

《心理科学进展》审稿意见与作者回应

题目：安全学习与焦虑障碍：三阶段神经机制框架

作者：董展鹏；张婕；张亚蕊；雷怡

第一轮

编委 1 初审意见：

本文从安全学习的角度探讨焦虑障碍的神经机制，视角较为新颖。然而，目前提出的框架与既有的负性情绪(威胁)加工理论在本质上差异有限，显得与传统的模型高度重合。

回应：感谢编委的审阅与意见。我们认同您所指出的，本研究提出的“三阶段”框架在结构上与既有的恐惧学习模型确实存在一定相似性。这种相似主要源于二者在逻辑框架上的延续性：无论是恐惧还是安全学习，都可以被分解为信息获取、联结形成及行为调控等基本过程，因此框架层面难免存在一定的平行结构。但安全学习在范式和机制上均展现出区别于恐惧学习的特异性，本文的独特之处在于系统性地强调了“安全学习的缺陷”作为焦虑障碍的核心机制，这一点是传统模型较少系统阐述的。因此系统整合了“安全学习”视角下的神经机制证据，不仅帮助解释了威胁加工模型难以完全涵盖的现象(如安全辨别困难、正性情绪缺失及恐惧抑制无效)，也对新型干预策略(例如增强安全信号学习)具有启示意义。为突显与传统威胁学习的差异性，在文中进行了下述补充：

①首先，我们在引言部分对恐惧学习视角的局限与安全学习的独立机制进行了补充，并在2.1部分对安全学习中的“安全”的概念进行了阐述，强调其并非恐惧学习中单纯的“威胁缺失”，而是一个独立且具有积极价值的适应性过程，补充内容如下：

引言：

尽管识别安全信号的能力对适应性行为至关重要，但传统上对焦虑障碍神经机制的理解主要源于恐惧学习研究。该类研究认为焦虑障碍源于恐惧反应的过度泛化或持续存在，并在神经机制层面主要聚焦于杏仁核(amygdala)、背侧前扣带回(dorsal anterior cingulate cortex, dACC)、岛叶(insula)等在恐惧表达中过度活跃的区域(Graham & Milad, 2011; Sevenster et al.,

2018, 雷怡等,2018; 申思怡等, 2023), 并为暴露疗法等临床干预提供了重要理论依据。然而, 单一依赖“恐惧中心”视角也存在局限, 难以完全解释焦虑障碍中常见的安全辨别困难、正性情绪缺失及恐惧抑制无效等现象, 因为它忽视了个体主动习得并利用安全信号抑制恐惧的适应性机制。

近年来, 研究者逐渐认识到安全学习具有独立的神经机制, 而非简单的恐惧对照。Pollak 等人(2008)的小鼠研究首次表明, 习得性安全(主动学习到的安全信号)不仅能抑制条件性恐惧(冻结反应), 还能在升高十字迷宫实验中显著减少回避行为, 表现为对先天性焦虑的缓解。同时在神经机制层面促进海马新生神经元存活并上调 BDNF, 从而引发类似服用氟西汀(常用抗抑郁药)相当的神经可塑性过程, 提示安全学习可作为一种行为性抗抑郁模式; 随后的研究认为, 安全学习是一种与恐惧学习这一兴奋性过程相对的抑制性学习机制, 即个体通过学习 CS 预测“无威胁”, 从而主动抑制恐惧反应(Christianson et al., 2012; Kong et al., 2014); Sangha 等人(2020)认为安全学习与恐惧学习同样重要, 都是适应性行为所必需的。并强调不同任务程序(辨别、消退和主动回避)可能影响对安全线索的解释。此外, 虽然威胁和安全学习都依赖杏仁核-前额叶-海马体回路, 但激活的亚区不同, 例如小鼠实验中 PL 更偏向恐惧表达与辨别, IL 更偏向恐惧抑制与安全线索处理。以上证据表明, 安全学习并非恐惧学习的简单对立, 而是依赖特定神经机制的相对独立过程……

2.1 安全的定义

在威胁学习视角下, “安全”通常被界定为威胁结果的缺失。而在安全学习视角下, “安全”被视为一种更为主动和相对独立的过程, 重点关注“刺激-安全”联结。Tashjian 等人(2021)认为, 安全并非单纯是威胁的缺失, 而是个体对当前与未来生存可能性的感知, 存在于从“零安全”到“完全安全”的连续体上。安全可以表现为情境安全(如庇护所中)、消退性安全(原本危险刺激被学习为安全)、安全信号(威胁被外部因素减缓)以及前瞻性安全(通过行为预期未来安全)。这些过程不仅体现在恐惧消退过程中对威胁记忆的削弱, 还包括安全辨别、安全信号学习等对安全刺激本身学习的任务。基于上述认识, 下面将梳理安全学习的常见实验范式, 以揭示其在情绪调节与适应中的独特作用。

②此外, 我们会还进一步补充了用于建立“刺激 - 安全”联结的安全学习范式, 以突出安全学习在实验设计方法上的独立价值与独特贡献:

2.2 安全学习的研究范式

……

除此之外，还存在一些涉及将线索与安全结果相关联的实验范式。例如，主动回避任务中(Kleine et al., 2023; Fisher & Urcelay, 2024)，个体可以通过学习特定的行为规避威胁，并且成功回避后能体会到一种“解脱感”；在后向恐惧条件任务(Andreatta et al., 2010; Gründahl et al., 2022)中，威胁刺激 US 出现在 CS 之前，CS 代表 US 终结后随之而来的安全期。安全评估研究中，探究了个体在威胁情境中如何综合外部威胁线索与自我保护信息，从而对未来安全性进行动态预测(Tashjian et al., 2025)。这些范式进一步丰富了安全学习的研究视角，表明安全学习不仅能抑制恐惧，还能产生积极的情绪和动机价值以指导后续行为。由此可见，学习“安全”并非单一的恐惧抑制过程，而是一个复杂的包含认知评估，动机驱动以及情绪调节的适应性机制。

③对于神经机制框架，我们作出了相应修改，补充了与安全学习更为相关的研究，并进一步梳理了与安全信号加工更为相关的脑区与亚区：

1) 在感知阶段的修改中，我们认为安全信号的加工更依赖自上而下的认知调控，而非威胁加工中常被提出的丘脑—杏仁核快速通道，因此安全学习并不存在同类的“快速通路”。同时，我们补充了岛叶、vmPFC 等在安全信号中的特异性作用，以区别于恐惧加工主要依赖的丘脑—杏仁核回路，从而突出了安全学习在涉及脑区和加工方式上的独特性。

另外，虽然安全与威胁信号在感知阶段涉及的回路相似，但在具体脑区的亚区层面上存在差异，例如：

“啮齿类动物研究表明，在区分 CS+(威胁线索)与 CS-(安全线索)时尽管都依赖杏仁核与前额叶，但在不同亚区上存在功能分化……这些发现表明，安全识别不仅依赖情绪通路，还涉及前额叶对刺激意义的认知性评估，且与威胁识别存在差异。同时，最近的一项人类多体素研究(Tashjian et al., 2025)进一步区分了 vmPFC 亚区在威胁和安全评估上的差异……可见在对安全的感知评估中，前 vmPFC 对安全的主观表征的关键区域。”

2) 在学习阶段的修改中，我们强调中脑—纹状体回存在功能性分工：VS 更侧重于安全价值和积极情感的形成，DS 则更多参与规则转换与“无威胁”情境下的更新。安全学习的习得不仅是对威胁信息的重新评估，而是涉及安全价值及其伴随的积极情感形成，因此在机制上具有独特性。

3) 在表达阶段，我们认为安全学习的表达不仅是“恐惧记忆的削弱”，更是“安全记忆的

提取与利用。根据安全类型的不同区分了两条回路：其一是 aHC-vmPFC-杏仁核 回路，主要支持消退性安全学习；其二是 aHC-dACC 回路，参与安全信号学习，并在儿童与青少年阶段发挥核心作用。通过强调不同学习类型和发展阶段的差异，我们突出了安全学习在神经机制上的独立性与特异性。

.....

编委 2 意见：

当前综述结合了大量近年来的实证研究，并对 Laing(2022)的框架进行了扩展式阐释，有一定的创新性，但其主要贡献在于概念性细化与动态过程的系统化整理，且对神经机制的解读具有过度因果推论性，存在一定的局限性。另外，摘要后半段的总结，以及文中对于科学问题的凝练存在不足，建议作者修改相关说法后送审。

回应：

1. 感谢编委的审阅与意见。针对您提出的“对神经机制的解读具有过度因果推论性，存在一定的局限性”的问题，我们已对神经机制的解读进行了修改，以避免不当推断，并补充相应的文献确保推理的合理性。其中文中的橙色字体为修改内容，蓝色字体为观点小结。例如：安全的感知评估是个体在接触到外部刺激时，判断该刺激是否不具威胁性的过程，丘脑—杏仁核感觉通路在此过程的加工中具有重要作用。早期啮齿类动物研究发现，恐惧学习主要涉及丘脑的内侧膝状体(medial geniculate body, MGB)(Romanski & LeDoux, 1992).....而 Heldt 和 Falls(2006)在小鼠的条件抑制任务中，探讨了这一路径对安全信号的加工。结果发现，单独损毁在恐惧条件化起重要作用的内侧膝状体(medial geniculate body, MGB)或听觉皮层不影响安全信号对恐惧的抑制作用，但损毁更大范围的听觉丘脑(包括 MGB+周围核团)会破坏安全信号的作用；(小结)以上结果表明，安全与威胁信息的感知加工虽然依赖相似的感知通路，但二者在具体的神经结构上可能存在差异。

.....

