**传递性，灵活性，联合认知和海马体的关联**

摘要：提出了一种计算神经网络模型，解释了海马如何对老鼠中观察到的传递性推断性能做出贡献（Dusek和Eichenbaum，1997。Proc Natl Acad Sci USA 94：7109-7114; Van Elzakker et al。，2003。Hippocampus 12：这个问题）。 与通过灵活地将先前编码的记忆相关联来强调海马体的贡献的现有理论相反，我们发现海马体通过在学习期间改变个体刺激元素的元素相关权重来做出贡献。 我们使用此模型来计算一系列现有数据，并进行一系列独特的预测，明确对比这两种观点。

**介绍**

老鼠，鸽子和灵长类动物表现出所谓的传递性推论。在日常用语中，当一个人被告知John比比弗雷德高的比特高时，使用传递推理，然后从逻辑上推断约翰也比弗雷德高。如在实验室中研究的那样，在受试者接受关于共同刺激的一系列同时辨别问题（例如，A B，B C，C D，D E）的训练之后出现这种现象，其中A B，B C，C D，D E指定奖励选择，并指定无奖励选择。在这样的训练之后，然后用新颖的组合BD测试受试者，并且当受试者可以“推断”因为B被选择而不是C，并且选择C而不是D时，证明了传递性推论，那么必须选择B而不是D在神经完整的受试者中已经多次获得该结果（例如，von Fersen等人，1991; Dusek和Eichenbaum，1997; Van Elzakker等人，2003）。

本文讨论的核心问题是海马损伤老鼠可以解决训练问题，但BD转移试验失败（Dusek和Eichenbaum，1997）。 Dusek和Eichenbaum（1997）将他们的结果解释为强烈支持海马体为表现灵活性提供基质的观点 - 存储在海马体中的信息可以在新的情况下被检索和适当使uiloll 用。具体而言，他们提出完整的老鼠编码了问题的各个要素A B C D E的关系表示，这使得能够在它们之间进行比较。因此，如果面对新颖的组合BD，老鼠比较B和D在有序表示上的位置，并推断如果B C和C D，则B D。这种比较导致B的选择。

然而，我们在我们的同伴文章（Van Elzakker et al。，2003）中指出，有充分的理由质疑基于五个问题辨别集（AB，BC，CD， DE，EF）。这个问题的五个版本允许两个传递性测试，即BD和BE。然而，我们的老鼠在用BE测试时仅表现出传递性。相比之下，代表性的灵活性理论似乎可以预测：当用新的对测试时，老鼠应该表现出传递性。我们对这些结果的分析使我们得出结论，传递性不是由表达灵活性所需的推理过程所介导的，而是由于老鼠根据各个测试线索的绝对联想兴奋强度做出选择的结果。关联强度的概念代表刺激引起选择反应的能力。

从本质上讲，我们的分析意味着先前的结论是老鼠正在使用推理过程来做出选择，因为新的测试不满足选择线索具有相同的联想强度的要求，因此没有提供传递推理的真实检验。我们的理论与解决这个问题的其他理论家有共同点（例如，von Fersen等，1991; Siemann和Delius，1998）。

我们的分析和结论提出了传递理论的两个问题。首先，为什么对一组前提对（AB，BC，CD，DE，EF）的训练导致表面上应具有相等值的单个刺激之间的关联强度的分级水平（例如，B，D和E） ）？其次，海马体如何促成这一结果？本文的目的是解决这两个问题。为此，我们使用了先前应用于各种现象的海马和新皮质的基于生物学的模型（O'Reilly和Rudy，2001）。

本文分几个阶段进行。首先，我们确立了这样的立场：传递性是分级关联强度与在获取过程中形成的个体线索的产物，而不是表示灵活性的产物。然后，我们详细描述如何实现计算模型。然后我们展示（1）该模型产生可以调节传递性的个体线索的分级联想强度，以及（2）海马体如何有助于分级联想强度。我们还从模型中推导出几个可测试的预测，并最后讨论了我们的模型与其他视图的关系。

**关联强度的梯度**

表现灵活性理论假设非语言主体的传递行为是这种类型的逻辑推理的产物：如果是A>B和B>C，那么A>C。因此，对该假设的有效检验需要个体刺激（例如，B和D）的奖励关联值必须相等。否则，选择对象可能取决于这些元素不相等的关联值，而不是依赖灵活性地推导。这正是边缘刺激A和E不用于传递性测试的原因——A总是得到奖励而E永远不会得到奖励，因此对于老鼠来说选择A而不是E是显然的。

然而，相关文章的结果（Van Elzakker，2003）和von Fersen先前对于鸽子实验的发现表明了内部刺激（包括B和D）具有相同的关联强度的假设是错误的。相反，这些发现提出了一个更简单的解释：在测试训练中基础的判别训练程序产生了我们称之为联想强度的梯度（与von Fersen，1991; Siemann和Delius提出的类似观点一致）1998。这意味着训练会为个体刺激产生不等的奖励关联值。正是刺激关联价值的不同，而不是逻辑推理，决定了传递性测试中的选择行为。具体来说，我们表明A的强烈奖励关联（总是得到奖励）可以有效地渗透到B中（如下一节中更多的机械术语所解释的），这样B也有一些积极的联想。类似地，与E的负相关（从未得到奖励）可以渗透到D中，这样它也具有负关联。因此，当给出BD测试时，受试者跟AE情况一样不依赖于比较过程，因为单个元素的关联强度足以确定Bd的选择。我们指A和E这类重头刺激产生的渗透效应为锚定效应。

当用5对判别对（A B，B C，C D，D E，E F）训练老鼠时，在同类文章中出现了对关联强度梯度概念的经验支持。该训练在两组内部刺激对（BD和BE）之间进行了两种传递性测试。代表性灵活性理论预测受试者相比于D和E，应该选择B。但是，老鼠并没有这样做：B优于E，但不优于D。尽管这些结果难以与代表性灵活性理论协调，但它们和关联强度梯度的观点一致（参见von Fersen等，1991）。简单地说，B的关联强度值大于E，但与D相近。特别地，我们认为F的负关联强度渗透到了E，但不足以渗透到D。

因此，我们的计算模型必须实现两个目标：（1）它必须解释训练如何产生与数据一致的关联值梯度，以及（2）它必须阐明海马体对传递性的贡献。表现灵活性理论认为其贡献是在测试期间运行的检索过程。我们的理论将显示在训练阶段，海马体在各个元素的关联值被学习时就已经对传递性作出了贡献。

**传递推理范例在模型中的实现**

我们通过提供老鼠解决前提判别问题的各种过程的高级描述来开始我们对模型的解释。然后，我们在海马和新皮质的计算神经网络模型中形式化这些过程（O'Reilly和Rudy，2001）。

**实验范例**

在实际的实验中（Dusek和Eichenbaum，1997; Van Elzakker等，2003），在两个相互靠近的沙子填充杯中给老鼠提供了两种不同的气味。奖励被埋在其中一个杯子中，老鼠可以通过从沙子中挖出它来获得它。因此，当老鼠接近一个杯子时，它有两种选择：它可以挖掘或转移到另一个杯子。我们使用五个问题（前提）集（A B，B C，C D，D E，E F）来模拟问题，因为这五个问题集允许多个传递性测试：BD，BE和CE。

