

学习
Learning

肖承丽
南京大学社会学院心理学系

1

学习是什么？

- 定义？
- 类型？
- 机制？

迷你翻转课堂-2
人类的学习
vs
机器学习

下次课上分享，5-10分钟

2

为了改变自己或他人的行为，你曾经用过什么样的好方法？

小红薯688875A6

3

定义“学习”

- 由于经验或实践所带来的相对永久性的行为改变
 - “相对永久性”是指人们在学习时，某些脑区会发生改变，以此来记录已学到的东西，即记忆
 - 经验或实践
 - 负性的：儿童学会不去触碰滚烫的炉子
 - 积极的：解答难题后的成就感
 - 并非所有的改变都伴随着学习，如身高或脑容量的改变——成熟
 - 达到成熟后，经验和实践才能发挥重要作用
 - 例如，婴儿单靠实践本身是学不会走路的。他们之所以能够走路，是因为他们的神经系统、肌肉强度、平衡感都达到了走路所需的程度

4

学习的类型

- 1 经典条件反射 (Classical Conditioning)**
- 2 操作性条件反射 (Operant Conditioning)**
- 3 认知学习理论 (Cognitive Learning Theory)**
- 4 观察学习 (Observational Learning)**

行为 → 认知 → 社会

5

1 经典条件反射 Classical Conditioning

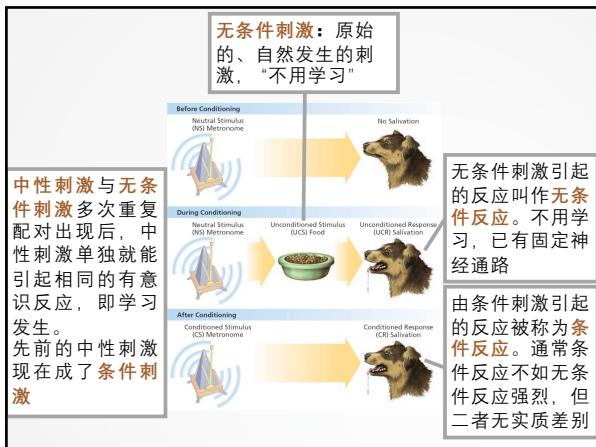
学会对通常产生反应的原始自然刺激以外的刺激做出非自主 (involuntary) 反应

关键要素

- 无条件刺激 (unconditioned stimulus, UCS)
- 无条件反应 (unconditioned response, UCR)
- 中性刺激 (neutral stimulus, NS)
- 条件刺激 (conditioned stimulus, CS)
- 条件反应 (conditioned response, CR)

