

2. 将“...因此，研究补偿或改善此类患者的认知/记忆缺陷的有效策略至关重要...”**(引言部分第 2 段第 4 行)**。修改为：因此，探索能改善此类患者认知或记忆缺陷的有效策略，具有重要的临床意义。

3. 将“...也未能整合近年来快速增长的、涵盖更广泛临床人群的新证据...”**(引言第 3 段第 2 行)**修改为：也未能梳理近年来涉及更丰富病种的新证据

4. 将“...其效果表现存在怎样的特异性和效应边界...”**(引言第 3 段第 4~5 行)**修改为：其效果存在怎样的特异性和效应边界

5. 删除“为临床康复实践提供基于证据的指引，并最终为未来研究指明创新方向”**(引言部分第 3 段最后一行)**。并将其上一句“并进一步指出现有研究存在的薄弱或空白领域。”修改为：“并进一步指出现有研究存在的薄弱和未来研究方向，最后为临床干预方案的实施提出指导性意见。”

6. 将“...从 2020 年 1 月到 2025 年 7 月”修改为“2020 年 1 月至 2025 年 7 月”**(正文第 3 页 2.1 文献检索第 1 段第 1 行，及其他相关部分)**

7. 将“...项目成功提取后停止重学, 且持续提取练习...”**(第 6 页第 2 段第 6 行)**，修改为：...初步掌握后继续提取，不再重学...

8. 将“...但 1 个月延时测试中提取练习组在的回忆表现仍显著优于无提取练习组...”。**(在 3.1.3 阿尔茨海默症中第 1 段第 6 行，正文第 8 页)**修改为...但提取练习组在延迟测试中的表现仍优于无提取练习组...

9. 将“...Wang 等(2020)的研究综合证据表明，间隔提取练习的效果显著优于集中提取练习，且单词对学习效果优于图片对...”**(在 3.1.4 失语症中第 1 段第 2 行，正文第 8 页)**。修改为 Wang 等(2020)的证据表明，间隔提取练习的效果显著优于集中提取练习，且词语学习效果优于图片。

10. 将“...尽管这些发现富有洞察力，但 PWA 患者的亚型颇多，患者之间的特异性较大，前人研究中的被试量均偏少。此外，提取练习在 PWA 患者跨领域联想学习效果和深层认知机制，以及提取练习在更复杂任务中的具体效果仍是未来研究需要填补的空白...”**(在 3.1.4 失语症中第 7~9 行，正文第 8 页)**修改为：“...尽管这些发现富有洞察力，但 PWA 患者亚型多样、个体间差异显著，这可能限制了结论的适用范围。同时，提取练习对该群体在复杂任务及跨领域联想学习中的干预效果及其深层认知机制，仍是当前研究中的薄弱环节....”

11. 将“...Lázaro(2024)的实验发现即时和间隔提取均能显著优于无提取条件，有效促进 DHH 儿童的 9.单词学习与记忆表现”**(在 3.1.5 听力障碍中第 1 段第 2 行，正文第 8~9 页)**。修改为：...de Diego-Lázaro(2025)的实验发现 DHH 儿童即时和间隔提取的记忆效果显著优于无提取条件。(我们投稿时该文章为在线提前出版，当时文献信息不全，且我们本轮修改发现了作者姓氏错误的问题，遂进行了更正)

12. 将“...以上两种疾病类型仍需更大样本研究进一步研究...”(该句发现了重复表述。**第 11 页第 4 段第 6 行**)修改为：“...以上两种疾病类型仍需更大样本进一步研究...”

13. 将“...可促进语义信息的精细化组织，并增强提取练习的效果，与提取练习联合提升记忆维持能力(Pan et al., 2025)**(在 3.2.1 精神分裂症部分，正文第 10 页第 2 段第 8~9 行)**。修改为...可促进语义信息的精细化组织，进而增强提取练习的干预效果(Pan et al., 2025)...

14. 将“...提取练习的补偿作用可能存在天花板效应或必须依赖其他因素，才能发挥长久的作用...”。(第十三页第一段第三行)修改为：...提取练习的补偿作用可能存在一定的上

限，或需借助其他支持性条件，才能发挥长久的作用...**(在 4.2.1 疾病核心认知缺陷的类型部分，正文第 13 页第 1 段第 5~6 行)**

15. 将“...为提升认知治疗在复杂病例中的实效提供科学支撑与实践指引...”。修改为：...为提取练习在复杂病例中的应用提供科学支撑与实践指引...**(在 5.1 提取练习在共病人群中的研究部分的第 3 段第 6 行)**

第三轮

审稿人 1 意见：

经过对修改稿的仔细审阅，我认为作者已经充分、妥善地回复了第二轮审稿意见。文章的质量得到显著提升，逻辑结构更加清晰，论证更加充实，行文更加流畅。我对稿件当前版本没有新的意见。

回应：衷心感谢您的仔细审阅与前两轮所提的宝贵建议！

审稿人 3 意见：

本文经过两轮修改后，整体结构已有明显改善，但从系统综述的严谨性和论证力度来看，仍有一些关键问题尚未充分解决。以下意见供作者进一步修改参考。

回应：非常感谢您的审阅！为提升研究的严谨性和论证力度提出了宝贵的意见。我们已逐一参考您的意见进行修改，由衷感谢您的指导(为和文中已有的蓝色、红色和绿色区分，我们将本次修改后的文字在正文标记为了紫色)。

意见 1：当前稿件在方法部分对系统综述实施过程的描述仍不够充分。文中仅写到“评估由三位作者共同完成”，但未进一步说明标题摘要筛选、全文筛选、数据提取和偏倚评估是否为独立进行，也未说明分歧如何处理、是否评估一致性。这些信息对于系统综述的可信度十分重要。相关问题主要见于 2.2 排除和纳入标准及 2.3.1 评估方法。

回应：感谢您的意见。我们在之前的确未详细报告这一部分，为了提高本研究的严谨性，我们做了如下工作：

第一，重新筛选文献，在筛选过程中，两位筛选者的结果与各自的第一次文献筛选结果基本一致。出现的分歧由第三位作者介入并讨论，主要的分歧为：是否将 Wilschut 等(2025)纳入本研究。该研究的研究目的为，基于打字和基于语音的两种提取练习反应形式是否对患有发展性阅读障碍的学习者有差异影响。但在研究设计中，该研究仅比较了上述两种提取练习反应形式，未涉及到提取练习效应。因此本研究的一位作者意见为排除该文献，这也正是

