《心理科学进展》审稿意见与作者回应

题目：空间导航中方向感的多模态信息整合认知神经机制
作者：张军恒 黄雷 李奎良 王靖 姬鸣

第一轮

审稿人 1 意见：

空间导航能力是人类和动物在三维世界中赖以生存的特殊基本能力，在当今 AI 热议空间智能和具身智能的背景下，对已有的人类和动物的空间导航机制的研究进行梳理不仅对本领域、也对跨学科研究具有较为重要的价值。本文从行为、神经两个方面对空间导航时的视觉线索和自我运动线索加工进行了梳理，视角较为新颖、内容较为全面。但是，也存在着核心概念定义有待厘清，行为-神经研究结果有待整合，语言的准确性和可读性有待进一步提高等问题，以下是进一步的修改建议：

意见 1：首先，本文的核心关键词是空间导航中的“方向感”，它和空间导航“定位”有何区别？方向感是否意味着个体只需要知道自己大概位于什么位置朝向哪里就可以了，无需有特别精确的定位？从文章综述涉及的研究内容来看，很多研究都是和精确的空间定位有关，而不仅仅局限于方向感。因此，建议作者进一步思考本研究以“方向感”作为关键概念是否恰当。

回应：非常感谢审稿专家对文章核心概念“方向感”提出的一些问题，下面我们将逐一解答审稿专家的疑问，所有意见涉及的修改都在文中以“蓝色”字体标注。

首先，空间导航中的“方向感”和“定位”有何区别？方向感是指个体对自己所处位置的意识以及通过与周围环境(地标、边界等)的关系来确定自己所处位置的方向的能力(Kozlowski & Bryant, 1977; 许琴等, 2010)。它拥有动态的、连续的空间表征与更新机制，不仅包括个体对“我在哪里”的认知表征，还包括“我正面向哪里”以及“我应当往哪儿走”的认知表征。因此，方向感代表了个体在空间中感知、表征和更新自身朝向与位置关系的综合能力。在认知神经层面，方向感依赖于 HD 细胞、位置细胞和网格细胞的协同放电以及相关神经信号在前庭-视觉通路的整合与投射。然而，定位是指个体判断自身在环境中的绝对位置或相对于其他客体的相对位置。与方向感相比，定位更强调静态、离散的位置表征且通常只包括个体对“我在哪里”的认知表征。在认知神经层面，定位的神经基础为位置细胞和网格细胞的协调放电以及相关神经信号在海马-内嗅皮层之间的传导(Burgess et al., 2002; Moser et al., 2008)。对于两者之间的关系，徐琴等(2010)曾在综述中提到：“方向感是一种定位当前位置和选择前往目的地的路线及方向的能力，包括对地标的辨认和记忆，即对空间地点的表征，对地标间的空间关系的辨认和记忆即对空间关系的表征，以及对这些表征的整合进而确定自己位置和将要行动的朝向。”这一观点的提出是基于以往研究通常将复杂导航环境中的返航、对地标的再认、对空间地图的描绘、对路径距离的估计及对视野外目的地方向的指认等任务绩效视为空间导航中方向感的具体行为表现(Cornell et al., 2003; Li, 2007)。因此，不难看出方向感包含并超越了定位，即定位能够提供“我在哪里”的信息，而方向感能够在此基础上引入朝向和路径整合机制，让个体能够实现路线规划等复杂空间导航行为。换句话说，定位是方向感的基础，而方向感是定位在时间维度上的连续延伸与动态整合，两者的关系可以类比为“坐标点”与“坐标向量”的关系。举个例子，当一个人走进一个陌生的建筑，当他进入大厅时知道自己正处在“一楼大厅靠近门口的位置”（定位），当他转身

走向电梯时，他可能知道自己“面朝西方、电梯在他的左前方十几米远的地方”（方向感），当他到达其他楼层从电梯走出之后他仍知道自己“刚从南侧上来，现在面向北方”（方向感的动态维持与更新）。与之对应的，大脑的定位系统其实是方向感系统的一部分。定位功能的实现依赖于海马-内嗅皮层的功能联结，位置细胞负责表征当前位置，而网格细胞提供空间度量坐标。然而，方向感功能的实现不止于此，还需要头向信号传递网络的参与，即早期前庭通路、ADN、PoS、RSC、LMN 等区域共同编码个体朝向信息，以 RSC 和海马之间的双向神经信号传到通路为媒介，与海马-内嗅皮层中编码的位置信息产生交互，从而建立、维持和更新方向感。

综上，方向感与定位的区别在于前者是时空整合性的导航机制，后者是空间点位识别机制：方向感需要精确的定位信息作为基础，从而进一步整合朝向、运动和空间记忆表征，用以维持稳定的空间认知与导航行为。因此，在研究空间导航的神经机制时，以“方向感”为核心比“定位”更能体现导航系统的综合性与动态特征。

其次，方向感是否意味着个体只需要知道自己大概位于什么位置朝向哪里就可以了，无需有特别精确的定位？基于对上一问题的解答，不难发现方向感并非模糊的空间表征，而是一个整合空间和方向信息的精确认知系统，区别于我们日常用语中的“模糊朝向感”，即知道东西南北方向的大致感觉。相反，在心理学和认知神经科学领域，方向感是一种由多模态信息（视觉、前庭觉、本体觉等）整合而成的，对位置和朝向信息的精准的、动态的空间表征，而且具有时空连续性。本文综述了大量心理学、认知神经科学领域的研究，例如，心理学领域的多模态信息贝叶斯整合机制就是基于三角返航任务中被试返航到起点的方向和位置误差推理得出的；认知神经科学领域更是通过探讨多类编码朝向信息、位置信息的细胞协作来阐述方向感的认知神经网络模型。这些研究均以“方向感是对位置和朝向信息的整合和更新机制”为核心内容开展的，而本文则是以这些研究的核心发现为基础展开综述的。基于此，我们认为以“方向感”作为核心关键词能够准确反映本综述的研究主题与理论价值。

参考文献：

许琴, 罗宇, 刘嘉. (2010). 方向感的加工机制及影响因素. *心理科学进展*, *18*(8), 1208-1221.

Kozlowski, L. T., & Bryant, K. J. (1977). Sense of direction, spatial orientation, and cognitive maps. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(4), 590.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.

Moser, E. I., Kropff, E., & Moser, M. B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.*, *31*(1), 69-89.

Cornell, E. H., Sorenson, A., & Mio, T. (2003). Human sense of direction and wayfinding. *Annals of the association of American Geographers*, *93*(2), 399-425.

Li, R. (2007). *Human wayfinding and navigation in a large-scale environment: cognitive map development and wayfinding strategies* (Doctoral dissertation, University of Saskatchewan).

意见 2：其次，文章从行为、神经两个方面对空间导航时的视觉线索和自我运动线索加工进行了梳理，但是这两部分的关系并不紧密。最直接的表现就是图 2 和图 1 之间很难建立联系，也即从图 2 上难以得出哪些神经通路和机制分别与空间导航时的视觉线索和自我运动线索加工相关，而这个或许是读者最希望能够看到的内容。以研究 AI 空间智能和具身智能的同行作为读者为例，他/她可能最希望了解的是人和动物是如何实现空间导航的，要充分利用好视觉线索和自我运动线索需要设计哪些具体的信息加工环节（即模拟人和动物的神经机制）。倘若作者能够梳理出具有指导价值的神经机制图，本文的学术影响潜力将有大幅度的

提升。

回应：非常感谢审稿专家指出对行为层面和认知神经层面研究的综述关系不够紧密的问题，最直接的表现就是图 2 和图 1 之间比较难以建立联系，而且审稿专家提到如果能够在图 2 上看到哪些神经通路和机制分别与空间导航时的视觉线索和自我感知运动线索加工相关，将能够解决这一问题，并大幅提升本文的影响潜力。我们非常赞同审稿专家的观点，事实上，我们在构思图 2 时也是如此设计的。如图 2 所示，“→”代表视觉信号与基础自我运动信号的输入；“→”代表自我感知运动信号的传导通路（虚线代表的通路有待证实）；“→”代表 HD 信号的传导通路(包含自我感知运动信号和视觉信号)；“→”代表视觉信号的传导通路；红黄双色箭头代表 HD 信号、视觉信号及其整合后的复杂神经信号；蓝色环形箭头代表环形吸引子回路。基于此，我们就可以捕捉到哪些神经核团、脑区或神经通路与加工视觉线索或自我感知运动线索有关。

