

Translation

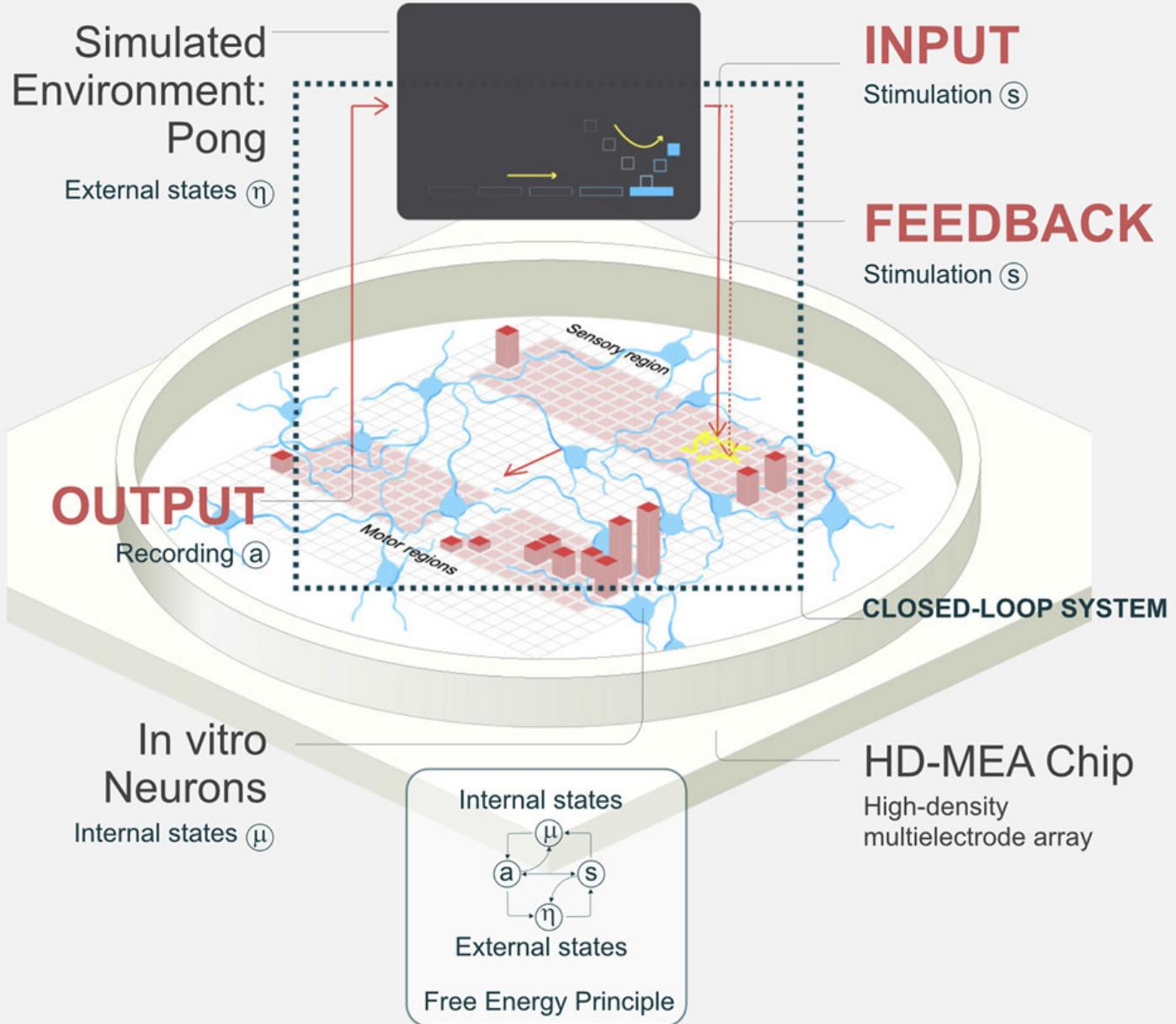
原文链接： [In vitro neurons learn and exhibit sentience when embodied in a simulated game-world](#)

体外神经元在模拟游戏世界中学习并表现出感知能力

摘要

将神经元整合到数字系统中可能会实现硅单独无法实现的性能。在这里，我们开发了 DishBrain 系统，该系统利用结构化环境中神经元的固有自适应计算。体外来自人类或啮齿动物来源的神经网络通过高密度多电极阵列与硅内计算相结合。通过电生理刺激和记录，培养物被嵌入到模拟游戏世界中，模拟街机游戏“乒乓球”。应用主动推理理论的含义，通过自由能原理，我们发现在实时游戏中的五分钟内出现了明显的学习，而在对照条件中没有观察到。进一步的实验表明，通过随时间推移引发学习的闭环结构化反馈的重要性。培养物展示了对其行为后果的稀疏感知信息的自组织活动的能力，我们将其称为合成生物智能。未来应用可能会进一步深入了解智能的细胞对应物。

Integrating neurons into digital systems may enable performance infeasible with silicon alone. Here, we develop DishBrain, a system that harnesses the inherent adaptive computation of neurons in a structured environment. In vitro neural networks from human or rodent origins are integrated with *in silico* computing via a high-density multielectrode array. Through electrophysiological stimulation and recording, cultures are embedded in a simulated game-world, mimicking the arcade game ‘‘Pong.’’ Applying implications from the theory of active inference via the free energy principle, we find apparent learning within five minutes of real-time gameplay not observed in control conditions. Further experiments demonstrate the importance of closed-loop structured feedback in eliciting learning over time. Cultures display the ability to self-organize activity in a goal-directed manner in response to sparse sensory information about the consequences of their actions, which we term synthetic biological intelligence. Future applications may provide further insights into the cellular correlates of intelligence.



Neural activity changes in real-time to minimise environmental unpredictability

引言

利用生物神经元的计算能力创造合成生物智能 (SBI)，以前仅限于科幻领域，现在可能已经在人类创新的范围内。生物计算的优越性已被广泛推测，并尝试开发支持类神经计算的仿生硬件 (Kumar等, 2020年)。然而，在生物神经元之外的任何人工系统都无法支持至少三阶复杂度（能够表示三个状态变量），这是重新创建生物神经网络 (BNN) 的复杂性所必需的 (Izhikevich, 2006年；Kumar等, 2020年)。尽管在体内神经计算的映射方面取得了重大进展，但在体外探索这一领域存在技术限制 (Barron等, 2020年)。在这里，我们旨在建立从胚胎啮齿动物和人类诱导多能干细胞 (HiPSCs) 到高密度多电极阵列 (HD-MEAs) 上的功能性体外 BNNs，以证明这些神经培养物能够表现出生物智能，如在模拟游戏环境中的实时学习来改变活动方式，这在实时游戏中被证明 (图1)。提出这些神经培养物将符合感知的正式定义，即通过自适应内部过程对感觉印象作出反

应 (Friston 等 , 2020 年) 。实例化 SBIs 可能会引领研究进入生物智能的范式转变 , 包括伪认知反应作为药物筛选的一部分 (Kagan 等 , 2022 年 ; Myers , 2017 年) , 弥合单细胞和群体编码方法对理解神经生物学的探索 (Ebitz 和 Hayden , 2021 年) , 探索 BNN 如何计算以指导机器学习方法 (Mattar 和 Lengyel , 2022 年) , 并潜在地催生超越现有纯硅硬件性能的硅 - 生物计算平台。从理论上说 , 普遍的 SBI 可能在人工通用智能 (AGI) 之前到来 , 这是由于生物系统的内在效率和进化优势 (Buchanan , 2018 年) 。

Harnessing the computational power of living neurons to create synthetic biological intelligence (SBI), previously confined to the realm of science fiction, may now be within reach of human innovation. The superiority of biological computation has been widely theorized with attempts to develop biomimetic hardware supporting neuromorphic computing (Kumar et al., 2020). Yet no artificial system outside biological neurons is capable of supporting at least third-order complexity (able to represent three state variables), which is necessary to recreate the complexity of a biological neuronal network (BNN) (Izhikevich, 2006; Kumar et al., 2020). While significant progress has been made in mapping *in vivo* neural computation, there are technical limits to exploring this *in vitro* (Barron et al., 2020). Here, we aim to establish functional *in vitro* BNNs from embryonic rodent and human-induced pluripotent stem cells (hiPSCs) on high-density multielectrode arrays (HD-MEAs) to demonstrate that these neural cultures can exhibit biological intelligence—as evidenced by learning in a simulated gameplay environment to alter activity in an otherwise arbitrary manner—in real time (Figure 1). It is proposed that these neural cultures would meet the formal definition of sentience as being ‘‘responsive to sensory impressions’’ through adaptive internal processes (Friston et al., 2020). Instantiating SBIs could herald a paradigm shift of research into biological intelligence, including pseudo-cognitive responses as part of drug screening (Kagan et al., 2022; Myers, 2017), bridging the divide between single-cell and population-coding approaches to understanding neurobiology (Ebitz and Hayden, 2021), exploring how BNNs compute to inform machine-learning approaches (Mattar and Lengyel, 2022), and potentially giving rise to *silico*-biological computational platforms that surpass the performance of existing purely silicon hardware. Theoretically, generalized SBI may arrive before artificial general intelligence (AGI) due to the inherent efficiency and evolutionary advantage of biological systems (Buchanan, 2018).

这个系统被称为 DishBrain , 它可以利用神经元共享电活动的 “语言” 这一固有属性 , 通过电生理刺激和记录将硅和 BNN 系统连接起来。鉴于硬件和细胞 (湿件) 的兼容性 , 有必要研究当 BNN 通过闭环系统具体化时 , 会导致智能 (目标导向) 行为的过程。智能系统中需要两个相互关联的过程来产生感知行为。首先 , 系统必须学习外部状态如何通过感知影响内部状态 , 以及内部状态如何通过行动影响外部状态。其次 , 系统必须从其感知状态推断 , 确定何时采取特定活动以及其行动将如何影响环境。

This system, termed DishBrain, can leverage the inherent property of neurons to share a ‘‘language’’ of electrical activity to link silicon and BNN systems through electrophysiological stimulation and recording. Given the compatibility of hardware and cells (wetware), it is necessary to investigate what processes would result in intelligent (goal-directed) behavior when BNNs are embodied through a closed-loop system. Two interrelated processes are required for sentient behavior in an intelligent system. Firstly, the system must learn how external states influence internal states via perception and how internal states influence external states via action. Secondly, the system must infer from its sensory states when it should adopt a particular activity and how its actions will influence the environment.

为了解决第一个要求 , 开发了定制软件驱动程序 , 创建了低延迟闭环反馈系统 , 通过电刺激模拟 BNNs 与环境的交互。闭环系统通过提供有关细胞培养行为影响的反馈 , 使体外培养物 “具象化” 。具象化需要将内部状态与外部状态分开 , 其中提供了有关行为对给定环境影响的反馈。先前的研究 , 无论是体内还是体外 , 都表明电生理闭环反馈系统会引发显著的网络可塑性 (Bakkum 等 , 2008a ; Chao 等 , 2008) 。在体内 , 通过破坏小鼠初级视觉皮层中视觉反馈与运动输出之间的闭环耦合 , 进一步支持了这一观点 , 突显了反馈与 BNNs 功能行为发展之间的联系 (Attinger 等 , 2017) 。

To address the first imperative, custom software drivers were developed to create low-latency closed-loop feedback systems that simulated exchange with an environment for BNNs through electrical stimulation. Closed-loop systems afford an *in vitro* culture ‘‘embodiment’’ by providing feedback on the causal effect of the behavior from the cell culture. Embodiment requires a separation of internal versus external states where feedback of the effect of an action on a given environment is available. Previous works, both *in vitro* and *in silico*, have shown that electrophysiological closed-loop feedback systems engender significant network plasticity (Bakkum et al., 2008a; Chao et al., 2008). Further support is found *in vivo* by disrupting the closed-loop coupling between visual feedback and motor outputs in the primary visual cortex of mice (Attinger et al., 2017), highlighting the link between feedback and the development of functional behavior in BNNs.

为了满足第二个要求，DishBrain系统测试了智能行为可能如何产生的理论框架。智能系统如何在环境中具象化时产生智能行为的一个假设是通过自由能原则 (FEP) 的主动推理理论 (Friston等, 2012)。FEP提出了一个可测试的含义，即在每个时空尺度上，任何与其环境分离的自组织系统都寻求最小化其变分自由能 (VFE) (Friston, 2010; Palacios等, 2020; Parr和Friston, 2019)。模型预测与观察到的感觉之间的差距 (“惊讶”或“预测误差”) 可以通过两种方式最小化：通过优化有关环境的概率信念，使预测更接近感觉，或者通过对环境采取行动，使感觉符合其预测。该模型随后暗示了行动和感知的一个共同客观函数，评分内部模型与外部环境之间的匹配度。根据这一理论，BNNs对世界状态持有“信念”，其中学习涉及更新这些信念以最小化它们的VFE或积极改变世界，使其更少令人惊讶 (Parr和Friston, 2018, 2019)。如果属实，这意味着通过简单呈现与“错误”行为相关的不可预测反馈，应该可以塑造BNN行为。从理论上讲，BNNs应该采取行动来避免导致不可预测输入的状态。通过开发一个允许神经培养物在模拟游戏世界中具象化的系统，我们不仅能够测试这些细胞是否能够在动态环境中进行目标导向学习，还能够探究智能的基础。

To address the second requirement, a theoretical framework for how intelligent behavior may arise was tested by the DishBrain system. One proposition for how intelligent behavior may arise in an intelligent system embodied in an environment is the theory of active inference via the free energy principle (FEP) (Friston et al., 2012). The FEP suggests a testable implication that at every spatiotemporal scale, any self-organizing system separate from its environment seeks to minimize its variational free energy (VFE) (Friston, 2010; Palacios et al., 2020; Parr and Friston, 2019). The gap between the model predictions and observed sensations (‘‘surprise’’ or ‘‘prediction error’’) may be minimized in two ways: by optimizing probabilistic beliefs about the environment to make predictions more like sensations or by acting upon the environment to make sensations conform to its predictions. This model then implies a common objective function for action and perception that scores the fit between an internal model and the external environment. Under this theory, BNNs hold ‘‘beliefs’’ about the state of the world, where learning involves updating these beliefs to minimize their VFE or actively change the world to make it less surprising (Parr and Friston, 2018, 2019). If true, this implies that it should be possible to shape BNN behavior by simply presenting unpredictable feedback following ‘‘incorrect’’ behavior. Theoretically, BNNs should adopt actions that avoid the states that result in unpredictable input. By developing a system that allows for neural cultures to be embodied in a simulated game-world, we are not only able to test whether these cells are capable of engaging in goal-directed learning in a dynamic environment, but we are also able to investigate the foundations of intelligence.