巴甫洛夫, 俄罗斯生理学家

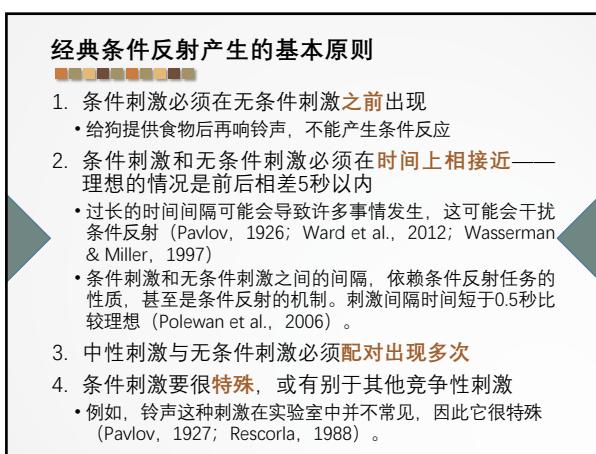
6



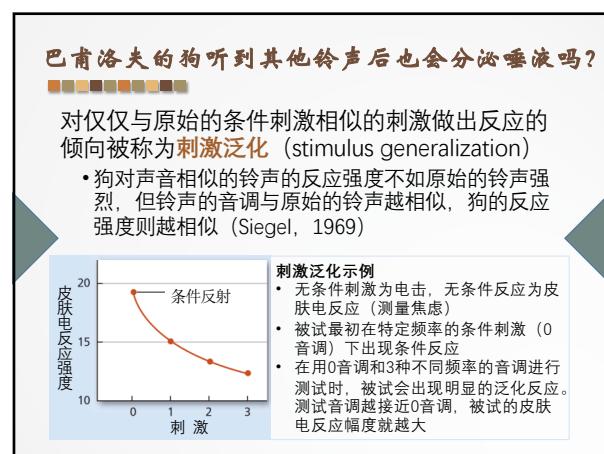
7



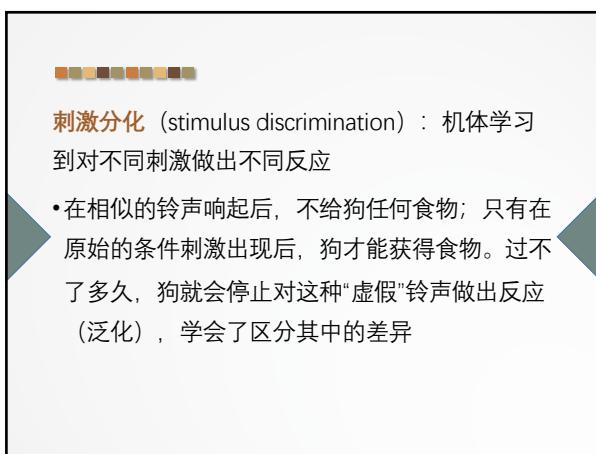
8



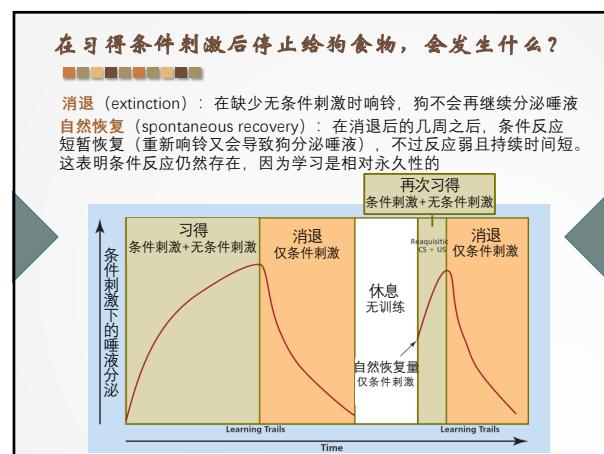
9



10



11



12

高级条件反射 (higher-order conditioning)

一种条件刺激引起另一种条件刺激



13

经典条件反射为什么会起作用?

罗伯特·雷斯科拉 (Robert Rescorla) 发现, 为了实现条件反射, 条件刺激必须提供一些关于无条件刺激即将到来的信息。即, 条件刺激必须能预测无条件刺激的到来

- 大鼠组1: 听音调, 播放中的某一时刻电击大鼠。很快, 在音调呈现后, 大鼠变得焦虑不安并表现出恐惧反应 (颤抖和尖叫)
- 大鼠组2: 只在音调停止后才对大鼠进行电击。这一组老鼠对音调停止表现出恐惧。

通过将音调或音调缺失与电击配对, 使得大鼠产生了特殊的期望, 从而决定了它们的特殊反应。由于这种解释涉及有意识地期待某事发生的心理活动, 因此它是认知视角解释经典条件反射的一个例子

14

适用于人类行为的经典条件反射

恐惧症 (Phobias)

- 一种特定类型的经典条件反射——**条件性情绪反应** (conditioned emotional response, CER)
- **替代性条件反射** (vicarious conditioning): 通过观看他人对刺激的反应, 也能形成经典条件反射
 - 儿童观察到母亲对狗的恐惧反应, 她对所有的狗也产生了一种无理由的、强烈的恐惧