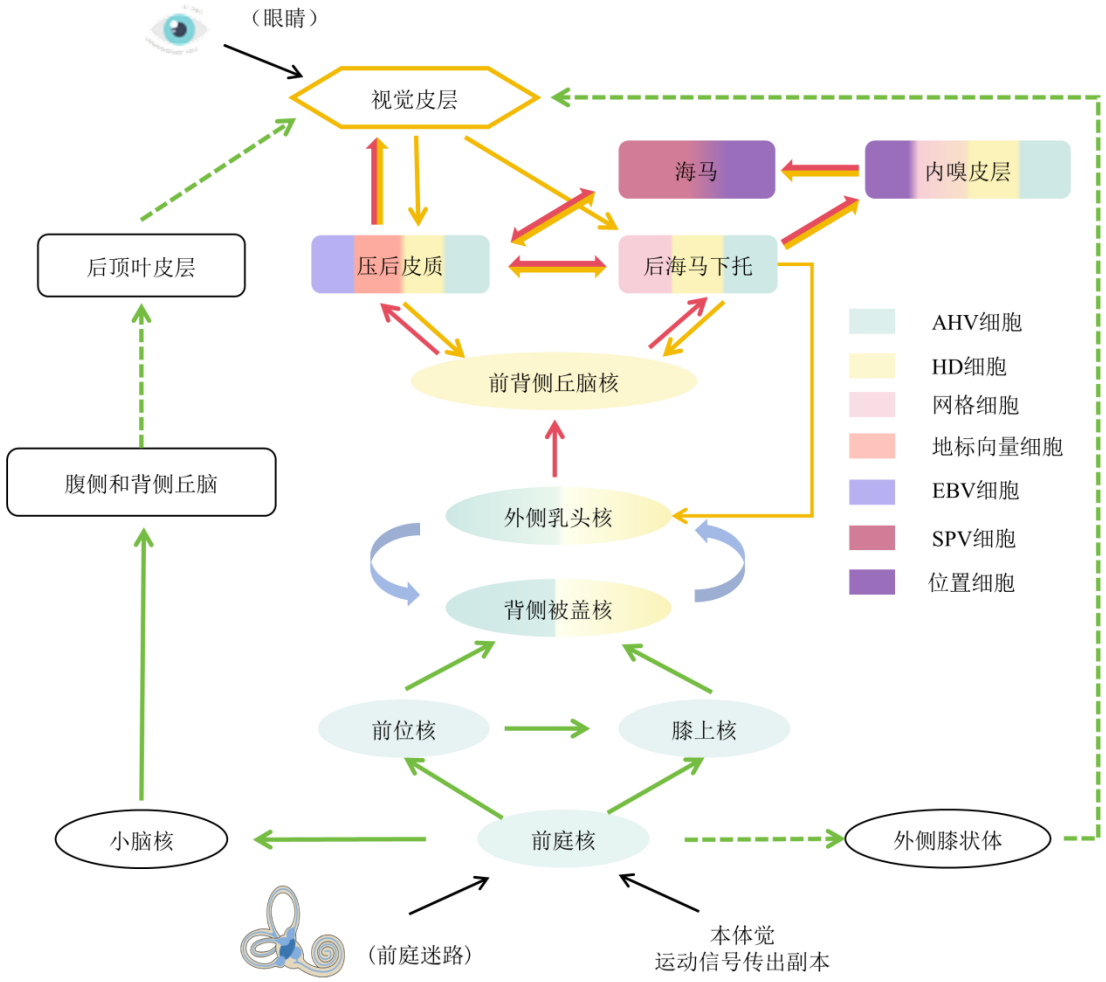

图 2 方向感的多模态信息整合神经网络模型（修改前）

具体来讲，前庭迷路和本体觉感受器对自我感知运动线索进行神经编码后“自下而上”传导自我感知运动神经信号，经“→”代表的自我感知运动神经信号传导通路传递至前庭核，再由前庭核传递至前位核以及膝上核，这一信息传导通路对应了文中所描述的“自我感知运动信号传导的早期前庭通路”，在行为层面对应着个体在空间导航中对自我感知运动线索的认知加工。

自我感知运动神经信号传导至背侧被盖核和外侧乳头核会与由视觉皮层对视觉线索进行神经编码后产生且经“→”代表的视觉神经信号传导通路传递至外侧乳头核的视觉神经信

号进行整合，产生 HD 信号，这一信息传导通路对应了文中所描述的“HD 信号产生的环形吸引子回路”，在行为层面对应着个体在空间导航中对视觉线索的认知加工以及初步整合视觉线索与自我感知运动线索建立方向感的认知加工过程。

HD 信号（包含自我感知运动信息和视觉信息）“自下而上”经“→”代表的 HD 信号传导通路投射到前背侧丘脑核，经前背侧丘脑核上行传递至压后皮质、后海马下托，HD 信号传递至这两处脑区后会再与视觉信号进行整合，形成成分复杂的神经信号（红黄双色箭头代表着 HD 信号、视觉信号及其整合后的复杂神经信号），这些神经信号会在压后皮质与后海马下托之间相互传导，或经压后皮质传递至海马、经后海马下托传递至内嗅皮层，内嗅皮层也会将这些复杂神经信号传递至海马，经海马再传递至压后皮质，这一系列神经信号的传导对应文中“HD 信号的投射网络及其与视觉信号的深度整合”的内容，在行为层面对应着基于复杂视觉线索或自我感知运动线索整合的认知加工过程，例如基于编码时空背景信息、边界信息、精确位置信息以及对时空信息的记忆表征的更新方向感或建立方向感预期。

另外，通过观察图 2 中神经核团或脑区的主要细胞群结构，也可大致判断这些神经核团或脑区与哪些空间导航线索的认知加工有关。例如，AHV 细胞明显只处理自我感知运动信息，所以这些细胞主要集中在早期前庭系统，到背侧被盖核、外侧乳头核以及前背侧丘脑核就开始逐渐“去 AHV 细胞化”（外侧乳头核中 AHV 细胞仅占比），出现能够同时编码两种线索的 HD 细胞初步建立方向感，随后到高级皮层区域的细胞群则功能更加复杂，用于编码更为精准的位置和朝向信息，如网格细胞、位置细胞、地标向量细胞等，这些细胞能够整合多模态的线索相关神经信号且获得精准的方位信息，以帮助个体在复杂的导航环境中维持和更新方向感。此外，还有两条未验证的通路与自我感知运动线索的认知加工有关，如图 2 最外侧两条“→”（虚线）所示。一是从前庭系统开始，经前庭核、外侧膝状体后直接投射到视觉皮层，二是从前庭系统开始，经前庭核、小脑核、腹侧和后侧丘脑、后顶叶皮层投射到视觉皮层，这两条通路被推测都只传递纯粹的自我运动信号，在行为层面对应着哪些认知过程还有待进一步验证。

由此可见，图 2 基本可以说明在空间导航中哪些神经通路和机制分别与空间导航时的视觉线索和自我运动线索加工相关，只是还需要在图注中给予更多解释，才能让读者更清晰地了解多模态线索的认知加工与哪些神经核团、脑区、神经信号或传导通路的对应关系。另外，我们在图中还增加了模拟导航场景示意图，以分别代表个体从外界获取视觉线索或自我感知运动线索，也更正了一个小问题：对空间视角细胞的简称通常有两种 SVC 细胞和 SPV 细胞，但 SPV 细胞为不常用的简称，因此将其修改为了 SVC 细胞。修改后的图 2 如下(正文 p10)：

