

《心理科学进展》审稿意见与作者回应

题目：视听知觉训练对老年人认知能力的促进及其机制

作者：杨伟平，李睿智，李胜楠，林金飞，任艳娜

第一轮

审稿人 1 意见：

作者针对视听训练对于老年人的认知能力的促进做了综述梳理，论文具有较好的理论总结贡献。但对于普通的读者而言，需要具体解释作者所采用的指标的含义。通读全文，更多的是涉及跨通道时间感知的研究，未见全面的、按特定标准分类的其它“视听”研究范式，这对于一篇良好的综述是不充分的。另外，关于一些术语的科学描述，我附上原文的批注，请作者修订；标识黄底的语句，需要重新斟酌合适的中文表达以及逻辑（推论）合理性。

意见 1：文中“研究表明跨视听通道信息相互作用存在促进和补偿效应，以缓解老年人单一感觉通道知觉能力的下降”是特异性？其他群体呢？

回应：非常感谢审稿专家的建议。跨视听通道信息相互作用存在促进和补偿效应不仅存在老年人群中，在其他群体中也存在，如 MCI 和 AD 患者群体(Wu, Yang, Yu, et al., 2012)。现已对文中表述进行修正：

在老年人群中，跨视听通道信息相互作用存在促进和补偿效应，以缓解单一感觉通道知觉能力的下降。（摘要部分，绿色字体）

Wu, J., Yang, J., Yu, Y., Li, Q., Nakamura, N., Shen, Y., Ohta, Y., Yu, S., & Abe, K. (2012).

Delayed audiovisual integration of patients with mild cognitive impairment and Alzheimer's disease compared with normal aged controls. *Journal of Alzheimer's Disease*, 32(2), 317-328.

意见 2：Mishra 等（2015）使用的“视觉辨别任务”是什么任务？

回应：非常感谢您的建议。视觉辨别任务是“对光栅刺激动态扩张或收缩的判断任务”。已在原文添加（2.1 第二段，绿色字体）。

意见 3：“未来研究可设置更丰富的视觉感知材料，更利于提高个体的训练兴趣和更好地达到训练效果”，为什么采用丰富的材料就能提高训练效果？

回应：非常感谢审稿专家的建议。由于有研究表明单一简单的刺激容易导致老年人训练过程

中较高的脱落率，使用更丰富的视觉感知材料可提升老年人的训练兴趣，增强注意，提升训练参与度，从而提高训练效果(Ren et al., 2020)。因此，未来研究可考虑设置丰富的材料。

Ren, Y., Guo, A., Xu, Z., Wang, T., Wu, R., & Yang, W. (2020). Age-related functional brain connectivity during audio-visual hand-held tool recognition. *Brain and behavior*, 10(9), e01759.

意见 4：“脑电研究结果显示，经过听觉认知训练后顶叶和额叶的 P3b 振幅下降且潜伏期变短” 这个成分有什么特定的认知含义？

回应：非常感谢您的问题。P3b 振幅下降表明注意力资源分配效率的提高；P3b 潜伏期变短表明处理速度的提高(O'Brien et al., 2017)。

O'Brien, J. L., Lister, J. J., Fausto, B. A., Clifton, G. K., & Edwards, J. D. (2017). Cognitive training enhances auditory attention efficiency in older adults. *Frontiers in Aging Neuroscience*, 9, 322.

意见 5：“听觉训练可以改善老年人衰退的听觉能力，提升噪音环境中感知语音的能力，对听力损失老年人感知信息起到促进作用。”建议按这个总结句，对以上的研究例子描述归类。

回应：非常感谢您的建议。已按照此总结句，对上述例子进行归类（2.2 第一段，绿色字体）。

意见 6：“并且可能通过增强与多感觉信息整合相关的联合皮质改善老年人对视听刺激的敏感程度” 增强...联合皮质，不管从语法还是意义上看，都说不通。

回应：非常抱歉我们的表达有误，现已进行修改。

并且可能通过靶向训练增强多感觉信息整合相关联合皮质的激活程度（3 第一段，绿色字体）

意见 7：“跨视听通道知觉训练对认知老化的影响”从内容上看，这里的知觉训练应该都是比较广泛的内容，但后面聚焦在时间感知方面，需要扩充或者做必要的约束。

回应：非常感谢审稿专家的建议。已将“跨视听通道知觉训练”改成“跨视听通道时间知觉训练”，对内容进行约束。（3 标题，绿色字体）

意见 8：“老年人对同时呈现的视听觉信息反应显著快于单独视觉或听觉信息，即表现出冗余效应” Redundancy effect; Coactivation; 这两种机制都可能起作用。