我们假设老鼠表现出的选择行为由两个学习过程确定：（1）刺激选择过程，和（2）响应选择过程（图1）。例如，当老鼠遇到第一个锚选择刺激（AB）时，它将对两种气味刺激进行大致相等的采样组合，并对接近哪种刺激进行初步评估。据推测，这种评估是基于哪种刺激更有可能提供奖励。这是刺激选择过程。如果老鼠接着接近了A，则与刺激B相比，它将增加从A接受的刺激，我们将其表示为Ab（其中较近的刺激是大写的，而另一个是小写的）。相反，如果它接近B，它将经历Ba情况。然后响应选择过程对输入的这种配置进行操作：老鼠决定在近的刺激处挖掘奖励或切换到其他刺激（在这种情况下，老鼠在这个其他刺激的杯中挖掘）。因此，在Ab的情况下，老鼠可以在A处挖掘或在切换到b处挖掘，并且强化结果将是这种选择的直接结果（即，它将在A处挖掘而不是切换到b）。在训练期间经常观察到转移行为 - 老鼠接近一个杯子然后不进行挖掘，而是选择在另一个杯子中挖掘。

这两个过程通过五个训练对（A B，B C，C D，D E，E F）有差异地进行，这给老鼠带来了质量上不同的问题。锚定问题（A B和E F）可以仅仅通过刺激的差异强化历史来解决，因此更加依靠刺激选择过程。在A B情况下，A总是得到奖励，但B仅在一半试验中得到奖励，因此老鼠可以相对容易地选择在刺激选择过程中接近A。类似地，在E F情况下，E在一半试验中得到奖励，但F从未得到奖励，因此老鼠只需通过学习在刺激选择阶段避免F来解决这个问题。这些锚定问题可与内部问题（B C，C D，D E）形成对比。因为对于这些内部问题中的每个个体刺激，老鼠相同概率地得到奖励，所以老鼠似乎没有直接的解决方案。因此，刺激选择过程不能解决这些问题，而是将其留给响应选择过程，以便最终决定挖掘的位置。

总之，锚定问题和内部问题在质量上有所不同，因为锚定问题允许元素解，因此可以简单地基于附加到各个元素的关联值来解决。相反，除非主体构建了能够解决其联想价值模糊性的刺激选择的联合表示，否则内部问题无法解决（Rudy和Sutherland，1995; O'Reilly和Rudy，2001）。总之，刺激选择过程充当针对相对容易的锚定问题的优化机制（即，导致更快的性能），但不针对内部问题。这个事实将在我们的模拟中发挥关键作用。

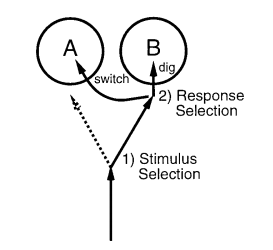


图1。模型中表示的两个选择阶段。 在刺激选择阶段，选择两个刺激中的一个接近。 在响应选择阶段，决定挖掘所选（接近）刺激或切换到其他（远）刺激。 A和B代表两种气味。

**刺激选择过程的模拟**

我们将刺激选择和响应选择过程分开，将刺激选择作为一种机制来控制输入模式的概率，然后将其呈现给我们的新皮质和海马的标准神经网络模型（O'Reilly和Rudy，2001）。 因此，神经网络模型构成响应选择机制（如下一节所述），并且刺激选择机制简单地控制呈现给网络的输入模式的频率。 例如，刺激选择过程可以将A B问题变成两个不同的响应选择问题，Ab（A比b更接近）或Ba（B比a更接近）（图2）。 与我们对老鼠行为的分析一致，我们将这些模式中每一种的刺激选择概率基于两种选择刺激的相对强化值：



因此，刺激选择过程在A B问题中选择Ab的概率仅仅是A [Rew（A）]的相对奖励强度，其由A和B的总奖励强度归一化。

最初，所有刺激的奖励强度都设置为0.5，因此概率也是0.5。由于老鼠的模拟神经网络在选择不同刺激时经历不同的奖励模式，因此奖励强度被修改，进而改变了问题的两个替代呈现的概率。具体地，当任何给定的试验得到奖励时，近端和远端刺激的奖励值都被修改。例如，如果刺激选择过程在A B问题上的选择接近A刺激（Ab），并且响应选择过程选择响应于Ab而进行挖掘A并获得奖励。此奖励将增加A [Rew（A）]的奖励强度一个小的固定金额（0.035），并将B的奖励强度减去相同的金额（即，选择A和避免B的决策都会得到奖励，因为两者都导致相同的结果，并且两方没有偏向性）。根据等式0，这将增加Ab下次出现的概率（并降低Ba的概率）。同样，如果刺激选择过程选择Ba并且响应过程决定转换而不是挖掘，那么这也会得到回报并且奖励值会发生一样地改变。如果没有奖励发生（因为不正确的挖掘或转移），那么我们不改变这些值。我们还尝试了一种替代方案，其中值在与奖励案例相反的方向上进行调整，但是幅度只有一半，并且结果没有显着差异，因此我们使用了更简单的奖励模型。总而言之，元素X的奖励强度的变化是



为了了解为什么刺激选择是模型的重要部分，我们推断出A B和E F的锚定情况会发生什么。因为A总是得到奖励，所以奖励力Rew（A）将接近1。Rew（B）将保持在0.5左右，因为它有无奖励概率相等。预计结果是刺激选择过程中会有三分之二的次数倾向于接近A（即，产生Ab）。相反，Ba的选择只会在三分之一的时间内发生。最终结果是响应选择模型主要学习挖掘Ab模式，并且在面对Ba时不必学习更多的切换。因此，模型更容易学习在B C问题中挖掘Bc模式，因为它不必克服这种强烈的偏置来响应B刺激而切换。于是，刺激选择提供了锚定效应对邻近刺激的“渗透”手段。在E F情况下，这些效果甚至更强，因为F的奖励强度为0，而E保持在0.5左右。这意味着在经过足够的训练后，几乎100％的时间将选择Ef。这相当于大鼠总是避免F，这比Ab的情况更极端。因为老鼠仍然有三分之一的时间接近B。因为内部问题（BC，CD和DE）的元素组成部分通常与奖励和非奖励等概率关联，所以不会出现主要的刺激选择模式，并且与每个问题相关的两种模式将等概率地呈现给响应选择网络。

虽然刺激选择过程对于模拟的采集阶段至关重要，但是在测试阶段，当模型应用于BD和BE时，它不会起作用。因此，例如，当老鼠面对BE时，测试性能由网络在呈现Be模式时产生挖掘响应和在呈现Eb时产生切换响应的趋势来确定。这仅仅是因为响应选择模型没有随机分量并且可以通过单个测试进行评估，而刺激选择过程是随机的并且需要采样以最小化测量中的附加噪声。此外，因为测试元素通常具有大致相等的关联值，所以刺激选择过程在任何情况下都不会显着影响结果。

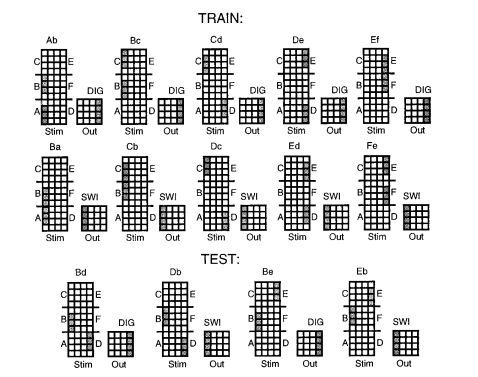


图2。训练模式：刺激输入对和训练有素的输出响应。 每个刺激具有四个相关联的单位，较低的三个单位代表刺激的近端版本（刺激选择过程最初接近的刺激），并且上三个单位代表远端版本。