先前的研究支持体外神经元网络能够在开环环境中通过与自由能原理 (FEP) 一致的状态依赖性海伯恩可塑性执行盲源分离 (Isomura et al., 2015; Isomura and Friston, 2018)。我们试图在这项工作的基础上，测试主动推理理论，该理论将自由能原理应用于不仅适应其环境，还能对环境进行作用以使环境适应自身的有知觉系统。因此，我们假设当在DishBrain系统中提供模拟经典街机游戏“乒乓” (Pong) 的结构化外部刺激时，生物神经网络 (BNN) 将修改内部活动，以避免采用与不可预测的外部刺激相关联的状态。这种输入不可预测性的最小化将表现为在这个简化的模拟“乒乓”环境中对模拟“球拍”的目标导向控制。

Previous work supports that in vitro neuronal networks can perform blind-source separation in an open-loop environment via state-dependent Hebbian plasticity consistent with the FEP (Isomura et al., 2015; Isomura and Friston, 2018). We sought to build upon this work to test the theory of active inference, which applies the FEP to sentient systems that not only adapt to fit their environment, but also act upon their environment to fit it to themselves. We therefore hypothesize that when provided a structured external stimulation simulating the classic arcade game ‘‘Pong’’ within the DishBrain system, the BNN would modify internal activity to avoid adopting states linked to unpredictable external stimulation. This minimization of input unpredictability would manifest as the goal-directed control of the simulated ‘‘paddle’’ in this simplified simulated ‘‘Pong’’ environment.

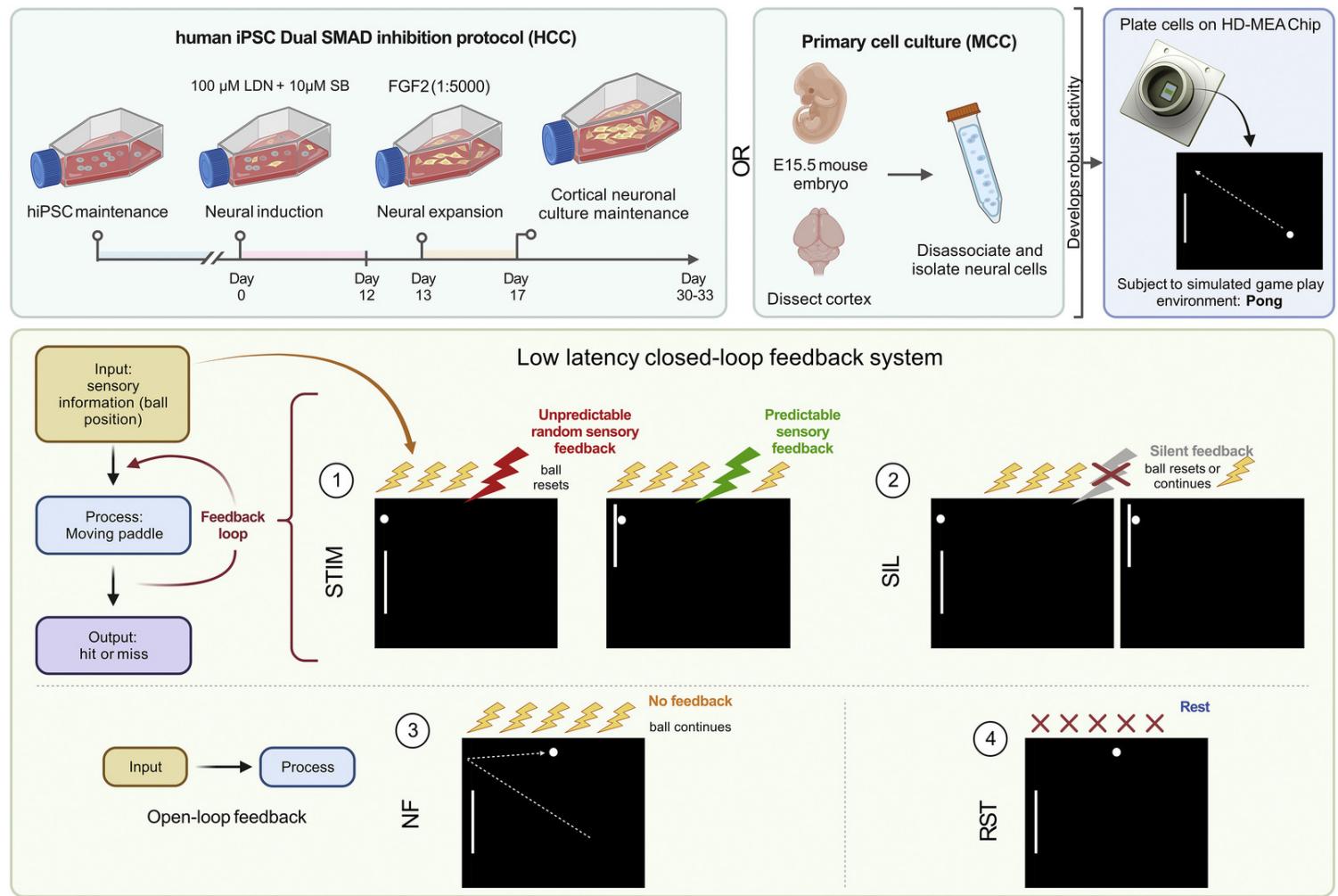


图 1：盘中之脑系统 (DishBrain System) 和实验方案图解

结果

用于计算的神经元“湿件”的生长

来自啮齿胚胎解剖皮层的皮质细胞可以在营养丰富的培养基中生长在微电极阵列上，并保持数月 (Bardy等, 2015年；Lossi和Merighi, 2018年)。这些培养物将发展出具有复杂形态的结构，具有许多树突和轴突连接，形成功能性BNNs (Kamioka等, 1996年；Wagenaar等, 2006年)。从胚胎15.5天 (E15.5) 的小鼠胚胎中培养的原代神经培养物在图2A中显示。HiPSC被诱导分化为活跃的异质性皮层神经元单层，已显示出成熟的功能性特性 (Denham等, 2012年；Denham和Dottori, 2009年；Shi等, 2012年)。使用双SMAD抑制 (DSI) (Denham等, 2012年；Fattah等, 2015年)，我们培养出了长期形成与支持胶

质细胞密集连接的皮层神经元（图2B和2C）。最后，我们计划使用不同的HiPSC分化方法 - NGN2直接重编程（Pak等，2018年；Zhang等，2013年）- 用于我们研究的最后部分，研究反馈机制。这种高产方法导致细胞显示出泛神经元标记物（图S1A和S1B）。这些细胞通常显示出高比例的兴奋性谷氨酸能神经元，使用qPCR定量，如图2D所示。通过扫描电子显微镜（SEM）确认了这些神经元培养物在维持了3个月的细胞上的HD-MEAs上的整合（图2E）。可以观察到在形成交织网络跨越MEA区域的神经元培养物中的密集连接的树突网络（图2F）。这些神经元培养物似乎很少遵循MEA的拓扑结构，更有可能形成连接细胞的大团簇，具有密集的树突网络（图2G和2H）。这可能是由于MEA中单个电极的大尺寸，以及可能的趋化效应，这些效应可以帮助抵消基质地形对神经纤维投射的影响（Mattotti等，2012年）。

Cortical cells from the dissected cortices of rodent embryos can be grown on MEAs in nutrient-rich media and maintained for months (Bardy et al., 2015; Lossi and Merighi, 2018). These cultures will develop complicated morphology with numerous dendritic and axonal connections, leading to functional BNNs (Kamioka et al., 1996; Wagenaar et al., 2006). Primary neural cultures from embryonic day 15.5 (E15.5) mouse embryos were cultured, with representative cultures shown in Figure 2A. HiPSCs were differentiated into monolayers of active heterogeneous cortical neurons, which have been shown to display mature functional properties (Denham et al., 2012; Denham and Dottori, 2009; Shi et al., 2012). Using dual SMAD inhibition (DSI) (Denham et al., 2012; Fattah et al., 2015), we developed long-term cortical neurons that formed dense connections with supporting glial cells (Figures 2B and 2C). Finally, we aimed to expand our study using a different method of hiPSC differentiation—NGN2 direct reprogramming (Pak et al., 2018; Zhang et al., 2013)—used in our final part of this study investigating feedback mechanisms. This high-yield method resulted in cells displaying pan-neuronal markers (Figures S1A and S1B). These cells typically display a high proportion of excitatory glutamatergic cells, quantified using qPCR, shown in Figure 2D. Integration of these neuronal cultures on the HD-MEAs was confirmed via scanning electron microscopy (SEM) on cells that had been maintained for 3 months (Figure 2E). Densely interconnected dendritic networks could be observed in neuronal cultures forming interlaced networks spanning the MEA area (Figure 2F). These neuronal cultures appeared to rarely follow the topography of the MEA, being more likely to form large clusters of connected cells with dense dendritic networks (Figures 2G and 2H). This is likely due to the large size of an individual electrode within the MEA and potentially also chemotactic effects that can contribute to counteract the effect of substrate topography on neurite projections (Mattotti et al., 2012).

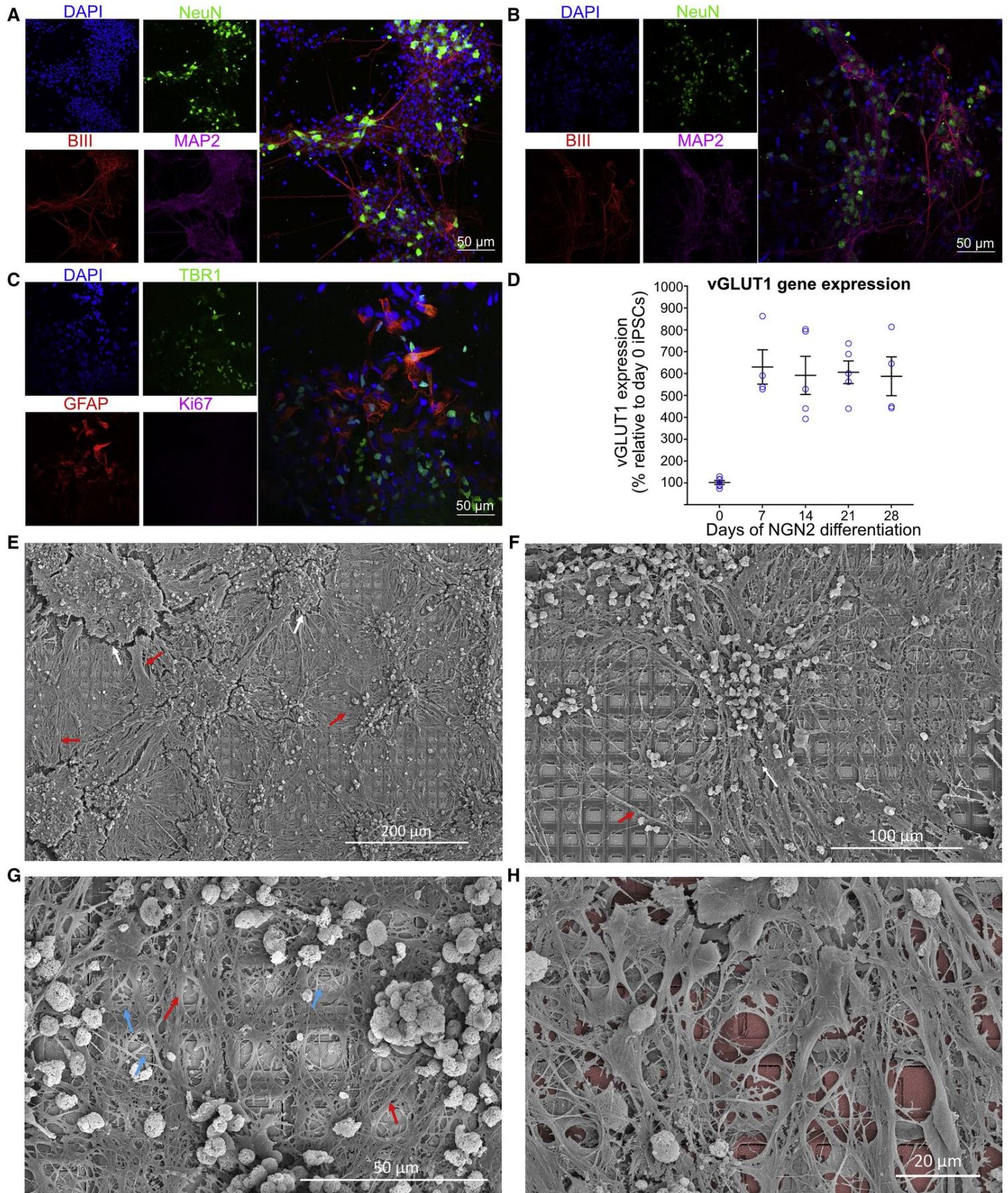


图 2：皮层细胞形成密集的相互连接的网络

神经细胞显示出具有时间发展规律的特征性自发动作电位

我们以高空间和时间分辨率映射了体外培养的神经系统中电生理活动的发展。在E15.5啮齿动物的原代皮质细胞中，发现在培养天数14天（DIV）时出现强劲的活动（图3A和3E），在这里定期观察到同步爆发活动，正如以前展示的（Kamioka等，1996年；Wagenaar等，2006年）。作为对比，与以前的报告类似（Shi等，2012年），使用DSI分化的HiPSC背景的皮质细胞直到DIV 73才显示出同步爆发活动（图3A和3F）。使用NGN2直接重编程分化的HiPSC表现出更早的活动，通常在第14到24天之间（图3A和3G）。通过每日活动扫描监测电生理成熟性。在测试期间，所有细胞类型的最大放电频率常会增加并在测试期间保持相对稳定（图3B），但在测试天数中观察到了平均放电频率（图3C）和放电频率方差（图3D）的变化；特别是，使用NGN2直接重编程方法分化的在测试期间的平均放电频率和放电频率方差均显著增加。

In vitro development of electrophysiological activity in neural systems at high spatial and temporal resolution was mapped. Robust activity in primary cortical cells from E15.5 rodents was found at days in vitro (DIV) 14 (Figures 3A and 3E) where bursts of synchronized activity were regularly observed, as previously demonstrated (Kamioka et al., 1996; Wagenaar et al., 2006). In contrast, similar to previous reports (Shi et al., 2012), synchronized bursting activity was not observed in cortical cells from an hiPSC background differentiated using DSI until DIV 73 (Figures 3A and 3F). HiPSCs differentiated using NGN2 direct reprogramming showed activity much earlier, typically between days 14 and 24 (Figures 3A and 3G). Electrophysiological maturation was monitored with daily activity scans. While max firing rate typically increased and remained relatively stable over time for all cell types during the testing period (Figure 3B), changes were observed in both the mean firing rate (Figure 3C) and variance in firing rate (Figure 3D) over the days of testing; in particular, hiPSCs differentiated using the NGN2 direct reprogramming method showed a considerable increase in mean firing rate and the variance in firing over days of testing.