在第一次文献筛选时将其排除的原因。但在本次讨论中，另一位作者认为，基于打字和基于语音的提取练习可能存在区别，也属于两种不同的学习方式，该研究的结果也证明了该观点。因此，可认为该文献符合本研究的主题。在由第三位作者介入并讨论后，决定将其纳入本研究。本研究所纳入的文献最终为 22 篇。

同时我们也考虑到原稿对纳入标准第 4 条(在学习方式的变量上，需要有两个及以上的不同水平，如提取练习和重复学习)的阐述不够清晰，才会出现上述争议，因此我们对此做了完善，将其修改为“4)在学习方式的变量上，需有两个及以上的不同水平(如提取练习和重复学习)，或对提取练习的不同形式进行比较(如间隔提取与即时提取)。”。

第二，我们细化了对文献筛选和信息提取部分的阐述，新增了“2.3 文献的筛选及信息提取”部分。原“2.3 质量评估”修改为“2.4 质量评估”。原“2.3.1 评估方法”修改为“2.4.1 评估方法与过程”。原“2.3.2 评估结果”修改为“2.4.2 评估结果”。在 2.3 文献的筛选及信息提取”部分，我们报告了文献筛选及信息提取是否独立，并参考徐建平和张厚粲(2005)提出的“归类一致性”公式，检测文献筛选和信息提取的一致性。经计算 2 名作者在题目和摘要初筛、全文筛选及信息提取时的归类一致性分别为 0.97、0.93、0.92。随后，筛选和信息提取的分歧，则与第 3 位作者讨论并由其作出最终决定。

第三，我们细化了对文献质量评估部分的阐述，同时我们也对质量评估的结果进行了检查，评估过程中存在的分歧之处再次进行了深入讨论。在“2.4.1 评估方法与过程”部分，我们报告了评估的过程、是否独立、分歧如何处理，并计算了评估一致性(受常见一致性指标 kappa 局限性的影响，我们选用了更为直观的一致性比例作为一致性指标，计算方法为：不同评分者将对象分类到同一类别中的次数除以总样本量。该指标为评分者之间观察到的一致性，可以用作一致性的估计，参考文献：Flight, L., & Julious, S. A. (2015). The disagreeable behaviour of the kappa statistic. *Pharmaceutical statistics, 14*(1), 74–78.)。具体修改如下：“2 名作者独立对所有纳入研究的文献进行偏倚风险评估，偏倚评估的一致性为 93.51%(Flight & Julious, 2015)。评估分歧之处，由第 3 名作者介入讨论并作出最终决定。”

第四，在“2.4.2 评估结果”部分我们细化了纳入文献的质量报告，具体为“参考魏依婷 等(2024)中的类似评估方法与标准，本文纳入的高质量文章(6～7 个方面为低风险)占比为 27.27%(6 / 22) ，中等质量的文章(3～5 个方面为低风险)占比 68.18%(15 / 22)，低质量文章(1～2 个方面为低风险)占比 4.55%(1 / 22)。最终评估结果显示(见附录表 2)，纳入的 22 项研究总体质量中等。”

相关部分的具体修改如下：

2.3 文献的筛选及信息提取

本研究的文献筛选过程遵循 PRISMA 标准(Moher et al., 2009)。由 2 名作者独立开展文献筛选和信息提取工作。完成后，这两人交叉核对彼此的筛选结果和提取的信息，提取的信息包括：标题、作者、发表时间、被试、学习材料、学习方式、最终测试时间点、结论。2

名作者在题目和摘要初筛、全文筛选及信息提取时的归类一致性分别为 0.97、0.93、0.92(徐建平, 张厚粲, 2005)。分歧之处，则与第 3 名作者讨论并由其作出最终决定。最终，通过题目和摘要初筛获得 24 篇文献，经全文筛选后最终共纳入 22 篇文献。具体的筛选过程见图 1。

2.4 质量评估

2.4.1 评估方法与过程

为了评估纳入研究的方法学质量，本研究进行了偏倚风险评估。根据 de Lima 等人(2020)依据 Higgins 等(2019)和 Ding 等(2015)的模板开发的 7 方面检查表进行了评估。该清单共包含 7 个方面，分别评估：随机化过程(方面 1)、缺失结果数据(方面 2)、结果测量(方面 3)、计分(方面 4)、选择性报告(方面 5)、多次最终测试(方面 6)和初始学习(方面 7)。该工具从随机化、数据缺失、评分、报告等维度进行评价，符合心理学研究的一般标准。其中方面 6 和方面 7 则针对提取练习实验设计，适用于本领域研究。各方面的操作定义和评分标准详见补充材料。2 名作者独立对所有纳入研究的文献进行偏倚风险评估，偏倚评估的一致性为 93.51%(Flight & Julious, 2015)。评估分歧之处，由第 3 名作者介入讨论并由其作出最终决定。

2.4.2 评估结果

参考魏依婷 等(2024)中的类似评估方法与标准，本文纳入的高质量文章(6～7 个方面为低风险)占比为 27.27%(6 / 22) ，中等质量的文章(3～5 个方面为低风险)占比 68.18%(15 / 22)，低质量文章(1～2 个方面为低风险)占比 4.55%(1 / 22)。最终评估结果显示(见附录表 2)，纳入的 22 项研究总体质量中等。其中，随机化过程有 21 项研究为低风险，1 项研究风险不明。最终测试方面无高风险研究，但 9 项研究因细节报告不足而风险不明。选择性报告方面全部为低风险。

证据的不确定性主要体现在：第一，多次最终测试方面共有 12 项高风险研究，因测试本身构成额外练习，这可能高估提取练习的长期效果。第二，计分偏倚方面，10 项研究风险不明，这可能多因电子自动计分而未阐述是否盲评或报告评分者信度。第三，缺失数据方面，1 项高风险研究、6 项风险不明，部分研究未报告退出者与完成者的系统差异。第四，初始学习信息方面，2 项高风险研究、2 项风险不明，学习阶段报告不足可能影响对提取练习效应的归因。