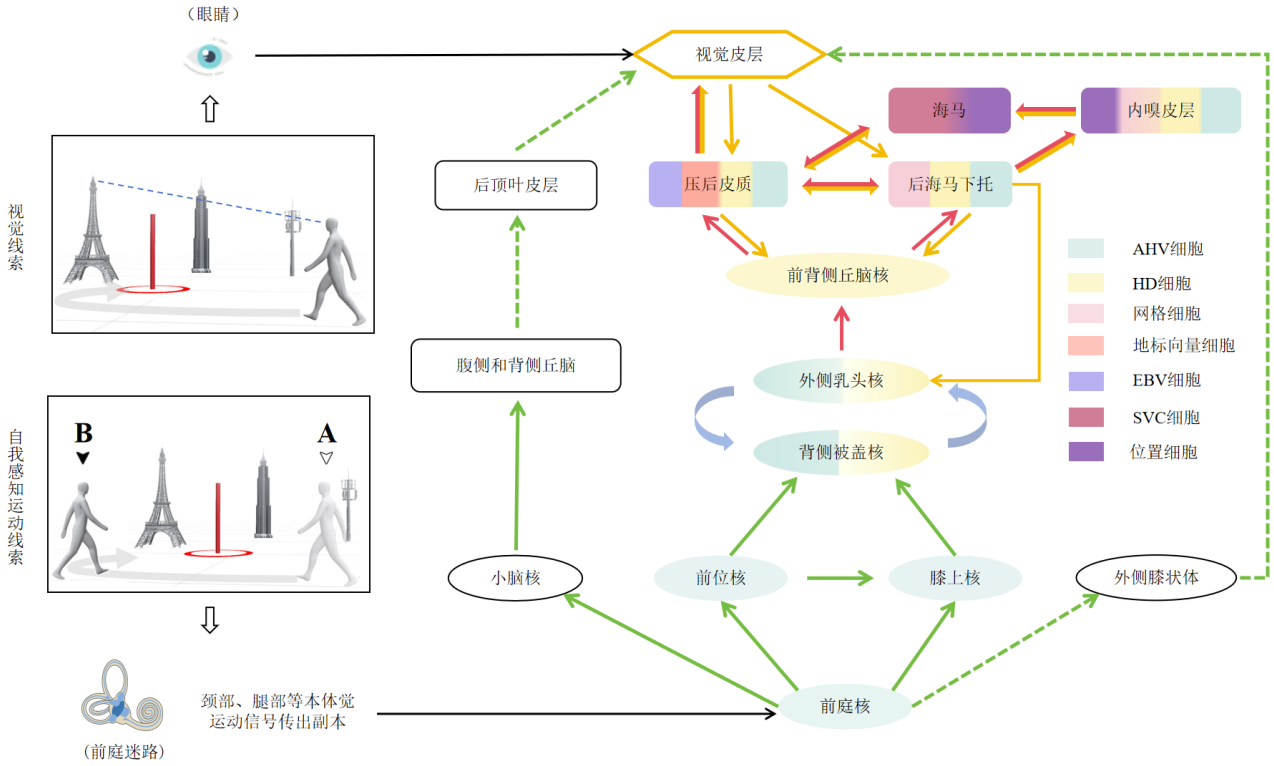

图 2 方向感的多模态信息整合神经网络模型

注：视觉线索代表个体在空间导航环境中通过视觉获取的地标、边界等信息；自我感知运动线索代表个体从“A”点到“B”点运动产生的前庭觉、本体觉或运动信号传出副本等信息；“→”代表视觉信号与基础自我运动信号的输入；“→”代表自我感知运动信号的传导通路（虚线代表的通路有待证实）；“→”代表 HD 信号(包含自我感知运动信号和视觉信号)的传导通路；“→”代表视觉信号的传导通路；红黄双色箭头代表 HD 信号、视觉信号及其整合后的复杂神经信号；蓝色环形箭头代表环形吸引子回路。前庭核对前庭觉、本体觉、运动信号传出副本等自我感知运动线索进行编码产生自我感知运动信号，经“→”传导至背侧被盖核和外侧乳头核，对应着个体在空间导航中对自我感知运动线索的认知加工。由视觉皮层对视觉线索编码后产生的视觉信号经“→”传递至外侧乳头核与自我感知运动信号整合产生 HD 信号，对应着个体在空间导航中对视觉线索的认知加工以及初步整合视觉线索与自我感知运动线索以建立方向感的认知加工过程；HD 信号经“→”投射到前背侧丘脑核，又经前背侧丘脑核上行传递至压后皮质、后海马下托，与视觉信号整合为成分复杂的神经信号，这些复杂的神经信号以压后皮质为信息联络皮层，在后海马下托与海马之间之间交互传导，后海马下托亦可通过内嗅皮层单向传递这些复杂信号至海马，这一 HD 信号的投射及其与视觉信号深度整合的神经网络激活对应着个体在空间导航中基于复杂视觉线索、自我感知运动线索及其相关记忆表征的整合以维持和更新方向感的认知加工过程。

综上，通过在图 2 中增加代表个体从外界获取视觉线索或自我感知运动线索的空间导航示意图，以及在图注中详细介绍相应通路及其行为层面对应的认知加工过程，较好地综合以往研究在行为层面和认知神经层面的研究结果。如此也可将较为粗放的行为层面研究与较为细致的认知神经层面研究对应起来，方便读者联系两条综述主线的内容。

意见 3：第三，文章关注空间导航时的视觉线索和自我运动线索，但对于具体是哪些线索内容以及各线索内容究竟在哪些情境下对空间导航有重要作用还有待进一步阐明。比如，视觉线索包括哪些，视觉流、地标、边界、空间轴向、朝向依赖的快照……为什么运动信号传出副本是必须的？即只靠视觉和前庭本体等信号为什么不行？

回应：非常感谢审稿专家指出应当进一步阐明视觉线索和自我感知运动线索具体包含了哪几类线索，以及这些线索在空间导航中发挥了怎样的作用。目前我们已经在“2.1 多模态信息编码的独立性”部分专门增加了一个段落来详细阐述这些线索以及它们的作用，具体内容如下（正文 p3-p4）：

“视觉线索和自我感知运动线索是个体从导航环境获取空间信息的两个主要来源。视觉线索包括地标、边界、环境的几何结构等(郝鑫等, 2022; Cheng & Newcombe, 2005; Hamburger, 2020)。其中，地标是指环境中具有显著视觉特征且易于被感知和记忆的、稳定的外部参照物(Presson & Montello, 1988; O’Keefe & Nadel, 1978)，通常被作为空间参照点来确定自身位置和方向，如果在空间导航中长时间缺少地标作为参照物，路径整合误差会逐渐变大，导致个体逐渐迷失方向(Anastasiou et al., 2023)。边界是指在个体的视野中占据较大比例且具有立体拓展平面的障碍物，也是个体在空间导航中实现路径整合的关键参照物(Lee, 2017; Lew, 2011)，而且在对物体进行定位时，边界具有较高的参照优先级(Lee, 2017)。环境的几何结构通常指空间整体的形状、比例、对称性及各边界之间的角度关系，这些几何属性能够为个体提供稳定的空间参照，从而在缺乏显著地标的空间导航情境中确定自己的位置(Gallistel, 1990)，特别是在小尺度空间内的定向个体尤其依赖环境几何结构(Chiandetti & Vallortigara, 2008)。自我感知运动线索包括前庭信号、本体觉信号、运动信号传出副本、视觉流(optic flow)等(Berthoz et al., 1995; Srinivasan et al., 1997; Newman et al., 2023)。前庭信号是指前庭系统对头部角向加速度、线性加速度以及重力方向的感知，能够帮助个体在空间导航中感知身体姿态、维持躯体平衡，并为方向感知提供动态参照(Angelaki & Cullen, 2008; Cullen, 2012)。本体觉信号是指个体通过肌肉、肌腱和关节感受器对身体各部分的位置、姿势和运动状态的感知，与前庭信号结合可以帮助个体实现对身体姿态变化的连续估计，为空间定向与路径整合提供躯体参照(Proske & Gandevia, 2012)。在空间导航中缺少前庭觉和本体觉会导致个体无法准确更新自我位置和方向信息，路径回溯能力也会显著下降(Waller et al., 2004; Brandt et al., 2005)。运动信号传出副本是感觉–运动整合和维持方向感连续性的关键机制，通常指运动指令发出时大脑同时向感觉系统发送的复制信号，从而建立由躯体运动引发的感觉预期以区分自我运动与被动运动(Crapse & Sommer, 2008)。例如，个体在转动头部时视野会发生变化，但并没有产生周围事物在旋转的感觉，这是因为视觉系统接收了与头部运动相关的运动信号传出副本。研究发现，当运动信号传出副本缺失或延迟会导致视觉稳定性降低或产生错觉性运动感，从而使个体的方向感紊乱，对目标方向的估计产生明显偏离(Medendorp et al., 2003)。视觉流是个体相对环境运动时图像在视网膜上移动所产生的动态视觉信息(Warren & Hannon, 1988)，缺少视觉流不仅会让个体无法正确估计运动速度和方向，还会导致路径整合误差随时间增加(Harris et al., 2000; Wall & Smith, 2008)。”

增引文献：

郝鑫, 袁忠萍, 林淑婷, 沈婷. (2022). 边界促进空间导航的认知神经机制. *心理科学进展*, *30*(7), 1496.

Cheng, K., & Newcombe, N. S. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychonomic Bulletin & Review*, *12(*1), 1-23.

Presson, C. C., & Montello, D. R. (1988). Points of reference in spatial cognition: Stalking the elusive landmark. *British Journal of Developmental Psychology*.

Lee, S. A. (2017). The boundary-based view of spatial cognition: a synthesis. *Current Opinion in Behavioral Sciences*, *16*, 58-65.

Lew, A. R. (2011). Looking beyond the boundaries: time to put landmarks back on the cognitive map?. *Psychological Bulletin*, *137*(3), 484.

O'keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford university press.

Gallistel, C. R. (1990). *The organization of learning*. The MIT Press.

Chiandetti, C., & Vallortigara, G. (2008). Spatial reorientation in large and small enclosures: comparative and developmental perspectives. *Cognitive processing*, *9*(4), 229-238.

Berthoz, A., Israël, I., Georges-François, P., Grasso, R., & Tsuzuku, T. (1995). Spatial memory of body linear displacement: what is being stored?. *Science*, *269*(5220), 95-98.

Srinivasan, M. V., Zhang, S. W., & Bidwell, N. J. (1997). Visually mediated odometry in honeybees. *Journal of Experimental Biology*, *200*(19), 2513-2522.