2. 针对您提出的“摘要后半段的总结，以及文中对于科学问题的凝练存在不足”的问题，进行了相应修改。本文的科学问题为“基于安全学习三阶段神经机制框架，从感知评估、习得和行为表达三个阶段阐明焦虑及其相关障碍个体的异常。”以此帮助解释传统威胁加工模型难以完全涵盖的现象，同时探讨基于安全学习未来如何更好地临床干预。

- 1) 因此在摘要中,我们首先强调了安全学习在焦虑个体中的重要性及现有不足,接下来阐述基于系统整合“安全学习”视角下的神经机制研究的结论,最后展望未来研究如何更好地推进临床干预。修改后如下:

摘要:安全学习是指个体通过识别安全信号(预示无危险的线索)来抑制恐惧的适应性机制,与焦虑及其相关障碍的发生发展密切相关。基于现有研究,本研究提出安全学习三阶段神经机制框架阐释焦虑发生发展的潜在机制。研究发现,安全的感知评估依赖丘脑-皮层-岛叶-杏仁核回路,安全的习得由中脑-纹状体和前海马-前额叶回路支持并伴随积极情感体验,安全的行为表达则通过海马-前额叶-杏仁核及海马-前扣带皮层回路实现安全记忆提取与恐惧抑制。并且焦虑及其相关个体在安全学习中存在阶段性神经机制异常。未来研究应通过更高生态效度的实验范式与多模态神经成像技术,检验三阶段框架的有效性,并探索安全学习不同阶段的治疗神经靶点,以推动安全学习机制在焦虑障碍干预中的临床转化。

- 2) 文本中科学问题的提出,放在引言部分最后一段,通过强调安全学习的重要性及神经机制区别与恐惧学习,从而提出本文的研究问题“焦虑及其相关障碍个体的安全学习缺陷究竟体现在哪些神经机制环节?”。修改后如下:

综上所述,安全学习并非仅仅是恐惧学习的反面,而是揭示了一条独立的适应性行为的神经通路,探究安全学习的神经机制对于理解焦虑及其相关障碍具有重要意义。然而,焦虑及其相关障碍个体的安全学习缺陷究竟体现在哪些神经机制环节?为此,本文在梳理既有研究的基础上,提出“感知-习得-表达”三阶段安全学习神经机制框架,以更全面揭示安全学习的过程及焦虑个体在各阶段的异常特点,从而为焦虑障碍的精准干预和机制研究提供新的思路。

新增参考文献(质检报告已更新):

- 曹杨婧文,李俊娇,陈伟,杨勇,胡琰健,郑希付. (2019). 条件性恐惧记忆消退的提取干预范式及其作用的神经机制. *心理科学进展*, 27(2), 268-277.
- 雷怡,梅颖,张文海,李红. (2018). 基于知觉的恐惧泛化的认知神经机制. *心理科学进展*, 26(8), 1391-1403.
- 申思怡,梅颖,王金霞,戴雨芊,吴奇,雷怡. (2023). 条件学习视角下的恐惧与厌恶*[J]. *心理科学* 46(4), 825-832.
- Anderson, M. C., & Floresco, S. B. (2022). Prefrontal-hippocampal interactions supporting the extinction of emotional memories: The retrieval stopping model. *Neuropsychopharmacology*, 47(1), 180–195.
- Andreatta, M., Mühlberger, A., Yarali, A., Gerber, B., & Pauli, P. (2010). A rift between implicit and explicit conditioned valence in human pain relief learning. *Proceedings. Biological Sciences*, 277(1692),

2411–2416.

- Atlas, L. Y., Doll, B. B., Li, J., Daw, N. D., & Phelps, E. A. (2016). Instructed knowledge shapes feedback-driven aversive learning in striatum and orbitofrontal cortex, but not the amygdala. *eLife*, 5, e15192.
- Bezmaternykh, D. D., Melnikov, M. Ye., Petrovskiy, E. D., Mazhirina, K. G., Savelov, A. A., Shtark, M. B., Vuilleumier, P., & Koush, Y. (2025). Attenuation processes in positive social emotion upregulation: Disentangling functional role of ventrolateral prefrontal cortex. *iScience*, 28(2), 111909.
- Fisher, C. T. L., & Urcelay, G. P. (2024). Safety signals reinforce instrumental avoidance in humans. *Learning & Memory*, 31(8), a053914.
- Grill, F., Guitart-Masip, M., Johansson, J., Stiernman, L., Axelsson, J., Nyberg, L., & Rieckmann, A. (2024). Dopamine release in human associative striatum during reversal learning. *Nature Communications*, 15(1), 59.
- Gründahl, M., Retzlaff, L., Herrmann, M. J., Hein, G., & Andreatta, M. (2022). The skin conductance response indicating pain relief is independent of self or social influence on pain. *Psychophysiology*, 59(3), e13978.
- Heathcote, L. C., Timmers, I., Kronman, C. A., Mahmud, F., Hernandez, J. M., Bentley, J., Youssef, A. M., Pine, D. S., Borsook, D., & Simons, L. E. (2020). Brain signatures of threat-safety discrimination in adolescent chronic pain. *Pain*, 161(3), 630–640.
- Kleine, R. A. D., Hutschemaekers, M. H. M., Hendriks, G. J., Kampman, M., Papalini, S., Minnen, A. V., & Vervliet, B. (2023). Impaired action-safety learning and excessive relief during avoidance in patients with anxiety disorders. *Journal of Anxiety Disorders*, 96, 102698.
- Milad, M. R., & Quirk, G. J. (2012). Fear Extinction as a Model for Translational Neuroscience: Ten Years of Progress. *Annual Review of Psychology*, 63(1), 129–151.
- Pattwell, S. S., Liston, C., Jing, D., Ninan, I., Yang, R. R., Witztum, J., Murdock, M. H., Dincheva, I., Bath, K. G., Casey, B. J., Deisseroth, K., & Lee, F. S. (2016). Dynamic changes in neural circuitry during adolescence are associated with persistent attenuation of fear memories. *Nature Communications*, 7, 11475.
- Pessoa, L. (2017). A Network Model of the Emotional Brain. *Trends in Cognitive Sciences*, 21(5), 357–371.
- Pollak, D. D., Monje, F. J., Zuckerman, L., Denny, C. A., Drew, M. R., & Kandel, E. R. (2008). AN ANIMAL MODEL OF A BEHAVIORAL INTERVENTION FOR DEPRESSION. *Neuron*, 60(1), Article 1.
- Tashjian, S. M., Zbozinek, T. D., & Mobbs, D. (2021). A Decision Architecture for Safety Computations. *Trends in Cognitive Sciences*, 25(5), 342–354.
- Widegren, E., Vegelius, J., Frick, M. A., Roy, A. A., Möller, S., Kleberg, J. L., Hoppe, J. M., Hjorth, O., Fällmar, D., Pine, D. S., Brocki, K., Gingnell, M., & Frick, A. (2025). Fear extinction retention in children, adolescents, and adults. *Developmental Cognitive Neuroscience*, 71, 101509.
- Wiemer, J., Leimeister, F., Gamer, M., & Pauli, P. (2024). The ventromedial prefrontal cortex in response to threat omission is associated with subsequent explicit safety memory. *Scientific Reports*, 14(1), 7378.

文献比例更新后如下：

		文献年代		
		2021~2025	2016~2020	2016 年以前
文献类型	理论性综述	19(45.2%)	14(33.3%)	9(21.4%)
	实证研究报告	35(61.4%)	11(19.3%)	11(19.3%)

第二轮

审稿人 1 意见：

本论文是一篇围绕安全学习的综述研究，扩展了现有关于安全学习神经机制的理论框架，将安全学习过程细化为安全的感知评估、安全的习得与安全的行为表达三个阶段，强调这是一个动态交互的过程，更系统地反映了安全学习神经活动的动态特性。同时，作者提出，安全学习并非恐惧学习的简单对立，而是依赖特定神经机制的相对独立过程。以下建议供作者参考：

意见 1：关于安全学习与经典恐惧条件反射范式中的恐惧消退(fear extinction)是否相同，建议进一步讨论这两个概念的异同。例如，恐惧消退通常指已建立的恐惧反应由于条件刺激不再伴随厌恶结果而逐渐减弱，而安全学习可能涉及更广泛的、主动建立安全信号的过程。此外，恐惧消退还包括下一阶段——恐惧消退回忆(recall)，那么安全学习是否也存在类似的安全学习回忆？建议对此进行探讨。

回应：感谢审稿人的细致审阅与宝贵的意见，我们将分成 2 点回答上述问题。

①关于安全学习和恐惧消退的异同，我们认识到在安全学习和恐惧消退的概念解释上本文存在不足。现有研究认为，恐惧消退(fear extinction)是一种抑制性学习过程，即个体学习到条件刺激(CS)不再预示厌恶性结果(US)，从而形成“CS→无 US(安全)”的新联结。因此，恐惧消退也被研究者看作是安全学习(safety learning)的一种形式，例如：

Kausche, F. M., Carsten, H. P., Sobania, K. M., & Riesel, A. (2025). Fear and safety learning in anxiety- and stress-related disorders: An updated meta-analysis. *Neuroscience & Biobehavioral Reviews*, 169, 105983.