对于近端刺激，模型必须使用Dig或Switch作出响应。 请注意，每个刺激的两个版本重叠，只有一位信息区分近距离和远距离表示。

**响应选择的实现**

响应选择过程通过神经网络模型（图3）模拟，该模型通过接收选定的输入模式，并产生挖掘或开关输出模式来作为刺激选择过程的接口（图2）。这个模型是我们海马和新皮层的“标准模型”（O'Reilly等，1998; O'Reilly和Munakata，2000; O'Reilly和Rudy，2001; Rudy和O'Reilly，2001）该模型的基本机制吸收了在一个被称为Leabra的连贯框架中的被广泛接受的皮质功能概念（O'Reilly，1998; O'Reilly和Munakata，2000）。海马组件使用这些相同的机制，同时结合了许多理论家所强调的海马的专门解剖学和生理学（Marr，1971; McNaughton和Morris，1987; Rolls，1989; O'Reilly和McClelland，1994; McClelland等）。 al。，1995; Hasselmo，1995）。

该模型学习将输入模式与输出模式相关联。这种学习由Hebbian和error-driven学习相结合，这是Leabra算法的核心特征。error-driven学习使用双向激活传播来传递错误信号，并且与突触修改的已知属性兼容（O'Reilly，1996; O'Reilly和Munakata，2000）。皮质具有刺激输入的元素表示（图3中的层Elem）和可以学习刺激元素的更高阶表示的关联层（图3中的层Assoc）。 Error-driven学习可以响应许多学习试验的明确任务要求，在关联层中形成表示来编码刺激的联合表示（O'Reilly和Rudy，2001），这是学习内部问题所需要的。输出层从元素层和关联层接收，并且具有两个有效模式：激活层右侧的所有单元意味着“挖掘”，而左侧意味着“切换”。响应层提供简单的缩减用于编码响应的分布式输出层的输出：输出层本身必须被分配以在海马系统中提供实质表示。

海马系统从整个皮质系统（包括输出层）接收输入（通过内嗅皮层（EC）层EC\_in），并将相互投影发送回这些相同的层（通过层EC\_out）。海马体区域（齿状回（DG），区域CA3和区域CA1）在EC输入上形成联合编码，并且可以通过将输入提示的模式与输出层上的激活模式相关联来贡献输出响应。 。 DG，CA3和CA1层的稀疏激活产生了该系统编码刺激连接的自然趋势（Marr，1971; O'Reilly和McClelland，1994; O'Reilly和Rudy，2001）。区域CA3的循环附属物支持模式完成，其中部分输入模式（例如，Ab输入模式）可以触发先前联合编码的整个相关模式的调用（例如，“挖掘”输出模式）（McNaughton和Morris， 1987; Rolls，1989）。

最初通过提出试验块中的问题来训练模型：A B块，然后是B C块，依此类推E F。在模型在阻塞阶段表现良好之后，问题被随机交错。该训练序列捕获了老鼠使用的训练范例的本质（Dusek和Eichenbaum，1998; Van Elzakker等，2003）。我们的基本模拟使用了两个块，每个块有七个试验，然后是交错条件下的10个时期（通过所有训练案例）。训练后，用新的刺激组合（BD，BE和AF）测试模型。为了分析它们对性能的影响，每块的试验次数，块总数和交错试验的数量都是不同的。完整的模型和缺乏海马电路的不完整模型都经过训练。

至关重要的是要认识到模型的性能，即其输出表示（图3），是由三个系统的综合影响决定的：（1）元素皮层，（3）联想皮层，和（3）海马。我们模型的结果可以理解为架构的这个特征与在整个网络中运行的错误驱动学习的交互的产物。请注意，响应层是驱动此学习的错误信号的来源。如果给定试验没有错误，则不存在这种学习形式。由于模型的输出由三个系统的总和决定，因此任何一个系统产生的正确输出将消除所有系统所经历的误差信号。换句话说，网络可以表现出类似于众所周知的Reslola和Wagner（1972）Pavlovian调节模型所描述的阻挡效果。该模型假设关联强度的增量是训练试验中出现的所有线索的总关联强度与无条件刺激（US）将支持的关联强度的渐近水平之间的差异的产物。在两个刺激AB无条件地放在一起时（Kamin，1968），当A由于先前训练的刺激阻止老鼠感应另一个刺激元素（B）时，阻滞效应会发生。因为无条件情况下A已经成功被预测，所以能够增加关联强度的预测误差被删除。因此，对于这一步训练线索，B的关联强度几乎没有增加。

模拟的主要目标是了解海马体如何促进传染性。为了预览结果，模拟表明海马体的输出产生了“阻塞效应”，减少了模型中其他系统学到的东西。

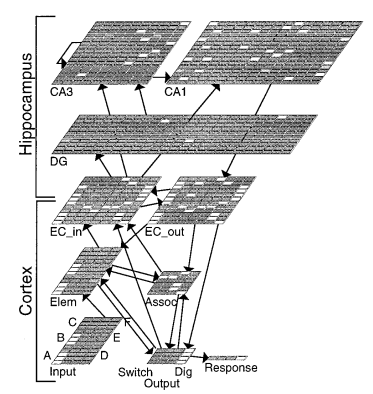


图3。响应选择过程的神经网络模型，具有皮质和海马组分。 显示的激活模式（白色的活动单位）显示Ab输入模式，模型的响应为Dig。

**基本传递性发现的模拟**

我们首先描述完整和海马损伤模型在基本传递推理任务中的表现。图4显示了在使用BD，BE和AF探针刺激测试模型之前，在交错训练的最后阶段期间完整模型的性能。损伤的模型性能与完整模型无法区分。请注意，虽然内部训练问题需要联合表示，但是受损模型解决了这些问题。这个结果已经在早期的模型中得到了证明（O'Reilly和Rudy，2001）。此外，两种模型都显示出强大的锚定效应，其性能更好地解决了A B和E F的锚定问题（几乎是完美的）。请注意，AB与EF问题的不对称性与前面描述的刺激选择不对称性一致，因为B具有与A竞争的一些奖励值，但是F具有零奖励值，因此不提供与E的竞争。这些结果是相似的在老鼠和鸽子中发现的那些（Dusek和Eichenbaum，1997; von Fersen等，1991; Van Elzakker等，2003）。

图5显示了完整和损伤模型在三个转移问题上的性能：BD，BE和AF。图5展示了两个重要现象：

1。无论是完整的还是损伤的模型都没有证明对BD问题的传递性推论 - 他们选择B超过D仅几乎不超过50％的机会水平。这模拟了Van Elzakker等人的研究结果。 （2003）并清楚地表明，在逻辑推理方面，人们无法解释这项任务的表现。