回应：我们非常同意您的观点，冗余效应（Redundancy effect）和共同激活（Coactivation）

两种机制都可能起作用。由于本文强调视听信息相对于单模态信息具有促进优势，现已将“冗余效应”更改为“视听促进效应”。（1 引言第二段，绿色字体）

意见 9：“即感知能力得到显著提升(Near transfer, 近迁移)”相对于“远迁移”？“近迁移”这个词的含义是什么？

回应：非常感谢审稿专家的建议。近迁移是指视听感知训练迁移到其他未经训练的视听感知任务上，即表现为视听感知能力的提升；远迁移是指视听感知训练影响了与视听感知相关联的其他认知功能，例如：提升了工作记忆、流体智力等。

意见 10：“但未发现 SJ 训练对视听整合能力的影响(Powers Iii et al., 2016) ”用什么作为效标？

回应：非常感谢审稿专家的建议。此处使用声音诱发闪光错觉(Sound-induced flash illusion, SIFI)来评估视听整合能力。已在文中进行了补充（3.1 第二段，绿色字体）。

意见 11：“降低裂变和融合错觉”普通读者可能不知道这两个术语的含义。

回应：感谢审稿专家的建议。裂变错觉指单个视觉闪光与两个听觉刺激同时出现时，单个视觉闪光被误判为两个的错觉；融合错觉指单个听觉刺激与两个视觉闪光同时出现时，两个视觉闪光被误判为一个的错觉。现已在文中进行了补充（3.1 第二段，绿色字体）。

意见 12：“MCI、AD、PD”这些都需要中文全称。

回应：感谢审稿专家的提醒，我们已添加了中文全称。轻度认知障碍(Mild Cognitive Impairment, MCI)、阿尔茨海默病(Alzheimer's Disease, AD)、帕金森病(Parkinson's Disease, PD)（3.1 第三段，绿色字体）。

意见 13：“未来研究可从认知功能损害最早、最敏感的视觉和听觉层面”为什么说最早、最敏感？怎么界定？

回应：感谢审稿专家的建议。由于不同心理机能老化的早晚和速率不同，最早衰退的是听觉，许多人不到 60 岁就有非常明显的听觉衰退；其次是视觉，55 岁后开始出现急剧衰退，且在成年晚期个体的各种认知活动中，感知觉的衰退最明显(林崇德, 2018)。因此，说视觉和听觉是认知功能损害最早、最敏感。

林崇德. (2018). *发展心理学（第三版）*. 北京: 人民教育出版社, 491-492.

意见 14: “上述视听知觉训练均采用单调的视觉闪光(Flash)和听觉纯音(Pure Tone), 由于刺激材料比较单一无趣”是自己检索的问题, 还是真的无类似的研究?

回应: 非常感谢审稿专家的建议。经过再次详细检索, 尚未检索到类似研究, 已对表述进行了修改。(小结与展望, 第三段, 绿色字体)

意见 15: 原文标识黄底的语句需要重新斟酌合适的中文表达以及逻辑(推论)合理性。

(1) “窄化”、“靶向增强”。

回应: 现已将“窄化”改为“缩小”, “靶向增强”改为“增强效应”。(摘要处, 绿色字体)

(2) “强化”

回应: 现已将“强化”改为“提高”。(2.1 第二段, 绿色字体)

(3) “听觉训练后向未训练的其他认知功能迁移是有限的”

回应: 非常感谢您的建议, 已修改为听觉训练对于未经训练的其他认知功能的迁移效果有限(2.2 第二段绿色字体)。

(4) “构建时间绑定窗口(Temporal binding window, TBW, 即感知到非同步视觉和听觉的时间间隔)模型”

回应: 非常感谢审稿专家的建议, 已对表达进行修改。

以视听时间辨别任务得到的数据构建时间绑定窗口模型 (3.1 第二段, 绿色字体)

.....