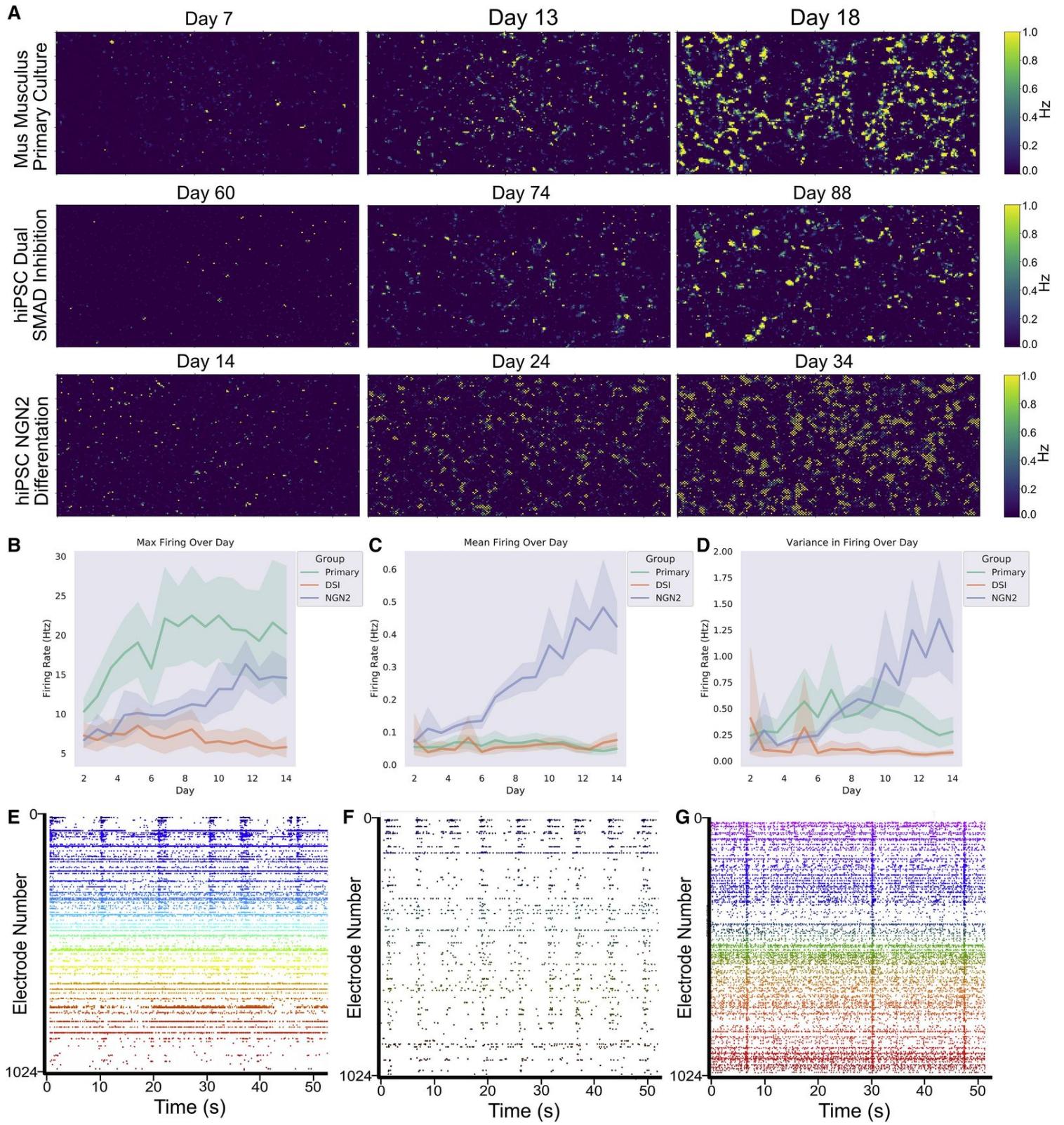


图 3：皮层细胞显示出自发的电生理活动

构建模块化实时平台，利用神经元计算

我们开发了 DishBrain 系统，以利用神经元计算并与嵌入模拟环境的神经元进行交互（STAR方法；图4A；视频S2）。DishBrain 环境是一个低延迟、实时系统，可与供应商 MaxOne 软件进行交互，使其可以以扩展其原始功能的方式使用（图4B）。该系统可以记录神经元培养物的电活动，并通过电生理刺激提供“感觉”（非侵入性）电刺激，与神经网络的活动产生的

动作电位相似 (Ruaro 等 , 2005 年) 。使用 STAR 方法中描述的编码方案 , 外部电刺激传达一系列信息。为了我们的目的 , 我们选择了三种不同的信息类别 : 可预测、随机和感觉 (STAR 方法 , 图 4C) 。 DishBrain (图 S2) 旨在将这些功能集成到一个闭环系统中 , 以 “ 读取 ” 神经元培养物的信息 , 并在实时闭环系统中 “ 写入 ” 感觉数据 , 使神经元 “ 动作 ” 影响未来到来的 “ 感觉 ” 刺激。意图是将 BNNs 体现在虚拟环境中 , 并量化可证明的学习效果。

The DishBrain system was developed to leverage neuronal computation and interact with neurons embodied in a simulated environment (STAR Methods; Figure 4A; Video S2). The DishBrain environment is a low-latency, real-time system that interacts with the vendor MaxOne software, allowing it to be used in ways that extend its original functions (Figure 4B). This system can record electrical activity in a neuronal culture and provide ‘ ‘sensory’ ’ (non-invasive) electrical stimulation comparably to the generation of action potentials by activity in the neuronal network (Ruaro et al., 2005). Using the coding schemes described in STAR Methods, external electrical stimulations convey a range of information. For our purposes, we opted for three distinct information categories: predictable, random, and sensory (STAR Methods, Figure 4C). DishBrain (Figure S2) was designed to integrate these functions to ‘ ‘read’ ’ information from and ‘ ‘write’ ’ sensory data to a neural culture in a closed-loop system so neural ‘ ‘action’ ’ influences future incoming ‘ ‘sensory’ ’ stimulation in real time. The intent was to embody BNNs in a virtual environment and to quantify demonstrable learning effects.

通过 DishBrain 的初步验证是模拟经典街机游戏 “ 乒乓球 ” , 通过在 8 个电极的预定义感觉区域提供输入来实现 (图 4D) 。电极的排列方式允许一种粗略但拓扑一致的位置编码 , 与体内系统一致 (参见 STAR 方法) (Baranes 等 , 2012 年 ; Patel 等 , 2014 年 ; Shlens 等 , 2006 年) 。在实时收集的定义的运动区域的电生理活动中 , 移动一个挡板。如果这种活动没有导致挡板拦截球 , 将提供一个不可预测的刺激 (150mV 电压 , 5Hz , 持续 4 秒 ; 参见 STAR 方法) , 之后球刺激将重新开始 , 并沿着随机矢量进行。相反 , 如果成功拦截发生 , 将在所有电极上同时以 100Hz 持续 100ms 提供可预测的刺激 (短暂中断常规感觉刺激) , 然后游戏会继续可预测地进行。初步调查比较了不同的运动区域配置 , 以验证运动区域设置是否仅通过输入刺激引入偏差 (与球位置对齐的挡板移动) (STAR 方法 ; 图 S3) 。皮质细胞的实验培养物显示了更高的击中 - 失误比率 , 我们将其定义为平均回合长度 , 在平衡的分割运动配置中 (图 4D) , 而作为对照组使用的仅填充介质的 MEA 显示了最小的偏差。不同的区域被定义为 “ 运动区域 ” , 其中运动区域动作 1 的活动将挡板移动 “ 向上 ” , 而运动区域动作 2 的活动将挡板移动 “ 向下 ” 。这种固定布局意味着通过自组织 , 需要采取不同的发射模式 , 这些模式是随机分布的 , 与 “ 运动 ” 配置相对应 , 这引发了一个问题 , 即在多大程度上会发生这种自组织。

The initial proof of principle using DishBrain was to simulate the classic arcade game ‘ ‘Pong’ ’ by delivering inputs to a predefined sensory area of 8 electrodes (Figure 4D). Electrodes were arranged in a manner that would allow a coarse, yet topographically consistent, place coding, consistent with in vivo systems (see STAR Methods) (Baranes et al., 2012; Patel et al., 2014; Shlens et al., 2006). The electrophysiological activity of defined motor regions was gathered—in real time—to move a paddle. If this activity did not result in an interception of the ball by the paddle, an unpredictable stimulus was delivered (150mV voltage at 5Hz for 4 seconds; see STAR Methods), after which time the ball stimulation would recommence on a random vector. In contrast, if a successful interception occurred, a predictable stimulus was delivered across all electrodes simultaneously at 100Hz for 100ms (briefly interrupting the regular sensory stimulation) before the game continued predictably. Preliminary investigations compared different motor region configurations to verify that motor region setup did not introduce bias (paddle movement that aligned to the ball position) from input stimulation alone (STAR Methods; Figure S3). Experimental cultures of cortical cells showed a higher hit-miss ratio, which we defined as the average rally length, on counterbalanced split-motor configurations (Figure 4D), where media-only-filled MEAs used as a control group also showed minimal bias. Distinct areas were defined as ‘ ‘motor regions,’ ’ where activity in motor region action 1 moved the paddle ‘ ‘up’ ’ and activity in motor region action 2 moved the paddle ‘ ‘down.’ ’ This fixed layout means that monolayers of cells—with a random distribution that is arbitrary in relation to the

'motor' configuration—will need to adopt distinct firing patterns through self-organization (and raises the question to what extent this self-organization will occur).

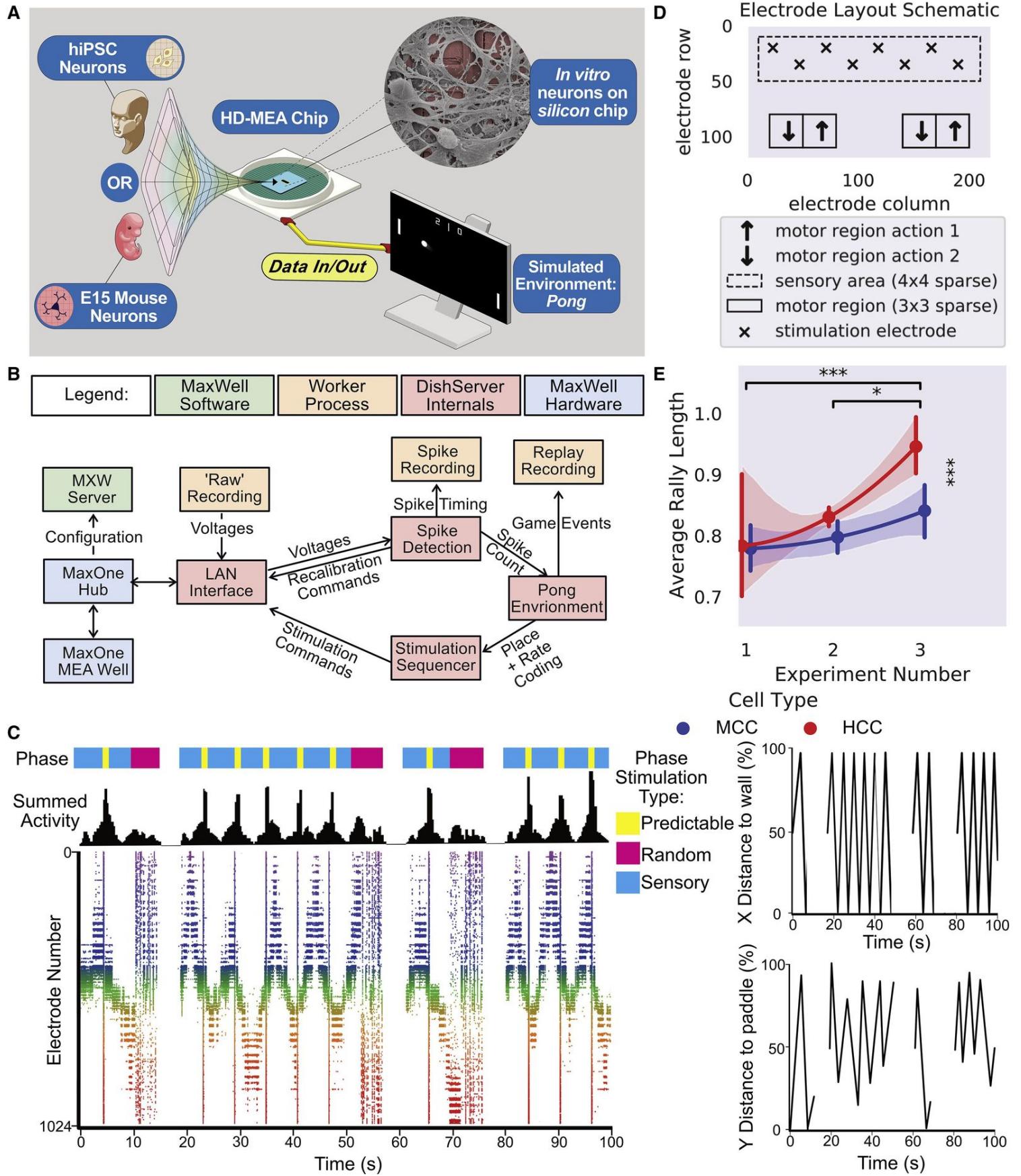


图 4：信息密度不断增加的示意图和试点测试



视频 S2: 交互式SpikeStream可视化工具的代表性电影和系统设置概述，与图 5相关

<https://www.cell.com/cms/10.1016/j.neuron.2022.09.001/attachment/db2f7e55-0486-47dc-a440-86d54864fbc1/mmc3.mp4>