2。完整模型很好表现了BE问题的传递性，选择B的概率很大。相反，受损网络对BE问题的抉择概率相近。这一结果模拟了Dusek和Eichenbaum（1997）发现的海马结构损伤的老鼠在传染性测试中偶然发挥作用的现象。应当理解，Dusek和Eichenbaum在四个问题集（A B，B C，C D，D E）上训练老鼠并在BD上测试它们。然而，就像E是用于训练模型的五个问题集中的锚定问题E F的成员一样，D in是Dusek和Eichenbaum使用的四个问题集中的锚定问题D E的成员。我们将证明锚定问题是如何学习的，这对于理解海马体的贡献至关重要；因此，根据我们的模型，Dusek和Eichenbaum的BD测试在功能上与我们的BE测试一样。

图5还显示完整和损伤模型在AF试验中大概率地选择A。该结果正确地模拟了Van Elzakker和Dusek、Eichenbaum（1997）等人报道的结果（2003）。因为与A和F相关的训练历史不同，所以这是预期的结果；对A的选择总是得到奖励，但选择F却从未得到奖励。

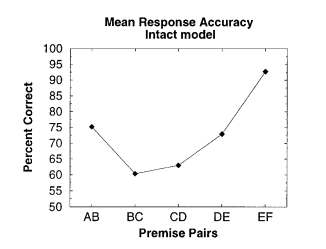


图4。交错训练期间的错误。 注意，性能曲线与Eichenbaum研究中的行为老鼠具有相同的形状。 具体而言，交错训练期间的表现最适合E对F对; 由于锚定效应较大，因此冲突最小。

**初步结果分析**

在测试试验中，该模型被新的刺激组合进行测试。因此，测试对的联合表示是不可用的，并且我们假设模型必须使用元素关联系统的输出来生成其选择。因此，AF，BD和BE测试案例的性能一定决定于各个元素与挖掘或转移的关联程度。

**联想权重**

为了检验响应其他测试案例的基础，我们检查了元素刺激表示与挖掘或开关的输出响应之间的连接强度（图6）。当一个元素在刺激选择过程中被选中时（即紧密刺激），通过比较刺激表示与挖掘响应和转移响应之间的强度差异，我们可以评估模型涉及这些刺激时的测试表现。对于刺激的远表示，发现了重量的互补模式（即，接近A与挖掘响应强烈相关，而远离A与转移响应强烈相关）。

我们可以根据这些联想直接解释上面列出的两个主要发现：

BD性能差：完整和损伤的模型都具有D刺激的模糊权重 - 它与挖掘或开关响应没有强烈关联，因此可以产生任一响应。尽管B刺激与挖掘更强烈相关，但这还不足以克服对D刺激的可变响应。因此，两个网络都没有在BD测试案例上表现出可靠的“传递推理”性能的整体模式。

差分BE性能：完整网络与E刺激中的挖掘权重相比产生更强的切换。这意味着它将倾向于在BE测试中避免E，并且如所观察到的选择B而不是E。相反，受损网络与挖掘和切换具有更加模糊的关联，仅具有弱的整体切换关联。因此，它在BE测试中表现不佳。此外，完整模型产生比损伤模型更强的B挖掘关联，这也有利于B超过E的传递推断选择。因此，理解海马系统对这个问题的贡献的关键是理解它如何有助于更强大的Eswitch和Bdig联系

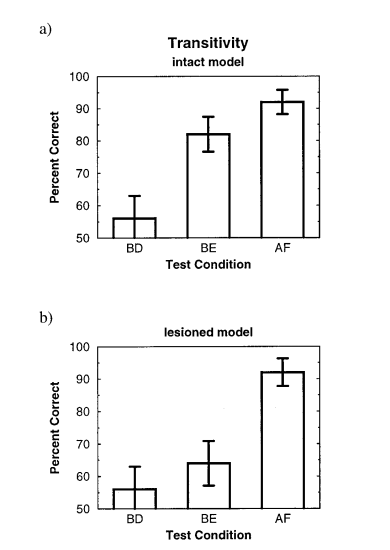


图5。完整（a）和损伤（b）模型的传递推理传递对性能。 两种模型都在BD上表现近，并且在AF上接近完美（由于A总是得到奖励而F从未得到过，因此得到强有力的基本关联支持）。 关键的发现是完整模型在BE传递性测试中的表现要比损伤模型好得多。

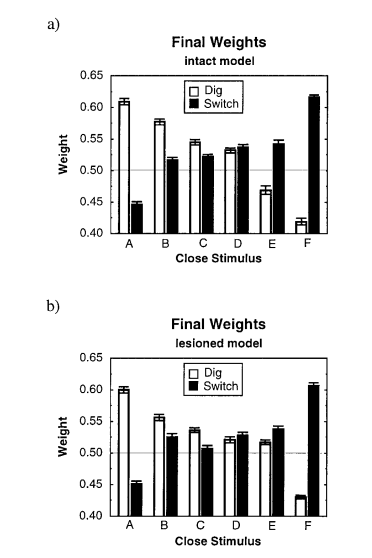


图6。在完整（a）和损伤（b）模型训练后，近端刺激中挖掘和切换的平均最终重量。 两种模型在BD上表现不佳可以通过注意模型喜欢挖掘和切换到D刺激来产生随机行为，即使它稍微倾向于挖掘B。 BE问题的差异性能出现了 因为完整模型从E学习净开关关联（即，开关明显强于挖掘），这有利于选择B超过E。相反，受损模型学习大致相同的强挖掘和切换关联E。同样，完整模型 与损伤模型相比，B开发了更强的挖掘关联，这也在完整模型中产生更好的BE测试性能。

**海马阻滞效应**

我们注意到阻断效应对于理解海马体如何有助于老鼠显示传递性的能力至关重要。根据我们对权重的分析，我们可以得出结论，这种阻塞效应主要发生在EF试验上(当网络正确地选择挖掘E时)：E和dig之间的本该产生的强烈关联被阻断效应阻止。然而受损模型没有这种阻塞效应，因此它开发了更强的E-dig关联，从而阻止它在BE测试中选择Be。此外，刺激选择过程会使老鼠更频繁地接近E（因为F永远不会得到奖励），因此E-dig关联会有很多机会得到加强。

为什么海马体会阻止这个E-dig关联？首先，与皮层不同，海马体快速自动地构建所有刺激模式的联合表示（O'Reilly和Rudy，2001）。所以，即使解决E F锚点问题时不需要联合表示，完整模型的海马系统也会自动生成一个。例如，当刺激选择过程向网络呈现Ef模式时，完整的海马体构造该模式的联合表示并将其与挖掘响应相关联。因此，由于这种海马联合表示需要在这些试验中产生正确的响应，它将阻止使得E-dig正向关联的错误信号（图7）。

除了这种基本的阻塞逻辑之外，结合E还是D E训练对的成员这一事实，还能得到两个重要的暗示。首先，DE训练模式会提高E-switch关联程度，将与E F训练产生的E-dig关联相冲突。在完整神经网络中，当E-dig关联被阻止时，这会为E生成E-switch关联。其次，这个D E案例使得E的刺激强度本身模糊不清，这在一个错误驱动的学习规则下会使网络更强烈地依赖于Ef的海马联合表现，从而增强阻挡效果。

总之，E-dig关联的学习被海马体对Ef和挖掘间的联合表示所阻断。受损模型不会自动构建Ef的联合表示，因此不存在阻塞的来源。这样不同关联权重然后导致不同的结果，完整模型会明显地选择Be，而损伤模型则不然。