总体而言，上述结果表明，关于提取练习对临床人群具有积极作用的结论是相对稳健的。纳入研究在随机化、结果测量和选择性报告等核心方法学维度上普遍质量较好，增强了证据的可信度。但提取练习在临床人群中长期效果的确定性可能受多次测试的制约。未来研究应控制多次测试干扰，并详细报告计分流程、缺失数据及学习阶段表现。

意见 2：当前稿件对不同病种的结论在文字上权重相近，但实际上各病种的证据基础差异较大。有的领域已有多篇研究、结论相对集中，有的领域则仅有 1-2 项研究。现有写法容易给人一种各病种支持程度较为接近的印象，缺少对证据强弱的区分。

回应：感谢您的宝贵意见。部分只有少量研究的疾病领域，我们的阐述也与其他研究的权重较为接近，这样的确可能给读者造成误区。因此，我们决定缩减部分疾病的阐述，并合并一些只有少量研究(2 项及以下)的疾病阐述。主要修改涉及两方面：(1)将听力障碍(2 项)、脑卒中(2 项)、发展性阅读障碍(1 项)合并为：“3.1.5 其他神经系统疾病”。此外，虽然阿尔茨海默症、失语症在本研究纳入的文献数量不多，但我们考虑到这两种疾病已有相关综述，证据比较充分，因此我们保留了原有的权重，并且合理引用了相关综述，本研究只对已有综述以外的文献进行阐述。(2)将原“3.2 提取练习在精神障碍人群中的作用”修改为“3.2 提取练习在精神分裂症中的作用”，因为考虑到早期精神病(EP)的唯一一项研究中涉及的被试仍然主要是精神分裂症，其余被试也是涉及精神分裂症谱系和伴有精神病性症状的双相Ⅰ型等疾病(Ragland et al., 2020)，因此我们将该文章合并到精神分裂症的板块中做了简单阐述。

相关具体修改如下：

3.1.5 其他神经系统疾病

听力障碍(DHH)的两项研究均支持提取练习的积极作用。de Diego-Lázaro(2025)的实验发现 DHH 儿童即时和间隔提取的词汇再认表现均优于被动学习(无提取)条件。Reimer 等(2024)指出共病(如瓦登伯格综合征)是影响 DHH 儿童提取练习效果的重要因素。尽管上述发现为 DHH 儿童的词汇干预提供了初步证据，但该领域的研究仍较欠缺，未来需在临床与教育实践中系统评估其应用效果。脑卒中(SS)领域目前仅少量研究(Evans et al., 2020)。初步证据提示，提取练习可能增强患者对姓名-面孔的记忆能力，且在短期与长期保持中均具优势。未来需在不同认知任务中验证其效果。发展性阅读障碍(DD)的研究，发现基于语音的提取练习可缓解 DD 学习者因打字产生的学习劣势，其问题主要在于编码与产出的时间延迟而非记忆本身(Wilschut et al., 2025)。但对该人群的研究仍有较大拓展空间。

3.2 提取练习在精神分裂症中的作用

.....

此外，一项研究关注早期精神病(EP)，该研究中同时包含精神分裂症、伴有精神病性症状的双相Ⅰ型和重性抑郁障碍等疾病(Ragland et al., 2020)。该研究发现，提取练习可显著提升 EP 患者对临床相关信息的长期记忆保持，且阴性症状较轻者获益更明显。但由于仅有单篇报道且样本异质性较高，提取练习对不同症状维度及诊断亚型的效果尚不明确，未来需在更大样本中重复验证。

....

意见 3：局限部分尚未充分回应更核心的方法学和证据层面限制。当前第 6 部分本研究的不足与局限主要强调了纳入文献数量较少、未做元分析等问题，但对检索式和数据库覆盖可能遗漏文献、不同病种与任务之间高度异质、临床功能外推有限等问题涉及不足。局限部分与前文讨论出的证据复杂性之间仍不够匹配。

回应：感谢您对局限部分的意见，本研究的确存在您所提到的局限，我们已根据您的意见进行了修改。在局限部分我们首先回应了检索式和数据库可能遗漏的文献。其次，在临床适用范围方面也存在一定的局限。我们将临床适用范围方面的局限归纳为两方面原因(1)纳入的研究及病种存在高度异质性、(2)实验室任务与真实临床情景存在差异，由此以期与讨论部分的证据复杂性相匹配。具体修改如下：

6 本研究的不足与局限

本研究主要存在以下不足与局限。第一，应更广泛的纳入参考文献。本研究纳入的文献只有 22 篇不够全面，其原因可能是以下两方面。(1)检索策略与时间限制。2020 年前的相关文献已有类似综述进行了研究(de Lima et al., 2020)，因此本研究只检索了 2020 年至 2025 年的相关文献。此外本研究仅检索了六个数据库，未纳入预印本或联系相关研究者获取可能存在的非显著性结果，且在检索式设置上未充分利用截词符，因此可能遗漏了部分相关研究。(2)根据纳入和排除标准，在纳入的文献中，学习方式变量需要有两个及以上的不同水平(如提取练习和重复学习)，或对提取练习的不同形式进行比较(如间隔提取与即时提取)，因此只单纯研究提取练习的文章则被排除(如：Smith et al., 2024)。第二，可进行定量的元分析。本研究是对当前研究领域的系统综述与未来研究展望，未来可对相关研究进行提取练习效应的元分析，进一步分析提取练习在临床人群中的效应量大小。第三，结论的临床适用范围有限，其原因可能是以下两方面。(1)纳入的研究及病种存在高度异质性。不同研究在疾病类型与阶段、学习材料、因变量指标上存在显著差异。(2)实验室任务与真实临床情景存在差异。多数任务为实验室条件下的词汇或句子学习，与真实临床场景中的功能性记忆(如服药安排、日常规划)存在一定差距，且缺乏长期随访数据。

意见 4：结果部分虽然增加了纳入研究“总体特征”，但目前该段概括仍偏表面。当前主要概括了疾病分布和干预内容，而未进一步总结更关键的研究设计性信息。

回应：感谢您对纳入研究“总体特征”部分的建议。我们之前在描述总体特征时，偏向于单纯概括疾病种类和干预内容，但未涉及到具体的实验设计。为此，我们添加了这一部分，从研究设计的类型(实验、准实验以及被试间、被试内等)、样本量范围、干预目标、因变量指标、随访周期等方面进行了总结，具体修改如下：