审稿人 2 意见:

该文综述了老年人视听知觉训练的相关研究, 着力探讨了跨通道视听训练的干预效应、大脑神经加工的变化, 提出了对未来老年人视听知觉训练的发展方向的思考, 选题在认知老化特别是老年人认知干预领域是比较有新意的, 文献丰富, 能为同行提供一定的参考。但综述整体目标还不够清晰, 行文术语较多且缺乏解释, 文献引用更偏罗列而非梳理, 在综述写作上还需要进行大的修改。具体意见如下:

意见 1: 综述的目的在于目前版本中呈现的还不够清晰。首先, 作者关注老年人视听知觉训练的作用, 是关注其对老年人视听整合能力的作用, 还是关注其对非感知觉的认知能力的作

用？文章前半部分有的研究可能仅探讨了视听训练对老年人视听知觉本身的作用，有的可能考察了对加工速度、工作记忆等认知能力的迁移，而文章后半部分，即小结和展望中，似乎又关心的是如何提升老年人的视听整合能力。第二，如文章开篇所说，“通过视听知觉训练对老年人的视听整合能力及补偿机制进行干预研究，揭示大脑认知加工原理、增强神经退行性疾病的预防能力、探索科学精准的知觉学习的行为学训练方法，为改善老年人的认知能力提供科学解决方案。”这里提到的补偿机制、精准训练在文中都没有比较明确的阐述。何谓补偿，补偿机制如何被干预；何谓精准，之前的训练研究都没有体现精准？第三，从行文逻辑看，作者意在强调跨通道视听训练是比单视觉或单听觉训练更有效的训练，但从引用文献来看，并未体现出这种趋势。比如单通道训练也都表现出了显著的干预效应，跨通道训练也同样存在迁移效应小、长期效应弱等问题。那么跨通道视听训练的优势到底是什么？如何回应老年人视听整合能力甚至优于年轻人这个现象？

回应：非常感谢审稿专家的建议。

首先，本文主要关注的是视听训练对老年人视听整合能力的作用，但部分视听训练研究在考察训练效应的同时，也发现了对加工速度和工作记忆等未训练能力的迁移。虽然本文也对迁移效应进行了梳理，但关注核心仍是视听训练对于视听知觉本身的作用。

第二，补偿是指与单模态刺激相比，老年人对跨模态刺激的反应会额外激活特定的大脑区域或降低大脑半球的特化程度来补偿单一感觉加工的功能减退，因此老年人单一感觉的衰退可通过增强视听觉整合能力进行补偿。未来的研究可以根据补偿策略的脑机制研究，对补偿相关的大脑区域活动进行靶向训练，增强补偿机制的功能，以进一步延缓老化。

精准是指未来的研究可以使用提升视听整合能力的任务（如：视听辨别任务）直接训练老年人的视听整合，而前人的研究是通过改善视听时间敏感性，间接增强老年人视听整合的能力。

第三，已有研究对单通道训练和双通道视听训练进行直接对比发现，视听训练相较于单通道视觉训练有更高的效率(Seitz et al., 2006)，且与单通道听觉训练组相比，跨视听通道训练组被试对噪声中听觉语音识别能力有更大改善(Pilling M, 2011)，甚至有研究发现，仅视听训练能促进被试对噪声中听觉语音识别，而单通道听觉训练却无法促进(Lidestam et al., 2014)。因此，通过直接的对照研究发现，视听训练最鲜明的优势在于更高的效率和更好的训练效果。老年人对跨模态刺激的反应会额外激活特定的大脑区域或降低大脑半球的特化程度来补偿单一感觉加工的功能减退，因此，补偿机制使老年人的视听整合能力优于年轻人。

Lidestam, B., Moradi, S., Pettersson, R., & Ricklefs, T. (2014). Audiovisual training is better than auditory-only

training for auditory-only speech-in-noise identification. *The Journal of the Acoustical Society of America*, 136(2), EL142–147.

Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology*, 16(14), 1422–1427.

Pilling, M., & Thomas, S.M. (2011). Audiovisual Cues and Perceptual Learning of Spectrally Distorted Speech. *Language and Speech*, 54, 487–497.

意见 2: 文献综述整体偏罗列，而非梳理。以“2.1”的内容为例，2.1 的第一段和第二段都在讲视觉训练的效果，两段内容的区别是什么？作者是否想在第一段阐述视觉训练对视知觉本身的效果，第二段阐述视觉训练对其他认知能力的效果？如果是，建议用段首句概括本段落主题。类似的，3.1 的第一段和第二段也在内容主题上区分不明显，没有反映出对文献的梳理。另外，在每一部分的文献综述之后，作者用蓝色字标出的分析性文字，大多和前文列出的文献缺少对应关系，即读者读完作者引用的那些文献，并不能从作者的行文中比较顺利地得出蓝色字表达的那些分析和结论。比如 2.1 的蓝色字提到的，“神经机制研究不完善”是指前述研究没有关注认知神经加工过程吗？可是前述文献关心了白质分数各向异性、N1 成分等，作者认为的不完善是指什么？再如，为什么要设置更为丰富的视觉感知材料？前述研究因为训练材料单一所以弱化了干预效果吗？靶向性训练指什么？前述文献实施的训练不是有针对性(针对视觉能力)的训练吗？又为什么提出要用大数据平台对 MCI 老年人实施训练？大数据平台指什么，又是否具有可行性或者合理性？仅以此为例来说明全文在文献梳理与分析方面的问题，其余各处蓝色文字的问题类似，不一在此列举。