但安全学习不完全等同于消退，还包括条件抑制、辨别性恐惧条件化、主动回避等能够带来安全结果的学习形式。因此我们在文中补充相应内容对两者关系进行说明，主要内容如

下：

1.在引言部分阐明恐惧消退是安全学习的一种形式：

尽管识别安全信号的能力对适应性行为至关重要，但传统上对于焦虑障碍神经机制的理解主要集中在恐惧学习研究。该类研究认为焦虑障碍源于恐惧反应的过度泛化或持续存在，且在神经机制层面主要聚焦于杏仁核(Amygdala)、背侧前扣带回(dorsal anterior cingulate cortex, dACC)、岛叶(Insula)等在恐惧表达中过度活跃的区域(Graham & Milad, 2011; Sevenster et al., 2018; 雷怡等, 2018; 申思怡等, 2023)，该类研究中常通过恐惧消退来研究人们是如何习得安全的，恐惧消退指将威胁化的条件刺激(conditional stimulus, CS)与“威胁缺失”(安全)建立新的联结，从而使得个体对威胁的预测进行更新和抑制恐惧反应，这一过程为暴露疗法等临床干预提供了重要理论依据。因此，恐惧消退可被视为安全学习的一种特殊形式。然而，单一依赖“恐惧中心”视角去理解人类如何获得安全可能存在一定局限，难以完全解释焦虑障碍中常见的安全辨别困难、正性情绪缺失及恐惧抑制无效等现象，因为它忽视了个体主动习得并利用安全信号抑制恐惧的适应性机制。

2.在 3.1 安全学习的定义部分，进一步讨论两者的关系，主要内容如下：

在恐惧学习视角下，“安全”并不携带明确的学习价值，通常被界定为“威胁”结果的对照。而在安全学习视角下，“安全”的建立被视为一种积极的学习过程，具有多种表现形式。Tashjian 等人(2021)认为，安全并非单纯是威胁的缺失，而是个体对当前与未来生存可能性的感知，存在于从“零安全”到“完全安全”的连续体上。安全可以表现为情境安全(如庇护所)、消退性安全(原本危险刺激被学习为安全)、安全信号(威胁被外部因素减缓)以及前瞻性安全(通过行为预期未来安全)。换言之，安全学习的核心表现在于线索与安全结果(如无厌恶刺激或正性结果)相联结。在现有研究中，这种线索-安全结果的联结学习不仅体现于恐惧消退过程中对原有威胁记忆的抑制，还包括安全辨别任务中对威胁与安全线索的主动区分，以及在条件性抑制和回避任务中，个体通过安全信号实现对潜在威胁结果的抑制或规避(Grasser & Jovanovic, 2021; Sangha et al., 2020)。基于上述定义，下面将梳理安全学习的常见实验范式，以揭示其在适应性行为中的独特作用。

②对于您所提出“安全学习是否也存在类似的安全学习回忆？”的问题，我们在文中主要做出以下两点补充：

1. 在本文中，我们认为习得安全后的行为表达依赖于安全记忆的顺利提取(回忆)，我们在 3.3 部分进一步修改，论述了“安全”的回忆可能涉及的神经机制，主要内容如下：

...相较于传统单变量方法，MVPA 能够更敏感地识别边缘系统和 vmPFC 等关键区域

在消退学习后的表征变化与记忆重现,从而为理解人类消退学习在回忆阶段的神经机制提供了更精细的证据。例如, Hennings 等人(2022)采用编码—提取相似性(encoding–retrieval similarity, ERS)分析, 分别标记个体在恐惧习得与消退阶段对刺激形成的神经表征, 并在 24 小时后的记忆测试中检测这些表征是否被重新激活。结果显示, 在恐惧消退 24 小时后, vmPFC 与 aHC 均显示对消退记忆的神经重现(neural reinstatement), 表明这两个脑区可能是消退学习后共同支持安全记忆的回忆关键区域; Bauer 等人(2025)的研究采用 ERS, 进一步检验了标准消退与反条件化(counterconditioning)消退(消退阶段 CS+与积极结果联结)的神经重现。在消退学习 24 小时后, 个体在 vmPFC 中对反条件化安全线索的神经表征出现了显著的重现, 而对标准消退刺激和威胁刺激未观察到相同效果。综上, vmPFC 可能是支持消退学习回忆的关键脑区, 而积极价值更强的安全记忆更容易在 vmPFC 被重新激活。

2. 通过梳理现有文献, 我们发现目前关于安全学习的回忆(recall)研究主要还是集中在消退学习, 而关于条件抑制等任务尚缺乏回忆阶段的研究。这可能是由于条件抑制范式涉及多重刺激的交互, 安全刺激需要与威胁刺激同时呈现以测量其抑制性, 在回忆阶段要确保安全记忆的独立测量较为困难, 我们针对这点在总结与展望部分的 **“5.1 加强条件抑制范式在焦虑个体安全学习中的应用”**对未来安全学习的回忆机制研究进行展望:

另一方面, 焦虑相关障碍个体通常存在安全学习后回忆阶段的异常。然而, 目前针对“安全记忆”的回忆研究仍主要集中于恐惧消退任务。相比之下, 条件抑制作为一种更主动、也更稳定的安全学习形式, 仍不清楚焦虑个体在其回忆过程中是否也存在功能受损, 或该机制是否可能弥补其在消退回忆中的缺陷, 对此需要未来研究进一步探讨。此外, 条件抑制在回忆阶段的研究之所以有限, 可能也与其范式本身的复杂性有关: 由于抑制性安全信号需要与威胁刺激在共现情境下发挥抑制效应, 因此在回忆阶段独立测量其记忆表征将更具挑战性。一项近期的神经生物学研究(Tashjian et al., 2025) 为探索抑制性安全学习的回忆机制提供了启示。该研究通过单独呈现保护线索和威胁线索, 让个体进行配对并联想“战斗场景”以评估威胁程度, 并在任务结束后再次单独呈现刺激, 测量个体对不同类型线索的神经表征。未来可在该范式基础上进行改进, 并结合 RSA 等方法, 比较焦虑个体与健康个体在条件抑制安全学习的回忆阶段是否也存在神经表征上的差异, 从而进一步揭示焦虑个体在安全学习回忆阶段异常的神经机制。

意见 2: 关于安全学习的三阶段划分, 有哪些证据表明这三个阶段彼此之间值得区分? 建议

单独说明每个阶段的独立性及其神经基础,例如不同脑区或环路在三个阶段中的特异性激活模式。

回应: 感谢审稿人的细致审阅与宝贵的意见。我们在文中“3 安全学习三阶段的神经机制框架”后补充说明了安全学习“感知-习得-表达”三阶段划分的依据: 首先, Laing 等人(2022a) 根据安全关联形成与行为表达过程在神经机制上的差异, 将安全学习区分为习得阶段和表达阶段。在此基础上, 我们通过梳理近年来研究发现, 在安全习得阶段也包含对积极情绪的加工和安全记忆的建立, 涉及海马、纹状体与前额皮层的加工。因此在本文中进行了补充。此外, 一系列研究表明个体在刺激呈现后的早期阶段便会对安全与威胁线索表现出不同的神经反应模式, 且在神经回路上区别于习得与表达过程。因此, 本文将安全学习的神经机制框架进行了调整, 并增加了感知阶段, 以进一步补充说明安全学习的神经机制。主要内容如下:

神经生物学研究表明, 安全学习是一个包含多阶段、多神经系统协同作用的复杂过程。Laing 等人(2022a) 根据安全联结的形成和之后的行为表达两个过程在神经机制上的差异, 将安全学习区分为习得阶段和表达阶段: 安全习得阶段是一个由预测误差驱动的学习过程, 该过程依赖于中脑-纹状体回路以支持安全关联的形成; 安全表达阶段更多体现习得安全结果后, 积极情感的获得和安全记忆对威胁反应的抑制, 该过程主要涉及前额叶与边缘系统之间的交互作用。但现有研究表明, 与积极情绪加工相关的多巴胺回路(Grill et al., 2024)以及 与记忆存储相关的海马-vmPFC 回路活动(Meyer et al., 2019) 同样参与安全的学习过程。换言之, 安全习得过程不仅涉及对威胁预测的修正, 也可能包括对积极情绪与安全记忆的加工。因此, 安全习得应被视为一个同时整合安全联结形成与积极情绪加工的过程。

