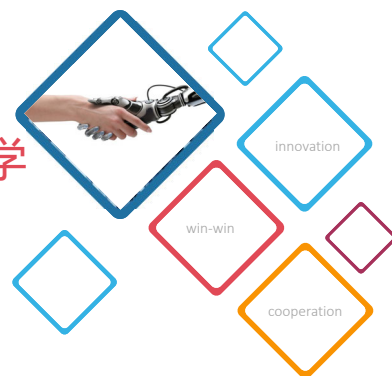


第15讲

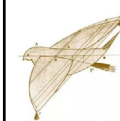
人机交互中的心理学

南京大学社会学院心理学系

肖承丽



机器人发展历史的简要回顾

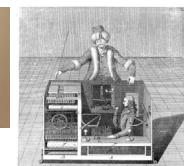


历史&传说

- 公元前1000年,《列子·周穆王篇》,偃师造能唱会跳的机器人,外形似真人
- 公元前400年,古希腊的阿契塔,蒸汽动力的机械鸽

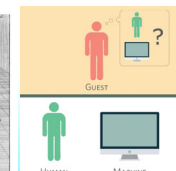
机械骑士

由达芬奇设计,骑士身穿中世纪盔甲,靠风能或水利驱动,可以坐、站、摆动手臂以及摇头



Turk机器人

由冯·肯佩伦发明,是设计巧妙的机械工程,会下国际象棋,但实际由人在暗中操纵,有人将其称为“机器人史上的最大骗局”



图灵测验

基于“图灵测验”的思想,研究者开始设计不需要人类持续操作的自动化机器人,为以后机器人的发展奠定了基础



作为工具的机器人

工业机器人:最初的机器人应用于汽车行业,是数字化编程的机械臂

军事机器人:用于执行监视、拆弹、自动化武器等任务

1495年

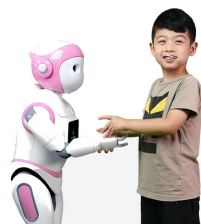
18世纪晚期

1950年

20世纪50年代

社交机器人

- 过去20年,社交机器人逐渐成为研发热点
- 与人频繁交互,作为人工的陪伴者或帮助者
- 应用于家庭、医院、学校、商场等场所
- 模仿人类——外形、思维、情绪表达、行为



- 人-机器人交互与人-人交互有何异同?
- 机器人的外形、思维、情绪表达、行为应如何模拟人?不同的模拟对人的认知有何影响?
- 从人-机器人交互中能揭示人自身的哪些认知特点?

大纲

01

当前社交机器人的应用领域

02

当前机器人研究的关注方向

03

从制造机器人中学习人

04

人对机器人的观点感受与行为

01 > 当前社交机器人的应用领域

- 老年人的保健与护理
- 自闭症儿童的训练
- 中风患者的康复治疗
- 远程遥控机器人
- 教育
- 商场导购



> 老年人的保健与护理

健康保健类的机器人可以从4个方面为老年人提供帮助：

- **肢体功能**（走路、搬运、取物、洗澡等）
- **认知问题**（提醒做事、记忆训练等）
- **健康管理**（鼓励锻炼、测血压、检测是否摔倒等）
- **社会心理支持**（提供陪伴、互动娱乐等）

老年人
自闭症儿童
中风患者

> 老年人的保健与护理

案例——日本著名的社交机器人Paro

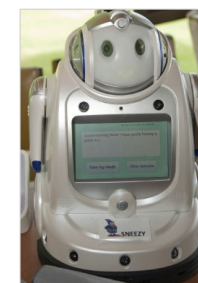
- 海豹幼崽造型，可以移动和模拟海豹叫声，对声音触摸等做出反应
- 选择海豹而非常见宠物（如猫）的样子，是因为人们非常见动物的行为期待比较少
- 可减少**孤独感**、增加居家人群或独自在医院就医患者的**社会互动感**(Robinson et al. 2013)
- 通过与Paro互动可降低痴呆患者的**焦虑感和易抑郁感**(Jørganson et al. 2015);
- 在准实验中 Paro还可以为精神衰老的患者解决**行为困难问题**(Bemelmans et al. 2015)
- Paro**对痴呆患者的帮助效果**也得到了相关照料人员的认可(Robinson et al. 2012)。