回应: 感谢审稿专家的建议。

①正如您所说，我们想表达的是：在 2.1 的第一段阐述视觉训练对视知觉本身的效果，第二段阐述视觉训练对其他认知能力的效果。现已在 2.1 第一段、第二段加入了概括性语句用以区分段落间的差异。

②3.1 的第一段和第二段的区别在于，第一段梳理了使用视听同时性判断任务 (Simultaneity judgment task) 进行视听训练的研究，第二段梳理了使用视听顺序判断任务 (Temporal order judgment task) 进行视听训练的研究。由于两段均为视听时间知觉训练，为了读者更好地理解，现已整合为一段，并添加了介绍性段落和段尾小结。(3.1 绿色字体)

③蓝色字体中与文献梳理对应方面的问题，均已在原文中进行了修改(绿色字体)。

④虽然前述研究也关注了认知神经加工过程，这里的神经机制研究不完善主要指的是视觉训练迁移效应的神经机制仍待探究。

⑤关于为什么未来研究要设置更为丰富的视觉感知材料,因为研究表明单一简单的刺激较容易导致老年人训练过程中较高的脱落率,使用更丰富的视觉感知材料可以提升老年人的训练兴趣,增强注意,提升训练参与度,从而提高训练效果(Ren et al., 2020)。因此,未来研究可考虑设置丰富的材料。

⑥前述文献实施的训练有针对视觉能力的训练,但缺乏针对视觉信息具体属性的训练,这里的靶向性训练主要指根据老年人对视觉信息具体属性感知的脑机制,设计神经靶向性视觉训练,对老年人大脑特定区域活动进行干预,改善老年人的视觉能力。

⑦非常抱歉对大数据平台的表述有误,原本想表达的是扩大样本种类和数量,现已进行修正(2.1 第三段,绿色字体)。

意见 3: 建议作者增加视听训练为何能提升认知能力的理论基础介绍。有关感知觉在认知老化中的作用,早在 1994 年,柏林老化研究就已经有了探讨,相关的理论假说也不少。增加理论基础介绍,可以让读者更明确地理解视听训练对于提升老年人认知能力的意义。

回应: 非常感谢审稿专家的建议。已经补充了 1994 年柏林老化研究相关理论假说。

柏林老化研究(BASE)探讨了老年人感知觉的衰退与认知老化的紧密联系,1994 年 Baltes 和 Lindenberger 提出的共因假说(Common hypothesis)认为,感觉功能的衰退导致了老年人的认知老化,且感官功能的敏感程度是衡量老年人大脑衰老程度的指标(Lindenberger et al., 1994; 申继亮 et al., 2003)。因此,对老年人的感知训练可能通过增强老年人的感知能力,延缓老年人大脑衰老和认知下降,从“感觉功能影响认知”的角度,探究视听训练对于老年人认知能力的提升同样具有重要的理论意义。(1 引言第二段,绿色字体)

意见 4: 建议作者在梳理和分析文献时,按照一般干预研究的要素来介绍研究结果,如干预方案、干预时间和频率、干预效应、近迁移效应、远迁移效应、长期效应。这样便于对各个研究进行比较,也便于得出结论。

回应: 感谢审稿专家的建议。在文章的表述中已按照一般干预研究的要素来介绍研究结果,为了方便读者更好地理解,新增附表对文中干预研究进行了梳理(见附表 1)。

意见 5: 行文术语、英文缩写太多,使得阅读理解的负荷大增,同时也降低了文章的可理解性。核心术语需要作者有解释,如视听整合能力、适应性代偿、视听信息绑定窗口、时间绑定窗口。PD、AD 没有给出中文名称,SJ 这类的非大众熟知的英文缩写应避免。同时也注

意用语的严谨性，比如什么叫“基本认知能力 p300”？

回应：感谢审稿专家的建议。已对核心术语添加了具体解释，补充了 PD 与 AD 的中文名称，并把英文缩写 SJ 更改为中文全称“视听同时性判断任务”。

非常抱歉对“基本认知能力 P300”的表述欠妥，已将其改为“基本认知能力（例如：注意）”。（3.1 与 3.2，绿色字体）

意见 6：3.2 中谈到的跨通道视听训练的神经机制究竟指什么？目前文章呈现出来的只是一些认知神经加工的结果，而非机制。机制是说训练补偿了、加强了、替代了、泛化了大脑的什么什么活动吗？