增加感官信息输入密度会提高性能

DishBrain 系统在三个预试验中得到了完善，每个预试验都增加了感官信息的密度。预试验1使用4Hz的刺激，仅涉及位置编码，其中刺激的位置对应于球在y轴上的位置。预试验2探索了不同的配置，并引入了基于活动的权重到运动区域，以解释细胞密度或活动差异。预试验3采用了图4D中的布局，并改用了结合频率(4-40Hz)和位置编码的数据输入方法。这种结合的频率和位置编码在概念上与啮齿动物的桶状皮层有引人注目的生物学相似性，表明这种编码在生理上是连贯的(Harrell等人，2020年；Ly等人，2012年；Petersen等人，2001年)。比较了每种培养类型在最后十五分钟的游戏表现(图4E和表S1)。培养物在第二次和最后一次预试验以及第一次和最后一次预试验之间显示出平均回合长度的显著增加。在不同的培养物之间，人类皮层细胞(HCCs)的平均回合长度显著长于带有小鼠皮层细胞(MCCs)的培养物(表S2)。总体而言，这些结果支持这样的观点：增加感官信息的量可以提高性能，即使在细胞培养特征保持不变的情况下也是如此。

The DishBrain protocol was refined over three pilot studies, each increasing the density of sensory information. Pilot study 1 operated with a 4Hz stimulation that only involved place coding, where the location of the stimulation corresponded to the position of the ball on the y axis. Pilot study 2 investigated different configurations and introduced activity-based weighting to motor regions to account for cell density or activity differences. Pilot study 3 adopted the layout in Figure 4D and changed to the combined rate (4–40Hz) and place-coding method of data input. This combined rate and place coding has compelling biological similarities conceptually to the rodent barrel cortex, suggesting this encoding is physiologically coherent (Harrell et al., 2020; Ly et al., 2012; Petersen et al., 2001). Gameplay for the final fifteen minutes for each culture type was compared (Figure 4E and Table S1). Cultures displayed a significant increase in the average rally length between the second and final pilot studies and the first and final pilot studies. Between cultures, human cortical cells (HCCs) had significantly longer average rally lengths than cultures with mice cortical cells (MCCs) (Table S2). Overall, these results support that increasing the amount of sensory information improved performance, even when cell culture features were kept constant.

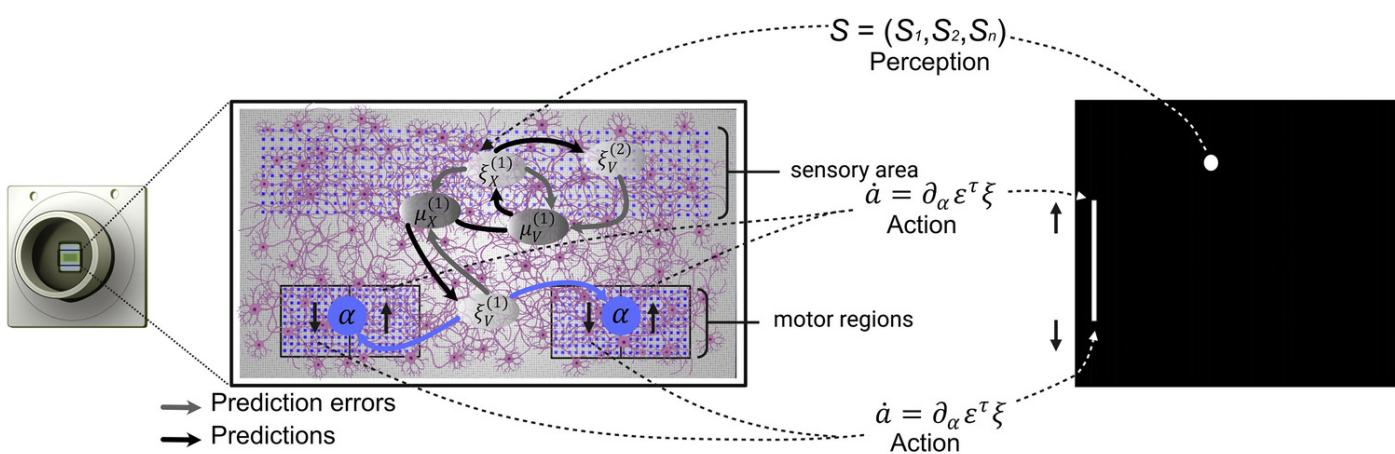
BNNs 在游戏环境中具有实体时，会随时间而学习

为了测试自由能原理(FEP)的预测(图5A)，使用选定的参数(STAR方法)，将皮层细胞(MCCs和HCCs)与仅有培养基的对照组(CTL)进行了比较；还有休息阶段(RST)，在此阶段，活跃的培养物控制了球拍但没有接收到任何感觉信息；以及在硅(IS)对照组，这些对照组模拟了游戏的所有方面，除了球拍是由随机噪声驱动，在399次测试会话中进行了比较(80-CTL[n=6 MEA]，42-RST[n=20培养物]，38-IS[n=3种子]，101-MCCs[n=9培养物]，138-HCCs[n=11培养物])。平均对打长度显示了组和时间(前5分钟和后15分钟)之间的显著交互作用(图5B和表S1)。只有MCC和HCC培养物显示出明显的学习迹象，随着时间的推移对打长度显著增加。此外，在游戏中发现在时间点1(T1)期间观察到关键显著差异(表S1)：HCC组表现明显比MCC、CTL和IS组差(表S2)。这表明HCC在初始体验环境时表现不佳，表明球拍的初始失调控制或可能是一种探索性行为。值得注意的是，在时间点2(T2)时，这一趋势发生了逆转；MCC和HCC组明显优于所有对照组，HCC组与MCC组相比稍微但显著表现更好(表S1和表S2)。这些数据表明实验组中存在显著的学习效应，而在对照组中不存在，同时证明了学习能力在老鼠和人类细胞之间存在差异，符合先前结果(视频S1)。

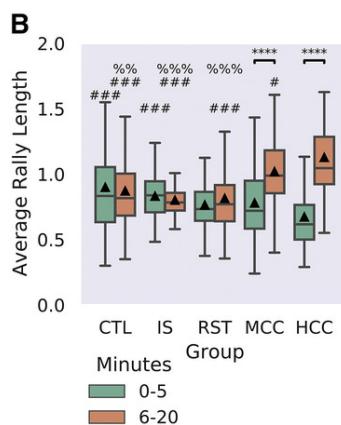
To test the predictions of the FEP (Figure 5A) using selected parameters (STAR Methods), cortical cells (MCCs and HCCs) were compared with media-only controls (CTL); rest sessions (RST), where active cultures controlled the paddle but received no sensory information; and in-silico (IS) controls that mimicked all aspects of the gameplay except the paddle were driven by random noise over 399 test sessions (80-CTL [n = 6 MEA], 42-RST [n = 20 cultures], 38-IS [n = 3 seeds], 101-MCCs [n = 9 cultures], 138-HCCs [n = 11 cultures]). The average rally length showed a significant interaction (Figure 5B and Table S1) between group and time (first 5 and last 15 min). Only the MCC and HCC cultures showed evidence of learning with significantly increased rally lengths over time. Further, it

was found that during gameplay in timepoint 1 (T1), key significant differences were observed (Table S1): the HCC group performed significantly worse than MCC, CTL, and IS groups (Table S2). This suggests that HCCs perform worse than controls when first embodied in an environment, suggesting an initial maladaptive control of the paddle or perhaps an exploratory behavior. Notably, at timepoint 2 (T2), this trend was reversed; the MCC and HCC groups significantly outperformed all control groups along with HCC showing a slight but significant outperformance over the MCC group (Tables S1 and S2). This data demonstrates a significant learning effect in both experimental groups absent in the control groups, along with evidence that the learning capabilities differ between mice and human cells in line with previous results (Video S1).

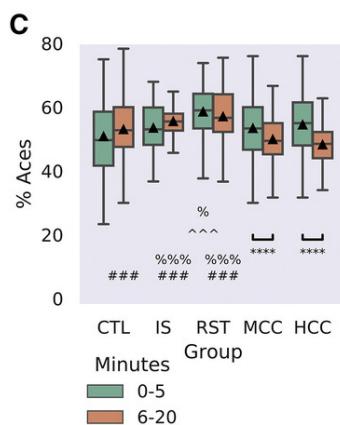
A



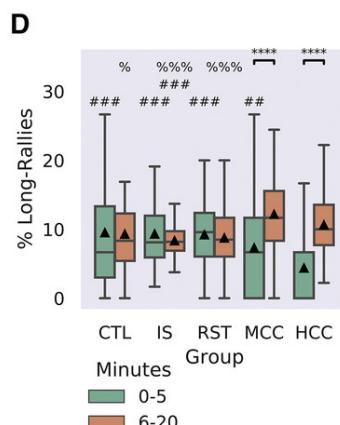
B



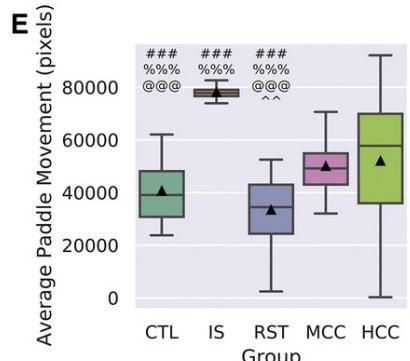
C



D



E



F

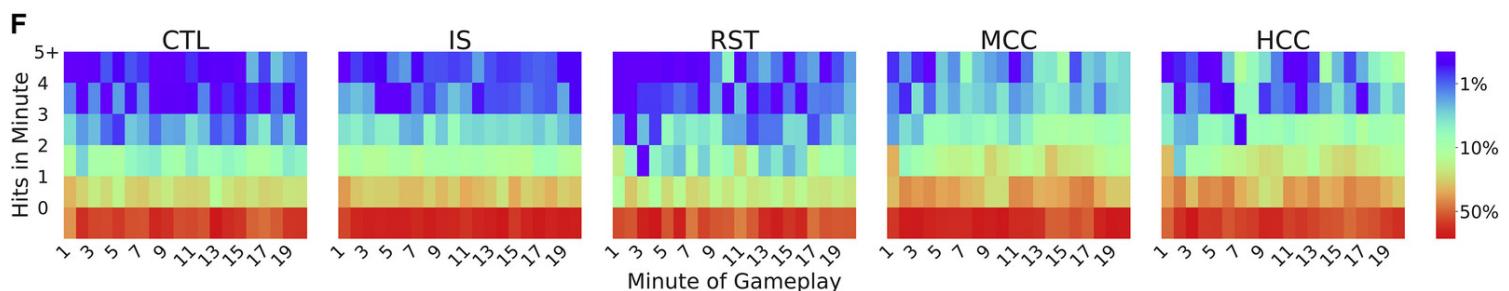


图 5：当嵌入虚拟游戏世界中时，实例化的皮层神经元在《乒乓》游戏中表现出显著提升的性能



视频 S1: DishBrain 系统运行中的代表性影片，与图 1 相关。

<https://www.cell.com/cms/10.1016/j.neuron.2022.09.001/attachment/bb13a45e-c582-424d-b623-426c35683bb3/mmc2.mp4>

BNNs 的学习效果可通过其他指标观察

还计算了其他关键游戏特征，如球拍未能拦截球而没有一击的次数定义为“ace”，以及游戏中连续三次以上击中的次数定义为“长对打”。与平均对打长度一样，对于ace和长对打，组和时间之间的显著交互作用被发现（表S1）。只有MCC和HCC组在T2中显示出与T1相比ace明显减少（图5C和表S2）。同样，只有MCC和HCC组在T2中显示出明显更多的长对打，与第一次相比（图5D和表S2）。总体而言，数据显示实验培养物（HCCs和MCCs）通过减少错失最初发球的频率和实现更多连续击中或更长对打来提高表现。

Other key gameplay characteristics, such as the number of times the paddle failed to intercept the ball without a single hit defined as ‘‘aces,’’ and the number of gameplays with greater than 3 consecutive hits defined as ‘‘long rallies,’’ were calculated. As with average rally length, significant interactions between groups and time were found for aces and long rallies (Table S1). Only the MCC and HCC groups showed significantly fewer aces in T2 compared with T1 (Figure 5C and Table S2). Likewise, only the MCC and HCC groups showed significantly more long rallies in T2 compared with the first (Figure 5D and Table S2). Collectively, the data shows that both experimental cultures (HCCs and MCCs) improved performance by reducing how often they missed the initial serve and achieving more consecutive hits or longer rallies.