除此之外, 研究表明, 个体在刺激呈现后的早期阶段便会对安全与威胁线索表现出不同的神经反应模式。来自事件相关电位(event-related potential, ERP)的研究显示, N1(约刺激出现后 100 ms)和 P2 振幅(约刺激出现后 200 ms)是反映刺激早期感知加工(Favero et al., 2023)的主要成分: N1 主要反映初级知觉加工, P2 则与刺激的初步意义评估有关。安全与威胁刺激在这两个早期 ERP 成分的振幅模式上均表现出显著差异。例如, 在后向掩蔽条件(刺激呈现后, 立即呈现一个掩蔽刺激将其覆盖, 使刺激难以进入意识层面)下, 相比于安全线索, 威胁线索能诱发更大的 N1(Mei et al., 2024); Dou 等人(2022)的研究进一步表明, 刺激的社会属性会调节安全的早期感知加工。研究发现, 相较于陌生人声音, 伴侣的声音作为安全线索时, 能显著降低个体的主观威胁感, 同时表现出更低的 P2 振幅。这些发现表明, 刺激在呈现的早期便受到差异化加工, 并且可能进一步影响后续的安全习得与表达。

另外, Grasser 和 Jovanovic(2021)在总结既有神经影像学研究的基础上指出, 个体对威胁与安全线索的感知能力与神经系统的成熟密切相关, 尤其依赖皮层下结构(如杏仁核、丘脑、基底节等)及其相关连接的发育。儿童期这些皮层下结构及其连接发育成熟, 从而支持儿童更有效地感知并区分环境中的威胁与安全线索。除此之外, Calhoon 和 Tye(2015)在焦虑神经环路模型中也提出, 威胁与安全线索的加工可被划分为“检测(Detection)”与“解释(Interpretation)”两个早期阶段。这两个阶段主要依赖丘脑、初级感觉皮层、岛叶以及杏仁核等区域, 对外界刺激的显著性与情绪意义进行初步分析和整合。焦虑可能使这两个阶段所涉及的神经回路出现异常活动。综上, 安全刺激的感知加工应被视为安全学习中一个独立且具有特异神经机制的环节。该阶段的异常可能影响后续的学习与表达过程, 并且与焦虑相关障碍的发生密切相关。

安全学习三阶段神经机制框架:

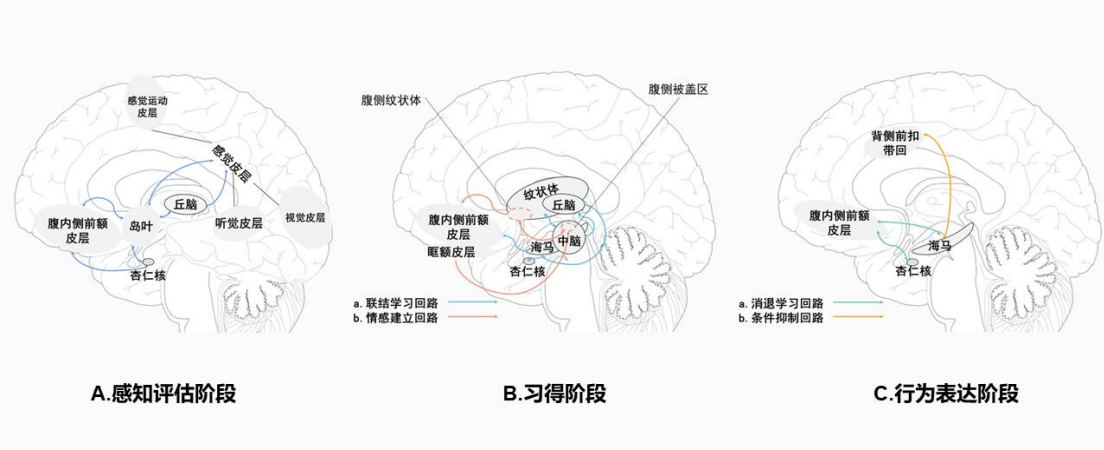


图1 安全学习三阶段神经机制回路

注: 图 1(A)中感觉皮层主要包括感觉运动皮层(位顶叶中央后回以及额叶中央前回)、视觉皮层(位于枕叶)以及听觉皮层(位于颞叶); 图 1(B)中丘脑在纹状体里侧, 中脑-纹状体涉及背侧纹状体-黑质以及腹侧纹状体-腹侧被盖区两条回路(腹侧被盖区和黑质位于中脑, 纹状体包含腹侧和背侧纹状体)。

新增参考文献:

Dou, H., Dai, Y., Qiu, Y., & Lei, Y. (2022). Attachment voices promote safety learning in humans: A critical role for P2. *Psychophysiology*, 59(6), e13997

Favero, J. D., Luck, C., Lipp, O. V., Nguyen, A. T., & Marinovic, W. (2023). N1-P2 event-related potentials and perceived intensity are associated: The effects of a weak pre-stimulus and attentional load on processing of a subsequent intense stimulus. *Biological Psychology*, 184, 108711.

Grasser, L. R., & Jovanovic, T. (2021). Safety learning during development: Implications for development of psychopathology. *Behavioural Brain Research*, 408, 113297. <https://doi.org/10.1016/j.bbr.2021.113297>

Mei, Y., Becker, B., Leppänen, P. H. T., & Lei, Y. (2024). Exploring the ‘black box’ of anxiety: An ER

意见 3: 作者提出, 安全学习并非恐惧学习的简单对立, 而是依赖特定神经机制的相对独立过程。建议系统对比这两种机制的差异, 例如在神经环路、分子机制及行为表现等方面的不同。如果能够通过表格或示意图进行对比, 将更有助于读者理解。

回应: 感谢审稿人的细致审阅与宝贵的意见。我们在引言进行了相应修改与补充, 从学习性质、行为表现、神经回路、分子与可塑性机制方面对安全学习和恐惧学习进行了差异论述, 并通过表格进行对比。主要内容如下:

近年来, 研究者认为, 安全学习是一种区别于恐惧学习的抑制性学习机制。在行为层面表现为个体通过学习 CS 预测“无威胁”, 从而主动抑制恐惧反应(Christianson et al., 2012; Kong et al., 2014)。与恐惧学习带来的冻结(Freeze)、回避等防御性行为(LeDoux & Pine, 2016)不同, 安全学习表现为恐惧反应的缓解。例如, Pollak 等人(2008)的小鼠研究发现, 在安全信号呈现时, 小鼠的冻结反应(即条件性恐惧行为)显著减少, 表明安全信号能够抑制条件性恐惧的表达。研究者还进一步采用高架十字迷宫(Elevated Plus Maze, EPM)范式发现, 安全学习组的小鼠在开放臂中的进入次数和停留时间均增加, 这一结果反映了小鼠焦虑行为水平的降低和探索性行为的增加。人类研究也提供了与此相一致的证据, 安全学习任务中学习到的安全信号能够显著降低个体的主观威胁评估(Odriozola et al., 2024), 这些发现为开发焦虑干预策略提供了重要的理论与方法基础(Laing et al., 2024)。

在神经回路层面, 虽然威胁和安全学习都依赖杏仁核-前额叶-海马回路, 但激活的亚区不同, 例如小鼠实验中前边缘皮层(prelimbic cortex, PL)—通常被认为与人类的 dACC/背内侧前额叶皮层 (dorsomedial prefrontal cortex, dmPFC)同源—更偏向恐惧表达与辨别, 下边缘皮层(infralimbic cortex, IL)—通常被认为与人类的腹内侧前额叶皮层(ventromedial prefrontal cortex, vmPFC)同源—更偏向恐惧抑制与安全线索处理(Sangha et al., 2020)。Duvarci(2024)进一步指出, 恐惧联结的建立主要由导水管周围灰质/背缝核(periaqueductal gray/dorsal raphe, PAG/DR)-中央杏仁核(central amygdala, CeA)多巴胺回路支持, 该回路是威胁联结形成所必要的。而安全联结的建立依赖中脑腹侧被盖区(ventral tegmental area, VTA)-前内侧伏隔核(anteromedial nucleus accumbens, amNAc) 回路的多巴胺能活动, 该回路在安全学习过程中活动增强, 从而支持安全联结的建立。

在更微观的突触与分子水平上,安全与恐惧联结的建立与保持与神经元中的谷氨酸 NM DA 受体(N-methyl-D-aspartate receptor) 和 AMPA 受体 (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor) 参与的长时程增强(long-term potentiation, LTP)有关,但两种学习机制的 LTP 发生在不同类型的神经元中:恐惧联结形成的 LTP 主要发生在杏仁核外侧核(lateral amygdala, LA)的兴奋性神经元上,以支持恐惧记忆的形成与巩固(Palchadhuri et al., 2024); 然而,安全联结形成的 LTP,更多发生在基底外侧杏仁核(basolateral amygdala, BLA)和 CeA 之间的 γ -氨基丁酸能间插细胞(GABAergic intercalated cells),通过抑制 CeA 以降低恐惧输出(Asede et al., 2022); 此外, Harb 等人(2021)的研究表明,脑源性神经营养因子(brain-derived neurotrophic factor, BDNF)缺失并不会显著影响恐惧联结的形成,但会损害安全联结的建立。这一结果提示,相较于恐惧学习,安全学习可能更依赖 BDNF 来支持神经相关回路的突触可塑性与学习更新过程;以上证据表明,安全学习并非恐惧学习的简单对照过程,而是一种在神经机制和功能指向上呈现相对独立性的适应性学习机制(如表 1)。