回应：感谢审稿专家的建议。认知神经加工的结果可反映出神经加工的机制，视听训练的神经机制涉及神经元之间的连接改变，以及大脑区域的功能重组。因此，正如您所说，视听训练机制主要指视听训练对老年人相关脑区活动的补偿、加强、替代、泛化等影响。

意见 7：“4 小结和展望”与前面部分不太能够照应。这部分的“第一”、“第二”、“第三”都比较难从前文的文献综述中得出这样的小结。与意见 2 同类，这都与作者缺少对文献的梳理和分析有关。

回应：感谢审稿专家的建议。我们对文献进行了进一步的梳理和分析，已在文中增加了与“4 小结和展望”对应的相关内容。

然而，目前研究仍存在一些不足：第一，以上研究主要通过视听时间知觉训练缩小老年人视听绑定窗口，进而影响视听整合，而非直接干预老年人的视听整合能力。同时，缺少视听训练对老年人视听补偿机制的影响。更重要的是，大部分研究采用简单的视觉闪光(Flash)和听觉纯音(Pure Tone)刺激，这些单一信息的重复呈现可能会使被试感到枯燥，从而影响训练效果。因此，目前对老年人视听整合的训练效率有待提升。第二，以上研究的样本量大多为 20 人左右，受到样本量的限制，无法根据人口学等信息深入分析导致训练效果差异的因素，且未对训练的长期效应进行追踪考察。因此，尚不清楚视听训练是否具有群体特异性和保持性。第三，以上研究训练的频率和时长存在较大差异，且未对训练负荷进行横向对比研究，因此无法考察干预负荷与脑功能的相关性。（3.1 第二段，绿色字体）

意见 8：一些段落安排需要调整，如 2.2 的最后一段中跨通道训练的文字应该挪到下一个标题。3.2 最后有关动物的研究比较突兀，与上文和最后一句总结性的话没什么逻辑关联。

回应：感谢审稿专家的建议。已将 2.2 最后一段中跨通道训练的文字挪到了下一个标题。

对 3.2 最后有关动物的研究增加了与上文的关联，并修改了最后一句总结性的话（3.2 第三段，绿色）。

第二轮

审稿人 1 意见：

作者对关心的问题做了解答，论文做了必要的修改，建议接收。

审稿人 2 意见：

“共因假说”认为存在一个共同因素在影响着感觉功能和认知功能，使得它们的关系表现出较强的共变性。这个共同因素反映了中枢神经系统的机能，会对整个信息加工系统产生广泛影响，认知系统的老化可以被描述为这样一个退行性发展过程：神经信号的模糊化、加工通路去独立性以及大脑表征的去分化。作者在文中对共因假说的观点阐述可能不够严谨，不是感觉功能下降导致认知功能老化（这是因果关系的说法）。如果以共因假说为理论基础，那么为什么训练视听功能可以带来认知能力的改善？即视听训练并不是在改变那个“共因”。视听训练的理论基础还需要修正和完善。

回应：非常抱歉我们对“共因假说”的表述欠妥，已在文中进行修正，并增加了“信息退化假说”以完善视听训练的理论基础。（引言第二段，紫色字体）

共因假说(Common-cause hypothesis)认为存在一个影响着感觉功能和认知功能的共同因素，使得感觉功能与认知功能表现出较强的共变性，这个共同因素反映了中枢神经系统的机能，会对整个信息加工系统产生广泛的影响(Lindenberger & Baltes, 1994)，此假说表明老年人感知觉的衰退与认知老化具有紧密联系。同时，信息退化假说(Information-degradation hypothesis)认为，感知处理和认知加工共用一个容量有限的处理资源系统，老化导致的感知觉衰退会增加感知处理阶段的资源需求，使后续认知加工阶段可用的处理资源减少，进而影响老年人认知加工能力(Schneider & Pichora-Fuller, 2000; Slade et al., 2020)。因此，通过感知训练可能增强老年人的感知能力，以减少感知处理阶段需要调动的处理资源，进而释放更多处理资源供后续认知加工阶段使用，从“感觉与认知密切相关”的角度，探究视听训练对于老年人认知能力的提升具有重要的理论意义。

第三轮

审稿人 2 意见：同意发表。

编委 1 复审意见：同意发表。

编委 2 复审意见：稿件 jz23-246 已经详细修改，具有参考价值。建议发表。

主编意见：根据编委和审稿专家的意见，建议发表。