在T1时，ace和长对打的组间差异被发现（表S1）。RST条件显示的ace数量明显多于CTL和MCC组（表S2），表明细胞在从游戏中休息时表现出一定程度的偶发行为。当调查T1时的长对打数量时，发现只有HCCs的长对打明显减少（表S2）。这一发现与上述减少的平均对打长度相一致。T2时也发现了ace和长对打的组间显著差异（图5C和5D和表S1）。值得注意的是，HCC组的ace数量明显少于CTL、RST和IS组（表S1）。MCC组与RST和IS组相比也显示出明显较少的ace，但与CTL组相比没有（表S2）。相反，对于长对打，MCC组明显多于CTL、RST和IS组（表S2），但HCC组与IS组相比只显示出明显更多的长对打，而与RST或CTL相比没有（表S2）。

Differences between groups at T1 were found both for aces and long rallies (Table S1). The RST condition displayed significantly more aces than the CTL and MCC groups (Table S2), suggesting a degree of sporadic behavior that the cells exhibit when initially introduced to the rest period from gameplay that results in this behavior. When the number of long rallies at T1 was investigated, it was found that only HCCs had significantly fewer long rallies (Table S2). This finding complements the reduced average rally lengths discussed above. Significant differences between groups at T2 were also found for aces and long rallies (Figures 5C and 5D and Table S1). Notably, the HCC group showed significantly fewer aces than CTL, RST, and IS groups (Table S1). The MCC group also showed significantly fewer aces than RST and IS groups, but not the CTL group (Table S2). In contrast, for long rallies, the MCC group showed significantly more than the CTL, RST, and IS groups (Table S2), yet the HCC group only showed significantly more long rallies compared with the IS group, but not RST or CTL (Table S2).

在电活动不活跃的非神经细胞（HEK293T细胞）和仅介质对照中未发现学习效应（图S4A–S4C）。此外，发现MCCs和HCCs的ace百分比与长对打百分比之间存在显著负相关，表明表现不是由于恶性行为（例如将球拍固定在一个角落）所致（图S4D）。还研究了仅通过刺激可能导致球拍更大移动并导致观察到的学习效应的情况。正如图5E所示，虽然条件之间观察到显著差异（表S1），但对于CTL和RST，这导致相对于其他组显著较低的移动，其中RST是所有组中移动最少的（表S2）。IS对照组显示的球拍移动比所有其他组都要多，但与其他对照组（CTL和RST）的性能指标没有实质性差异（表S2）。此外，图5F强调MCC和HCC在T2中显示出较少的ace和更多的长对打，与T1相比，进一步强调了随时间观察到的学习效应。这也在线性回归中看到（图S4F），只有MCC和HCC组显示出平均对打长度与游戏持续时间之间存在显著正相关。

No learning effect was found in electrically inactive non-neuronal cells (HEK293T cells) and media-only controls (Figures S4A–S4C). Further, a significant negative correlation between percentage of aces and percentage of long rallies of both MCCs and HCCs was found, suggesting that the performance was not arising from maladaptive behavior such as fixing the paddle to a single corner (Figure S4D). Whether stimulation alone may cause greater

movement of the paddle and that this may result in the observed learning effects was also investigated. As Figure 5E shows, while there were significant differences observed in paddle movement between conditions (Table S1), for the CTL and RST, this resulted in significantly lower movement relative to the other groups, with the RST being the lowest movement of all groups (Table S2). The IS control group showed significantly more paddle movement than all other groups yet displayed no meaningfully different performance metrics to the other control groups (CTL and RST) (Table S2). Additionally, Figure S4E shows no significant correlation between paddle movement and average rally length was observed, supporting that movement alone of the paddle does not explain the observed learning effects. Wholistically, Figure 5F emphasizes that both MCCs and HCCs showed fewer aces and more long rallies in T2 compared with T1, reiterating the observed learning effect over time. This can also be seen in linear regressions (Figure S4F), where only the MCC and HCC groups showed a statistically significant positive relationship between average rally length and duration of gameplay.

BNNs 需要反馈才能学习

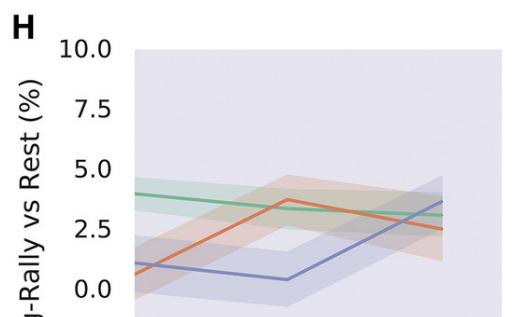
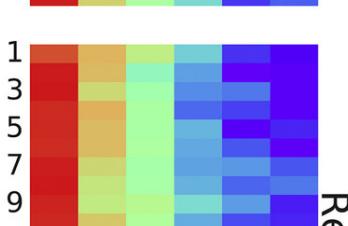
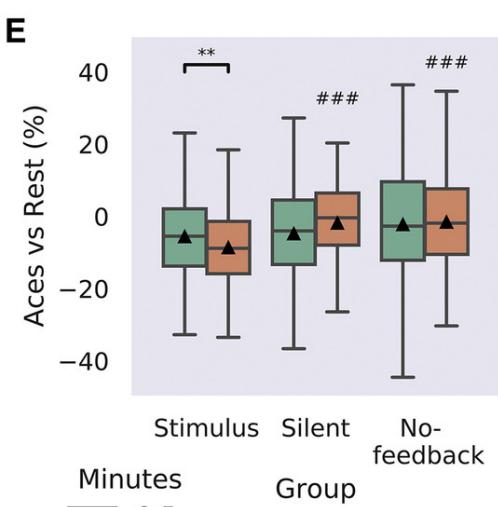
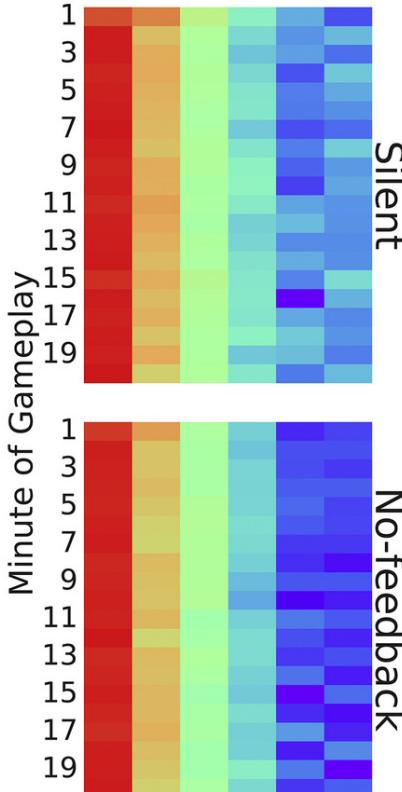
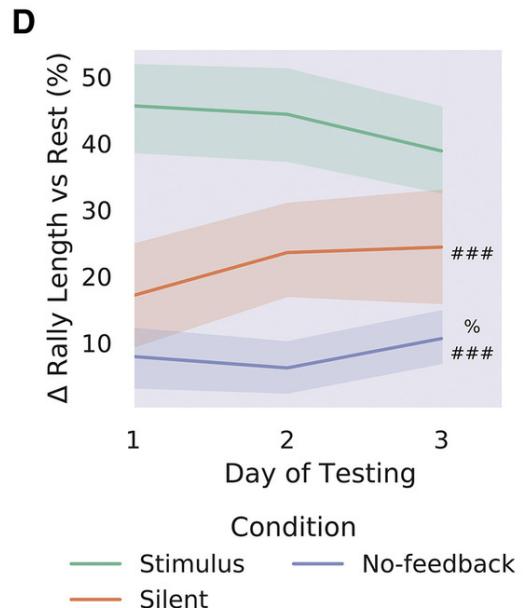
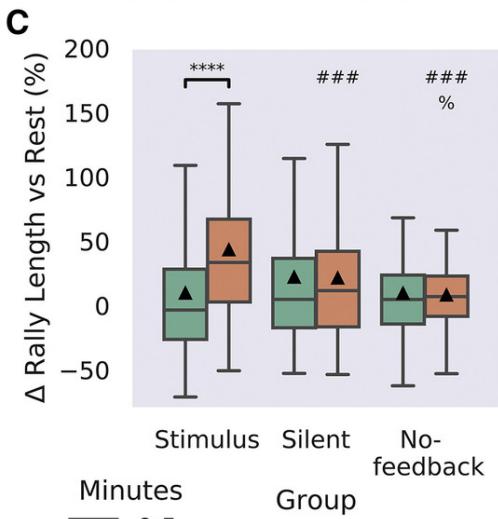
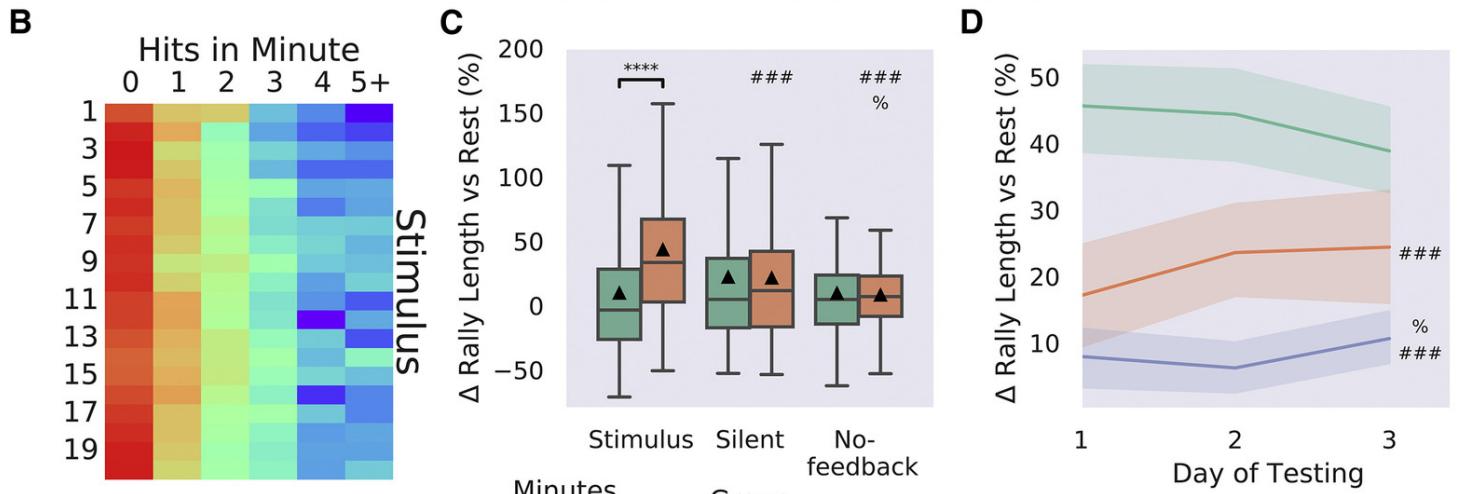
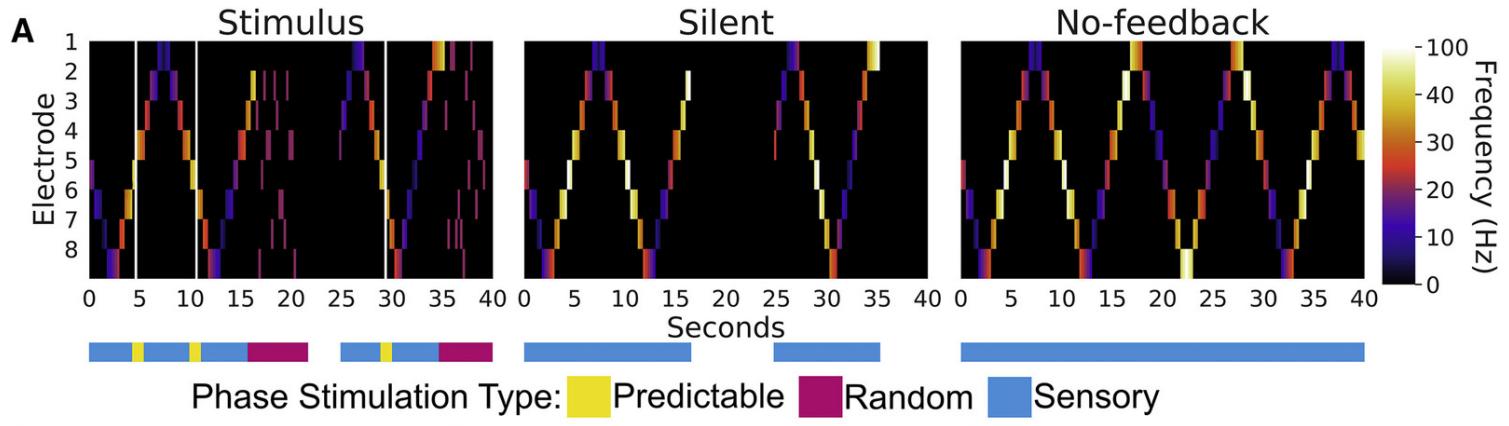
为了研究学习中反馈类型的重要性，培养物，包括 MCCs 和 HCCs，在3天内进行了3种条件的测试，每天3次会话（3局游戏），共进行了486次会话。条件1（刺激；n = 27）模拟了上述使用的情况，即在培养物表现良好或不佳时分别施加可预测和不可预测的刺激。条件2（静默；n = 17）涉及刺激反馈被替换为一个同样长的时间段，这个时间段中所有刺激被停止，时间段结束之后，游戏重新开始，球从中心以随机方向开始移动。条件3（无反馈；n = 15）在错失时没有重新开始。当球拍未成功拦截球时，球会弹起继续移动而不中断；仍提供报告球位置的刺激。这些条件之间的差异在图6A中有所说明。还收集了休息期间的活动，并用于基于每次会话的性能进行标准化，以考虑未刺激活动的差异（图1）。

To investigate the importance of the feedback type for learning, cultures, both MCCs and HCCs, were tested under 3 conditions for 3 days, with 3 sessions per day resulting in a total of 486 sessions. Condition 1 (Stimulus; n = 27) mimicked that used above, where predictable and unpredictable stimuli were administered when the cultures behaved desirably or not, respectively. Condition 2 (Silent; n = 17) involved the stimulus feedback being replaced with a matching time period in which all stimulation was withheld, after which the game restarted with the ball beginning in a random direction. Condition 3 (No feedback; n = 15) removed the restart after a miss. When the paddle did not successfully intercept the ball, the ball would bounce and continue without interruption; the stimulus reporting ball position was still provided. The difference between these conditions is illustrated in Figure 6A. Rest-period activity was also gathered and used to normalize performance per session basis to account for differences in unstimulated activity (Figure 1).