老年人
自闭症儿童
中风患者

> 老年人的保健与护理

- 使用iRobi和 Cafero两种有健康保健功能的机器人提醒老年人吃药、测血压及血氧浓度等，并辅助提供娱乐和认知刺激，初步研究结果表明效果与人工照料一样，为降低人工照看的成本提供了可能 (Broadbent et al. 2014, Orejana et al. 2015)。



iRobi



Cafero

老年人
自闭症儿童
中风患者

老年人的保健与护理

- 机器人的**外貌**应与其执行的任务相匹配：老年人更喜欢**毛茸茸**的机器人作为陪伴，而更喜欢**偏机械化**的机器人提醒自己吃药(Broadbent et al. 2009)；
- 体积：不能过大，导致在家庭环境中移动不方便，且吓人；不能过小，导致不易被看见。
- 人们并不总是喜欢外观更像人类的机器人：不像人的托举机器人更受欢迎、**外观拟人化程度**高的远程医疗机器人更受偏爱(Broadbent et al. 2012)；
- 人们更喜爱**声音**和人类相似的医疗保健机器人(Tamagawa et al. 2011)。

担忧：

- 机器人的使用会造成人工照料者失业、使被照料者缺乏真实的社会接触(Broadbent et al. 2012)；
- 使用机器人潜在的事故危害，机器人的可靠性，传递信息错误或延迟，使用者的隐私泄露，复杂场景中机器人的定位导航困难

自闭症儿童的训练

- 自闭症儿童一般都有明显的社会交往障碍和沟通障碍，以及局限性、刻板性和重复性行为。
- 机器人的行为模式、反应规则等更简单和容易预测，因而**比人类更好相处**，与机器人相处为自闭症儿童习得人人交互的社会行为提供了很好的机会。
- 当前应用于自闭症儿童的机器人一般具有以下功能(Diehl et al. 2012)：
 - 激发儿童的行为（意图是诊断个体症状或增加其亲社会行为）；
 - 教授患者社交行为模式、做行为互动练习；
 - 提供及时的反馈和鼓励。



机器人Kaspar

自闭症儿童的训练

● 该领域的研究尚处在探索阶段

- 已有的研究存在样本量少、交互时间短、质性研究、无对照组、患者诊断信息缺失等问题
- 所用机器人的种类繁多，交互大多由治疗师控制并在同一房间进行，而非自发产生交互(Diehl et al. 2012)
- 研究结果也存在很大的个体差异(Diehl et al. 2012)，仅有部分儿童更喜欢机器人而非其他玩具
- 自闭症患者从人机交互中学到的行为模式是否能迁移到其他情境中也尚无定论
- 自闭症的机器人辅助疗法是否真实有效需要未来**临床医生和心理学家**合作展开研究

中风患者的康复治疗

目前用于治疗中风患者的大多数康复机器人都**不会被归类为社交机器人**，因为它们一般具有可穿戴性、仅提供物理辅助功能（比如增强患者的肢体力量）

- 机器人能在各个康复阶段提升中风患者的**手和手臂**功能(Basteris et al. 2014)，促进患者的**步态恢复**，但效果是否优于其他方法尚不明确(Basteris et al. 2014)；
- 尽管机器人会存在一些技术和用户体验问题，有些患者还是更喜欢机器人而非人类辅助 (Cherry et al. 2016)，因为机器人更**便利**，可自行规划练习时间，让患者感觉到较强的**独立性、提升情绪**；
- 中风患者的**使用反馈**集中于如何改善训练系统，很少评价其社交功能 (Hughes et al. 2011)；



中风患者的康复治疗

也有研究表明，康复机器人的**社会属性应该被更多的应用到**中风的治疗过程中以增强疗效。

- 当康复机器人的行为（常规或有挑战性的指示语）与患者的**性格**（内向或外向）**相匹配**时，患者与机器人交互的时间更长 (Tapus et al. 2008)。

老年人
自闭症儿童
中风患者
医疗康复领域

远程遥控机器人

- 当人们跨距离交流或办公时会用到远程机器人，这种情境下远程机器人通常仅仅充当一种工具而不是有社会性质的个体；
- 当前远程机器人大多依附在移动设备上以屏幕形式呈现，当然，也有某些远程操作机器人的**外观更像人类** (Kristoffersson et al. 2013)。
- 在一些特殊环境中远程机器人也可以大显身手，如太空、深海等；也可**应用到**远程会议、远程医疗等领域。
- 当前的远程遥控机器人仍然存在通讯迟滞、无法开门或上下楼梯，巡航navigation和归位docking困难等问题



远程操控

远程遥控机器人

医疗保健

- 使用远程机器人进行术后病房巡查的**患者满意度**与医生亲临检查相同 (Ellison et al. 2007)；
- 在重症监护病房中使用网真机器人可以提升医生的工作效率，并可以缩短患者**住院时间和住院成本** (Vespa et al. 2007)。
- 使用远程机器人而非电话，可使医生快速获得关键的视觉信息，保证快速诊断和决策