“...第二，在研究设计方面，22 项研究均为实验或准实验设计，其中 16 项采用混合实验设计，3 项为被试间设计，3 项为被试内设计。样本量范围为 11 至 92 人(不含健康对照组)。所有研究均设置了提取练习条件或其不同形式(间隔提取、产出式提取等)。干预目标多为促进陈述性记忆，少数研究探讨了情景记忆及复杂认知任务。因变量指标主要为即时或延迟回忆/再认的正确率，测试周期从即时测试到 3 个月不等。大多数研究证明提取练习(或其不同形式)相比对照学习条件更具优势，也有研究比较了不同提取形式的差异性效果。这一发现初步验证了该干预手段在临床人群中的应用潜力，下文将按疾病类别对研究结果进行详细阐

述。”

意见 5：在 4.2 提取练习的人群特异性与效应边界部分，作者提出了“疾病核心认知缺陷类型”“学习材料的复杂性和生态性”“多治疗手段支持”三个维度，这一框架较有启发性。但从前文结果来看，这三个维度目前都仍主要基于少数病种和少量研究进行归纳，整体上更接近一种解释性框架，而非已有研究充分支持的稳定结论。当前写法容易给人一种这些边界条件均已较为稳定成立的印象，而这一点与前文实际证据基础之间仍存在一定距离。

回应：感谢您对这一部分提出的建议。在上一轮修改中，我们扩展了对此部分的阐述，但却未充分意识到该框架仍然是需要后续研究深入验证，而不是十分确定性的结论。因此我们对这一部分的相关表述进行了修改，使其成为一种探索性、假设性的框架。同时，我们也相应的对摘要和结论部分进行了修改。具体修改如下：

摘要部分：

摘　要　.....基于已有研究，本综述初步提出提取练习的效果可能受以下三方面因素的调节：(1)疾病核心认知缺陷的类型；(2)学习材料的复杂性和生态性；(3)是否有其他治疗手段。未来可对上述调节因素进行深入验证，并同时关注共病人群、提取练习的间接效应及个体差异等问题。...

Abstract: .....However, in some diseases, the effectiveness of retrieval practice may be moderated by the core cognitive deficits associated with the condition, the complexity and ecological validity of the learning materials, as well as the presence of other therapeutic interventions. Future studies could further investigate the moderating factors and focus on interventions for comorbid conditions, indirect benefits, and the role of individual differences.....

结论部分：

.....在部分疾病中其效果可能受到疾病核心认知缺陷的类型、学习材料的复杂性和生态性，以及是否结合其他治疗手段的影响。....

正文部分：

4.2 提取练习的人群特异性与效应边界

综合现有证据，提取练习在临床人群中的作用可能是一种条件性的认知补偿策略，其潜在的特异性和效应边界初步体现在以下三方面。这三方面是基于现有研究的初步框架，旨在为未来研究提供可检验的假设。

4.2.1 疾病核心认知缺陷的类型

现有研究提示，提取练习的临床效应可能受到疾病核心认知缺陷类型的调节，这一推测主要基于 AD 与 SP 患者的对比分析，尚需更多病种证据支持。

Stamate 等(2020)指出 AD 患者的真正挑战在于编码困难。针对这一特征，提取练习虽能部分补偿记忆缺陷，但初步证据表明其补偿作用可能存在上限，或需借助线索等外部支持

才能维持长期效果(Small & Cochrane, 2020)。相比之下，SP 患者的认知缺损可能更多表现为干扰抑制功能下降(Weisbrod et al., 2000)。Tan 等(2024)和 Pan 等(2025)的研究发现，提取练习能通过减少前摄干扰、促进语义组织能力，进而改善记忆效果。这说明，对于 SP 患者，提取练习的有效性可能在于优化信息表征与抗干扰能力，从而有助于调动其尚存但效率较低的认知控制资源。

据此可提出一个待继续验证的假设：当缺损集中于编码层面时，提取练习可能需要外部支持才能发挥持久作用。而当缺损集中于提取与干扰控制层面时，提取练习本身或可作为一种相对有效的康复手段。这一假设的跨病种适用性(是否适用于其他以编码缺陷或提取缺陷为主的疾病)仍有待未来研究检验。

4.2.2 学习材料的复杂性和生态性

DLD 儿童与 ADHD 的研究证据初步表明，材料复杂性与生态性可能共同调节提取练习的干预效果。在复杂性方面，当 DLD 儿童学习动词时，提取练习能促进其学习效果，但相较于名词学习时的效应量更小，且提取练习未能改善其语法形态转换能力(Leonard et al., 2023)。这可能与学习动词和形态转换时需要的认知负荷有关。这一结果提示，当材料复杂度超出患者的认知负荷承受范围时，提取练习的额外认知成本可能超过其记忆巩固效果(Zheng et al., 2023)。然而，目前关于材料复杂性的系统操纵研究仍然有限，不同病种的最佳难度范围尚不明确。

在生态性方面，Stern 和 Halamish(2023)在 ADHD 患者中使用高生态性的文字材料未发现提取练习优势，而 Souto 等(2025)将提取练习嵌入绘本阅读则显示出良好效果。这些初步发现提示，提取练习在符合日常学习方式的情境中更易发挥促进作用，但生态性材料在何种条件下削弱或增强提取练习效应，仍需更多研究加以厘清。

4.2.3 多治疗手段的支持

ADHD 和 SP 的研究初步提示，提取练习的效果可能受到其他治疗手段的调节。首先，在药物治疗方面，Stern 和 Halamish(2023)要求 ADHD 被试在实验前停药，这可能是导致提取练习未表现出干预效果的原因之一。反之，Minear 等(2023)的研究证明，服药患者在最终测试时表现更佳。这可能是因为药物改善了 ADHD 患者在学习阶段的注意、动机和努力水平，从而影响患者的记忆编码能力，间接提高提取练习的干预效果(Addicott et al., 2019; Rubinson et al., 2019)。这一对比提示，药物治疗所提供的认知准备状态可能是提取练习生效的重要前提条件之一。