刺激和静默条件显示了比休息和无反馈条件更高的平均对打长度（图6B）。当测试组在平均对打长度上的百分比增加与匹配休息对照组之间的差异时，发现了显著的交互作用（图6C和表S1）。只有刺激条件在随时间的推移中显示出平均对打长度的显著增加。虽然在T1没有发现差异，但在T2中发现了组之间的显著主效应，在那里刺激条件的平均对打长度显著高于静默和无反馈条件（表S2）。有趣的是，静默条件在T2中也明显优于无反馈条件，尽管效果较小（表S2）。重要的是，这表明信息本身是不足够的；学习需要形成一个闭环学习系统。在T2的日历水平上进行后续跟踪（图6D），未观察到随时间的显著差异，但观察到了与上述组间差异相同。在检查T2时的ace的总和（图6E）和测试的各天之间（图6F）时发现，刺激组在T1时显示出比静默和无反馈条件明显更少的长对打，而在T2时，刺激组显示出比无反馈条件更多的长对打（图6G）。在跨天的后续测试中未发现差异（图6H）。综合来看，这些结果表明，BNNs中的自适应行为可以作为与环境互动并隐含地对环境进行建模的自发属性。

Stimulus and Silent conditions showed an overall higher average rally length compared with Rest and No-feedback conditions (Figure 6B). When testing for differences between groups in the percentage increase of average rally length over matched rest controls, a significant interaction was found (Figure 6C and Table S1). Only the Stimulus condition showed a significant increase in average rally length over time. While no differences were found for T1, a

significant main effect of group was found at T2, where the Stimulus condition had a significantly higher average rally length than the Silent and No-feedback conditions (Table S2). Interestingly, the Silent condition also significantly outperformed the No-feedback conditions, although with a smaller effect size (Table S2). Importantly, this demonstrates that information alone is insufficient; feedback is required to form a closed-loop learning system. When followed up at the level of day for T2 (Figure 6D), no significant differences over time were observed, but the same between-group differences as above were observed. This trend was similar when looking at aces both summed (Figure 6E) and across days of testing (Figure 6F). The Stimulus group at T1 showed significantly fewer long rallies compared with the Silent and No-feedback condition, being reversed at T2 with the Stimulus group showing significantly more long rallies compared with the No-feedback condition (Figure 6G). No difference was found when this was followed up across days (Figure 6H). Collectively, these results suggest that adaptive behavior seen in BNNs altering electrophysiological activity can be an emergent property of engaging with and implicitly modelling the environment.



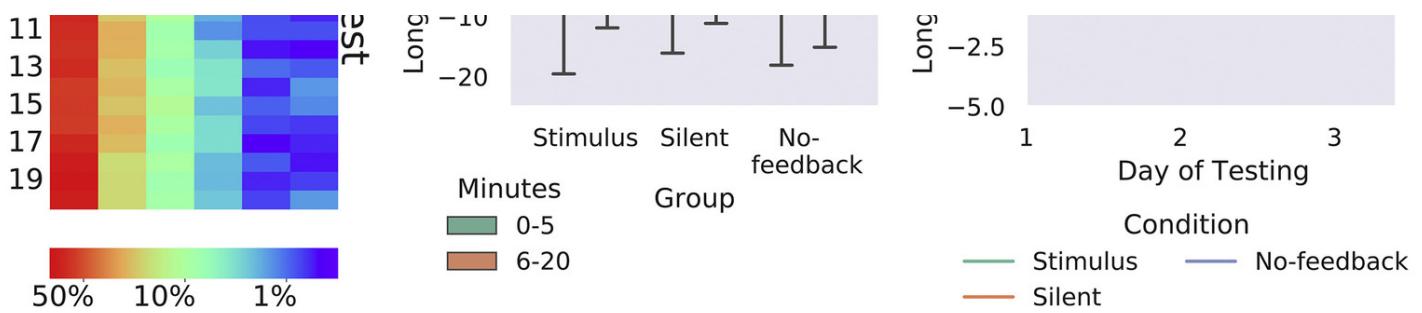


图 6 : 反馈对于学习的重要性

电生理活动的动力学显示出连贯的连通性

我们分析了处于刺激条件下的培养物在游戏过程中的电生理活动，以确定功能连接性（Mohseni Ahooyi等人，2018）。在休息和游戏中，以100毫秒时间段内发射的交叉相关性显示出显著强烈的正相关性，表明感觉区域和两个运动区域的活动之间存在显著的正相关（图7A–7D）。然而，当对这些相关性进行每个时间段的计算并取平均时，在游戏阶段观察到的相关性显著更强，而不是在休息时（图7E）。如果在游戏中感觉区域的活动直接与系统范围动态自组织中的运动区域的活动相关联，那么这种更高程度的连接性是可以预期的。与此一致，当每秒计算独占性运动区域活动的数量时——查找发生在运动区域1或运动区域2中任一区域中的高于噪声水平的活动事件，但不是同时发生在两个区域中——在培养物参与游戏时与休息时相比发现这些事件的显著增加（图7F）。这种内部调制与这些培养物的观察表现是一致的；在运动区域之间的独占性活动变化是适应性游戏所需的。最后，为了进一步支持这些结果，发现两个运动区域之间的相关性随时间变化显著不同（图7G）。在运动区域之间的100毫秒时间段内的相关性的线性回归发现，直到大约5分钟的游戏时间，相关性显著减少（ $R^2 = 0.013$, $F(1, 2049) = 27.51$, $p = 1.72 - 7$, $b = -1.18$, $p < 0.001$ ）。在此点之后，几乎没有进一步的变化（ $R^2 = 0.00$, $F(1, 5181) = 2.19$, $p = 0.139$, $b = -0.55$, $p = 0.139$ ），表明存在一定程度的稳态。这些差异不影响整体平均培养物放电，在整个游戏中保持稳定（图7H）。

Electrophysiological activity during gameplay was analyzed from cultures subjected to the stimulus condition to determine functional connectivity (Mohseni Ahooyi et al., 2018). The cross correlations of firing in 100ms-time bins revealed significant, strong positive correlations between activity in the sensory region and both motor regions during Rest and Gameplay (Figures 7A– 7D). However, when these correlations were calculated per bin and averaged, significantly stronger correlations were observed when cultures were in the Gameplay phase than at Rest (Figure 7E). This higher degree of connectivity would be expected if activity in the sensory region during gameplay was directly related to activity in motor regions through dynamic self-organization at the system-wide level. In line with this, when the quantity of exclusive motor region activity was calculated per second—looking for events where above-noise-level activity occurred in either motor region 1 or motor region 2, yet not both simultaneously—a significant increase in these events was found when cultures were engaged in gameplay versus rest (Figure 7F). This type of internal modulation is coherent with the observed performance of these cultures; exclusive activity changes among motor regions would be required for adaptive gameplay. Finally, to further support these results, the correlation between the two motor regions was found to vary substantially over time (Figure 7G). A linear regression of the correlation in 100ms-time bins between motor regions was found to decrease with time significantly until approximately 5 min of gameplay ($R^2 = 0.013$, $F(1, 2049) = 27.51$, $p = 1.72 - 7$, $b = -1.18$, $p < 0.001$). After this point, little further change was observed ($R^2 = 0.00$, $F(1, 5181) = 2.19$, $p = 0.139$, $b = -0.55$, $p = 0.139$), suggesting a degree of homeostasis. These differences do not affect the overall average culture firing that remains stable throughout the gameplay session (Figure 7H).

由于已经显示神经组织的电刺激可以改变神经元活动（Bakkum等，2008a，2008b；Chao等，2008），因此在游戏中培养物的功能可塑性与休息时进行了评估，如STAR方法中所述。图7I表明，在游戏中进行的闭环训练显示出与训练前休息时

基线可塑性相比显着增加的可塑性，表明在游戏过程中可塑性得到提升（表S1）。为了测试学习是否反映了BNNs内部自由能的降低，我们使用神经元响应的信息熵作为平均惊异（也称为自信息）的代理，这是由VFE上限界定的（请参见STAR方法）。我们预测在游戏过程中的学习过程中信息熵会降低。我们进一步预测在不可预测（随机）反馈后会出现熵增加，反映和随后的“惊异”状态（以及隐含的高VFE）相对于反馈之前的状态。在图5中报告的研究中，发现游戏过程中的平均信息熵低于休息时的信息熵，无论是在不可预测的反馈刺激之前还是之后（图7J和表S1）。在游戏过程中发现的平均信息熵在反馈后相对于反馈前的时间点有显著增加，但在休息时对应的时间点没有反馈，信息熵得到显著降低。由于熵的变化可以取决于反馈前的感觉活动水平，我们通过脉冲数对平均信息熵进行了标准化。这种关系得到了保留（图7K和表S1），在游戏过程中观察到了标准化平均熵的显著增加，但在休息时对应的时间点没有刺激。简而言之，如理论预测的，游戏过程中通过降低信息熵来进行可预测的与环境交换，而在游戏过程中不可预测的反馈会增加熵。

As electrical stimulation of neural tissue has been shown to modify neuronal activity (Bakkum et al., 2008a, 2008b; Chao et al., 2008), the functional plasticity of cultures during Gameplay was assessed compared with when at Rest as described in STAR Methods. Figure 7I suggests that closed-loop training during Gameplay displays significantly increased plasticity compared with baseline plasticity measured at Rest before training, indicating that functional plasticity was upregulated during gameplay (Table S1). To test whether learning reflects a reduction in VFE within BNNs, we used the information entropy of neuronal responses as a proxy for the average surprise (a.k.a. self-information), which is upper-bounded by VFE (see STAR Methods). We predicted a reduction in information entropy during the learning of gameplay. We further predicted an increase in entropy following unpredictable (random) feedback, reflecting and ensuing state of ‘surprise’ (and, implicitly, high VFE), relative to pre-feedback states. For the studies reported in Figure 5, the mean information entropy was found to be lower during Gameplay than during Rest, both before and after the unpredictable feedback stimulation (Figure 7J and Table S1). There was a significant increase in mean information entropy found post-feedback relative to pre-feedback timepoints during Gameplay, but not in the corresponding timepoints during Rest where no feedback occurred. As the change in entropy can depend on the level of sensory activity pre-feedback, we normalized the mean information entropy by the number of spikes. The relationship was conserved (Figure 7K and Table S1), where a significant increase in normalized mean entropy was observed during Gameplay, but not at the corresponding timepoint during Rest where no stimulation occurred. In short, as predicted theoretically, gameplay reduced information entropy during predictable exchanges with the environment, while unpredictable feedback increased entropy during gameplay.

我们在图6中报告的不同反馈机制的后续研究中重复了这项分析。重要的是要注意，培养物的内部信息熵不一定直接与施加到培养物中的外部（即感觉）信息熵相关联，看到培养物如何对不同的反馈协议做出反应是有趣的。如图7L所示，在标准刺激条件下，标准化平均信息熵的变化被复制（表S1）。有趣的是，在静默条件下，神经培养物的标准化平均信息熵甚至比刺激条件在反馈后更高。然而，在无反馈条件下，相对于应用反馈时期，标准化平均信息熵没有变化，而在反馈后期有显著较低的得分（表S2）。

We repeated this analysis on the follow-up study of different feedback mechanisms reported in Figure 6. While it is important to note that the internal information entropy of the culture is not necessarily and directly tied to the external (i.e., sensory) information entropy of the stimulus being applied into a culture, it is interesting to see how cultures respond to different feedback protocols. As shown in Figure 7L, the change during the stimulus condition between the normalized mean information entropy was replicated for the standard Stimulus condition (Table S1). Of interest is the finding that during the Silent condition, the neural cultures had a higher normalized mean information entropy than even the stimulus condition post-feedback. However, the No-feedback condition showed no change relative to the period when feedback would have been applied, with a significantly higher normalized mean information entropy score than either of the other two conditions pre-feedback, yet a significantly lower score post-feedback (Table S2).