可以应用相同的推理来理解为什么完整模型具有比损伤模型更强的B-dig关联。具体而言，海马体快速学习A+ B-问题，因此阻止B-switch（远离A）的关联。这种阻塞会允许来自B C问题的B-dig关联占主导地位，产生观察到的更强的B-dig关联。

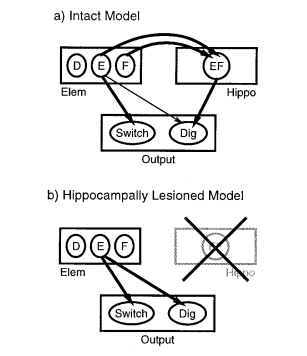


图7。 a：EF训练案例的快速海马编码可以阻止E和挖掘输出响应之间的元素关联（由较细的线显示）的图示，导致整体平衡有利于Eswitch关联，这不会受到影响 从这种阻挡效应，因此更强（如较粗线所示）。 B：在海马损伤模型中，不存在来自海马体的这种阻断效应，并且E和挖掘输出之间的元素关联大致等于开关输出的那些。 仅显示了紧密元素表示之间的临界权重。

内部问题缺乏海马阻滞

剩下的一个问题需要解决：如果海马体有助于学习锚定问题，为什么它不能有助于学习内部问题（B +C-，C +D-，D+ E-）？这个问题很重要，因为如果海马参与提供内部问题的快速联合表示，那么这些表示应该阻止所有刺激的关联学习，而不仅仅是B和E。这正是海马对这些刺激元素的不同影响，导致了老鼠神经网络的类似的传递行为。

我们的答案是内部问题是非线性判别问题。每个问题的构成完全与另一个问题重叠。例如，B+C-和C+D-共用C;C+D-和D+E-共用D;因此，每个刺激都是完全模棱两可的，经常相同概率地得到奖励和无奖励的结果。只有当动物或模型构建刺激对的联合表示时，才能解决这些问题。因此，人们可能会期待海马体的强大贡献（参见Sutherland和Rudy，1989）。然而，当我们之前详细探讨我们的模型如何解决非线性判别问题时，我们发现海马体通常不会促进更快的学习（O'Reilly和Rudy，2001）。这与文献一致（参见Rudy和Sutherland，1995，供审查）。我们将引用O'Reilly和Rudy（2001）的报告，报告中详细讨论了为什么会出现这种情况，却只有一个简短的总结。关键点在于，除了编码新的记忆之外，海马体还必须回忆以前存储的记忆，这些记忆通过CA3复发性抵押物支持的模式完成而发生。模式完成和模式分离（导致分离的联合表示的新编码）从根本上相互冲突（O'Reilly和McClelland，1994）;哪一个占主导地位取决于输入模式重叠的程度。因此，高度重叠的内部问题将倾向于触发回忆而不是新的编码。因此，海马将检索其他训练问题的答案，而不是学习当前问题的正确答案，这可以防止它产生阻塞效果。

认知重叠分析

为了进一步测试我们对网络行为的分析，我们测量了完整和损伤模型的关联皮质层中认知的重叠程度，以查看海马体是否影响这些表示。具体而言，我们预计在完整模型中通过海马体快速学习Ef连接将导致该Ef连接的认知与单个元素E和F的认知分离（即，通过海马模式分离来影响皮质）。实际上，这种分离使得模型能够使Ef训练问题正确（即，生成挖掘响应），即使E本身在完整模型中具有网络转换关联（图6）。相比之下，受损模型具有与挖掘更强烈关联的元素E权重，因此它仅在使用E和f的元素关联的条件下执行正确。为了验证这些想法，我们计算了Ef的关联层表示与元素E和f之间的欧几里德距离，并且发现在完整模型中该距离比损伤的距离大25％。因此，海马模式分离加强了单个元素皮质认知与连接的分离。

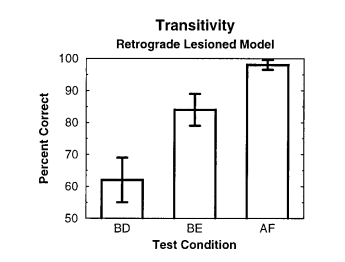


图8。逆行海马损伤对模型的试验性能的影响。 请注意，性能与完整模型的性能类似（图5）。 正如Dusek和Eichenbaum（1997）所论证的那样，这支持了我们的分析，海马通过阻断效应在训练期间做出贡献，而不是通过测试期间的灵活检索。

模型预测

我们对完整和损伤模型的模拟和分析提供了Van Elzakker等人报告的基本结果的解释。 （2003）和Dusek和Eichenbaum（1997）。 此外，该模型生成了许多预测，可以对其进行评估以进一步确定其有效性。 这些预测基于以下操作：当相对于训练，训练方案和海马亚区域的更多选择性损伤进行损伤时。

顺行与海马损伤的逆行效应

海马体通过其对学习的影响有助于在传递性测试中表现的想法是模型如何解释数据的基础。因此，正如我们所示，训练前的海马损伤（顺行性损伤）应该影响测试性能。然而，这种想法的另一个后果是训练后海马体的损伤（逆行损伤）不应影响测试性能。这是因为元素关联（确定测试性能）是在训练期间建立的，并且即使在受损的受试者或模型中也应该可用于影响测试性能。

我们通过在对五个问题进行训练之后但在BD，BE和AF上测试之前去除模型中的海马体来测试这种一般预测。图8显示在训练后移除海马回路不影响测试性能（参见图8和5）。

请注意，模型的预测与代表性灵活性理论的预测完全相反（Dusek和Eichenbaum，1997）。后一种观点假设海马体通过提供灵活的检索策略来支持传递性。因此，它预测海马的逆行损伤应该消除传染性。因此，对传染性测试中海马损伤的逆行效应的行为研究将对这两种观点提供强有力的检验。

用提示替换海马

根据我们的分析，海马体通过对EF和AB问题的阻断作用促成传递性推断。我们推断我们可以通过简单地提供额外的外部线索来模拟损伤模型中海马体的影响，所述外部线索可以产生与完整动物中通常产生的海马结合认知相同的阻断效应。具体来说，假设一个独特的提示，如形状或颜色（称为X）被添加到含有气味F的杯子中。这个X提示将为EF问题的表现提供独特，可靠的基础，就像我们一样认为联合海马表现在完整的老鼠中。特别地，该X提示将导致该EF问题的快速错误减少，并且因此阻止E刺激与挖掘响应的关联，导致元素权重的模式应当有利于BE传递推理问题的性能。图9显示该结果是在损伤模型中获得的，其中训练期间提供了X cue。通过简单地激活F刺激的两个附加单元来模拟该提示（这些单元之前从未在任何其他刺激中使用）。

使用提示改善BD性能

在本文和同伴文章（Van Elzakker等人，2003）中研究了五对训练问题（A B，B C，C D，D E，E F），BD上的表现总是比BE差。根据模型，这是因为B和D的挖掘响应的关联权重是相似的。然而，使用在先前预测中应用的相同类型的阻塞逻辑，我们可以通过在D E试验中向E杯添加独特的X提示来改变BD和BE的相对性能。这种操作使得与挖掘响应的D关联被X刺激阻挡，从而使D成为网络切换关联。这有利于在BD测试中选择B over D，如图10所示（注意，这种效果适用于完整和损伤的模型）。有趣的是，有了这个独特的X线索，完整的模型现在无法在BE测试中表现良好，因为EF测试中没有X线索，因此没有X线索的E获得净负（挖掘）关联BE表现。