其次，物理治疗的介入则暗示，提取练习的干预效果在一定程度上依赖于特定神经环路功能的有效性。Pan 等(2025)采用 tDCS 对精神分裂症患者的 L-DLPFC 皮层进行刺激，显著提升了提取练习的干预效果。这一促进作用可能来源于两方面机制：一方面，tDCS 通过影响 γ-氨基丁酸的抑制性功能并促进长时程增强(Long-term Potentiation)，为提取练习提供了更有利的神经环境。另一方面，tDCS 刺激并激活的 L-DLPFC 脑区，可能与提取练习所涉及

的神经机制相吻合(Zhuang et al., 2022)。这一发现提示，外部物理手段优化特定脑区功能可能为提取练习提供有利的神经环境。但其可重复性、作用机制及在不同病种中的适用性均有待进一步检验。

综上，已有证据初步勾勒出提取练习作为一种条件性认知补偿策略的可能性，其有效性或需部分依赖于其他治疗手段所提供的认知与神经层面的基础支持，但这一结论属探索性假说，需在更大范围的疾病种类中进一步验证。

意见 6：结果部分虽已按疾病类别和病种分节，并增加了小结，但小结部分仍可进一步加强。尤其是 3.1.7 不同神经系统疾病个体的提取练习差异和 3.2.3 不同精神障碍个体的提取练习差异，目前仍偏重继续讨论，而未能非常清楚地回答：这一类研究总体支持什么、证据主要集中在哪些任务、还缺什么。

回应：感谢您对结果部分小结的建议。我们对此做了认真考量并修改。

首先，我们重写了原“3.1.7 不同神经系统疾病个体的提取练习差异”，并将其修改为“3.1.6 神经系统疾病人群提取练习的证据与不足”。我们重点回应了这一类研究目前总体的结果如何，是否支持本研究的论点、以往研究的提取练习设计的记忆任务涉及哪些方面，以及当前这一类研究的不足之处。

其次，结合您所提的意见 2，我们调整了对早期精神病的权重(因为只有 1 项研究)，且由于该研究的主要被试仍属于精神分裂症谱系，因此我们撤销了原标题“3.2.2 早期精神病”，将其相关内容合并“3.2 提取练习在精神分裂症中的作用”中阐述。基于此，我们也同时撤销了原小标题“3.2.1 精神分裂症”和“3.2.3 不同精神障碍个体的提取练习差异”。直接在“3.2 提取练习在精神分裂症中的作用”的最后一段进行小结，阐述了当前这一类研究的总体结论和证据所支持的任务，以及研究不足。具体修改如下：

3.1.6 神经系统疾病人群提取练习的证据与不足

现有证据总体证明提取练习能够改善部分神经系统疾病患者的记忆功能，尤其在词汇学习和陈述性记忆任务中表现出一定的促进作用。且当前证据高度集中于词汇学习任务，少数研究涉及面孔-姓名联想记忆、情景记忆和语义记忆。提取练习的操作方式以产出式提取和间隔提取为主。然而，仍存在以下研究不足。第一，疾病覆盖不均衡。发展性语言障碍、失语症、阿尔茨海默症证据较为充分，但注意缺陷多动障碍、脑卒中、发展性阅读障碍仅有少量研究。第二，认知机制的疾病特异性尚不明确。提取练习是通过补偿编码缺陷、强化提取路径，还是巩固记忆痕迹来提高记忆效果，在不同疾病中可能机制不同，但目前缺乏跨疾病的比较研究。第三，生态效度与迁移效果有限。多数研究采用虚构词或实验室任务，在复杂任务和长期效果(如语法规则应用、日常功能改善)中的证据薄弱。第四，个性化干预参数缺乏指导。最优提取间隔、任务形式、是否需联合药物或编码策略，尚无基于某一疾病类型的明确推荐指南。

3.2 提取练习在精神分裂症中的作用

以往的研究显示提取练习可有效引起精神分裂症患者的前向测试效应(The Forward Testing Effect)，改善其对新学空间信息的记忆(Li et al., 2024; Tan et al., 2024)。而针对左背外侧前额叶(Left Dorsolateral Prefrontal Cortex，L-DLPFC)的连续周期性经颅直流电刺激(Transcranial Direct Current Stimulation，tDCS)可促进语义信息的精细化组织，进而增强提取练习的干预效果(Pan et al., 2025)。这些发现共同揭示出一种层次化的干预模式：行为层面的提取练习可以优化记忆提取效果，而神经调控技术则可能通过影响神经网络与认知组织策略，为记忆提取和巩固提供支持性条件。但其协同作用的内在神经机制仍待深入探索，如tDCS 如何更为精准地调控与提取练习相关的认知与神经环路，这可为发展更具个体化的联合干预方案提供重要方向。

此外，一项研究关注早期精神病(EP)，该研究中同时包含精神分裂症、伴有精神病性症状的双相Ⅰ型和重性抑郁障碍等疾病(Ragland et al., 2020)。该研究发现，提取练习可显著提升 EP 患者对临床相关信息的长期记忆保持，且阴性症状较轻者获益更明显。但由于仅有单篇报道且样本异质性较高，提取练习对不同症状维度及诊断亚型的效果尚不明确，未来需在更大样本中重复验证。

现有证据总体支持提取练习能够改善精神分裂症患者的记忆功能，且研究主要采用空间记忆任务、语义记忆任务，并结合经颅直流电刺激探索认知增强的神经机制。此外，有一项研究关注提取练习对早期精神病的陈述性记忆的作用。但当前研究仍有以下不足。首先，病种覆盖不足，除精神分裂症外，其他存在记忆缺陷的精神障碍(双相情感障碍、重性抑郁障碍)的提取练习效应尚无系统研究。其次，任务类型较为单一。现有研究多关注空间记忆任务和词汇记忆任务，对复杂记忆任务或高生态性记忆任务的考察相对缺乏。最后，干预交互作用研究不足。在精神分裂症等精神障碍的临床治疗中常采用多手段治疗，但提取练习与药物治疗、物理治疗等常规干预的协同效应需加以验证。

意见 7：文中若干事实、引文和术语表述仍存在问题，建议全面核查。例如，正文中出现 Wen 等(2025)，但参考文献列表中似乎未见其对应条目；“阿尔兹海默症”与“阿尔茨海默症”两种译法前后并存；参考文献目前仍有期刊、卷号未斜体、页码间符号格式不一致等。

回应：十分感谢您的细致审阅。我们对全文进行了全面核查，主要修改如下：

1. “Wen 等(2025)”应该为“Pan 等(2025)”，这是由于我们在上一轮修改中对作者姓名的错误理解而发生的错误，原作者全名应为 Pan Wen，在上一轮修改中，我们在部分引用之处错误的将名理解为了姓。已将其修改。

2. 参考“进展”最新有关文章中的译法(薛莹琦 等, 2026)。已将“阿尔兹海默症”统一修改为“阿尔茨海默症”。

参考文献：薛莹琦, 张瑶, 赵海潮, 何清华, 刘佳丽. (2026). 路径整合能力老化及其神经机制. *心理科*

学进展, 34(5), 890–905.