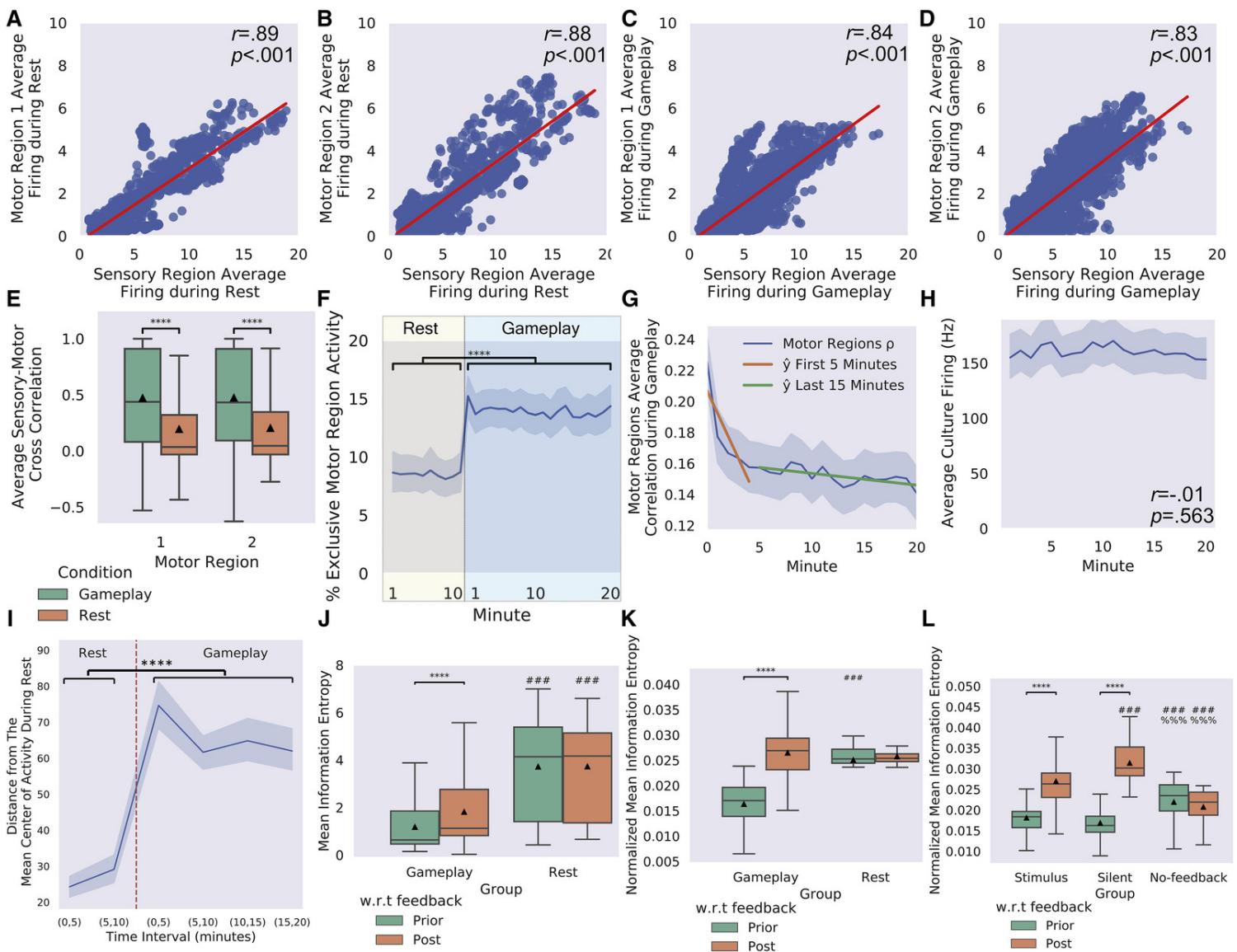


图 7：游戏和休息期间的电生理活动

电生理活动与更高的平均连续命中次数相关

我们对关键电生理活动指标和平均对打长度进行了初步未校正的Pearson相关分析。发现平均对打长度与平均发射率（图8A）和最大放电频率（图8B）之间存在显著正相关。同样，感觉区域与运动区域1（图8C）和2（图8E）之间的交叉相关性与表现显著正相关，进一步表明强大的连接性与更好的游戏表现相关。为了进一步研究活动的拓扑分布与表现之间的关系，使用四个离散余弦变换（DCT）系数的绝对值，这些系数被归一化为平均活动，用于总结自发活动的空间模式并评估活动的对称性（图8E）。DCT 0,1，用于测量水平平面上的活动（图8F），和DCT 2,0，用于测量水平边缘与水平中心的活动差异（图8I），与平均对打长度之间存在显著负相关。然而，DCT 0,2显示垂直边缘和垂直中心的活动差异（图8G），以及DCT 1,0，用于测量垂直平面上的活动（图8H），则没有显著相关。鉴于配置布局，可以认为游戏表现与电生理活动的对称性偏差密切相关。为了确认对称性的重要性，对游戏过程中的两个运动区域的电生理活动进行了分析，并计算了远离对称性的标准化偏差。与平均对打长度呈显著负相关，任何超过约1个偏差的不对称性似乎完全阻止了高于控制组观察到的表现（图8J）。这表明，如果细胞培养质量不均匀，那么培养物能够自组织自发活动的程度存在一定限制。最后，与上述结果一致，游戏中感觉区域（图8K）、运动区域1（图8L）和运动区域2（图8M）的活动水平与更高的平均对打长度相关联。

Exploratory uncorrected Pearson's correlations were computed for key electrophysiological activity metrics and average rally length. A significant positive correlation was found between average rally length with mean (Figure

8A) and max (Figure 8B) firing. Likewise, the cross-correlations with the sensory region for both motor region 1 (Figures 8C) and 2 (Figure 8E) were significantly positively correlated with performance, further suggesting that robust connectivity is linked with better gameplay outcomes. To further investigate whether the topographical distribution of activity correlated with performance, the absolute values of four discrete cosine transform (DCT) coefficients normalized to mean activity were used to summarize spatial modes of spontaneous activity and assess the symmetry of activity (Figure 8E). DCT 0,1, which measures activity across the horizontal plane (Figure 8F), and DCT 2,0, which measures activity on the horizontal edge versus the horizontal center (Figure 8I), were significantly negatively correlated with average rally length. Yet, DCT 0,2, which shows difference between activity on the vertical edges and the vertical center (Figure 8G), and DCT 1,0 which measures activity across the vertical plane (Figure 8H), did not significantly correlate. Given configuration layout, it is coherent that gameplay performance is closely linked to deviations in symmetry of electrophysiological activity. To confirm the importance of symmetry, gameplay electrophysiological activity was analyzed for both motor regions, and the normalized deviation away from symmetry was calculated. As deviation away from symmetry resulted in a significant negative correlation with the average rally length, any asymmetry exceeding approximately 1 deviation appeared to completely prevent performance above that observed in controls (Figure 8J). This suggests a limit to which cultures can self-organize spontaneous activity if cell culture quality is uneven. Finally—in line with the results above—higher activity in the sensory region (Figure 8K), motor region 1 (Figure 8L), and motor region 2 (Figure 8M) during gameplay was also correlated with higher average rally lengths.

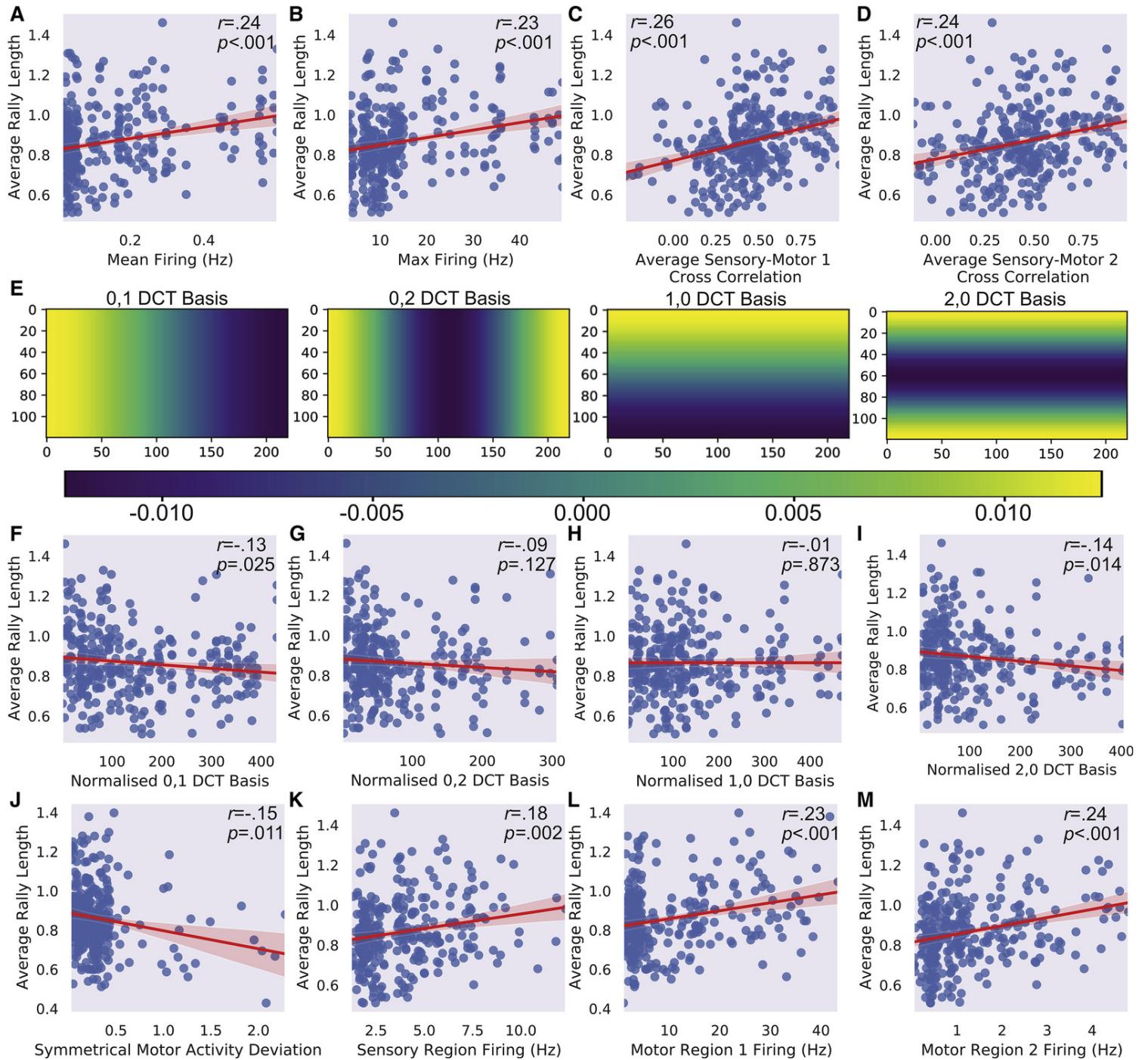


图 8：电生理活动与平均对打长度之间的关系

讨论

在这里，我们介绍了 DishBrain 系统，这是一个能够在虚拟环境中体现来自各种来源的BNNs并实时测量它们对刺激的反应的系统。神经元，尤其是在组装中，对外部刺激做出自适应响应的能力在体内已经被充分确立，因为这构成了所有动物学习的基础 (Attinger等, 2017)。然而，这项工作是首次在体外为目标导向行为建立了这种基本行为。我们能够利用这个硅-生物系统来研究生物神经计算的基本原理。简而言之，我们介绍了第一个能够实时展示自适应行为的SBI设备。这个系统本身提供了扩展以前神经行为的硅模型的机会，比如测试海马和环回细胞模型在解决空间和非空间问题时的情况 (Whittington等, 2020)。对 DishBrain 平台、选定的细胞类型、药物管理和反馈条件进行轻微变化，将使体外测试能够获取关于细胞如何处理和计算以前无法获得的信息的数据。

Here, we present the DishBrain system, a system capable of embodying BNNs from various sources in a virtual environment and measuring their responses to stimuli in real time. The ability of neurons, especially in assemblies, to respond to external stimuli adaptively is well established in vivo as it forms the basis for all animal learning (Attinger et al., 2017). However, this work is the first to establish this fundamental behavior in vitro for a goal-directed behavior. We were able to use this silico-biological system to investigate the fundamentals of biological neuronal computation. In brief, we introduce the first SBI device to demonstrate adaptive behavior in real time. The system itself offers opportunities to expand upon previous in silico models of neural behavior, such as where models of hippocampal and entorhinal cells were tested in solving spatial and non-spatial problems (Whittington et al., 2020). Minor variations on the DishBrain platform, selected cell types, drug administration, and feedback conditions would enable an in vitro test to garner data on how cells process and compute information that was previously unattainable.

最重要的是，这项工作展示了在为BNNs创建闭环环境方面的重大技术进步（Bakkum等，2008a；Chao等，2008；Wagenaar等，2004）。我们强调了在神经系统中实现目标导向学习所需的体现性。这在实验中的相对表现中得到体现，更密集的信息和更多样化的反馈影响了性能。同样，当没有提供反馈但提供了关于球位置的信息时，培养物表现出显著较差的性能和没有学习。特别值得注意的是，当刺激性反馈被移除并替换为静默反馈（即暂时移除所有刺激）时，培养物仍能够超越没有反馈的情况，尽管程度较小。一种解释是，打“Pong”产生了比不打“Pong”更可预测的结果，通过减少不确定性。请注意，“失误”导致不可预测的结果，因为球重新开始并其随后的运动是不可预测的。就传递的刺激的信息熵而言，虽然不可预测的刺激会产生高熵，但是静默条件仍相对于成功的游戏具有更高的熵，因为球以随机方向重新开始。这与我们的结果一致，即结果越不可预测，观察到的学习效果越大——因为BNN学会避免不确定性。

Most significantly, this work presents a substantial technical advancement in creating closed-loop environments for BNNs (Bakkum et al., 2008a; Chao et al., 2008; Wagenaar et al., 2004). We have emphasized the requirement for embodiment in neural systems for goal-directed learning to occur. This is seen in the relative performance over experiments, where denser information and more diverse feedback impacted performance. Likewise, when no feedback was provided yet information on ball position was available, cultures showed significantly poorer performance and no learning. Of particular interest was the finding that when stimulatory feedback was removed and replaced with silent feedback (i.e., transient removal of all stimuli), cultures were still able to outperform those with no feedback as in the open-loop condition, albeit to a lesser extent. One interpretation is that playing ‘‘Pong’’ generates more predictable outcomes than not playing ‘‘Pong’’ by reducing uncertainty. Note that a ‘‘miss’’ results in unpredictable outcomes because the ball resets and its subsequent motion is unpredictable. In terms of the informational entropy of the stimulus being delivered, while an unpredictable stimulus would have high entropy, the silent condition still entails higher entropy relative to successful play as the ball restarts in a random direction. This is consistent with our results, as the more unpredictable an outcome, the greater the observed learning effect—as the BNN learns to avoid uncertainty.

然而，值得注意的是，BNN活动的内部信息熵并不完全反映外部刺激的信息熵：虽然不可预测的刺激增加了内部熵，但静默条件反馈也是如此。然而，为了使BNN能够根据反馈改变活动，必须对其感觉输入进行可观察的系统变化，这可以与其先前的活动相关联。这与开环/无反馈条件中的学习缺失一致，因为这种情况本质上不提供学习机会，并且同样显示出比其他两种反馈条件更高的内部信息熵。这支持了一个观点，即单纯的刺激是不足以推动学习的：必须有一种影响（外部）可观察刺激的学习行为的动机。面对不可预测的感觉输入时，成功地打“Pong”可以作为一种最小化自由能的解决方案。即使系统的内部信息熵在反馈后增加并且外部信息熵更低（例如，静默反馈），这可能不会提供相同的学习动力。这些发现与马尔可夫链的提议作用一致，为系统提供了一个统计边界（马尔可夫边界），将其分隔为内部状态和外部状态。（Kirchhoff等，2018；Palacios等，2020）。然而，简单地最小化熵（即平均惊异）可能提供了对自适应行为过于简化的解释：主动推理的一个关键方面是选择最小化在执行该

行动后预期的惊异或自由能的行动。虽然这些结果很有趣且支持性，但并不具有决定性，未来的工作需要探索BNN行为与生成模型。

It is interesting to note, however, that the internal information entropy of BNN activity does not exactly mirror the information entropy of the external stimulation: while the unpredictable stimulus increased internal entropy, so did the Silent condition feedback. However, for a BNN to alter activity in response to feedback, there must be a change to its sensory input observable by the system that can be associated with its previous activity. This is consistent with the absence of learning in the open-loop/No-feedback condition, which by its nature affords no opportunity for learning, and likewise showed higher internal information entropy than the other two feedback conditions. This supports the thesis that stimulation alone is insufficient to drive learning: there must be a motivation for learning behaviors that influence the (external) observable stimulus. When faced with unpredictable sensorium, playing ‘‘Pong’’ successfully acts as a free energy-minimizing solution. Even if the internal information entropy of a system is increased following feedback and has lower external information entropy (e.g., silent feedback), this may not provide the same impetus for learning. These findings accord with the proposed role of a Markov blanket, providing a statistical boundary of the system to separate it into internal and external states (Kirchhoff et al., 2018; Palacios et al., 2020). Yet simply minimizing entropy (i.e., average surprise) may offer an overly simplified account of adaptive behavior: a key aspect of active inference is the selection of actions that minimize the surprise or free energy expected on following that action. While these results are interesting and supportive, they are not conclusive, and future work is required, including exploring BNN behavior with a generative model.