表 1 安全与恐惧学习的差异

维度	安全学习 (Safety Learning)	恐惧学习 (Fear Learning)
学习性质	建立“CS → 无威胁/安全”联结(抑制性学习)	建立“CS → 威胁”联结(兴奋性学习)
神经回路	VTA → amNac 多巴胺回路编码“威胁缺失”的预测误差信号; vmPFC 在安全线索处理和恐惧抑制中更为关键	PAG/DR→CEA 多巴胺回路编码“威胁出现”的预测误差信号; dmPFC/dACC 在恐惧表达中更为重要
分子与可塑性机制	安全学习的 LTP 更多发生在 BLA-CeA 之间的 GABA 能间插细胞(ITCs); BDNF 缺失损害安全联结形成	恐惧学习的 LTP 主要发生在 LA 的兴奋性神经元, 依赖 NMDA/AMPA 受体; 受 BDNF 影响较小
行为表现	恐惧、焦虑缓解以及趋近行为增加	恐惧、回避、警觉等行为增强

新增参考文献：

Asede, D., Doddapaneni, D., & Bolton, M. M. (2022). Amygdala Intercalated Cells: Gate Keepers and Conveyors of Internal State to the Circuits of Emotion. *The Journal of Neuroscience*, 42(49), 9098–9109

Harb, M., Jagusch, J., Durairaja, A., Endres, T., Leßmann, V., & Fendt, M. (2021). BDNF haploinsufficiency induces behavioral endophenotypes of schizophrenia in male mice that are rescued by enriched environment. *Translational Psychiatry*, 11(1), 233

LeDoux, J. E., & Pine, D. S. (2016). Using Neuroscience to Help Understand Fear and Anxiety: A Two-

意见 4:图 1 展示了安全学习三阶段的神经机制回路，对读者理解整体过程有较好帮助。如果能够补充一张清晰的对比图，说明焦虑个体可能存在的回路异常，将有助于读者更深入理解焦虑障碍中安全学习可能出现的功能失调。

回应:感谢审稿人的细致审阅与宝贵的意见。根据您的建议，我们在 4.1 部分新增了一张“图 2 焦虑个体在安全学习三阶段的神经机制异常”图，用以更直观地对比正常个体与焦虑个体在安全学习神经回路上的差异，从而更清晰地呈现焦虑障碍中可能存在的功能失调。图片如下所示：

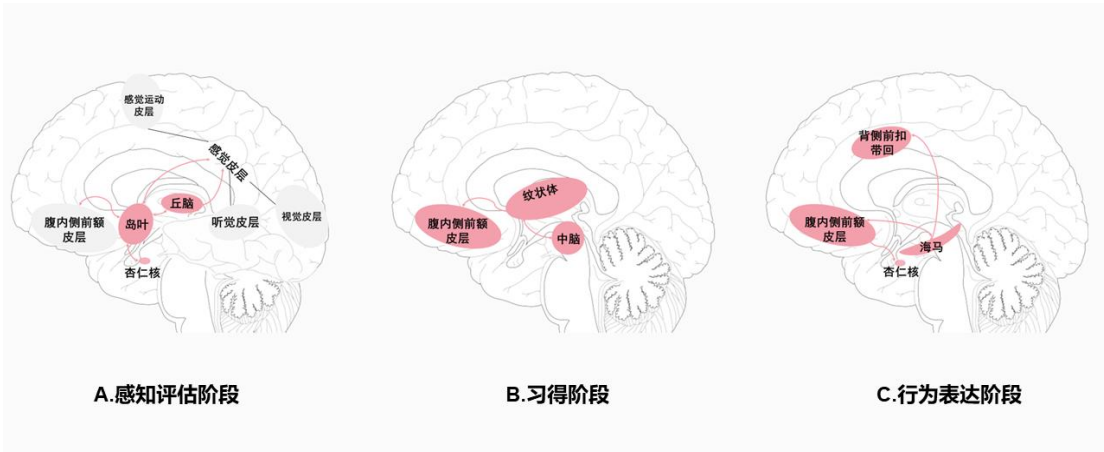


图 2 焦虑个体在安全学习三阶段的神经机制异常

注：图中标红部分表示焦虑及相关障碍个体在安全学习不同阶段中可能出现的脑区及其功能连接异常。

意见 5:在“3 安全学习三阶段的神经机制框架”部分，建议在 3.1 节之前增加一段总结性文字，统领三阶段的主要内容及其相互关系，以增强章节的逻辑连贯性。

回应:感谢审稿人的细致审阅与宝贵的意见。根据您的建议，我们在 3.1 节之前增加了一段总结性文字，用以概述感知、习得与表达三个阶段的核心内容及其逻辑关系：

.....

基于上述认识，我们进一步将安全学习的神经加工过程概括为“感知-习得-表达”三个阶段：(1)安全的感知评估阶段涉及个体对外界刺激的检测与初步解释，为安全联结的建立以及随后的反应调节提供支持。其神经基础主要依托丘脑、皮层、岛叶与杏仁核之间的功能性连接网络，共同实现对刺激安全价值的评估；(2)在安全的习得阶段，可进一步细分为两个

子环节：其一是安全联结的建立与保持，即个体通过预测误差驱动的学习形成刺激与无威胁结果之间的稳定关联；其二是积极情感的获得，即在安全结果的反复出现中形成对安全刺激的积极情绪价值赋予。在神经机制层面，中脑-纹状体环路在该阶段发挥核心作用，支持安全联结的形成和积极情感体验的产生。而安全联结的长期保持与巩固则依赖于海马-前额叶皮层回路的协同作用，为后续的安全记忆提取与行为表达奠定基础；(3)安全的行为表达阶段，体现为个体在习得安全后，能够利用安全信号有效抑制恐惧反应。在神经机制层面，主要涉及边缘系统、前额皮层以及颞叶皮层的交互作用，这些区域的协同活动共同支持安全记忆在不同情境下的提取，从而实现对恐惧反应的有效抑制。下文将详细阐述安全学习三个阶段所涉及的关键脑区及其相对独立的功能机制。

意见 6：一些语言表达方面的建议：

“这些范式进一步丰富了安全学习的研究视角”中，“进”应为“进一步”

“而最近的一项人类多体素研究”建议改为“而最近的一项人脑多体素研究”

回应：感谢审稿人的细致审阅与宝贵的意见。我们认识到本文在语言表述上还有待改进。根据您的建议，我们修改了文中上述不规范的表述，并对全文进行了检查和修改。例如：

这些范式进一步丰富了安全学习的研究视角，表明安全学习不仅能抑制恐惧，还能产生积极的情绪和动机价值以指导后续行为。

而最近的一项人脑多体素研究(Tashjian et al., 2025)进一步区分了 vmPFC 亚区在威胁和安全评估上的差异

意见 7：一篇很相关的近期发表文章，推荐阅读并加入参考文献：

Pacheco-Estefan, D., Bouyeure, A., Jacob, G. et al. Representational dynamics during extinction of fear memories in the human brain. *Nat Hum Behav* (2025). <https://doi.org/10.1038/s41562-025-02268-5>

回应：感谢审稿人提供这一重要且与本文高度相关的参考文献。我们已认真阅读 Pacheco-Estefan et al. (2025) 的研究，并在本文“3.2.1 安全的联结的建立与保持”补充了相应的讨论与引用，具体如下：

3.2.1 安全的联结的建立与保持

.....此外，最近的一项颅内脑电(intracranial electroencephalography, iEEG)研究(Pachec

o-Estefan et al., 2025), 采用 ABC 范式结合表征相似性分析(representational similarity analysis, RSA)发现, 随着消退学习进行, 杏仁核的 θ 振荡(由局部神经元群体同步活动产生的低频节律)在特定的时间窗(接近 US 呈现前的 1.18-1.75 秒)对“安全”刺激增强, 该结果提示杏仁核在消退学习中参与安全信息的编码。进一步分析显示, 杏仁核在消退阶段对刺激表征的稳定性与海马的表征稳定性显著相关, 表明消退记忆的形成, 可能依赖海马与杏仁核之间的协同作用。

新增参考文献:

Pacheco-Estefan, D., Bouyeure, A., Jacob, G., Fellner, M.-C., Lehongre, K., Lambrecq, V., Frazzini, V., Navarro, V., Güntürkün, O., Shen, L., Yang, J., Han, B., Chen, Q., & Axmacher, N. (2025). Representational dynamics during extinction of fear memories in the human brain. *Nature Human Behaviour*, 1–20.

.....