广告中的条件性情绪反应与替代性经典条件反射



在广告中通过某种客体或人使得观众产生一种特殊的情绪反应, 并希望这种情绪反应能与他们的产品联系起来

16

条件性味觉厌恶 (Conditioned taste aversion)

- 首先给老鼠一种甜味液体, 然后再给它们注射可导致恶心的药物或让它们暴露于可导致恶心的辐射中, 它们就不愿再接触这种液体了 (Garcia et al., 1989; Garcia & Koelling, 1966)
- 利用相似的方式, 在酒精成瘾的人饮酒时, 给他们使用一种导致恶心的药物, 他们将不再喝任何酒精饮料
- 癌症患者接受的化疗药物也能引起恶心, 这会导致他们形成对化疗前吃的任何食物的味觉厌恶 (Berteretche, 2004)
- 依靠视力发现食物的鸟类会避免任何容易引起恶心的物体或昆虫
 - 一种蛾子 (无毒) 颜色类似于黑脉金斑蝶 (有毒), 拟态伪装能使自己避免被鸟类吃掉



厌恶疗法

- 2006年1月, 杨永信成立了临沂市第四人民医院网络成瘾戒治中心, 提倡使用电击疗法治疗沉迷网络游戏的少年。该中心开办一年后, 受到许多家长的欢迎, 他们不仅将网瘾少年送去接受治疗, 还将家中不听话、叛逆的孩子送进该诊所



18

生理准备 (Biological preparedness)

- 某些刺激和反应之间的一些联系更容易形成，这在动物和人类身上都成立
- 哺乳动物对将味觉与疾病联系在一起有生理准备，而鸟类则是对将视觉和疾病联系在一起有生理准备 (Shapiro et al., 1980)
- 对危险事物产生生理准备对生存很有意义
 - 猴子通过观看其他猴子对刺激做出可怕反应的视频，很容易学会害怕玩具蛇或鳄鱼，不过，猴子从来没有学会用同样的方法害怕花或玩具兔子 (Cook & Mineka, 1989)

19

药物依赖 (Drug dependency)

- 使用药物所带来的兴奋，无论是来自阿片等兴奋剂、还是酒精等镇静剂，通常都发生在特定环境中，同时会有某些特定的人在场，甚至可能会用到特定的物品
- 这些人、环境和事物线索的存在会使人们更难抗拒药物，因为身体和大脑已对药物与线索的联系形成了经典条件反射。

在成瘾的后期，使用者会对药物产生强烈的渴望，而这种渴望往往是由药物相关的刺激所引发的。快克可卡因成瘾者可能感觉相对稳定，但一看到烟斗就会对毒品产生强烈的欲望。安非他命成瘾者在夜店的洗手间里兴奋不已，跳舞音乐甚至冲厕的声音都会引发他对安非他命的渴望。气味，比如注射前用勺子煮的海洛因的发霉味道，特别令人回味。在他感人的青少年海洛因成瘾自传《篮球日记》(The Basketball Diaries) 中，吉姆·卡罗尔 (Jim Carroll) 写道，他的一个朋友试图通过在年轻时的天主教堂中寻求精神慰藉来戒除毒瘾。然而，教堂香火的味道让他想起了冒着泡泡的海洛因的霉味，他感到一种无法抗拒的渴望，于是从弥撒中赶回家再次注射。 -<The Compass of Pleasure>

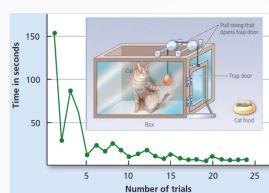
20

2 操作性条件反射 Operant Conditioning

通过反应所引发的愉快或不愉快后果所习得的自主行为 (voluntary behavior)



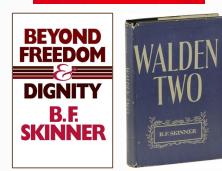
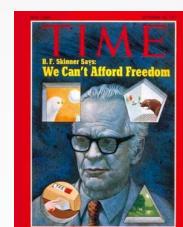
爱德华·桑代克
Edward Thorndike
(1874–1949)