从机制上讲，我们试图通过测试支持 FEP 的主动感知的基本原则来展示 DishBrain 的实用性。最接近的先前工作在神经培养物中进行了盲源分离的研究，但是在没有生理合理的训练的开环环境中进行了研究 (Isomura等 , 2015 ; Isomura和Friston , 2018)。我们展示了在出现 “违背训练目标” 的结果之后提供不可预测的感觉输入，以及在 “符合训练目标” 的结果之后提供可预测的输入，显著地塑造了神经培养物的实时行为。可预测的刺激也可以被视为稳定突触权重的过程，这与先前的研究一致，因为已经证明更高的放电率增加了短期和长期增强 (Pariz等 , 2018 ; Zhu等 , 2015)。另一方面，不可预测的刺激可以被视为，通过破坏不可取的自由能极小值来破坏连接性。这些结果可以被理解为相互作用层之间的赫布学习和稳态可塑性之间的动态交互的一部分，这可能导致增加在某些刺激模式后活动的可能性 (Ly等 , 2012 ; Pariz等 , 2018 ; Toyoizumi等 , 2014)。这与游戏中观察到的增加的功能可塑性一致，这与休息时相比。这可能是FEP对生物自组织的解释的一个潜在机制，有时以自组织不稳定性的形式讨论为 “自动退化” (Friston等 , 2012)。

Mechanistically, we sought to demonstrate the utility of the DishBrain by testing base principles that underwrite active sensing via the FEP. The closest previous work examined blind source separation in neural cultures, yet did so in an open-loop context without physiologically plausible training (Isomura et al., 2015; Isomura and Friston, 2018). We show that supplying unpredictable sensory input following an ‘‘undesirable’’ outcome and providing predictable input following a ‘‘desirable’’ one significantly shapes the behavior of neural cultures in real time. The predictable stimulation could also be read as a process of stabilizing synaptic weights in line with previous research as it has been shown that higher firing rates augment shortand long-term potentiation (Pariz et al., 2018; Zhu et al., 2015). In a complementary fashion, the unpredictable stimulation could be seen by destabilizing connectivity by destroying undesirable free energy minima. These results could be understood as part of a dynamic interaction between layers of interacting Hebbian and homeostatic plasticity that could lead to increasing the likelihood of activity following certain stimulation patterns (Ly et al., 2012; Pariz et al., 2018; Toyoizumi et al., 2014). This accords with the increased functional plasticity observed during gameplay versus during rest. This may be a potential mechanism behind the FEP account of biological self-organization, sometimes discussed in terms of self-organized instability termed ‘‘autovitiation’’ (Friston et al., 2012).

来自人类和小鼠细胞来源的活跃皮层培养物显示出与先前研究一致的同步活动模式 (Kamioka 等 , 1996 ; Sakaguchi 等 , 2019 ; Shi 等 , 2012 ; Wagenaar 等 , 2006)。重要的是 , 观察到了细胞来源之间的显著差异 , HCCs 在游戏特性方面平均优于 MCCs (带有细微差别)。尽管需要进一步的工作 , 因为这一发现是研究目的的辅助性发现 , 但这是第一项发现 , 提供了支持人类神经元比啮齿类神经元具有更优信息处理能力的假设的实证证据 (Beaulieu-Laroche 等 , 2018 ; Mihaljević 等 , 2020)。先前的研究提出 , 与小鼠细胞相比 , 人类细胞中的生物物理结构将产生不同的输入 - 输出特性 , 从而可能解释不同的计算能力 (Poirazi 和 Papoutsi , 2020)。在系统的初始开发阶段 , 我们无法实际和经验地测试所有关键方面 , 比如细胞亚型的差异、微观细胞结构或内神经元密度的差异。然而 , 未来的研究有机会专注于阐明这些差异。本文中描述的 DishBrain 系统可能为准确评估神经计算能力的差异提供了第一个途径 , 使其成为未来研究的一个激动人心的领域。

Active cortical cultures, from both human and mouse cell sources, displayed synchronous activity patterns in line with previous research (Kamioka et al., 1996; Sakaguchi et al., 2019; Shi et al., 2012; Wagenaar et al., 2006). Importantly, significant differences between cell sources were observed, with HCCs outperforming MCCs (with nuances), on average, in gameplay characteristics. Although further work is required as this finding was auxiliary to the aim of the study, this is the first work finding functional, albeit preliminary, empirical evidence supporting the hypothesis that human neurons have superior information-processing capacity over rodent neurons (Beaulieu-Laroche et al., 2018; Mihaljević et al., 2020). Previous work has proposed that biophysical structures in human cells compared with mouse cells would yield different input-output properties and may thereby explain different computational capacities (Poirazi and Papoutsi, 2020). When focusing on the initial development of the system, we could not feasibly and empirically test all key aspects, such as differences in cell sub-types, microscopic cell structure, or interneuron density. However, the opportunity exists for future studies to focus on elucidating these differences. The DishBrain system described in this work potentially offers the first avenue to accurately assess differences in neurocomputational ability, making this an exciting area of future research.

这项工作的另一个发现涉及固有的细胞网络组织 , 可以在运动区域的定义中看到。我们的早期试点研究 , 以及该领域的先前工作 (Bakkum 等 , 2008a) , 基于网络活动扫描绘制了运动区域。然而 , 我们对感觉和运动区域在培养物之间固定时自组织能够如何适应感兴趣。我们的发现表明 , 虽然活跃细胞的自组织活动可以发生 , 但当活动细胞在 MEA 上不均匀分布时 , 这种自组织受到限制。游戏过程中的活动变化与过去的工作一致 , 显示环境和行动之间的反馈对于体内神经发育是必要的 (Attinger 等 , 2017)。观察到的变化还表明 , 也许这种发展是基于细胞水平固有的特性发生的。虽然这些结论是暂时性的 , 因为控制实验之间的刺激统计数据有所不同 , 但数据突出了未来研究的方向。对于闭环环境对学习的重要性的进一步实验应包括增加读取神经活动并影响环境之间的延迟 , 或使用与环境脱钩的刺激。然而 , DishBrain 系统及其技术的未来改进提供了探索网络动态以更好地理解自组织的这一方面 , 并包括对 BNN 结构组织的研究的机会。

Another finding from this work relates to innate cell network organization, seen in the definition of motor regions. Our early pilot studies, along with previous work in this field (Bakkum et al., 2008a), mapped motor regions based on network activity scans. However, we were interested in the extent that self-organization would adapt if sensory and motor regions were fixed between cultures. Our findings demonstrate that while significant self-organization of activity can occur, this was limited when active cells were not evenly distributed across the MEA. The changes in activity during gameplay are consistent with past work showing that feedback between environment and action is required for proper *in vivo* neural development (Attinger et al., 2017). The observed changes also suggest that perhaps this development occurs based on properties inherent at the level of the cell. While these conclusions are tentative as the statistics of stimulations do differ between control experiments, the data does highlight future research directions. Further experimentation on the extent that the closed-loop environment is important for learning should include increasing the delay between reading neural activity and having it influence the environment or using stimulation decoupled from the environment. Nonetheless, the DishBrain system and future improvements of this technology do provide the opportunity to explore network dynamics to better understand this aspect of self-organization and include investigations into structural organization of BNNs.

由于目前硬件的限制，感觉刺激与甚至简单的体内生物相比要粗糙得多。这意味着在实时中，不可能区分神经元体或树突区域的刺激，这两者都可能被刺激。同样，在实时计算中，不可能分离来自不同神经结构的电学变化的处理，例如区分来自细胞体与树突的动作电位。改进这两个领域是未来研究的关键方向。此外，对于试图执行类似任务的体内生物而言，例如本体感知，或者将闭环系统解耦以测试时间延迟的影响是不可行的。此外，嵌入在单层格式中的相对较少细胞数量意味着驱动这种行为的神经架构在可能的连接数量方面非常简单，与拥有 3D 脑结构的小生物相比。然而，仅仅使用可预测和不可预测刺激的简单模式，该系统能够在几分钟的时间内展示系统化的行为。虽然会话内学习已经得到充分确立，但在多天之间的会话间学习并不稳健。培养物似乎会在每个新会话中重新学习关联。鉴于选择了皮质细胞，这是可以预料的，因为体内皮质细胞并不专门用于长期记忆（Rolls, 2018）。未来使用该系统可以研究其他神经元细胞类型和/或更复杂的生物结构的使用。

Due to current hardware limitations, the sensory stimulation is much coarser compared with that for even simple *in vivo* organisms. This meant that it was not possible to distinguish, in real time, between stimulation of neuronal somatic or dendritic domains and that both were likely stimulated. Likewise, it was not computationally possible in real time to separate processing electrical changes from different neuronal structures such as discriminating between action potentials from the soma versus dendrites. Improving both areas is a key direction for future research. Additionally, it was infeasible to meaningfully implement mechanisms that would be crucial for an *in vivo* organism attempting a comparable task, such as proprioception, or to decouple the closed-loop system to test the impact of time delays. Moreover, the relatively small number of cells embedded in a monolayer format means the neural architecture driving this behavior is incredibly simple in terms of the number of possible connections available compared with even small organisms that have a 3D brain structure. Nonetheless, using only simple patterns of predictable and unpredictable stimulation, this system was able to show systematic behavior in an order of minutes. While within-session learning was well established, between-session learning over multiple days was not robustly observed. Cultures appeared to relearn associations with each new session. Given that cortical cells were selected, this is to be expected as *in vivo* cortical cells are not specialized for long-term memory (Rolls, 2018). Future work with this system can investigate the use of other neuronal cell types and/or more complex biological structures.

结论

神经细胞显示出具有时间发展规律的特征性自发动作电位神经细胞显示出具有时间发展规律的特征性自发动作电位通过使用这个 DishBrain 系统，我们已经证明了一层体外皮质神经元可以自组织活动，展示出在模拟游戏世界中体现的智能和有感知的行为。我们已经展示，即使在没有对细胞活动进行实质性过滤的情况下，也可以观察到神经元培养物在它们感知的世界中随着时间的推移和对多个对照实验的比较中呈现出统计上显著的差异。这些发现展示了一个有希望的 SBI 系统的演示，该系统随着时间的推移以输入为导向以系统化的方式学习。该系统提供了一个完全可视化的学习模型的能力，可以开发独特的环境来评估BNNs正在执行的实际计算。这是长期以来一直在寻求的，并且超越了纯粹的硅模型或单纯预测分子途径的范围（Karr等，2012；Whittington等，2020；Yu等，2018）。因此，这项工作提供了可以用来支持或挑战解释大脑如何与世界互动以及智能的理论的实证证据（Friston，2010；Schwartz，2016）。最终，尽管仍需要大量的硬件、软件和湿件工程来改进 DishBrain 系统，但这项工作确实展示了活的神经元具有学习适应能力，能够与它们的感觉器官进行积极交流。这代表迄今为止实现 SBI 的最大进步，以响应外部定义的目标导向行为。

Using this DishBrain system, we have demonstrated that a single layer of *in vitro* cortical neurons can self-organize activity to display intelligent and sentient behavior when embodied in a simulated game-world. We have shown that even without a substantial filtering of cellular activity, statistically robust differences over time and against multiple controls could be observed in the behavior of neuronal cultures in their sensed world. These findings provide a promising demonstration of an SBI system that learns over time in a systematic manner directed by input. The system provides the capability for a fully visualized model of learning, where unique environments may

be developed to assess the actual computations being performed by BNNs. This is something that is long sought after and extends beyond purely in silico models or predictions of molecular pathways alone (Karr et al., 2012; Whittington et al., 2020; Yu et al., 2018). Therefore, this work provides empirical evidence that can be used to support or challenge theories explaining how the brain interacts with the world and intelligence in general (Friston, 2010; Schwartz, 2016). Ultimately, although substantial hardware, software, and wetware engineering are still required to improve the DishBrain system, this work does evince the computational power of living neurons to learn adaptively in active exchange with their sensorium. This represents the largest step to date of achieving SBI that responds with externally defined goal-directed behavior.

STAR★Methods

详细方法请参阅本文的在线版本，包括以下内容：

- 关键资源表
- 资源可用性
 - 主要联系人
 - 材料可用性
 - 数据和代码可用性
- 实验模型和受试者细节
 - 伦理声明
 - 动物繁殖和养护
 - 干细胞系
 - 干细胞生长和养护
- 方法细节
 - 主要细胞培养
 - 干细胞双SMAD分化
 - 干细胞NGN2直接分化
 - HEK293T细胞培养
 - MEA设置和准备
 - 在MEA上培养和维护细胞
 - 测量电生理活动
 - DishBrain软件平台
 - 游戏环境的表示
 - 与Maxwell API的接口
 - 初始试验
 - 使用EXP3算法进行试验
 - 输入配置
 - 输出配置

- 量化和统计分析
 - 样本量和盲化方案
 - 免疫细胞化学
 - 扫描电子显微镜
 - 宽场荧光显微镜
 - 数据分析
 - 信息熵的计算
 - 功能可塑性的计算
- 其他资源
- 附加资源 (补充信息)
 - 补充信息可在以下网址找到 : <https://doi.org/10.1016/j.neuron.2022.09.001>