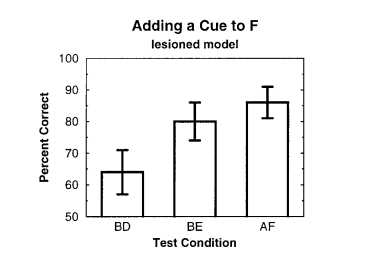


图9。在训练中向刺激F添加特殊线索时，模型在测试试验中的损伤性能。 该提示通过阻断E和dig之间的关联来模拟海马体在产生更好的BE测试性能中的假设作用，否则将在EF试验中学习。

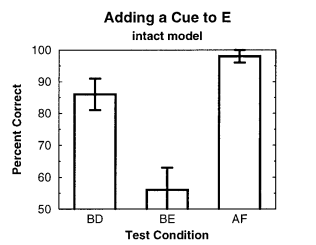


图10。当在D E训练试验中将独特的线索添加到刺激E时，模型对测试试验的完整性能。

该提示阻止D与挖掘响应的关联，给D一个有利于在BD测试中选择B的净开关关联。 请注意，BE上的表现现在是偶然的，因为在E F试验中E不会添加独特的提示，并且模型自身（没有提示）得到E是有益的。

CE测试案例

我们可以使用这些模型来预测CE测试问题的性能，这在老鼠中没有经过行为测试（Van Elzakker等，2003）。 基于图6中所示的关联权重，C与挖掘和切换具有大致中性的关联，只有轻微的挖掘偏好。 但是，完整模型对E具有强大的切换关联，这有利于在CE上选择C。 因此，我们预计CE上的表现将优于BD，但比BE差，因为B具有比C更强的净挖掘关联（即，C不会受益于A锚定效应）。 这被证实（图11）。 损伤模型中的CE性能比完整模型中的差，并且与BE基本相同，正如人们所期望的那样。

六对问题

因为模型中的所有测试探针性能都源于锚定效应而不是来自“真实”推断，因此如果网络在两个不直接受益于锚定的刺激上进行测试，则不应该执行上述机会。 这在迄今为止探索的五对问题中是不可能的，因为所有可能的新型测试探针都涉及B或E刺激，其直接受锚定现象的影响。 因此，我们将训练刺激扩展到从A到G的六对问题，使得锚定效应适用于B和F。因此刺激C，D和E应该相对不受影响，实际上我们发现对CE产生的测试产生了 低于模型的性能水平（图12）。 相反，正如锚定预测的那样，该模型的表现远高于BF试验的机会。

过度训练的效果

该模型对提供给模型的培训水平的影响做出了有趣的预测。在所有呈现到这一点的结果中，我们为网络提供了足够的培训，以在培训对上获得显着高于性能的水平。然而，当我们继续训练模型超出这一点时，会出现一个有趣的动态：我们发现传染性测试项目上的损伤模型和完整模型之间的差异消失了。回想一下，完整模型和损伤​​模型之间的完整差异在于阻塞效应，导致E表示不能获得EF训练模式的挖掘关联。相反，E表示与DE问题的切换相关联。然而，如果我们继续训练受损模型，EF问题比DE问题更容易，因为它确实具有明确的F刺激。因此，在某些时候，模型将很好地学习EF，并停止为E建立挖掘关联，同时仍然建立基于DE问题的切换关联。因此，我们预测，随着E获得更多的净开关关联，受损老鼠的BE测试性能将继续提高。 B挖掘关联将发生类似的动态。图13显示了损伤模型中的这种预测，其中培训量更大。

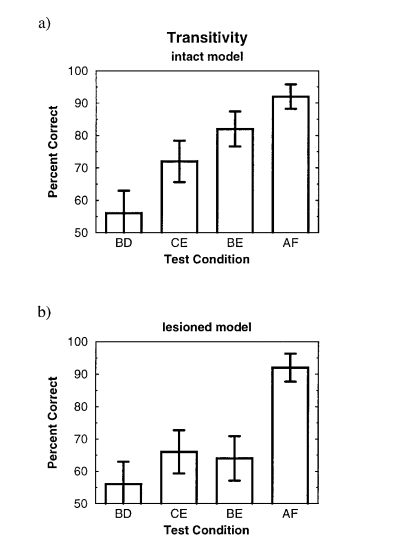


图11。与先前测试的问题相比，CE测试问题的性能。 基于关联权重（图6），E应该在完整模型中具有净开关关联，而C大致为中性。 因此，它应该落在BD（其中B具有小的网络挖掘关联并且D大致为中性）和BE（其中E具有网络切换关联，其与B的净挖掘关联以产生更好的性能）之间。 损伤的模型性能应该与BE大致相同。

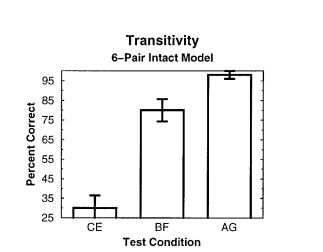


图12。六对问题的测试问题的结果。 只有在这个问题上我们才能测试刺激，CE，它们不会以任何方式从锚定效应中受益。 CE上的不良表现支持我们的分析，即之前在其他测试中的良好表现是由于锚定效应。

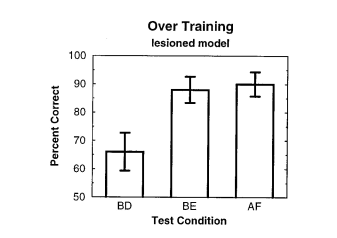


图13。额外训练对损伤模型的影响，这改善了BE和BD的测试性能。 这主要是由于锚定效应的放大，其中锚定问题得到充分学习并且停止驱动关联权重（即，E与EF试验中的挖掘强烈关联，并且B与AB试验的开关关联性较弱）。 这些更改可提高测试性能。

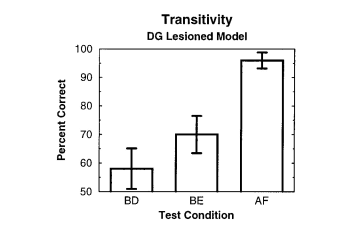


图14。模型中齿状回（DG）损伤的影响，其在与整个海马损伤相同的方向上产生损伤，这是由DG导致的模式分离的丧失引起的。

Dentate Gyrus的病变

EF和AB训练模式的海马模式分离对于阻塞效应是至关重要的，这导致完整网络相对于受损网络的良好测试性能。根据我们对海马系统的详细分析（O'Reilly和McClelland，1994），齿状回（DG）对于模式分离特别重要，因为其极其稀疏的激活和大尺寸。此外，该区域是海马体中唯一可以损伤的主要区域，而不会完全阻断通过海马回路的激活流动。因此，我们测试了DG病变是否产生与整个海马体病变观察到的相同性质的损伤。事实上，他们确实如此（图14）。这为我们解释海马对传染性的依赖提供了支持性证据。具体地，DG的解剖结构有助于分离必要的模式，使得老鼠可以在测试试验中选择B超过E。在没有DG的情况下，海马体中的模式分离减少，从而减少了在完整模型中看到的阻断效应。