审稿人 2 意见:

本论文在现有模型和相关研究的基础上, 提出了安全学习的三阶段神经机制模型, 并利用该框架来试图解释焦虑发生发展的潜在机制, 具有一定的创新性。论文中的文献比较翔实, 逻辑比较严谨。

本论文还存在一些有待改进的稳妥:

意见 1: 论文的题目为“焦虑及其相关障碍个体的神经机制异常: 基于安全学习的视角”; 但论文的主体都是在讲安全学习及其神经机制框架(2、3 两部分), 与焦虑并没有直接关系。因此, 存在文题不符的问题。

回应: 感谢审稿人宝贵的意见。我们认识到, 原稿标题过于强调“焦虑及其相关障碍的神经机制”; 容易使读者预期本文将主要围绕焦虑障碍的神经机制本身展开讨论。然而, 本文的主要贡献在于: 提出一个用于理解焦虑相关障碍的安全学习三阶段神经机制框架。因此本文综述的重点是“安全学习”; 为使标题更加准确的反应论文的主要内容, 我们对标题进行了调整:

原标题: 焦虑及其相关障碍个体的神经机制异常: 基于安全学习的视角

修改后标题: 安全学习视角下的焦虑相关障碍: 三阶段神经机制框架

意见 2: 摘要要做相应修改, 以突出焦虑及其相关个体在安全学习中的神经机制异常。

回应： 感谢审稿人的细致审阅与宝贵的意见。根据您的建议，我们对摘要进行了针对性修改，修订后的摘要在描述安全学习三阶段框架的同时，突出呈现了焦虑个体在感知、习得与表达三个阶段中的异常，修改后如下：

摘要：安全学习是指个体通过识别安全信号(预示无危险的线索)来抑制恐惧的适应性机制，与焦虑及其相关障碍的发生密切相关。基于现有研究，本研究尝试提出一个由“安全的感知评估、安全的习得以及安全的行为表达”三阶段构成安全学习神经机制框架，强调焦虑相关障碍个体在安全学习三阶段表现出相应的神经机制异常：在感知评估阶段表现为以岛叶-感知皮层为核心的相关回路异常，影响个体有效识别安全线索；在习得阶段表现为中脑-纹状体回路的异常，阻碍个体安全联结的建立与积极情绪体验的获取；在行为表达阶段表现为以海马为核心的抑制回路功能异常，使个体难以提取安全记忆以抑制恐惧。未来研究应通过更高生态效度的范式 and 测量手段，揭示焦虑个体在安全学习中的阶段性异常，以促进安全学习在临床干预的转化应用。

关键词：安全学习，焦虑，神经机制

意见 3：图 1 应该在首次引用后即呈现。

回应：感谢审稿人的细致审阅与宝贵的意见。我们已将图 1 调整至 3.1 部分首次引用之后，以确保阅读流程更加顺畅。修改后如下：

3.1 安全的感知评估：皮层-丘脑-岛叶-杏仁核回路

.....

综上所述，安全刺激的感知评估可能依赖于 vmPFC-丘脑-岛叶-杏仁核等区域之间的功能性回路(如图 1A)，该回路可能参与个体整合多模态感觉信息与认知性评估，并对外界线索作出安全判断的过程。

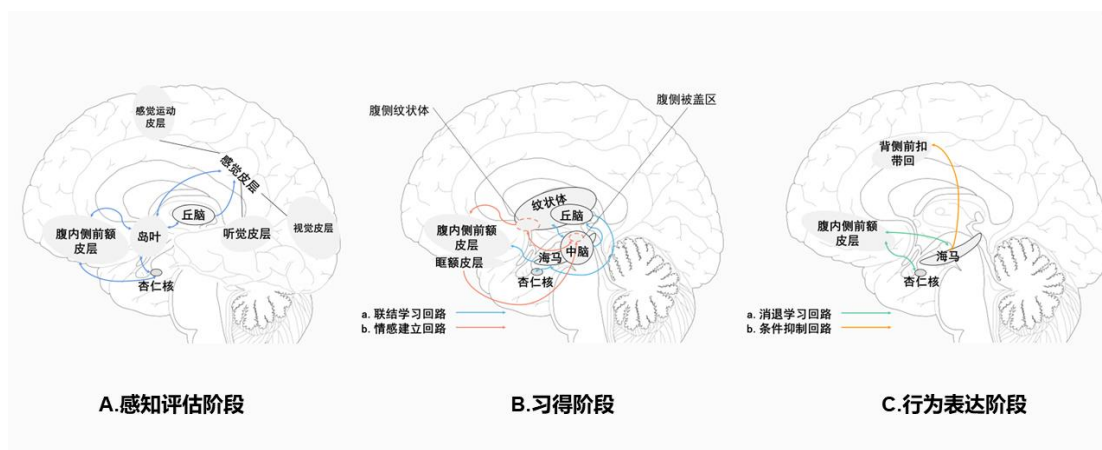


图 1 安全学习三阶段神经机制回路

注：图 1(A)中感觉皮层主要包括感觉运动皮层(位于顶叶中央后回以及额叶中央前回)、视觉皮层(位于枕叶)以及听觉皮层(位于颞叶)；图 1(B)中丘脑在纹状体里侧，中脑-纹状体涉及背侧纹状体-黑质以及腹侧纹状体-腹侧被盖区两条回路(腹侧被盖区和黑质位于中脑，纹状体包含腹侧和背侧纹状体)。

意见 4：“升高十字迷宫实验”表达不太清楚。

回应：感谢审稿人的细致审阅与宝贵的意见。我们认识到文中对“升高十字迷宫实验”描述不够清晰，可能造成阅读理解困难。Elevated Plus Maze 在中文研究中更多的被表述为“高架十字迷宫”实验。在该实验中，由开放臂和闭臂各两条组成，呈十字形交叉。其中开放臂被视为更具威胁性的区域，进入次数与停留时间减少代表更高的焦虑水平。我们在文中补充相应的内容进行了解释，修改后如下：

研究者还进一步采用高架十字迷宫(Elevated Plus Maze, EPM)范式发现，安全学习组的小鼠在开放臂中的进入次数和停留时间均增加，这一结果反映了小鼠焦虑行为水平的降低和探索性行为的增加。

意见 5：文中第一次出现的简称，要给出英文全称，必要时给出中文翻译，如 BDNF、dACC、dmPFC 等。

回应：感谢审稿人的细致审阅与宝贵的意见。我们检查了文中所有简称，并在其首次出现的位置补充了完整的英文全称及必要的中文解释。例如：

BDNF、dACC、dmPFC 第一次出现在引言部分，修改后为：

脑源性神经营养因子(brain-derived neurotrophic factor, BDNF)

背侧前扣带回(dorsal anterior cingulate cortex, dACC)

背内侧前额叶皮层 (dorsomedial prefrontal cortex, dmPFC)

此外对文中其他简称也进行了核对，包括：

长时程增强(long-term potentiation, LTP)

导水管周围灰质/背缝核(periaqueductal gray/dorsal raphe,PAG/DR)

条件刺激(conditional stimulus, CS)

无条件刺激(unconditioned stimulus, US)

多体素模式分析(Multivoxel Pattern Analysis, MVPA)

表征相似性分析(representational similarity analysis, RSA)

.....

意见 6：文中很多句子以“并”“表明”等开头，不合适，如“并强调不同任务程序(辨别、消退和主动回避)可能影响对安全线索的解释。”并基于该区分将人类的安全学习的神经机制划分为“习得”与“表达”两个阶段”表明消退性安全学习能力依赖于相关回路的发育成熟”

回应：感谢审稿人的细致审阅与宝贵的意见。我们已对全文中以“并”“表明”等词语置于句首的表述进行了检查与修改，避免无主语句式，保证论述更严谨，例如：

“并基于该区分将人类的安全学习的神经机制划分为“习得”与“表达”两个阶段 ” 修改后为：

Laing 等人(2022a) 基于个体在安全关联的形成与行为反应两个过程所表现出的功能与机制差异，将安全学习的神经机制划分为“习得”和“表达”两个阶段，为理解安全学习的神经机制提供了新的视角。

“表明消退性安全学习能力依赖于相关回路的发育成熟”修改后为：

上述结果表明，消退性安全学习能力依赖于相关回路的发育成熟，儿童和青少年可能缺乏成熟的 vmPFC 调控支持，更容易出现恐惧反应恢复。

意见 7：“2.1 安全的定义”这个标题太宽泛。

回应：感谢审稿人的细致审阅与宝贵的意见。为了使 2.1 小节标题更聚焦于本文的主题，我们已将“2.1 安全的定义”修改为“**2.1 安全学习的定义**”并对内容作出相应修改，使其更具体