Angelaki, D. E., & Cullen, K. E. (2008). Vestibular system: the many facets of a multimodal sense. *Annu. Rev. Neurosci.*, *31*(1), 125-150.

Cullen, K. E. (2012). The vestibular system: multimodal integration and encoding of self-motion for motor control. *Trends in Neurosciences*, *35*(3), 185-196.

Proske, U., & Gandevia, S. C. (2012). The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiological Reviews*, *92*(4), 1651 - 1697.

Waller, D., Loomis, J. M., & Haun, D. B. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review*, *11*(1), 157-163.

Brandt, T., Schautzer, F., Hamilton, D. A., Brüning, R., Markowitsch, H. J., Kalla, R., ... & Strupp, M. (2005). Vestibular loss causes hippocampal atrophy and impaired spatial memory in humans. *Brain*, *128*(11), 2732-2741.

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge circuits in the primate brain. *Current Opinion in Neurobiology*, *18*(6), 552-557.

Medendorp, W. P., Goltz, H. C., Vilis, T., & Crawford, J. D. (2003). Gaze-centered updating of visual space in human parietal cortex. *Journal of Neuroscience*, *23*(15), 6209-6214.

Warren Jr, W. H., & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature*, *336*(6195), 162-163.

Harris, L. R., Jenkin, M., & Zikovitz, D. C. (2000). Visual and non-visual cues in the perception of linear self motion. *Experimental Brain Research*, *135*(1), 12-21.

Wall, M. B., & Smith, A. T. (2008). The representation of egomotion in the human brain. *Current Biology*, *18*(3), 191-194.

意见 4：第四，文章中很多标蓝的部分都以"本研究认为"等非常笃定的语气来进行论述。一方面，本文作为综述文章，并非"研究"，因此自称为"本研究"不太妥当。另一方面，建议把"认为"这种非常肯定的语气降级为"推测"等更为保守的语气，毕竟作者的结论并非建立在经过自己实证研究检验的基础上。

回应：非常感谢审稿专家指出在文章写作中我们基于前人研究结果做出推理时的表述方式所存在的问题。我们非常认同审稿专家的观点：一方面，综述文章并非原创性文章，以"本研究"为主语的表述方式欠妥；另一方面，在没有实证研究结果作为支撑的条件下，基于前人研究结果做出的推理时，应当以保守的口吻展开表述。在最新的修改稿中，我们已经全面修改了相应的表述，全部删除了"本研究认为"等表述，改为"基于此可以推测或基于上述内

容可以推测”等表述。由于相关修改量大而简短，这里仅给出几处示例作为参考：

（1）原文：“研究发现，地标导航和路径整合是两个相对独立且功能互补的过程(Chen et al., 2017; Chen et al., 2019; Nardini et al., 2008; Sjolund et al., 2018; Zhang et al., 2020; Zhang & Mou, 2017)，本研究认为在空间导航期间对视觉线索和自我运动线索的信息提取与编码也是相互独立的。”

修改为：“研究发现，地标导航和路径整合是两个相对独立且功能互补的过程(Chen et al., 2017; Chen et al., 2019; Nardini et al., 2008; Sjolund et al., 2018; Zhang et al., 2020; Zhang & Mou, 2017)，基于此可以推测在空间导航期间对视觉线索和自我运动线索的信息提取与编码也是相互独立的。”(正文 p5)

（2）原文：“本研究认为多模态信息整合能否发生在出行阶段及其整合频率取决于地标等视觉线索的相对可靠性以及个体动机。”

修改为：“综上，可以推测多模态信息整合能否发生在出行阶段及其整合频率可能取决于地标等视觉线索的相对可靠性以及个体动机。”(正文 p6)

（3）原文：“可见，在空间导航中想要建立、维持和更新方向感必须依靠工作记忆，因此本研究认为多模态信息整合不仅存在于个体建立和维持工作记忆的过程，也发生在工作记忆向长时记忆的动态信息交互阶段。”

修改为：“可见，在空间导航中想要建立、维持和更新方向感必须依靠工作记忆，因此可以推测多模态信息整合不仅存在于个体建立和维持工作记忆的过程，也发生在工作记忆向长时记忆的动态信息交互阶段。”(正文 p8-p9)

（4）原文：“正如本研究提出的方向感的多模态信息整合神经网络模型，其结构十分复杂，因此想要直接构建完整的神经计算模型较为困难。”

修改为：“正如上述提到的方向感的多模态信息整合神经网络模型，其结构十分复杂，因此想要直接构建完整的神经计算模型较为困难。”(正文 p20)

意见 5：最后，文章涉及到大量的英文缩写，尽管作者给出了表 1，但是，表 1 并没有囊括全部的英文缩写，在后面的正文部分，又间或地出现新词语和缩写，对于不熟悉本领域的读者来说，需要不停地前后翻看寻找比对。从对读者友好的角度，建议文章中不太复杂或者使用不频繁的词语都直接采用中文。

回应：非常感谢审稿专家指出表 1 中的英文缩写没有囊括所有专有名词的问题，以及建议文章中不太复杂或者使用不频繁的词语直接使用中文。我们非常赞同审稿专家的观点，目前已经更新了表 1，以及将文章中不太复杂或使用不频繁的词语全文始终使用中文描述。在这里也简单解释一下，我们将一些专有名词列入表格的缘由。一方面，这些专有名词在认知神经科学领域有着独有的英文缩写，例如，“HD 细胞”通常专指头向细胞、“HPC”通常专指海马；另一方面，是这些专有名词在文章中出现的频率超过了 4 次。符合这两个条件的专有名词，在文中首次出现会被给予英文注释及英文缩写，随后都使用英文表述。不符合这两个条件的专有名词没有被列入表 1 中，一是以往研究中通常不以缩写的形式描述这些专有名词，文中便只在第一次出现时给予其英文注释，随后都使用中文表述，例如，网格细胞(grid cells)、位置细胞(place cells)等；二是以往研究中通常以英文缩写的形式表述但在本文中出现频率小于 4 次，文中便只在第一次出现时给予英文解释及英文缩写，随后都使用中文表述，例如，前位核(nucleus prepositus hypoglossi, NPH)、膝上核(supragenual nucleus, SN)等。

基于上述规则我们更改了表 1，在“脑区及相应神经核团”板块中删除了两个专有名词：前位核(nucleus prepositus hypoglossi, NPH)、膝上核(supragenual nucleus, SN)；在“细胞和神经元模块”中增加了一个专有名词：“空间视角细胞(spatial view cells, SVC)”。修改后的表 1 如下(正文 p2)：

表 1 专有名词中英文全称与简称对应列表

	中文全称	英文全称	简称
脑区及相应神经核团	前庭核	vestibular nuclei	VN
	背侧被盖核	dorsal tegmental nucleus	DTN
	外侧乳头核	lateral mammillary nucleus	LMN
	前背侧丘脑核	anterodorsal thalamic nucleus	ADN
	后海马下托	postsubiculum	PoS
	压后皮质	retrosplenial cortex	RSC
	初/高级视觉皮层	primary and higher-order visual cortices	V1/V2
	内嗅皮层	entorhinal cortex	EC
	内侧内嗅皮层	medial entorhinal cortex	MEC
	海马	hippocampus	HPC
细胞和神经元	头部角速度细胞	angular head velocity cells	AHV 细胞
	头部方向细胞	head direction cells	HD 细胞
	自我中心边界向量细胞	egocentric boundary vector cells	EBV 细胞
	环境中心边界向量细胞	allocentric boundary vector cells	ABV 细胞
	空间视角细胞	spatial view cells	SVC 细胞
	前庭-唯一神经元	vestibular-only neurons	VO 神经元
	前庭-眼反射神经元	vestibulo-ocular reflex neurons	VOR 神经元
	位置-前庭-暂停神经元	position-vestibular-pause neurons	PVP 神经元
	小脑叶状体靶向神经元	floccular targeting neurons	FTN 神经元

审稿人 2 意见：

本文较全面综述了多模态空间信息的整合的行为和神经机制的最新研究，并提出三个的研究方向。对未来的研究有一定的指导意义。

意见：在论述的过程中，作者时常提出自己的观点，并指出需要在今后的研究中检验自己的观点。建议作者更详细的说明自己的观点和文献中的已有观点有何重要不同，并如何用实验研究检测不同观点。结论部分提出的三个研究方向和前面的综述部分的脱节较大。三个研究方向本身都很有意义，但是和前文论述中作者自己的假设没有直接的关系。总之，本文对前沿研究的介绍很全面，但在讨论解决问题时需要提出和前文综述直接相关的，清晰的研究观点和实验研究的思路。

回应：首先，非常感谢审稿专家指出本文在提出自己的观点时应当进一步说明与前人观点的重要不同以及给出较为详细的研究思路去验证相关观点。我们非常赞同审稿专家的观点，给出详细的研究思路有助于对以往研究进行更为具体的延伸，并为未来研究提供更为具体的方向。其次，也非常感谢审稿专家指出总结展望部分和研究综述部分的内容关联性较低的问题，