促进社交

- 因疾病长期住院的学生可以通过远程机器人设备**参与学校活动**，总体来讲可以减少他们在社交、情感和学业方面的孤立感 (Fels et al. 2001)；



远程操控

远程遥控机器人

辅助手术

- 达芬奇外科系统 da Vinci surgical system
 - 使用微创的方法，实施复杂的外科手术



外科医生控制台、床旁机械臂系统、成像系统

- 元分析显示，机器人辅助手术和人工腹腔镜手术在**手术失血量、并发症发生率和住院时间**方面都没有显著差异，但机器人辅助手术要花费更长的时间 (Broholm et al. 2016)。
- 关于病人对手术机器人的知觉和态度的研究很少

远程操控

在教育领域的应用

- 当前机器人在教育领域的**主要应用**是科技和外语的学习(Mubin et al. 2013)
- 机器人通常可分为**两种角色**：提供激励的**同辈**或谆谆教诲的**老师**。

- 机器人**组装件**（如乐高）是用来帮助学生科技学习的很好的工具。
 - 带有轮子的组装件（比如车类）更适合年龄较小的学生，形状更像人型机器人的组装件更适合年长的学生(Basoeki et al. 2013);
- 对中低收入家庭的孩子来说，选用**人形机器人**作为学习工具比选用创意型工具组件更能增强动机、社交意识和自我表达能力(Han et al. 2015)



教育
辅助

在教育领域的应用

- 人形机器人Nao（如右图）曾以三种不同的行为模式（动机型、自我决定型、中立型）被应用到儿童的健康教育**测验**中(Blanson et al. 2013)。
 - 随着时间的增加，学生的知识相应增长
 - 儿童觉得很有趣，尤其是在最初的阶段
 - 儿童模仿机器人的交互模式，说明产生了社会榜样效应



教育
辅助

机器人导购

如今很多国家在购物商场、博物馆等公共场所都开始使用机器人向导，这类机器人在真实生活中通常都有一定的“自主行动能力”，能较好的完成向导任务；在实验研究中经常是有人遥控以便更好地控制实验条件。

- 日本某购物中心使用机器人导购长达三年以上，对其顾客的质性访谈表明，人们基于机器人导购的**外貌和行动方式**对其做出评判(Sabelli & Kanda 2015)。人们更多地将机器人视为商场的吉祥物而非设施
- 用机器人做向导、发传单会带来**好处**：如顾客能得到清晰明确的指导、顾客会更享受购物的乐趣、对一起购物的儿童更有益处(Satake et al. 2015)
- 调查显示：65%的人更喜欢机器人导购而不是真人导购，因为机器人**待人平等**，不会根据外貌把顾客分为三六九等；如果机器人过于像人会很恐怖；90%的人想再次使用导购机器人。



日本的美女机器人导购

购物
指导

02

当前机器人研究的关注方向



02

当前机器人研究的关注方向

服务
老年人

当今很多国家的人口呈现老龄化趋势，相关国家政府都高度关注老年人的医疗保健问题，服务老年人的机器人可以大大降低老年人的医疗成本，有广阔的发展前景。

机器人
对儿童的
影响

处在“永远在线”的大环境下，孩子们很难发展出享受独处和反思自我的能力；我们与移动设备之间的关系会影响人际关系的发展，手机的存在与移情能力和亲密感的降低有关(Misra et al. 2016) 机器人的应用过程中更应该考虑这些问题。

机器人的
接受度

人们会承认机器人在医疗保健的某些方面比人类具有优势(Broadbent et al. 2010)，也喜欢机器人的社会存在（至少在医疗保健领域）(Yoshikawa et al. 2011)，但当机器人在越来越多的领域比人类更有优势时，我们是应该高兴还是恐慌呢？

受欢迎
机器人的
特点

当机器人对患者点头、微笑时，患者会更喜欢机器人(Coeckelbergh et al. 2016)；机器人表示出同情时，人们对机器人的积极评价也更多；也有很多研究关注机器人的哪些外观特点更受人们欢迎。

伦理
问题

机器人只能模仿人类之间的情感，这种“感情欺骗”是不道德的，可能会对人们造成伤害(Sparrow & Sparrow 2006)；若使用机器人照顾照顾婴儿，可能会由于人类护理不足，导致婴儿无法获得正常的语言、社交或情感方面的能力(Sharkey & Sharkey 2011)；让机器人陪伴自闭症儿童是否符合道德伦理仍存在争议(Coeckelbergh et al. 2016)。

信任与
安全

- 2015年三星曾在BBC新闻报道中警告，尽量不要在智能电视前谈论私人事务，因为电视的语音识别软件可以录音并将其传输给第三方；
- 机器人有能力收集更多我们的个人数据，这些数据可能会被黑客入侵或以其他方式被他人访问；
- 智能音箱可能被黑，播放私人听见的信息(BBC news 2015)