- 效果律 (Law of Effect):** 如果某一行为会带来愉悦的结果，那么该行为会得到重复；而如果某一行为会带来不愉悦的结果，那么该行为不会得到重复

21

斯金纳 (B. F. Skinner, 1904–1990)

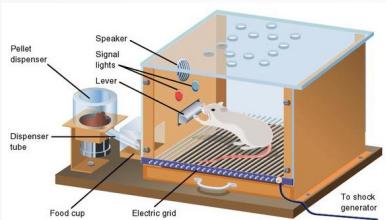


行为主义社会

22

强化 (reinforcement)

定义：伴随某反应的任何事物都能造成该反应更有可能再次出现。



斯金纳箱 (Skinner box), 或称操作性条件反射箱

23

- 初级强化物 (primary reinforcer)**：能够满足基本需求的强化物可以是任何食物 (饥饿驱动)、液体 (口渴驱动) 或接触 (愉悦驱动)，也可以是去除疼痛这种初级需求

- 婴儿~学龄前儿童和动物都比较容易通过初级强化物进行强化
- 次级强化物 (secondary reinforcer)**，需要与过去的初级强化物相联系
 - 过去曾得到钱的孩子意识到“这张纸”能用来买糖果和零食 (初级强化物)，因此钱本身变成了强化物
 - 如果一个人在拍孩子的头时 (拍头的动作是一种初级强化物) 表扬了他，那么表扬本身最终会使孩子感到愉悦

24

操作性条件反射

- 结果是增加已发生的反应的频率
- 反应是有意识的，由有机体发出
- 结果对联系的形成非常重要
- 强化必须即时
- 形成预期：正确反应后出现强化

经典条件反射

- 最终结果是对刺激形成新的反应，而该刺激通常无法引发这种反应
- 反应是无意识和自动的，由刺激引起
- 先行刺激对联系的形成非常重要
- 条件刺激必须紧接在无条件刺激之前发生
- 形成预期：条件刺激后出现无条件刺激

25

学习的神经基础

前扣带回皮层 anterior cingulate cortex (ACC) 与 伏隔核 nucleus accumbens 相连，后者属于药物依赖和奖赏回路的一部分，两个区域都参与多巴胺的释放——“上瘾” (Gale et al., 2016; Morita et al., 2013; Yavuz et al., 2015)

前扣带回皮层

Prefrontal cortex
Nucleus accumbens
中脑腹侧被盖区
Ventral tegmental area

奖赏回路

26

正强化 (positive reinforcement)：通过增加或体验愉悦的结果来强化反应，如奖赏或赞赏性地拍人后背

负强化 (negative reinforcement)：通过移除或摆脱不愉悦的事情来强化反应，如吃止痛药消除疼痛 (止痛药成瘾)

无论是正强化还是负强化，都会增加行为的可能性

1. 佩德罗的父亲唠叨着他去洗车，佩德罗讨厌唠叨，他只好乖乖地去洗车，好让父亲能停止唠叨。
2. 杰克发现，以一种有趣的声音说话能够吸引同学的注意，因此他现在常用这种方式说话。
3. 艾伦是餐厅的服务员，总是保持微笑和愉快的表情，因为这样做似乎能得到更多的小费。
4. 安娜在截止日期前将论文交给了老师，因为如果晚交一天，得分就会降一个等级。

在你上小学时，哪种强化方法最有效：正强化还是负强化？上高中、大学以后有变化吗？

27

强化程序

记住，每次他给你一颗小丸时，都要通过拉动杠杆来强化他的这种行为。

- 部分强化效应 (partial reinforcement effect)**
 - 在某些（但不是所有）正确反应后得到强化的反应比得到连续强化的反应（每一个正确的反应都有一个强化物）更耐消亡。
- 固定间隔强化 (fixed interval schedule of reinforcement)**
- 变化间隔强化 (variable interval schedule of reinforcement)**
- 固定比率强化计划 (fixed ratio schedule of reinforcement)**
- 变化比率强化 (Variable ratio schedule of reinforcement)**