预测摘要

基于上述模型的实验，我们对可在老鼠中直接测试的实验做出以下预测：

1。与顺行性病变相比，逆行性海马病变（训练后，试验前）应产生相对较小的效果。请注意，这一预测与文献中的其他几项发现相矛盾，证明逆行病变的效果大于顺行病变（例如，Kim和Fanselow，1992; Maren等，1997; Frankland等，1998; Richmond等。 。，1999; Anagnostaras等，1999; Rudy等，提交;）。这是因为海马体的缺失可以通过皮层学习得到补偿（O'Reilly和Rudy，2001），但如果在训练期间海马存在，老鼠将优先依赖它，因此表现将是逆行性病变受损。

2。通过在F刺激上用独特的提示（例如，添加的颜色或形状）训练，可以产生具有与完整动物类似的顺行海马损伤的BE测试案例上的敏感性推断性能。这种独特的线索模拟海马结合认知的影响，并产生阻碍效果，导致选择B超过E。

3。通过在DE训练试验中为E刺激添加独特的线索，可以逆转完整动物中BE比BD表现更好的模式。同样，这会产生阻塞效应，使关联权重转移到有利于B而不是B而不是B超过E。

4。在完整动物中对CE进行测试应该产生介于BD和BE之间的性能水平。这是基于关联权重梯度。

5。在完整的动物中延伸六个训练对并在CE上进行测试应该产生任何测试探针的最差测试性能。只有在这种情况下，两种刺激都不会从锚定效应中获益。

6。与受损动物的继续训练应该在BE和BD上产生稳定的更好的表现，因为这有效地放大了锚定效果。

只有齿状回的选择性病变应该产生与完整海马病变相似的行为障碍，因为该区域支持的模式分离对于海马对正常表现的贡献是至关重要的。

讨论

我们已经使用我们的海马和新皮质的计算模型来理解海马体如何有助于测试老鼠的传递性推断性能。与Dusek和Eichenbaum（1997）提出的观点（参见Bunsey和Eichenbaum，1996）相反，他们认为海马体通过在测试点灵活地比较记忆痕迹做出贡献，我们发现我们模型中的海马系统有贡献。通过塑造训练期间学到的元素相关权重。具体而言，快速学习的海马认知通过减少锚点问题（AB，EF，在五对训练集中）的误差信号产生阻塞效应，以增加B选择的方式塑造B和E刺激的元素关联。在BE传递推理测试中超过E。然而，D刺激不会受益于这种锚定效应，因此BD上的表现很接近。在伴随文献（Van Elzakker等人，2003）中描述的行为研究中观察到这种结果模式，并且与逻辑推理理论不一致。我们使用我们的模型来预测顺行海马病变对五对问题的影响，这与Dusek和Eichenbaum（1997）报道的四对结果一致。最后，我们使用该模型进行七个独特的预测，可以根据经验进行测试，包括训练参数（训练量，额外线索，六对问题），病变（逆行，齿状回）和其他测试病例（CE）的操作。我们正在开始测试我们实验室中的一些预测。

我们可以将现在的模型与早期的化身进行比较（O'Reilly和Rudy，2001）。在该早期模型中，该模型推动了同伴文章中报道的研究（Van Elzakker等，2003），我们没有在模型中引入自适应刺激选择阶段，这导致了更平等的关联权重分布。在这种情况下，传递推理可能发生的唯一方式是通过模式完成效应，其中在四对问题中，BD测试问题触发模式完成到BC或CD训练问题。如果重新激活BC，将产生正确的B响应。但是，如果CD被重新激活，相关的C响应不合格，所以我们想象这个C将支持BC的召回，再次导致B响应。虽然我们确实认为这种模式的某些完成可以在当前模型中进行，但它被锚定效应产生的微分关联强度的更大和更可靠的效果所淹没。因此，我们将当前模型视为我们初始工作的演变，其中考虑了我们的初始分析中缺失的关键部分。开发早期模型并从中推导出明确的可测试预测的过程证明了计算建模方法的一个重要优势 - 它可以澄清一组想法，如果不经验测试支持它们就可以被拒绝。相比之下，更模糊的言语理论并不总是那么容易被测试。

**与其他传递理论的比较**

也许与我们的工作最直接的比较可以在Levy及其同事开发的模型中找到（Levy和Wu，1997; Wu和Levy，2001）。 Levy和Wu（1997）使用基于序列学习的海马驱动计算模型来解决类似的传递推理问题 - 网络在80％的时间内选择B超过D。然而，他们提出的唯一机制是BD对激活C神经元射击，其暗示B的选择，因为B击败C，并且在C的上下文中未选择D。这种解释类似于上述的模式完成理论并且与象征性的遥远效果不一致。因此，理解同一模型如何解释后续研究中的符号距离效应有点困难，因为没有描述任何机制（Wu和Levy，2001）。此外，他们的模型纯粹是海马，因此无法评估海马相对于皮质的独特贡献。

另一种神经网络模型由Siemann和Delius（1998）开发。该模型将我们关注的焦点放在个体刺激项目的关联强度上，作为传递行为的基础，但它与特定的神经系统无关。因此，它无法说明海马与皮质的独特贡献。此外，Siemann和Delius（1998）在他们的模型中引入了一些复杂和专业的学习机制，而我们的模型已经表明，更通用的学习机制可以产生差异关联强度和传递行为。

我们已经强调了我们的理论与Dusek和Eichenbaum（1997）之间的对比。然而，文献中还有其他理论说法更接近我们自己。特别是，我们对模型的分析中出现的锚定理论在精神上与von Fersen等人提出的价值转移理论相似。 （1991年）。该理论认为联想或奖励价值被转移到序列中的相邻刺激。从本质上讲，有人认为，因为老鼠不需要学习太多关于E来接近F，所以它只需要学习更多关于D的信息，以便通过E来接近它。反过来，这意味着老鼠只需要学习更多关于C而不是D的知识，然后在D上接近C的问题也会稍微容易一些。价值转移理论意味着锚定的影响传播到所有刺激，形成线性有序的奖励关系。

在我们的模型（图6）中观察到的对应于元素刺激A到F的权重值的排序顺序似乎支持刺激项目之间的线性有序奖励关系（Trabasso和Riley，1975; von Fersen等，1991） ）。然而，我们假设锚定效应足以仅对锚附近的刺激产生优先益处。此外，尽管在接近元素刺激时存在排序的权重顺序，但是当在B对D测试时，完整模型仍然不能很好地执行。虽然接近B的权重比接近D的权重更强，但是接近的趋势D仍然存在。如果老鼠碰巧首先采样D（即D接近），它可能接近它并且不能通过测试。价值转移理论似乎适用于元素刺激的兴奋性，但价值的差异并不一定产生足够的效果，导致老鼠选择具有名义价值优势的刺激。简而言之，为了始终如一地选择正确的刺激，老鼠必须既偏向于接近它又偏向于避免不正确的刺激。

该模型还可以通过添加由Davis（1992）实施的不一致的F A训练对来解决延长训练周期的问题。有人认为这一对使得循环不可传递，因为A现在处于层次结构的高（A B）和低（A F）两端。我们发现，当在模型中实现这种前提对的包装时，为了正确训练需要更多的试验。这是可以预期的，因为通过包裹，有效地消除了锚固条件。没有“拐杖”可以使任何一对更容易解决，因为任务是完全非线性的。因此，没有真实的传递性测试，并且模型在任何未经过训练的测试对上进行测试时都会偶然执行。