3. 全面核查了参考文献引用是否合理，删除了如下参考文献(原因见括号内)，还有部分参考文献由于相关阐述的优化，也不再继续引用。以下三篇是相关表述仍在使用，但该文献不合适或非必要引用：

Dubois, P., St-Pierre, M.-C., Desmarais, C., & Guay, F. (2020). Young adults with developmental language disorder: A systematic review of education, employment, and independent living outcomes. *Journal of Speech, Language, and Hearing Research, 63*(11), 11–15.**(因其是综述范围为成年 DLD，引用在本文不恰当，且相关正文表述已有参考文献，此篇略显多余，则删除.)**

Wang, L., Lou, Y., Jiang, Q., Wang, H., Huang, S., Xie, Y., Wang, F., & Cao, S. (2025). Association between multimorbidity trajectories and subsequent cognitive decline: evidence from two nationwide longitudinal cohorts. *Journal of affective disorders, 391*, 119958. Advance online publication.**(相关正文表述已有参考文献，此篇略显多余，则删除)**

Quiñones, A. R., Allore, H. G., Botoseneanu, A., Newsom, J. T., Nagel, C. L., & Dorr, D. A. (2020). Tracking Multimorbidity Changes in Diverse Racial/Ethnic Populations Over Time: Issues and Considerations. *The journals of gerontology. Series A, Biological sciences and medical sciences, 75*(2), 297–300.**(相关正文表述已有参考文献，此篇略显多余，则删除)**

4. 核查了参考文献格式是否正确，做了如下修改(列举的是修改前，其后的括号内说明了如何修改)：

*Tan, Y., Li, T., Yang, C., Huo, X., Liu, M., Zhang, J., & Ma, X. (2024). Route learning in patients with schizophrenia: The role of the forward testing effect. Journal of Environmental Psychology, 96, 102305.**(“Journal of Environmental Psychology, 96”已修改为斜体)**

Ariel, R., & Karpicke, J. D. (2018). Improving self-regulated learning with a retrieval practice intervention. *Journal of Experimental Psychology: Applied, 24*(1), 43.**(已完善了页码“43–56”)**

Carpenter, S. K. (2023). Encouraging students to use retrieval practice: a review of emerging research from five types of interventions. *Educational Psychology Review, 35*(4), 17.**(“17”修改为了“96”)**

*Leonard, L. B., Deevy, P., Horvath, S., Christ, S. L., Karpicke, J., & Kueser, J. B. (2023). Can retrieval practice facilitate verb learning in children with developmental language disorder and their peers with typical language development? *Journal of Speech, Language, and Hearing Research : JSLHR, 66*(4), 21–25.**(“*Journal of Speech, Language, and Hearing Research : JSLHR, 66*(4), 21–25.”修改为了“*Journal of Speech, Language, and Hearing Research, 66*(4), 1309–1333.”)**

*Leonard, L. B., Deevy, P., Karpicke, J. D., Christ, S. L., & Kueser, J. B. (2020). After initial retrieval practice, more retrieval produces better retention than more study in the word learning of children with developmental

language disorder. *Journal of Speech, Language, and Hearing Research : JSLHR, 63*(8), 1–14.**(“*Journal of Speech, Language, and Hearing Research : JSLHR, 63*(8), 1–14.”修改为了“*Journal of Speech, Language, and Hearing Research, 63*(8), 2763–2776.”)**

OphuisCox, F. H. A., Catrysse, L., & Camp, G. (2023). The effect of retrieval practice on fluently retrieving multiplication facts in an authentic elementary school setting. *Applied Cognitive Psychology, 37*(6), 1463–1469.**(“OphuisCox”修改为了“Ophuis-Cox”，正文相关处也进行了修改)**

Olsson, C., Arvidsson, P., & Blom Johansson, M. (2024). Self-efficacy and resilience in severe aphasia - an exploratory cross-sectional study of two psychosocial factors and their relation to functional communication. *Disability and rehabilitation, 46*(21), 4988–5001.**(“*rehabilitation*”修改为了“*Rehabilitation*”)**

*Pan, W., Li, T., Ma, X., & Huo, X. (2025). Effects of transcranial direct current stimulation combined with retrieval practice on semantic memory in patients with schizophrenia. *BMC Psychiatry, 25*(1), 1–11.**(“*25*(1), 1–11.”修改为了“*25*(1), 214.”)**

Roediger, H. L. & Butler, A. C. (2010). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20–27.**(“2010”修改为了“2011”，2010 年为 Available online 时间，2011 年为 Version of Record 时间，正文处同时修改)**

Skodzik, T., Holling, H., & Pedersen, A. (2017). Long-term memory performance in adult ADHD. *Journal of Attention Disorders, 21*(4), 267–283.**(“Long-term memory performance in adult ADHD”修改为“Long-term memory performance in adult ADHD: A meta-analysis”)**

*Stern, P., & Halamish, V. (2023).Free-recall retrieval practice tasks for students with ADHD: whole-text versus section recall. *Frontiers in Psychology. 14*: 1301726**(“*Frontiers in Psychology. 14*: 1301726”修改为“*Frontiers in Psychology*, *14*, 1301726.”)**

Veras, A. B., Ribeiro, K. M., Peixoto, C., Horato, N., Kahn, J. P., & Nardi, A. E. (2025). Schizophrenia spectrum disorders with comorbid panic: A systematic review. *Schizophrenia Research, 277*, 33-41. **(“33-41”修改为了“33–41”)**