28

CUMULATIVE RESPONSES

TIME

- 无法产生快的反应速度
- 扇形效应：反应速度仅在强化之前上升，之后立即下降，直到接近下一次强化的时间
 - > 期末考试前拼命复习，体检前养生
- Fixed Interval
 - 反应速度缓慢而平稳
 - > 突击考试，钓鱼，打占线的电话
- Variable Interval
 - 反应速度非常快
 - > 在得到一个强化物后立即出现“停顿”（休息）
 - > 打卡10次得到1杯免费的
- Fixed Ratio
 - 反应速度非常快
 - > 在得到一个强化物后立即出现“停顿”（休息）
- Variable Ratio
 - 反应速度非常快
 - > 打卡10次得到1杯免费的
- 反应速度非常快
 - > 老虎机，幻想下一次会走运

29

惩罚

- 强化都会加强反应，无论是正强化还是负强化
- 惩罚是强化的反面，会削弱反应

	强化	惩罚
正性 (增加)	珍视或想要的事物 正强化 (如表现良好得到小红花)	不喜欢的事物 施加惩罚 (如不听话被打屁股)
负性 (消除/避免)	不喜欢的事物 负强化 (如系好安全带以防警报响)	珍视或想要的事物 消除式惩罚 (如不能玩游戏)

30

惩罚的问题

- 惩罚只能暂时压制或抑制行为（如阻止孩子跑向车流），通常并不能使不良行为彻底消失。
- 严厉的惩罚会造成儿童（或动物）回避惩罚者，而不是回避被惩罚的行为，因此儿童（或动物）会学到错误的反应。
- 严厉的惩罚可能会鼓励为了回避惩罚行为而说谎（一种负强化），这同样不是我们想要的。
- 严厉的惩罚可能导致恐惧、焦虑和无法促进学习的情绪反应。如果惩罚的目的是为了教育，那么这种结果毫无益处。
- 严厉的惩罚（如打孩子）本身就提供了攻击性的模仿对象
- 消除式惩罚只教孩子不要做什么，但没有教孩子要做什么

31

让惩罚更有效的规则：

1. 惩罚应该紧随需要被惩罚的行为
2. 惩罚应该具有一致性
 - 如果家长说要对孩子的某种行为进行惩罚，那么家长必须付诸行动，履行承诺
 - 对于特殊行为的惩罚应该保持相同的强度或轻微增加，而不是降低
3. 对错误行为的惩罚应该与正确行为的强化配对出现

32

操作性条件反射的相关概念

- **辨别刺激** (discriminative stimulus) 指的是任何能为有机体提供为获得强化而做出某一反应的线索的刺激
 - 如红灯还是绿灯，真警察还是假人
- **消退**：包含强化的去除



- 在商场哭闹的孩子
 - 家长妥协给孩子买玩具
 - 孩子——正强化
 - 家长——负强化
 - 让闹脾气这一行为停止的唯一方法是消除强化，也就是不给孩子买，也不理睬孩子。

33

泛化

- 正在学说话的孩子看到爸爸出现喊“爸爸”→孩子的爸爸的喜悦和关注→强化作用
- 孩子将喊“爸爸”反应泛化到任何男性身上→孩子的爸爸可能会感到尴尬不安，其他男性不强化这种反应→孩子会学会区别对待他人和“爸爸”

自然恢复

- 训练动物学习一种新把戏时，大多数动物都会试图通过表演旧把戏来获得强化物

34

应用：塑造和行为修饰

- **塑造** (shaping)：通过逐渐强化小的目标，逐渐渐进 (successive approximation)，直到完成最终目标
 - 训练狗跳圆环：从训练狗做自己能够做的行为开始，再慢慢塑造狗进行“跳”这种狗能够做但本身不会主动去做的行为
 - 将圆环放到地上，以食物作为诱饵，狗通过圆环后给予它食物（正强化）
 - 将圆环提高一点，……
 - 以此类推，……直到狗能够跳过圆环得到食物