地指向本研究关注的安全学习，为下文介绍“安全学习的范式”提供概念基础。**修改后如下：**

2.1 安全学习的定义

在恐惧学习视角下，“安全”并不携带明确的学习价值，通常被界定为“威胁”结果的对照。而在安全学习视角下，“安全”的建立被视为一种积极的学习过程，具有多种表现形式。Tashjian 等人(2021)认为，安全并非单纯是威胁的缺失，而是个体对当前与未来生存可能性的感知，存在于从“零安全”到“完全安全”的连续体上。安全可以表现为情境安全(如庇护所)、消退性安全(原本危险刺激被学习为安全)、安全信号(威胁被外部因素减缓)以及前瞻性安全(通过行为预期未来安全)。换言之，安全学习的核心表现在于线索与安全结果(如无厌恶刺激或正性结果)相联结。在现有研究中，这种线索-安全结果的联结学习不仅体现于恐惧消退过程中对原有威胁记忆的抑制，还包括安全辨别任务中对威胁与安全线索的主动区分，以及在条件性抑制和回避任务中，个体通过安全信号实现对潜在威胁结果的抑制或规避(Grasser & Jovanovic, 2021; Sangha et al., 2020)。基于上述定义，下面将梳理安全学习的常见实验范式，以揭示其在适应性行为中的独特作用。

意见 8：“年轻儿童”表述不恰当。

回应：感谢审稿人的细致审阅与宝贵的意见。我们认识到在该处使用“年轻儿童”的表述不恰当。**Kribakaran 等人(2024)的研究的样本年龄范围为 9-19 岁，采用平均年龄 ± 1 SD 分界，即小于 11.45 岁(younger youth)与大于 16.89 岁(older youth)。**因此我们在文中已将相关表述统一修改为“**年龄较小的青少年**”与“**年龄较大的青少年**”，修改后如下：

进一步的中介分析结果表明，在年龄较小的青少年(小于 11.45 岁)中，创伤暴露与更强的 aHC-下膝前扣带皮层(subgenual anterior cingulate cortex, sgACC)连接相关；而在年龄较大的青少年(大于 16.89 岁)中却表现出相反的模式，即创伤暴露与更弱的 aHC- sgACC 的功能连接相关。与此同时，该回路功能连接的削弱还进一步预测了 PTSD 症状的严重性(Kribakaran et al., 2024)。

意见 9：“焦虑相关障碍在跨诊断表明”不清楚。

回应：感谢审稿人的细致审阅。此处我们的用词表述不够清晰，这里想表达的意思为横跨不同焦虑障碍诊断类型的研究结果，修改后如下：

综合不同焦虑相关障碍的研究结果可以发现，持续的恐惧泛化以及行为抑制能力减弱可能与海马与前额叶皮层、杏仁核以及前扣带回之间神经回路的功能失调有关

意见 10:“5 总结与展望”中的 3 个点中的两个，都与焦虑的关系不太紧密。

回应：感谢审稿人的细致审阅与宝贵的意见。原文中“总结与展望”部分的前两个要点与焦虑相关障碍的联系的确不够明确。因此，我们在文中对 5.1 和 5.2 的内容进行了调整，以提升本节与焦虑的相关性，修改后如下：

标题修改前：

▲ 5总结与展望

5.1改进安全学习的实验模型

5.2优化安全学习中积极情绪的测量与评估

5.3构建神经机制导向的阶段性和发展性的临床干预策略

标题修改后：

▲ 5总结与展望

5.1加强条件抑制范式在焦虑个体安全学习中的应用

5.2增强焦虑个体在安全学习中积极情感体验

5.3构建神经机制导向的阶段性和发展性的临床干预策略

主要内容如下：

5.1 加强条件抑制范式在焦虑个体安全学习中的应用

近年来，条件抑制范式被认为具有表征安全学习的独特优势，能够避免威胁与安全记忆的竞争，有助于建立稳定的安全关联(Laing & Harrison, 2021; Laing et al., 2024)。然而，目前仍缺乏关于临床个体的条件抑制安全学习研究。另外，该类研究多使用几何图形等低语义刺激，该类刺激难以反映现实社会情境中个体对安全线索的复杂反应(Dunsmoor & Murphy, 2015)。这在一定程度上限制了研究成果的生态效度与推广性。因此，未来可引入具有语义或社会属性的自然刺激，并结 MVPA 等方法，揭示抑制性安全信号在临床焦虑个体与健康个体中的神经表征差异。

另一方面，焦虑相关障碍个体通常存在安全学习后回忆阶段的异常。然而，目前针对“安全记忆”的回忆研究仍主要集中于恐惧消退任务。相比之下，条件抑制作为一种更主动、也更稳定的安全学习形式，仍不清楚焦虑个体在其回忆过程中是否也存在功能受损，或该机制

是否可能弥补其在消退回忆中的缺陷，对此需要未来研究进一步探讨。此外，条件抑制在回忆阶段的研究之所以有限，可能也与其范式本身的复杂性有关：由于抑制性安全信号需要与威胁刺激在共现情境下发挥抑制效应，因此在回忆阶段独立测量其记忆表征将更具挑战性。一项近期的神经生物学研究(Tashjian et al., 2025) 为探索抑制性安全学习的回忆机制提供了启示。该研究通过单独呈现保护线索和威胁线索，让个体进行配对并联想“战斗场景”以评估威胁程度，并在任务结束后再次单独呈现刺激，测量个体对不同类型线索的神经表征。未来可在该范式基础上进行改进，并结合 RSA 等方法，比较焦虑个体与健康个体在条件抑制安全学习的回忆阶段是否也存在神经表征上的差异，从而进一步揭示焦虑个体在安全学习回忆阶段异常的神经机制。

5.2 增强焦虑个体在安全学习中积极情感体验

在安全学习过程中，“积极情感”的建立通常被视为判断安全学习是否成功的重要指标。然而，焦虑相关障碍个体可能难以从安全学习中获得足够的积极情绪体验，这一异常可能削弱安全学习的效果。未来以安全学习为基础的焦虑干预策略，应更加重视积极情绪体验的增强。例如，可通过将安全线索与奖励、正性意象或积极认知等建立联结，从而强化个体对安全线索的积极价值编码。

另外，当前研究对于从安全学习中获得积极情绪的测量仍过度依赖主观报告，存在主观性强、敏感性等不足(Laing et al., 2024; Odriozola & Gee, 2021)。未来应构建多模态、客观化且具生态效度的安全感评估体系。例如，通过同步采集 fMRI、EEG 与皮肤电信号并结合主观评估，识别和预测个体产生“安全”体验的神经生理特征，从而支持干预效果的精准评估，进一步推动基于安全学习的焦虑干预策略的发展。

意见 11：文中还存在一些表达不通顺的问题，如“这些范式进丰富了安全学习的研究视角”“Atlas 等人(2016)等人”“Pattwell 等人(2016)的研究啮齿类动物研究发现”“出生后第 60 天”“Kribakaran 等人(2022, 2024)的研究通过 SSL 任务在人类中发现发现”“位顶叶中央后回以及额叶中央前回”；请作者举一反三，认真修改。

回应：感谢审稿人的细致审阅与宝贵的意见。我们对上述表述不正确的地方进行了修改，并且认真检查了文稿中存在的语言表达不够通顺、重复或用词不当之处，对全文进行了逐句核对和修订。上述问题修改后如下：

1) “这些范式进丰富了安全学习的研究视角”修改后：

这些范式进一步丰富了安全学习的研究视角

2) “Atlas 等人(2016)等人” 修改后:

Atlas 等人(2016)

3) “Pattwell 等人(2016)的研究啮齿类动物研究发现” 修改后:

Pattwell 等人(2016)的研究发现

4) “出生后第 60 天” 修改后:

从小鼠的青春期(约出生后第 30 天)延续至成年(约出生后第 60 天)

5) “Kribakaran 等人(2022, 2024)的研究通过 SSL 任务在人类中发现发现” 修改后:

Kribakaran 等人(2022, 2024)的研究通过安全信号学习任务在人类中发现

6) “位顶叶中央后回以及额叶中央前回” 修改后:

位于顶叶中央后回以及额叶中央前回

意见 12: 文中还存在一些标点符号使用不当问题, 如“而是一个复杂的包含认知评估, 动机驱动以及情绪调节的适应性机制。”中间应该用顿号。

回应: 感谢审稿人的细致审阅与宝贵的意见。我们已对上述标点符号使用不当的问题进行了修改, 并且对文稿的标点符号进行了全面检查, 并对使用不当之处进行了修正。上述问题修改后如下:

“而是一个复杂的包含认知评估、动机驱动以及情绪调节的适应性机制。”

第三轮

审稿人 1 意见: 作者已经对审稿人的意见进行了较好的回复。

审稿人 2 意见: 作者已较好地处理了之前提出的意见。

编委 1 意见:

作者对审稿人意见进行了很好的回应和修改, 修改稿逻辑性良好, 语句通畅, 基本上达到了出版要求, 建议发表。

发表前作者就如下建议进一步思考修改为好:

意见 1：精神障碍专用名词使用的规范性问题：焦虑障碍是很早的概念，包括很多疾病，如在本文中使用的所用疾病，在既往都归属焦虑障碍。但近十年精神疾病分类变化很大，PTSD 从焦虑障碍中分离出来，成为创伤及相关障碍，强迫症也从焦虑障碍中独立出来成为单独疾病。目前焦虑及相关障碍就等同于恐惧症和焦虑症，属于狭义的焦虑障碍。在全文中建议统一。