以及感谢审稿专家对我们所提出的三个研究方向给出都很有意义的评价。我们十分认同审稿专家的观点，应当在提出研究方向时进一步结合在综述部分提出的相关假设，给出更加清晰的研究观点和实验研究思路。

审稿专家的意见原本可以拆分为两个问题，但我们决定在这里做出一个综合性的回复。一方面，是考虑到我们在综述部分提出的假设性观点与“总结与展望部分”的未来研究方向要保持高度一致，正如审稿专家所提到的，未来研究的方向应当与本文在综述部分所提出的假设性观点相对应；另一方面，在综述部分全面展开对未来研究的思路不仅会使综述部分的重点不够突出，而且类似的内容如果再出现在“总结与展望部分”就会产生重复论述的问题。因此，我们认为将用于验证综述部分假设性观点的详细研究思路写入“总结与展望”部分，不仅能够解决如何验证综述部分假设性观点的问题，也不至于未来研究方向相关的内容与综述部分相脱节。

下面，我们简要介绍一下相关回复的逻辑。首先，我们统计了文章中我们所表达的所有观点后将其分为三类：一是基于以往研究提出的假设性观点，没有任何研究证据作为支持；二是基于以往研究提出的假设性观点，但仅存在间接研究证据作为支持；三是以往研究中未提及，但我们基于以往大量实证研究结果提出的观点（多集中在“方向感的多模态信息整合神经机制”部分）。其中，需要进一步给出详细说明以及研究思路的观点属于第一类和第二类（主要集中在“方向感的多模态信息整合机制”部分）。其次，我们在综述部分保留了对这些观点的详细描述并进一步阐述了它们与以往研究观点有何重要不同。另外，为了方便审稿专家在回复中更快速地将假设性观点与研究思路相对应，我们在回复中阐述这些观点与以往观点有何重要不同之后，也随即给出了能够用于验证这些观点的详细研究思路，但实际上这些内容在最新的修改稿中被放在了“总结与展望部分”；最后，我们将这些内容进行了综合性的表述，放在了“总结与展望”中对未来研究方向的第一条中。具体涉及的观点及相应研究思路如下，涉及的内容修改在文中以“红色”字体标注：

（1）观点 1：“综上，可以推测多模态信息整合能否发生在出行阶段及其整合频率取决于地标等视觉线索的可靠性、可用性以及个体动机。在地标持续可用的条件下（可靠性较高且始终处在导航者视线范围内），多模态信息整合发生的阶段取决于个体的导航动机，如果个体只想在出行期间阶段性地矫正路径整合积累的误差，就会采取间歇式的整合策略(Zhang & Mou, 2017)，而当个体需要时刻维持方向感来构建认知地图时，就会采取连续式的整合策略(He & McNamara, 2018; Qi et al., 2021)。在地标并非持续可用的条件下（可靠性低且无法始终处在导航者视线范围内），个体只能使用间歇式的整合策略，且在任一时间进行信息整合的前提是存在可用的地标，如果出行期间没有可用地标，个体将无法进行多模态信息整合，只有在出行期间存在可用地标时，个体才有机会进行多模态信息整合，且整合策略受其导航动机影响。”

观点 1 是基于以往研究在多模态信息整合阶段、整合策略存在争议的基础上提出的，这一观点与以往研究观点的最大不同是强调了多模态信息整合的灵活性，推测多模态信息整合并不固定发生在出行阶段或返航阶段，其整合方式（连续 vs 间歇）也不是固定不变的，而是取决于视觉线索的相对可靠性（可靠性、可用性）以及导航动机。未来研究可以通过使用线索整合范式，在三角返航任务中通过操纵视觉线索的可靠性（高 vs 低）、是否持续可用（持续可见 vs 分段可见）并在出行和返航阶段每段路径的中间、末端连续施测，通过计算出行和返航阶段的指向误差、返航误差以及连续施测过程中误差曲线的变化模式（平滑下降-连续式整合、阶段性下降-间歇式整合）来探讨视觉线索相对可靠性对多模态信息整合发生阶段及其策略的影响。在此基础上，再操纵导航动机（地图学习动机 vs 目标抵达动机），亦可探讨导航动机与视觉线索相对可靠性对多模态信息整合发生阶段及其策略的共同影响及这些因素之间的交互作用。

基于此，修改后的观点 1 如下（正文 p6-p7）：

“综上，可以推测多模态信息整合能否发生在出行阶段及其整合频率取决于地标等视觉线索的相对可靠性以及个体动机，而不是以某种定式的整合策略固定发生于出行或返航阶段。在地标持续可用的条件下（可靠性较高且始终处在导航者视线范围内），多模态信息整合发生的阶段取决于个体的导航动机，如果个体只想在出行期间阶段性地矫正路径整合积累的误差，就会采取间歇式的整合策略(Zhang & Mou, 2017)，而当个体需要时刻维持方向感来构建认知地图时，就会采取连续式的整合策略(He & McNamara, 2018; Qi et al., 2021)。在地标并非持续可用的条件下（可靠性低且无法始终处在导航者视线范围内），个体只能使用间歇式的整合策略，且在任一时间进行信息整合的前提是存在可用的地标，如果出行期间没有可用地标，个体将无法进行多模态信息整合，只有在出行期间存在可用地标时，个体才有机会进行多模态信息整合，且整合策略受其导航动机影响。”

在“总结与展望部分”与观点 1 对应的未来研究方向如下（正文 p19-p20）：

“未来研究可以通过使用线索整合范式，在三角返航任务中通过操纵视觉线索的可靠性（高 vs 低）、是否持续可见（持续可见 vs 分段可见）并在出行和返航阶段每段路径的中间、末端连续施测，通过计算出行和返航阶段的指向误差、返航误差以及连续施测过程中误差曲线的变化模式（平滑下降-连续式整合、阶段性下降-间歇式整合）来探讨视觉线索相对可靠性对多模态信息整合发生阶段及其策略的影响。在此基础上，如果再操纵导航动机（地图学习动机 vs 目标抵达动机），亦可探讨导航动机与视觉线索相对可靠性对多模态信息整合发生阶段及其策略的共同影响以及这些因素之间的交互作用。”

（2）观点 2：“所以，视觉空间表征可能并非没有与自我运动表征建立耦合关系，只是它们建立耦合关系的过程发生在视觉增益的适应阶段，但是像 Tcheang 等(2011)和 Du 等(2020)的研究重点均在个体适应视觉增益的结果上，无法判断个体在适应视觉增益期间视觉空间表征与自我运动表征的耦合关系。因此，未来研究可以进一步探讨在个体适应视觉增益期间视觉空间表征和自我运动表征的耦合关系。”

观点 2 已经明确指出与以往研究结果的不同，即“视觉空间表征可能并非没有与自我运动表征建立耦合关系，只是它们建立耦合关系的过程发生在视觉增益适应阶段”，但是以往研究没能探讨个体在适应视觉增益期间视觉空间表征和自我运动表征的耦合关系。因此，未来可以通过在视觉增益训练期间，每经过几次训练就施加一个探测刺激（仅提供视觉线索或自我感知运动线索）让被试完成指向任务。如果在视觉增益训练期间仅提供视觉线索或自我感知运动线索的条件下被试指向表现均受视觉增益影响，则说明耦合是双向的；反之，如果仅呈现视觉线索条件下被试指向表现不受视觉增益影响，而仅呈现自我感知运动线索条件下被试指向表现受视觉增益影响，则说明耦合是自我感知运动表征向视觉表征的单向耦合。此外，还以通过与视觉增益训练后、自我感知运动增益训练期间及训练后的耦合模式对比，探讨耦合关系的稳定性以及耦合强度的不对称性问题。

基于此，修改后的观点 2 如下（正文 p8）：

“所以，视觉空间表征可能并非没有与自我运动表征建立耦合关系，只是它们建立耦合关系的过程发生在视觉增益的适应阶段，但是像 Tcheang 等(2011)和 Du 等(2020)的研究重点均在个体适应视觉增益的结果上，无法判断个体在适应视觉增益期间视觉空间表征与自我运动表征的耦合关系。”

在总结与展望部分与观点 2 对应的未来研究方向如下（正文 p19-p20）：

“未来研究可以通过在视觉增益训练期间，每经过几次训练就施加一个探测刺激（仅提供视觉线索或自我感知运动线索）让被试完成指向任务。如果在视觉增益训练期间仅提供视觉线索或自我感知运动线索的条件下被试指向表现均受视觉增益影响，则说明耦合是双向的；反之，如果仅呈现视觉线索条件下被试指向表现不受视觉增益影响，而仅呈现自我感知运动

线索条件下被试指向表现受视觉增益影响，则说明耦合是自我感知运动表征向视觉表征的单向耦合。此外，还以通过与视觉增益训练后、自我感知运动增益训练期间及训练后的耦合模式对比，探讨耦合关系的稳定性以及耦合强度的不对称性问题。”