5. 还包括但不限于如下修改：

(1)核对所有表格的内容并调整部分内容：

①根据文章正文中实际出现的顺序，调整了表号。将“表 3 中英文检索词对照表”调整为“表 1 中英文检索词对照表”，“表 4 纳入文献的质量评估”调整为“表 2 纳入文献的偏倚风险评估”，“表 1 纳入文献的基本情况”调整为“表 3 纳入文献的基本情况”，“表 2 中英文对照表”调整为“表 4 中英文术语对照表”

②在调整后的表 4 中英文术语对照表中增加了一列，以区分不同术语的类别(疾病种类/对照组、学习方式)

③将中英文术语对照表中“Health Controls”修改为“Healthy Controls”

④将“理解性/产出性提取练习”全文统一为“理解式/产出式提取练习”

⑤将学习材料具体化，例如部分“词对”修改为“斯瓦希里语-英语词对”，部分“新词”修改为“新异名词及其图片和意义”

(2)核对事实并修改了措辞：

①将“但提取练习组的表现仍优于无提取练习组”修改为“但重复提取组的表现仍优于单次提取组”**(正文第 8 页，3.1.3 阿尔茨海默症，第 4 行)**

②将“Knouse 等(2020)研究表明 ADHD 患者在提取练习条件下的表现比无练习条件更优。”修改为“Knouse 等(2020)研究表明 ADHD 患者在恒定标准的提取练习(提取成功次数有最低标准)条件下的表现比自我调节的提取练习(自我决定提取的次数和时间)更优。”**(正文第 8 页。3.1.2 注意缺陷/多动障碍，第 3 行)**

③将“de Diego-Lázaro(2025)的实验发现 DHH 儿童即时和间隔提取的记忆效果显著优于无提取条件。”修改为“de Diego-Lázaro(2025)的实验发现 DHH 儿童即时和间隔提取的词汇再认表现均优于被动学习(无提取)条件。”**(正文第 9 页，3.1.5 其他神经系统疾病，第 2 行)**

第四轮

审稿人 3 意见：

非常感谢作者对前一轮审稿意见的细致回应。本轮修改在方法学透明度(补充筛选与一致性指标)、论证强度的克制(将 4.2 节明确为探索性框架)、小结的清晰度(3.1.6 和 3.2 节末段)、以及引文与术语规范方面均有实质性改进，作者付出的努力值得肯定。仅余以下几处小问题，供作者参考。

回应：非常感谢您能再次审阅我们的稿件！为提升本文的严谨性和可读性再次提出了宝贵的意见。我们已根据您的意见进行完善修改，再次感谢您的指导(为和文中已有的蓝色、红色、绿色和紫色区分，我们将本次修改后的文字标记为了青色)。

意见 1：3.3 提取练习在 HIV 感染者和唐氏综合征儿童中的作用部分，仅用一段叙述两种疾病，且未像 3.1.6 神经系统疾病人群提取练习的证据与不足和 3.2 提取练习在精神分裂症中的作用末段那样附小结。考虑到结构对称性以及读者阅读体验，建议补一两句横向小结(如证据集中于词汇/健康信息记忆、样本量小、缺乏长期迁移证据等)，与前两类疾病的小结风格保持一致。

回应：感谢您的意见。确实，在之前的 3.3 提取练习在 HIV 感染者和唐氏综合征儿童中的作用部分，原有的阐述虽指出了当前研究的不足，但未考虑到全文结构的统一性和读者阅读体验，因此我们做了如下修改：删除“然而，以上两种疾病类型仍需更大样本进一步研究，并且提取练习在其中的长期迁移效果仍有待探索。”，并新增以下内容：“总体而言，提取练习在 HIV 感染者和唐氏综合征儿童中初步显示出积极效果，但相关研究和样本量均较少，证据集中于健康信息记忆或词汇记忆任务，且缺乏长期迁移效果及日常功能改善的证据。”

意见 2： 2.1 文献检索报告了“初步获得 407 篇”“手工检索补充 15 篇”“删重后 313 篇”，但从 313 篇到“题目摘要初筛获得 24 篇”这一关键筛选环节，正文并未交代(读者需要自行参考图 1 来补全)。按 PRISMA 报告规范，建议在 2.3 文献的筛选及信息提取补充一段相关说明，使流程在文字层面也完整。

回应：感谢您的意见！在之前的稿件中，文献筛选部分的文字层面报告的确不够详细，部分内容读者需要看图才能理解，这样也降低了本文的可读性。因此我们参照您的意见，在 2.3 文献的筛选及信息提取部分新增了一段，以说明文献筛选过程。具体修改如下：

2.3 文献的筛选及信息提取

本研究的文献筛选过程遵循 PRISMA 标准(Moher et al., 2009)。由 2 名作者独立开展文献筛选和信息提取工作。完成后，这两人交叉核对彼此的筛选结果和提取的信息，提取的信息包括：标题、作者、发表时间、被试、学习材料、学习方式、最终测试时间点、结论。2 名作者在题目和摘要初筛、全文筛选及信息提取时的归类一致性分别为 0.97、0.93、0.92(徐建平, 张厚粲, 2005)。分歧之处，则与第 3 名作者讨论并由其作出最终决定。

具体筛选流程及数量变化如下(图 1)：删除 109 篇重复文献后，剩余 313 篇文献进入题目和摘要初筛阶段。在此阶段，依据纳入与排除标准剔除不符合主题的文献 277 篇、综述文献 6 篇、学位论文 6 篇，剩余 24 篇进入全文筛选。全文筛选阶段进一步排除重复样本的文献 1 篇、不符合提取练习范式的文献 1 篇。 最终，经全文筛选后最终共纳入 22 篇文献。

意见 3：附录表 3 目前用“是/否”概括“提取练习是否有效/学习条件之间是否有差异”一栏，信息略显粗略。读者无法直接从表中区分：该研究究竟是支持“提取练习相较于重复学习更有优势”，还是“某种提取练习形式更优于另一种”。建议将该栏进一步具体化，简要