本能漂移 (instinctive drift)

- 动物回归到固有的行为，固有的行为比习得的行为更具有优势，即使这样会延迟食物的获得
- 用食物作为强化物，猪和浣熊很快就建立了捡玉米、把玉米带到某个地方、把玉米放进玩具筐的条件反射；然而，过了一会儿，动物开始做出一些不理想的行为

“猪停在路上，把玉米埋进沙子里，然后用嘴巴把它挖出来；浣熊用了很长时间玩耍玉米，做出它那著名的类似于清洗的动作。最初，这看起来很逗乐，但是最终，它太费时间了，让观众感觉表演不完善。从商业角度来看，这是一场灾难。”



35

36

行为矫正 (behavior modification)

- 通过操作性条件反射（有时是经典条件反射）来用于改变动物和人类的不良行为，创造令人满意的反应

老师通过行为矫正帮助学生集中注意力：

- 选择一个目标行为，比如让学生与自己保持眼神接触。
- 选择一个强化物。如可以在墙上的表格栏里给学生贴小红花。
- 将计划付诸行动。学生每进行一次眼神接触，就给他一朵小红花。对学生的不当行为（如上课走神儿）则不通过小红花进行强化。
- 一天结束时，对于获得一定小红花的学生给予特殊待遇或奖赏。该奖赏是事先决定好的，且与学生讨论过。



37

代币制 (token economy)

- 小红花：收集起来去“购买”特殊待遇

金融系统：

- 人们通过工作得到金钱奖励，以购买食物和保障等
- 信用卡公司提供奖励积分
- 航空公司提供飞行常客里程
- 快餐店饮料店打卡集好礼

计时孤立 (time-out)

- 行为失当的动物、儿童或成人被置于一旁，无人注意
- 轻度的消除式惩罚：他人对个体的关注（正强化）被“消除”
- 应用于儿童时，时长最好在1分钟，如果超过10分钟，儿童可能会忘记自己为什么会被计时孤立

38

应用行为分析 (applied behavior analysis, ABA)

- 通过当前行为分析和行为技术来解决与社会相关的问题
 - 如用糖果作为强化物，教孤独症儿童社交技能和语言
- 生物反馈 (biofeedback)**
 - 通过生理信息（如心率）反馈来创造放松的状态
- 神经反馈 (neurofeedback)**
 - 个人可以通过程序学习如何产生与特定认知或行为相关的脑电波或特定类型的大脑活动，如提高注意力、保持专注、放松意识等
 - 治疗注意缺陷多动障碍、控制慢性疼痛、治疗癫痫



39

3 认知学习理论 Cognitive Learning Theory

- 潜伏学习 (latent learning)**
- 顿悟学习 (insight learning)**
- 习得性无助 (learned helplessness)**

人工智能：
智能涌现？



“卫生间？当然，就在那个大厅的左边，向右走，向左走，再向左走，直接经过两个左转，然后向右走，就在第三个走廊的尽头，在你的右边。”

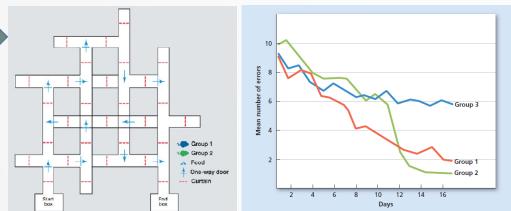
40

潜伏学习 (latent learning)

爱德华·托尔曼 (Edward Tolman)

不需要强化就能发生

- 第一组老鼠每天得到奖赏（强化）
- 第二组在第10天首次得到奖赏——认知地图cognitive map
- 第三组一直得不到奖赏



41

顿悟学习 (insight learning)

沃尔夫冈·苛勒 (Wolfgang Köhler)

- 在灵光一闪的时刻将问题的所有元素突然集合到一起



42

暖和感 (Metcalfe, 1986;
Metcalfe & Wiebe, 1987)