（3）观点 3：“另外，在熟悉的导航环境中地标导航重置路径整合的过程包括视觉线索与长时记忆中的多模态空间表征的匹配，以及将与视觉线索相关联的自我运动表征提取至工作记忆，最终重置路径整合(Etienne et al., 2004; Zhang & Mou, 2017)。这一过程也可能涉及多模态信息整合，因为个体若想有效地重置路径整合，必须从长时记忆提取与视觉线索相关联的自我运动表征，而想要判断自我运动表征与视觉线索的相关性，则需要将其与视觉线索提供的信息整合，并与长时记忆中相应的多模态表征进行匹配，若一致则代表相关，未来可以通过研究验证这一推测。”

观点 3 与以往研究所提出观点的最大不同在于，以往研究仅关注工作记忆建立和维持阶段的多模态信息整合，而我们推测多模态信息整合还可能发生在工作记忆与长时记忆动态信息交互的阶段（在观点 3 所在段落开头的第一句已有说明：可见，在空间导航中想要建立、维持和更新方向感必须依靠工作记忆，因此可以推测多模态信息整合不仅存在于个体建立和维持工作记忆的过程，也发生在工作记忆与长时记忆动态信息交互的阶段）。特别是在熟悉环境中需要重置路径整合误差时，从长时记忆中提取相应的自我感知运动表征与工作记忆中的视觉表征进行整合是最高效的方式。未来研究可以让被试学习包含稳定地标的虚拟导航路线，从而构建多模态的长时记忆空间表征，随后通过无地标路径整合任务干扰工作记忆表征，最后在视觉线索与自我感知运动线索匹配、不匹配或无视觉线索条件下分别要求被试完成指向任务。如果线索匹配条件相比其他条件能够显著降低指向误差，则表明视觉线索能够激活长时记忆中自我感知运动表征，并在工作记忆与长时记忆信息交互的过程中实现视觉表征与自我感知运动表征的整合，从而重置路径整合的误差。

基于此，修改后的观点 3 如下（正文 p9）：

“另外，在熟悉的导航环境中地标导航重置路径整合的过程包括视觉线索与长时记忆中的多模态空间表征的匹配，以及将与视觉线索相关联的自我运动表征提取至工作记忆，最终重置路径整合(Etienne et al., 2004; Zhang & Mou, 2017)。这一过程也可能涉及多模态信息整合，因为个体若想有效地重置路径整合，必须从长时记忆提取与视觉线索相关联的自我运动表征，而想要判断自我运动表征与视觉线索的相关性，则需要将其与视觉线索提供的信息整合，并与长时记忆中相应的多模态表征进行匹配。”

在总结与展望部分与之观点 3 对应的未来研究方向如下（正文 p19-p20）：

“未来研究可以让被试学习包含稳定地标的虚拟导航路线，从而构建多模态的长时记忆空间表征，随后通过无地标路径整合任务干扰工作记忆表征，最后在视觉线索与自我感知运动线索匹配、不匹配或无视觉线索条件下分别要求被试完成指向任务。如果线索匹配条件相比其他条件能够显著降低指向误差，则表明视觉线索能够激活长时记忆中自我感知运动表征，并在工作记忆与长时记忆信息交互的过程中实现视觉表征与自我感知运动表征的整合，从而重置路径整合的误差。”

此外，除了上述提到的观点所对应的未来研究方向之外，还有像多模态信息的贝叶斯最优整合等问题也有待进一步探讨。例如，在真实的空间导航环境中，视觉线索或感知运动线索的可靠性并非一成不变的，那么对应的贝叶斯最优整合的权重是否动态变化的。未来研究可以通过操纵在空间导航期间视觉线索或感知运动线索的可靠性（由高到低、由低到高或恒定高、恒定低），通过构建可靠性与权重关联的时间函数，分析 $w_v(t)$、$w_s(t)$变化曲线来判断多模态信息是否存在动态贝叶斯最优整合。总之，未来应从多模态信息的整合方式、整合阶段、耦合特征及记忆交互的角度进一步开展研究以完善方向感的多模态信息整合模型。在“总结与展望”部分增加的具体内容如下（正文 p19-p20）：

“第一，从多模态信息的整合方式、整合阶段、耦合特征及记忆交互的角度进一步开展研究以探讨方向感的多模态信息动态整合机制：

首先，对于多模态信息的整合方式，以往研究通常在固定视觉线索和自我感知运动线索可靠性的条件下，探讨个体在构建方向感时对多模态信息的赋权和整合方式(Sjolund et al., 2018; Zhao & Warren, 2015; Shettleworth & Sutton, 2005; Kessler et al., 2024)。然而，在现实的空间导航环境中，视觉线索或自我感知运动线索的可靠性并非一成不变的，视觉线索可靠性可能因光线变暗、障碍物遮挡等原因而发生变化，自我感知运动线索可靠性亦可能因交通方式在步行和乘车之间转换而发生变化，当这些线索的可靠性发生动态变化时，方向感的多模态信息最优贝叶斯整合模型是否仍适用还未可知。未来研究可以通过操纵在空间导航期间视觉线索或感知运动线索的可靠性（由高到低、由低到高或恒定高、恒定低），通过构建可靠性与权重关联的时间函数 $w_v(t)$、$w_s(t)$，分析相应变化曲线来判断与方向感相关的多模态信息是否存在动态贝叶斯最优整合模型。

其次，对于多模态信息的整合阶段，以往研究仍对信息整合发生在出行或返航阶段存在争议，对信息整合是否连续发生的结论也未能统一(Kalia et al., 2013; Newman & McNamara, 2020; Zhang & Mou, 2017; Zhang et al., 2020; Chen & Mou, 2025)。未来研究可以使用线索整合范式，在三角返航任务中通过操纵视觉线索的可靠性（高 vs 低）、是否持续可用（持续可见 vs 分段可见）并在出行和返航阶段每段路径的中间、末端连续施测，通过计算出行和返航阶段的指向误差、返航误差以及连续施测过程中误差曲线的变化模式（平滑下降-连续式整合、阶段性下降-间歇式整合）来探讨视觉线索相对可靠性对多模态信息整合发生阶段及其策略的影响。在此基础上，再操纵导航动机（地图学习动机 vs 目标抵达动机），亦可探讨导航动机与视觉线索相对可靠性对多模态信息整合发生阶段及其策略的共同影响以及这些因素之间的交互作用。

此外，对于多模态信息的耦合特征，以往研究仍存在双向耦合假说和单向耦合假说的争议(Du et al., 2020; Zhao & Warren, 2015; Tcheang et al., 2011)，这可能时因为大多数研究都侧重于建立视觉增益后信息耦合的结果，而忽视了建立视觉增益期间信息耦合的过程。未来研究可以通过在视觉增益训练期间，每经过几次训练就施加一个探测刺激（仅提供视觉线索或自我感知运动线索）让被试完成指向任务。如果在视觉增益训练期间仅提供视觉线索或自我感知运动线索的条件下被试指向表现均受视觉增益影响，则说明耦合是双向的；反之，如果仅呈现视觉线索条件下被试指向表现不受视觉增益影响，而仅呈现自我感知运动线索条件下被试指向表现受视觉增益影响，则说明耦合是自我感知运动表征向视觉表征的单向耦合。此外，还以通过与视觉增益训练后、自我感知运动增益训练期间及训练后的耦合模式对比，探讨耦合关系的稳定性以及耦合强度的不对称性问题。

最后，对于多模态信息的记忆表征，以往研究更多关注工作记忆建立和维持阶段的多模态信息整合，很少关注工作记忆与长时记忆动态信息交互的阶段是否存在多模态信息整合(Etienne & Jeffery, 2004; Wang, 2017; Yamamoto & Shelton, 2005; Huffman & Ekstrom, 2021; Sholl et al., 2006)。然而在熟悉环境中的空间导航时常涉及工作记忆与长时记忆的交互，尤其是需要重置路径整合误差时，从长时记忆中提取相应的自我感知运动表征与工作记忆中的视觉表征进行整合是最高效的方式。因此，未来研究可以让被试学习包含稳定地标的虚拟导航路线，从而构建多模态的长时记忆空间表征，随后通过无地标路径整合任务干扰工作记忆表征，最后在视觉线索与自我感知运动线索匹配、不匹配或无视觉线索条件下分别要求被试完成指向任务。如果线索匹配条件相比其他条件能够显著降低指向误差，则表明视觉线索能够激活长时记忆中自我感知运动表征，并在工作记忆与长时记忆信息交互的过程中实现视觉表征与自我感知运动表征的整合，从而重置路径整合的误差。”