总结和结论

在对一系列判别问题（例如，AB，BC，CD，DE）进行培训后，非语言生物显示出传递性能：当使用新组合BD进行测试时，他们选择B。重要的是理解为什么会出现这种现象，因为它对人们可以假设有机体可以用来适应其环境的认知过程有影响。一种可能性是老鼠，鸽子和猴子都可以从训练集（A B C D E）中提取一些有序表示并显示传递性能，因为它们在给出BD问题时使用该有序信息来选择B。这里提倡的更简单的替代方案是传递性能发生，因为在训练期间每个刺激获得一些绝对量的兴奋强度，并且当受试者面对新颖的组合（BD）时，它选择具有最大兴奋强度的刺激。

尽管生物体可能具有从这种训练序列中提取顺序的能力，但是不需要如此高水平的认知来产生观察到的行为，并且简约性有利于更简单的分级兴奋强度理论。至少，要声称行为传递性是由高级关系过程调解的，实验者应该提供独立证据，证明选择刺激具有相同的兴奋力。

为了支持分级兴奋强度理论，我们证明了我们的海马和新皮层的计算模型（O'Reilly和Rudy，2001）为个体​​线索产生了分级的兴奋强度并产生了适当的传递性。此外，通过分析网络的内部状态（例如，元素相关权重），我们能够重建这种情况的发生，并指出海马体在促成刺激的最终兴奋值方面所起的微妙且可能是短暂的作用。具体而言，海马体快速和自动地创建EF锚的联合表示的能力阻止了对E的兴奋性调节。在开发计算模型之前，我们无法解释这种贡献。因此，我们将此视为模型如何能够提供对认知神经科学中微妙和复杂问题的独特见解的示例。通过测试我们从该理论中得出的含义，我们希望确定这种洞察力是否正确。

**参考文献**

Anagnostaras SG, Maren S, Fanselow MS. 1999. Temporally graded retrograde amnesia of contextual fear after hippocampal damage in rats: withinsubjects examination. J Neurosci 19:1106.

Bunsey M, Eichenbaum H. 1996. Conservation of hippocampal memory function in rats and humans. Nature 379:255–257.

Davis M. 1992. The role of the amygdala in conditioned fear. In: Aggleton

JP, editor. The amygdala: neurobiological aspects of emotion, mem

ory, and mental dysfunction. New York: WileyLiss. p 255–305.

Dusek JA, Eichenbaum H. 1997. The hippocampus and memory for orderly stimulus relations. Proc Natl Acad Sci USA 94:7109 –7114.

Dusek JA, Eichenbaum H. 1998. The hippocampus and transverse patterning guided by olfactory cues. Behav Neurosci 112:762–771.

Frankland PW, Cestari V, Filipkowski RK, McDonald RJ, Silva AJ. 1998. The dorsal hippocampus is essential for context discrimination but not for contextual conditioning. Behav Neurosci 112:863– 874.

Hasselmo ME. 1995. Neuromodulation and cortical function: modeling the physiological basis of behavior. Behav Brain Res 67:1–27.

Kamin LJ. 1968. “attentionlike” processes in classical conditioning. In: Jones MR, editor. Miami symposium on the prediction of behavior. University of Miami Press.

Kim JJ, Fanselow MS. 1992. Modalityspecific retrograde amnesia of fear.

Science 256:675– 677.

Levy WB, Wu X. 1997. A simple, biologically motivated neural network solves the transitive inference problem. In: Proceedings of the IEEE international conference on neural networks. Vol I. IEEE. New York, NY. p 368 –371.

Maren S, Aharonov G, Fanselow MS. 1997. Neurotoxic lesions of the dorsal hippocampus and Pavlovian fear conditioning. Behav Brain Res 88:261–274.

Marr D. 1971. Simple memory: a theory for archicortex. Philos Trans R Soc Lond B 262:23– 81.

McClelland JL, McNaughton BL, O’Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychol Rev 102:419 – 457.

McNaughton BL, Morris RGM. 1987. Hippocampal synaptic enhancement and information storage within a distributed memory system. Trends Neurosci 10:408 – 415.

O’Reilly RC. 1996. Biologically plausible errordriven learning using local activation differences: the generalized recirculation algorithm. Neural Comput 8:895–938.

O’Reilly RC. 1998. Six principles for biologicallybased computational models of cortical cognition. Trends Cogn Sci 2:455– 462.

O’Reilly RC, McClelland JL. 1994. Hippocampal conjunctive encoding, storage, and recall: avoiding a tradeoff. Hippocampus 4:661– 682.

O’Reilly RC, Munakata Y. 2000. Computational explorations in cognitive neuroscience: understanding the mind by simulating the brain. Cambridge, MA: MIT Press.

O’Reilly RC, Rudy JW. 2001. Conjunctive representations in learning and memory: principles of cortical and hippocampal function. Psychol Rev 108:311–345.

O’Reilly RC, Norman KA, McClelland JL. 1998. A hippocampal model of recognition memory. In: Jordan MI, Kearns MJ, Solla SA, editors. Advances in neural information processing systems. Vol 10. Cambridge, MA: MIT Press. p 73–79.

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: variation in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. Classical conditioning. II. Theory and research. New York: AppletonCenturyCrofts. p 64 –99.

Richmond MA, Yee BK, Pouzet B, Veenman L, Rawlins JNP, Felden J, Bannerman DM. 1999. Dissociating context and space within the hippocampus: effects of complete, dorsal, and ventral excitotoxic hippocampal lesions on conditioned freezing and spatial learning. Behav Neurosci 113:1189 –1203.

Rolls ET. 1989. Functions of neuronal networks in the hippocampus and neocortex in memory. In: Byrne JH, Berry WO, editors. Neural models of plasticity: experimental and theoretical approaches. San Diego, CA: Academic Press. p 240 –265.

Rudy JW, O’Reilly RC. 2001. Conjunctive representations, the hippocampus, and contextual fear conditioning. Cognit Affect Behav Neurosci 1:66 – 82.

Rudy JW, Sutherland RW. 1995. Configural association theory and the hippocampal formation: an appraisal and reconfiguration. Hippocampus 5:375–389.

Rudy JW, Barrientos RM, O’Reilly RC. 2002. The hippocampal formation supports conditioning to memory of a context. Behav Neurosci (in press).

Siemann M, Delius JD. 1998. Algebraic learning and neural network models for transitive and nontransitive responding. Eur J Cogn Psychol 10:307–334.

Sutherland RJ, Rudy JW. 1989. Configural association theory: the role of the hippocampal formation in learning, memory, and amnesia. Psychobiology 17:129 –144.

Trabasso T, Riley CA. 1975. On the construction and use of representations involving linear order. In: RL. Solso, editor. Information processing and cognition: the Loyola symposium. Hillsdale, NJ: Erlbaum. p 381– 410.

Van Elzakker M, O’Reilly RC, Rudy JW. 2003. Transitivity, flexibility, conjunctive representations and the hippocampus. I. An empirical analysis. Hippocampus 13:292–298.

von Fersen L, Wynne CDL, Delius JD, Staddon JER. 1991. Transitive inference in pigeons. J Exp Psychol Anim Behav Processes 17:334 –341.

Wu X, Levy WB. 2001. Simulating symbolic distance effects in the transitive inference problem. Neurocomputing 30 – 40:1603–1610.