- 每隔15s快速报告自己觉得答案有多近（“暖”）或有多远（“冷”）

顿悟问题：“囚徒出塔”，囚徒在自己的囚室里发现了一根绳子，但其长度只够他下到塔的一半；他将绳子一分为二，接在一起，成功逃脱了，请问他是怎么做到的？

Time	1	2	3	4	5	6	7
Insight	~28	~10	~10	~10	~10	~10	~10
-15 seconds	~28	~10	~10	~10	~10	~10	~10
-30 seconds	~28	~10	~10	~10	~10	~10	~10
-45 seconds	~28	~10	~10	~10	~10	~10	~10
-60 seconds	~28	~10	~10	~10	~10	~10	~10

43

习得性无助 (learned helplessness)

马丁·塞利格曼 (Martin Seligman)

由于过去屡次失败而无法采取行动来摆脱困境的倾向

- 向一组狗呈现一个音调，之后伴随无害但会引起疼痛的电击，狗由于被束缚住，因此无法逃脱电击
- 不再束缚狗，但先前已经学习到“无法躲避电击”的狗不会尝试跳过障碍

44

• 中缝背核(the dorsal raphe nucleus)

- 释放5-羟色胺，激活杏仁核（产生害怕与焦虑），且降低负责“战或逃”反应脑区的活跃性（不逃脱或呆滞）
- 较为古老的脑区，无法确定哪种应激是可控的

• 腹内侧前额叶皮层 (ventromedial prefrontal cortex, vmPFC)

- 能够帮助确定哪些应激是可控的，可抑制脑干区域，平复杏仁核反应
- 从而使动物能够有效地对应应激源并表现出控制能力

(Maier & Seligman, 2016)

45

我认识一些人，他们的行为就像实验中的狗一样——他们生活在可怕的环境中，但不想离开。这是同一回事吗？？

NATIONAL BESTSELLER

LEARNED OPTIMISM

How to Change Your Mind and Your Life

WITH A NEW PREFACE

MARTIN E. P. SELIGMAN, Ph.D.

Author of Authentic Happiness

"Learned helplessness is the single most important concept I have ever learned in my field... No other academic discipline, perhaps not even psychology itself, has had such a profound impact on our understanding of human behavior." — Martin Seligman, The New York Times Best Seller

46

4 观察学习 observational learning

通过观察行为模型即他人正在做的行为，来学习新的行为

• 成人榜样忽视or攻击玩偶

阿尔伯特·班杜拉 (Albert Bandura)

学习/表现差异
learning/performance distinction:
学习可以在没有实际执行所学行为的情况下发生

47

媒体暴力⇒儿童和青少年攻击性提高？

• 儿童观看成人榜样击打波波玩偶的影片

- 榜样得到奖赏
- 榜样受到惩罚

• 随后将儿童置于一个有玩偶的房间

- 儿童击打玩偶
- 儿童不击打玩偶。不过，告诉儿童按影片中榜样的行为去做会得到奖赏，每个儿童都重复了榜样的行为

48

观察学习的4种要素AMID



• 注意 (attention)

- 人们可能更容易观察那些看上去与自己类似的人和更有吸引力的人。

• 记忆 (memory)

- 学习者必须能够保留已形成记忆

• 模仿 (imitation)

- 学习者必须能够再现和模仿榜样的行为。
- 眼睛：“看懂了，会了会了！”手：“不，你不会……”

• 渴望 (desire)

- 学习者必须有执行某种行为的渴望或动机
- 成功的榜样是人们模仿的绝佳对象，而对于失败或受惩罚的人，人们很少去模仿他们

49

Discussion

- What did you learn about “learning” so far?
- And what’s your opinion about the different types of learning?
- Any other types of learning?



50

Discussion

- Is reward always more effective than punishment?
- Any differences between animal learning and human learning? And machine learning?
- What are the roles of instinct and culture in human learning?
- Can intelligence be improved by learning?

51

THANKS



52