不过，如果仅给出上述研究思路作为“总结与展望”部分的内容过于单薄，它们可以解

决以往研究的一些争议，但无法完全代表未来研究的热门趋势，毕竟这些研究方向已经对应了相对具体问题，很难再延伸出一个与方向感的多模态信息整合认知神经机制相关的综合性研究体系。因此，我们决定在给出上述针对具体问题提出的研究思路作为未来研究方向的基础上，保留此前提出的三个未来研究方向。一方面，是因为这三个未来研究方向均能延伸出一个综合性的研究体系；另一方面，这三个研究方向与上述根据具体问题提出的研究方向相结合，能够让“总结与展望”部分的内容更具有层次感，既能聚焦于当下丞待解决的问题，又能给出符合未来研究热门趋势的方向。

此外，这三个研究并非完全与综述部分的内容脱节，都是从不同角度综述大量前人研究的基础上得出的理论意义和应用价值兼具的研究方向。其中，通过综述发现以往研究主要集中于视觉线索和自我感知运动线索的整合，忽略了听觉、嗅觉与触觉等感官信息在多模态信息整合中的作用，基于此我们提出未来应当拓展方向感研究中的多模态信息整合范畴，可以通过引入多感官冲突或协同的情境，系统检验个体如何赋权并整合各类感官信息，以构建更具生态效度的多模态信息整合模型。同时，通过探讨上丘-颞上回、嗅觉-嗅皮层、体感皮层与方向感的多模态信息整合神经网络的激活存在何种关联，来推测是否存在新的神经信号传导通路，从而完善现有神经网络模型。而且，我们已经通过对大量认知神经科学领域研究的总结，给出了 LMN-DTN 环形吸引子网络、RSC-EC-HPC 通路及视觉皮层等关键神经结构之间的协作路径，但总体上结构仍较为松散，很难通过算法模型的形式应用到一些新技术中。因此，需要对其进行全面的形式化神经网络建模与可计算模拟，基于分布式神经活动、贝叶斯整合原则与耦合机制，逐步构建从自我运动整合、HD 信号生成到高级皮层深度整合的模块化计算模型。基于此我们提出未来研究应当考虑构建方向感有关的多模态信息整合神经计算模型，如果能构建出这一模型，不仅能解释方向感建构的动态过程，验证认知层面的多模态信息整合模型，也能为航空航天、自动驾驶等领域的空间导航智能系统提供理论基础。最后，我们通过综述发现，以往研究多集中在视觉线索、自我感知运动线索如何影响个人方向感及导航行为，然而现实导航环境中人们有着从他人或导航软件中获取空间信息以建立方向感的需求，但个体如何将方向感以及相应的心理表征，特别是在协作导航任务中对路径整合的心理表征，以空间语言的形式传递给合作者，这其中的空间言语表述机制尚不明确，方向感的好坏对表述机制的影响也未可知。基于此我们认为探讨个体对方向感的空间言语表述机制，尤其是涉及多模态信息整合的表述机制，是理论意义和应用价值兼具的研究方向。而且，我们也修改了摘要中涉及未来研究展望的内容，即增加了一句概括性的内容：“未来研究应系统考察方向感的多模态信息动态整合机制”（正文 p1）。

参考文献

Sjolund, L. A., Kelly, J. W., & McNamara, T. P. (2018). Optimal combination of environmental cues and path integration during navigation. *Memory & Cognition*, *46*, 89-99.

Zhao, M., & Warren, W. H. (2015). How you get there from here: Interaction of visual landmarks and path integration in human navigation. *Psychological Science*, *26*(6), 915–924.

Shettleworth, S. J., & Sutton, J. E. (2005). Multiple systems for spatial learning: Dead reckoning and beacon homing in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *31*(2), 125–141.

Kessler, F., Frankenstein, J., & Rothkopf, C. A. (2024). Human navigation strategies and their errors result from dynamic interactions of spatial uncertainties. *Nature Communications*, *15*(1), 5677.

Kalia, A. A., Schrater, P. R., & Legge, G. E. (2013). Combining path integration and remembered landmarks when navigating without vision. *PLoS ONE, 8*(9), e72170.

Newman, P. M., & McNamara, T. P. (2020). A comparison of methods of assessing cue combination during navigation. *Behavior Research Methods*, *53*, 390-398.

Zhang, L., & Mou, W. (2017). Piloting systems reset path integration systems during position estimation. *Journal*

of Experimental Psychology: Learning, Memory, and Cognition, *43*(3), 472.

Zhang, L., Mou, W., Lei, X., & Du, Y. (2020). Cue combination used to update the navigator's self-localization, not the home location. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(12), 2314.

Chen, Y., & Mou, W. (2025). Disrupted orientation after path integration by absence of anticipated prevalent spatial views. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Du, Y., Mou, W., & Zhang, L. (2020). Unidirectional influence of vision on locomotion in multimodal spatial representations acquired from navigation. *Psychological Research*, *84*(5), 1284-1303.

Zhao, M., & Warren, W. H. (2015). How you get there from here: Interaction of visual landmarks and path integration in human navigation. *Psychological Science*, *26*(6), 915–924.

Tcheang, L., Bülthoff, H. H., & Burgess, N. (2011). Visual influence on path integration in darkness indicates a multimodal representation of large-scale space. *Proceedings of the National Academy of Sciences*, *108*(3), 1152-1157.

Etienne, A. S., & Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, *14*(2), 180-192.

Wang, R. F. (2017). Spatial updating and common misinterpretations of spatial reference frames. *Spatial Cognition and Computation*, *17*(3), 222–249.

Yamamoto, N., & Shelton, A. L. (2005). Visual and proprioceptive representations in spatial memory. *Memory & Cognition*, *33*(1), 140-150.

Huffman, D. J., & Ekstrom, A. D. (2021). An important step toward understanding the role of body-based cues on human spatial memory for large-scale environments. *Journal of Cognitive Neuroscience*, *33*(2), 167-179.

Sholl, M. J., Kenny, R. J., & DellaPorta, K. A. (2006). Allocentric-heading recall and its relation to self-reported sense-of-direction. Journal of Experimental Psychology: Learning, Memory, and *Cognition*, *32*(3), 516-533.

第二轮

审稿人 1 意见：

作者根据我的意见做了认真修改，我没有其他问题。

回应：非常感谢审稿专家对我们修改工作的认可。

审稿人 2 意见：

作者的回复和修改较好回答了上一轮提出的问题。如果可能，目前版本还可以再进一步细化论述或未来研究方向。

意见 1：行为实验发现了“地标(landmark)、边界(boundary)、环境的几何结构(geometric layout)等视觉线索获取信息能帮助个体在空间导航中确定空间参照系、估算距离以及感知方向变化”，那么各类线索在神经机制上是怎么体现的？目前的版本还可以进一步详细论述，建议在第 3 或第 4 部分有所体现。

回应：非常感谢审稿专家对我们此前修改工作的认可以及本次给出的进一步的修改意见。我们非常赞同审稿专家的观点，应当在认知神经机制部分（第 3 部分）详细论述“个体利用地标、边界以及环境的几何结构等视觉线索展开空间导航”的神经机制是怎样的。事实上，目前这些内容在文中均有体现，只是由于对不同视觉线索的处理涉及不同的细胞、脑区，这些内容被划分到了不同的章节内，下面是一些示例。

（1）地标是 HD 信号的锚定源，缺少地标 HD 信号就会发生漂移，对应行为层面路径整合的误差积累，HD 信号的校准对应着个体利用地标重置路径整合的误差，文中对应的具体内容：

“环形吸引子网络模型(ring attractor networks)常用于解释 HD 细胞如何基于自我运动信号和视觉信号建立并更新 HD 信号(Sit & Goard, 2023; Long et al., 2024; Turmer-Evans et al., 2020; Khona & Fiete, 2022)。根据该模型的假设，多个区域的 HD 细胞会组成一个建立、维持和更新 HD 信号的环形网络，通过对自我运动信号和视觉信号的整合产生 HD 信号，并动态调整 HD 信号以适应当前方向。环形网络激活后不需要外部持续刺激即可维持活动，并在环形网络内形成“活动丘”(activity hill)，“活动丘”会根据整合后的信息移动到环形网络的不同位置从而更新 HD 信号(Peyrache et al., 2015)。HD 信号由两部分构成：基于早期前庭通路整合的自我运动信号及其锚定源——从视觉皮层、RSC、后海马下托(postsubiculum, PoS)等脑区接收的视觉信号。视觉信号的作用是将 HD 信号锚定到特定的参考框架(Siegenthaler et al., 2024)，如果环境中缺少地标等视觉线索或视觉皮层等传递视觉信号的脑区受损，HD 信号就会发生漂移，对应着认知层面空间导航中缺少视觉线索时路径整合误差的迅速积累(Bjerknes et al., 2015; Ajabi et al., 2023)。”

（2）压后皮质中的地标向量细胞能够在处理 HD 信号时整合视觉信号（地标信息）为以自我为中心的方向表征提供参照点，自我边界向量细胞则能在此基础上编码自我到边界的距离，从而建立以自我为中心的方位表征，文中对应的具体内容：