回应：感谢审稿人的细致审阅与宝贵的意见，我们已经对相关名词进行统一：文中的焦虑障碍个体主要用于描述恐惧症与焦虑症，对于 PTSD 和 OCD 我们已经单独指出，而非划分为焦虑相关障碍。例如：

引言部分：

现有研究表明，社交焦虑(social anxiety disorder, SAD)、广泛性焦虑(generalized anxiety disorder, GAD)等焦虑相关障碍，以及创伤后应激障碍(post-traumatic stress disorder, PTSD)、强迫症(obsessive-compulsive disorder, OCD)等精神障碍均被认为与安全学习异常相关(Jovanovic et al., 2012; Cooper et al., 2018; Cooper & Dunsmoor, 2021; Lewis et al., 2024)，主要表现为难以区分安全与威胁刺激，安全习得效率低以及恐惧抑制失效等。

4.3 安全行为表达异常：

此外，在与焦虑具有高度共病性和相似病理特征的临床个体中，如 PTSD 和 OCD(American Psychiatric Association, 2022)，也发现该回路神经机制的异常。患有 PTSD 的青少年在安全记忆神经回路表现明显的发育缺陷，包括海马体积下降，杏仁核-前额叶耦合随着年龄的增长而减弱等异常现象(Herringa, 2017)。同样的，与健康对照组相比，OCD 患者的海马体积显著减小，且海马与前额叶、杏仁核的功能连接减弱(Boedhoe et al., 2017; Cooper & Dunsmoor, 2021).....

其余部分涉及只谈论恐惧症和焦虑症，均改为用焦虑障碍的表述

意见 2：由于疾病分类的变化，题目也需要完善。

回应：感谢审稿人的细致审阅与宝贵的意见，我们已经按照要求对题目做出相应修改，修改后为：**安全学习与焦虑障碍：三阶段神经机制框架**

意见 3: 安全与威胁是一个维度上的两个极端, 安全——积极体验——无恐惧反应, 威胁——恐惧——恐惧反应。是否是一个硬币的两面, 有很多是共同的神经环路机制?

回应: 感谢审稿专家的意见。下述将分为 2 点分别回复: ①安全与威胁维度问题及②共同神经环路问题。

①Tashjian 等人(2021)认为安全存在于一个从零安全(威胁不可避免必定造成伤害)到完全安全(无实际威胁, 没有预期的威胁)的连续体。并将“安全”定义为对生存可能性的感知, 而威胁是影响这种生存可能性的一个因素, 两者是可能同时存在并影响我们对生存状态的感知的(例如笼子+老虎), 而非是同一纬度的两个极端。而恐惧则是对这种生存状态的情绪反应。我们针对文中 2.1 部分关于“安全”的解释也进行了更清晰的阐述, 修改后如下:

Tashjian 等人(2021)认为, 安全并非单纯是威胁的缺失, 而是个体对当前与未来生存可能性的感知, 存在于从“零安全”到“完全安全”的连续体上。其中, “零安全”是指伤害已迫近且不可避免的危险状态, 即威胁不再是潜在或可缓解的; 在此情境下, 恐惧通常作为对此类危险状态的情绪反应而出现, 并可能随危险程度的增加而增强。而完全安全是指不存在实际的威胁, 也不存在可预期的未来威胁, 这个时候个体不存在恐惧反应。

参考文献如下:

Tashjian, S. M., Zbozinek, T. D., & Mobbs, D. (2021). A Decision Architecture for Safety Computations. *Trends in Cognitive Sciences*, 25(5), 342–354.

②另外, 现有研究表明, 安全与威胁在调节防御反应和情绪体验方面确实涉及部分重叠的神经回路(如杏仁核-前额叶-海马相关网络)的调节, 但安全学习的神经机制也并非单纯恐惧回路活动被抑制, 两者加工可能涉及不同的神经活动模式。例如本文所提到的: 前部海马可能负责安全记忆的加工, 而后部海马可能与恐惧记忆更为相关(Hennings et al., 2022); 前 vmPFC 在安全学习后对高安全属性的刺激有特异性的神经表征(Tashjian et., 2025); 杏仁核的 θ 振荡与安全信号加工密切相关(Pacheco-Estefan et al., 2025)。因此, 安全与威胁学习虽存在相似的神经回路基础, 但并非同一加工过程的正负两端, 而是依托不同亚区及特异性神经活动模式实现的、相对独立的两种学习机制。参考文献如下:

Hennings, A. C., Mason McClay, Drew, M. R., Lewis-Peacock, J. A., & Dunsmoor, J. E. (2022). Neural reinstatement reveals divided organization of fear and extinction memories in the human brain. *Current Biology*, 32(2), Article 2.

Tashjian, S. M., Cussen, J., Deng, W., Zhang, B., & Mobbs, D. (2025). Subregions in the ventromedial prefrontal

cortex integrate threat and protective information to meta-represent safety. *PLOS Biology*, 23(1), e3002986.

Pacheco-Estefan, D., Bouyeure, A., Jacob, G., Fellner, M.-C., Lehongre, K., Lambrecq, V., Frazzini, V., Navarro, V., Güntürkün, O., Shen, L., Yang, J., Han, B., Chen, Q., & Axmacher, N. (2025). Representational dynamics during extinction of fear memories in the human brain. *Nature Human Behaviour*, 1–20.

意见 4：文中必要的批注和文字修改(见附件)。

回应：感谢审稿专家的意见，我们已针对批注内容进行了相应修改完善，修改内容我们在正文中用绿色文本进行突出。

例如引言部分修改标注部分的表述后如下：

尽管焦虑的发生与安全学习缺陷密切相关,但过去对于焦虑如何产生的理解主要集中在恐惧学习研究。

.....可能不足以解释焦虑障碍中常见的多种异常表现，如安全识别困难、正性情绪缺失等问题。因此，有必要进一步关注个体对安全信号的学习及其在恐惧调节中的作用。

近年来，研究者认为，安全学习是一种区别于恐惧学习的抑制性学习机制。在行为层面表现为个体通过学习 CS 预测“无威胁”从而利用学习到的“安全”降低恐惧反应(Christianson et al., 2012; Kong et al., 2014).....

例如，Pollak 等人(2008)的小鼠研究发现，在安全信号呈现时，小鼠的冻结反应(即条件性恐惧行为)显著减少，表明安全信号能够削弱条件性恐惧的表达。

表 1 表述错误部分修改后如下：

维度	安全学习	恐惧学习
	(Safety Learning)	(Fear Learning)
学习性质	建立“CS→无威胁/安全”联结(抑制性学习)	建立“CS→威胁”联结(兴奋性学习)
神经回路	VTA→amNac 多巴胺回路编码“威胁缺失”的预测误差信号； vmPFC 在安全线索处理和恐惧抑制中更为关键	PAG/DR→CEA 多巴胺回路编码“威胁出现”的预测误差信号； dmPFC/dACC 在恐惧表达中更为重要
	安全学习的 LTP 更多发生在 BLA-CEA 之间的 GA	恐惧学习的 LTP 主要发生在 LA 的兴奋性神
分子与可塑性机制	BA 能间插细胞(ITCs)； BDNF 缺失损害安全联结形成	经元，依赖 NMDA/AMPA 受体； 受 BDNF 影响较小
行为表现	恐惧、焦虑缓解以及趋近行为增加	恐惧、回避、警觉等行为增强

3 安全学习三阶段的神经机制框架部分后如下：

第一段精简：

神经生物学研究表明，安全学习是一个包含多阶段、多神经系统协同作用的复杂过程。Laing 等人(2022a)根据安全联结的形成和之后的行为表达两个过程在神经机制上的差异，将安全学习区分为习得阶段和表达阶段(见 2.3)。但现有研究表明，与积极情绪加工相关的多巴胺回路(Grill et al., 2024)以及与记忆存储相关的海马-vmPFC 回路活动(Meyer et al., 2019)同样参与安全的学习过程。换言之，安全习得过程不仅涉及对威胁预测的修正，也可能包括对积极情绪与安全记忆的加工。因此，其应被视为一个同时整合安全联结形成与积极情绪加工的过程。

最后一段修改表述：

基于上述认识，我们进一步将安全学习的神经加工过程概括为“感知-习得-表达”三个阶段：(1)安全的感知评估阶段涉及个体对外界刺激的检测与早期加工，为安全联结的建立以及随后的反应调节提供支持。

2.1 安全学习的定义部分：增加对零安全，完全安全，恐惧关系的解释

其中，“零安全”是指伤害已迫近且不可避免的危险状态，即威胁不再是潜在或可缓解的；在此情境下，恐惧通常作为对此类危险状态的情绪反应而出现，并可能随危险程度的增加而增强。而完全安全是指不存在实际的威胁，也不存在可预期的未来威胁，这个时候个体不存在恐惧反应。

本次修改新增参考文献（自检报告已更新）：

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Publishing.

编委 2 意见：

质量有大幅度提升，同意发表。

主编意见：

文章具有显著的临床转化价值，同意发表。