说明主要的比较方向与结果。

回应：感谢您的意见！为了方便读者阅读，我们做了如下工作：

首先，将附表 3 中对被试和学习方式的英文缩写改为了中文。**其次，**以符号来说明不同干预方式之间的差异，“＞”表示前者干预效果优于后者；“＝”表示两者干预效果无显著差异，例如：“提取练习＞重复学习”代表提取练习干预效果优于重复学习。**最后，**删除了原附录表 4 中英文术语对照表，因为对学习方式的英文表述已在附表 3 中被替换为了中文，对于疾病类型的中英文表述已经在正文中有所阐述，如再增加表格，略有多余。**具体修改如下(正文第 31 页)。**

附表 3 纳入文献的基本情况

研究	被试	材料	最终测试时间点	学习方式及主要比较结果
Boundy et al., 2025	唐氏综合征儿童(n=11) 健康对照组(n=11)	新异名词及其图片和意义	5 分钟、1 周	重复即时提取＞重复学习
de Diego-Lázaro, 2025	听力障碍儿童(n=16) 健康对照组(n=26)	无意义单词及对应图片	即时、1 天、1 周	即时提取＞无提取 间隔提取＞无提取 即时提取=间隔提取
Evans et al., 2020	脑卒中患者(n=20) 健康对照组(n=20)	名字-面孔对	30 分钟、1 周	提取练习＞间隔重复学习 提取练习＞集中学习
Knouse et al., 2020	注意缺陷/多动障碍患者(n=58) 健康对照组(n=121)	心理学概念	1 周	恒定标准提取＞自我调节提取
Leonard et al., 2020	发展性语言障碍儿童(n=13) 健康对照组(n=13)	新异名词及其图片和意义	5 分钟、1 周	多提取少重学＞多重学少提取
Leonard et al., 2023	发展性语言障碍儿童(n=11) 健康对照组(n=12)	新异不及物动词及其视频	5 分钟、1 周	重复间隔提取＞重复学习
Leonard, Deevy et al., 2024	发展性语言障碍儿童(n=14) 健康对照组(n=13)	新异及物动词及多个视频	5 分钟、1 周	重复间隔提取＞重复学习
Leonard, Christ et al., 2024	发展性语言障碍儿童(n=14) 健康对照组(n=14)	新异名词及其图片和意义	5 分钟、1 周	扩展式间隔提取=等距式间隔提取
Li et al., 2024	精神分裂症患者(n=92)	3D 空间路线	分心任务后即时	中期测试＞重复学习
Middleton et al., 2022	失语症患者(n=12)	语义最小词对及图片	1 天 、1 周	产出式提取练习＞无错误学习 理解式提取练习＝重复学习
Minear et al., 2023	注意缺陷/多动障碍患者(n=36) 健康对照组(n=36)	斯瓦希里语-英语词对	2 天	提取练习＞重复学习
Pan et al., 2025	精神分裂症患者(n=52)	双字词	即时、1 天	提取练习＞重复学习
Ragland et al., 2020	早期精神病患者(n=20) 健康对照组(n=20)	关于精神病的知识	7 至 10 天	提取练习＞重复学习
Reimer et al., 2024	听力障碍儿童(n=16)	陌生名词及其图片	2 天	提取练习＞重复学习
Small & Cochrane, 2020	阿尔茨海默症患者(n=15)	近期事件细节(含视频与线索)	1 周、1 个月、3 个月	间隔提取＞不间隔提取
Souto et al., 2025	发展性语言障碍儿童(n=13) 健康对照组(n=14)	新异名词及其意义	5 分钟、1 周	重复间隔提取＞重复学习
Stamate et al., 2020	阿尔茨海默症患者(n=33) 健康对照组(n=42)	寓言故事	1 个月	多次测试＞单次测试
Stern & Halamish, 2023	注意缺陷/多动障碍患者(n=36) 健康对照组(n=36)	散文	2 天	全文回忆＞分段回忆 全文回忆=重复学习 分段回忆=重复学习
Tan et al., 2024	精神分裂症患者 (实验 1：n=64；实验 2：n=60) 健康对照组 (实验 1：n=64；实验 2：n=60)	实验 1：2D 平面路线 实验 2：3D 空间路线	分心任务后立即	中期测试＞重复学习
Wang et al., 2020	失语症患者(n=11) 年轻健康对照组(n=18) 老年健康对照组(n=18)	词对、图片对	10 分钟、2 天	间隔提取＞集中提取
Wilschut et al., 2025	阅读障碍患者(n=57) 健康对照组(n=61)	斯瓦希里语-英语词对	立即	言语提取＞打字提取
Woods et al., 2021	HIV 感染者(n=41)	虚构疾病的知识	20 分钟	带自我生成强化的间隔提取练习＞集中学习

注：“＞”表示前者干预效果优于后者；“＝”表示两者干预效果无显著差异

此外，参照《心理科学进展》已发表的文章，在正文中将“附录表 3”改为了“附表 3”，在附录中将“表 3”改为“附表 3”，附表 1 和附表 2 同理。

本次修改再次进行了全文通读，以润色修改部分表述，提高文章的可读性。包括但不限于：

(1)将 7 研究结论与贡献部分中的“...此外本研究为提取练习的临床转化与干预方案实施提供了思路和原则...”具体化为：“...此外，本研究从明确干预目标、加强指导与宣教、采用渐进式实施策略三方面，提出了提取练习临床转化的思路与原则...”**(正文第 19 页)**

(2)“基于语音的提取练习”修改为“言语提取练习”**(3.1.5 其他神经系统疾病的第 7 行，正文第 9 页)**

(3)“...该研究通过将抽象的干预目标具体化为产出式提取练习和理解式提取练习，不仅精准匹配了临床需求，也最终验证了提取练习在改善特定认知缺陷方面的干预效果。...”修改为“...该研究以此目标为导向，设计了产出式提取练习和理解式提取练习两种干预条件。通过对比提取练习与非提取的对照条件，该研究验证了提取练习在改善词汇-语义理解缺陷方面的优势。...”**(5.4 干预方案的临床转化与实施第 2 段第 3 行，正文第 17 页)**

其余细小修改可见正文青色文字。

第五轮

编委 1 意见：同意发表。

编委 2 意见：同意审稿人的意见，建议按期刊格式修改后，请主编终审。

主编意见：同意发表。