“RSC 处理 HD 信号时能够整合视觉信号为以自我为中心的方向表征提供参照点，EBV 细胞则能在此基础上编码自我到边界的距离，从而建立以自我为中心的方向和位置表征(Alexander et al., 2020)。其次，RSC 与 HPC、EC、海马下托复合体(subicular complex)之间存在紧密联系，后三者联合区域又称为“扩展海马体”，该区域包含环境中心边界向量细胞(allocentric boundary vector cells, ABV 细胞)，对应地能够将 HD 信号和视觉信号结合边界信息建立以环境为中心的方向和位置表征并将信号传递至 RSC，从而与 EBV 细胞建立的表征相互映射(Lever et al., 2009; Bicanski & Burgess, 2018)。最后，基于这种映射，个体可以根据需求实现空间参照系的转换：在感知驱动模式下，对客体的自我中心视角可被用于旋转并对齐现有的环境参照表征，而在记忆等由内在信息驱动的模式下，对客体的环境参照表征或空间记忆会被转换到个体当前对客体的自我中心视角(Alexander et al., 2023)，从而帮助个体在不同视角或尺度上获取方向感。”

（3）环境的几何结构通常指空间整体的形状、比例、对称性及各边界之间的角度关系，个体处在空间环境的特殊位置，网格细胞的放电会增强，此外对于某些对称的空间环境，某些 HD 细胞的放电机制会发生变化，有利于个体编码以环境为中心的空间信息，具体内容如下：

“研究发现，EC 中的网格细胞在个体处于环境中由等边三角形或六边形组成的网格节点时会增强放电，且这种放电模式受 HD 信号调谐(Gardner et al., 2022; Giocomo et al., 2014; Sargolini et al., 2006)。如果 ADN 功能受损导致 HD 信号中断，会改变网格细胞的周期性放电模式，而且网格细胞的活动水平会根据动物的移动路径、方向或与环境边界的距离进行更新(Burgess, 2008b; Hafting et al., 2005; Winter et al., 2015)，所以可以推测网格细胞编码以环境为中心的位置信息时，会参照以自我为中心的感觉输入或路径整合。例如，当视觉线索在边界的位置发生变化时，网格细胞的放电场随之发生变化(Savelli et al., 2017; Moon et al., 2022)；动物沿线性轨迹移动时，深层 EC 网格细胞的 θ 波相位会发生前移(Hafting et al., 2008)；动物靠近环境边界时网格细胞的放电规模会被压缩且表现出更准确的位置编码，反之远离边界时准确性会降低(Munn et al., 2020; Hardcastle et al., 2015)。而且，EC 中也包含了能以自我中心视角编码个体相对于边界距离和方向的 EBV 细胞(Solstad et al., 2008; Wang et al., 2020)，

因此可以推测 EBV 细胞与网格细胞之间的信号交互可能就是后者根据以自我为中心的感觉输入编码以环境为中心精确坐标位置的神经基础。另外，在内侧内嗅皮层(medial entorhinal cortex, MEC)中还发现了联合网格细胞(conjunctive grid cells)，其单向放电模式区别于普通网格细胞的全向放电，说明这类网格细胞能够独自整合方向和位置信息(Gerlei et al., 2020)。总之，以往研究直接或间接表明 EC 能够整合 HD 信号、EBV 细胞编码自我视角下相对于边界的距离和方向信息后以及网格细胞编码以环境为中心位置信息后产生的神经信号，可以帮助个体在不同尺度上建立和维持精确的方向感，并为构建认知地图提供度量信息(Jeffery & Burgess, 2006; Jeffery, 2024; Long et al., 2024)。”

总之，类似地内容文中还有许多，都是围绕着“不同脑区的特异性细胞对各类视觉线索、自我感知运动线索进行神经编码和信号整合”展开论述的，与行为层面的认知加工过程相呼应。

意见 2：贝叶斯最优整合的神经机制是什么？在哪里进行的整合？第 3 部分提到了一些神经位置是存在信息整合的，那么这些位置的整合都是贝叶斯方式的吗？或许可以在第 3 或第 4 部分论述一二。

回应：非常感谢审稿专家指出应当在第 3 或第 4 部分对贝叶斯最优整合的神经机制阐述一二。我们非常赞同审稿专家的观点，在第 3 或第 4 部分进一步阐述各脑区或细胞群对多模态线索是否存在贝叶斯最优整合式的神经编码，将有利于我们进一步理解方向感的多模态信息整合认知神经机制。

与心理物理学中行为层面的个体对多模态信息的贝叶斯最优整合相似，认知神经科学领域的贝叶斯假说认为大脑在多模态信息进行整合编码时会依据各感官线索的可靠性进行加权。具体而言，感官输入被转化为概率分布，大脑计算每种输入的似然函数，并结合先验信息形成后验分布，该后验反映对当前状态的最优估计(Angelaki et al., 2009; Friston, 2010)。这一过程的实现依赖于细胞群或神经元群体对信息的概率性响应或放电（放电率不仅表示刺激最可能的值，也表示刺激的不确定性; Ma et al., 2006），信息的可靠性高能够增益神经响应，不同群体神经元的响应分布可以被视为似然函数，大脑将不同细胞群体活动分布加权，最终产生后验分布的神经表征(Fetsch et al., 2012)。此外，贝叶斯整合并不是由某一脑区实现，而是通过一类神经计算规则（增益、权重调整、预测编码）在多个脑区网络中实现。根据层级预测框架，大脑中的贝叶斯整合是分布式的，根据所需要加工的信息发生在不同脑区、细胞或神经元群体，但并非所有脑区都存在贝叶斯最优整合。目前，与视觉线索和自我感知运动线索整合相关的神经机制研究较少（或有实质性发现的研究较少），其中只有少数研究发现猕猴的背内侧上颞区(dorsal medial superior temporal area, MSTd)的一些神经元对自我感知运动线索和视觉线索的神经编码符合贝叶斯最优整合规律(Gu et al., 2008; Fetsch et al., 2012)，也有少数研究发现腹侧顶叶皮层(ventral intraparietal cortex, VIP)与 MSTd 对视觉线索和自我感知运动线索的神经编码机制相似，但更倾向于赋予自我感知运动线索更高的权重(Chen et al., 2011)。此外，也有部分研究从宏观行为层面模拟和推断出大脑对视觉线索和自我感知运动线索编码的神经机制符合贝叶斯最优整合模型，但没有指定相应脑区(Dokka et al., 2015)。

由于这部分内容不足以独立成节，而且第 3 部分的论述是以“前庭-视觉”通路中 HD 信号产生与投射网络的为核心展开的，因此也很难在第 3 部分找到合适的“位置”插入这部分内容。基于此，我们将这部分内容放在了第 4 部分，还有一个原因是在第 4 部分的开始我们在以往研究发现“多模态信息贝叶斯整合”的基础上提出了“动态贝叶斯”的未来研究方向。所以，我们认为在此处也可以对“以往研究对视觉线索和自我感知运动线索的贝叶斯整合神经机制”的研究结果做一个总结，而后再指出未来可以探讨这些脑区或神经机制是否支持“动态贝叶斯整合”，这样也会让论述更有逻辑性，具体内容如下（正文 p20）：

“此外，以往研究发现猕猴背内侧上颞区(dorsal medial superior temporal area, MSTd)的一些神经元对视觉线索和自我感知运动线索的神经编码符合贝叶斯最优整合模型(Gu et al., 2008; Fetsch et al., 2012)，而且腹内侧顶叶皮层(ventral intraparietal cortex, VIP)对多模态线索的神经编码机制与之相似，但腹内侧顶叶皮层更倾向于赋予自我感知运动线索更高的权重(Chen et al., 2011)。未来研究可以结合“动态贝叶斯整合”的行为范式，探讨个体在根据多模态线索可靠性的动态变化对它们进行整合时，这些脑区的神经元是否相应地表现出对多模态信息的动态贝叶斯神经编码模式。”

参考文献

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology*, *19*(4), 452-458.

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nature Reviews Neuroscience*, *11*(2), 127-138.

Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, *9*(11), 1432-1438.

Dokka, K., DeAngelis, G. C., & Angelaki, D. E. (2015). Multisensory integration of visual and vestibular signals improves heading discrimination in the presence of a moving object. *Journal of Neuroscience, 35*(40), 13599-13607.

增引文献

Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011). Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *Journal of Neuroscience, 31*(33), 12036-12052.

Fetsch, C. R., Pouget, A., DeAngelis, G. C., & Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience*, *15*(1), 146-154.

Gu, Y., Angelaki, D. E., & DeAngelis, G. C. (2008). Neural correlates of multisensory cue integration in macaque MSTd. *Nature Neuroscience*, *11*(10), 1201-1210

第三轮

编委 1 意见：同意发表。

编委 2 意见：同意发表。

主编意见：稿件经过多位专家的审阅，作者进行了认真的修改，达到了发表水平，同意发表。