03

从制造机器人
中学习人

制造类人机器人

Creating Humanlike Robots

从制造机器人中
学习人

制造宠物机器人

Creating Pet-like Robots



制造类人机器人

Android人形机器人：特指外形和行为高度拟人化的机器人

Android科学是一个跨学科领域

- 工程师、计算机学家构建类人机器人
- 认知科学家将这些机器人与人进行测试，以进一步发现人的本质



日本大阪大学的石黑浩教授(左)和他制造的机器人(右)

- **Geminoid双子/替身机器人**：外形和行为与某一真人高度相似，由人遥控操作(Wizard of Oz fashion)，与真人拥有相同的动作、交谈能力、人格特质，——人造的身体+人类的思维，复制某一特定人
- **Telenoid遥控人形机器人**：只有头、躯干和简化的肢体，没有年龄、性别特征，遥控控制，以研究人的本质
 - 交互体验复杂，引发害怕情绪，人们宁愿和真人交互
 - 操作者（工程系学生）认为该机器人可中等程度地传递他们的意图，交互者体验到了中等程度的社交性



Telenoid R1
石黑浩教授制造

制造类人机器人

困境与探索

机器人学家们一直都在试图使机器人在外观和行为上与人类完全相同，以便人类可以在更直观、更自然的水平上与机器人交互。

然而实现这个目标面临着很多困难：

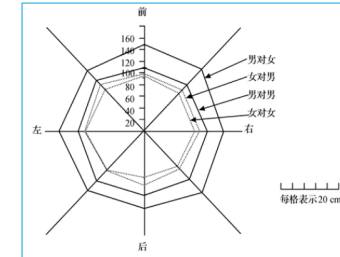
- **技术限制**：制造类人躯体实现类人感知和行为
- **复杂人类行为**：界定与建模
 - 建模人类行为可进一步揭示人类自身的特点
 - 机器人学家正在尝试建模的两种常见的社交行为：
 - **社交距离**
 - **注视行为**



制造类人机器人

社交距离：机器人应遵循人类的距离准则以更好地融入人类社会

- **个人空间圈**
 - 20世纪60年代心理学家沙姆最早提出
 - 指在心理上个体所需要的最小的空间范围，它就像一个“神秘的气泡”，会随我们身体的移动而移动
 - 一般来说女性比男性的个人空间圈小，这也是在日常生活中女性一般比男性更相互靠近的缘故。
- **心理学家和社会科学家**已经建立了关于社交距离的理论模型，并总结出影响社交距离的重要因素包括性别、年龄、好感度以及文化背景等。
- HRI研究者创建了使用相似接近行为的**机器人模型** (Avrunin & Simmons 2014)，发现人们在社交互动中对机器人的反应与对人的反应相似（尽管不完全相同） (Mumm & Mutlu 2011)。



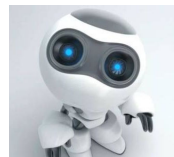
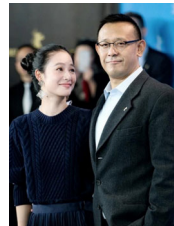
对空间距离需求均值比较示意图（采自杨志良等1989）



制造类人机器人

注视行为

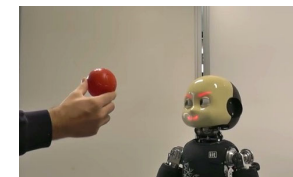
- 眼睛是社会互动的重要元素
 - 古罗马诗人奥维特曾说，“沉默的眼光中常有话语”
 - 已有很多学者研究了注视时间的长短、注视的方式、注视与其他身体动作的组合（如转动头部）等这些因素对人际互动的影响
- **眼睛注视的心理学功能**：
 1. **提供信息** (通过注视行为发现个体的情绪状态、专注程度等)；
 2. **维持互动** (比如在对话及讨论中起到一种提示轮流的作用)；
 3. **表达亲密** (当表达喜欢的感觉时，眼睛注视的频率会有所提高)；
 4. **社会控制** (说服、欺骗他人等)；
 5. **服务任务** (寻求社交互动的目标和结果，例如注视可以用于对个人评价信息的寻求、也可以通过促进有积极意图的交流从而形成进一步的合作)。



制造类人机器人

注视行为——HRI领域

- 根据研究目标和方法的不同，可以将HRI的注视研究分为三类：
 - **以人为中心** (着眼于人们对不同注视方式的反应)；
 - **以设计为中心** (研究机器人注视行为特点和外观的特征)；
 - **以技术为中心** (关注实现机器人不同社交注视方式的算法)。
- 已有HRI研究了社交环境中机器人的头、眼和身体朝向对人机交互的影响 (Srinivasan & Murphy 2011)，发现社交注视的功能包括：
 - 建立机器人代理人agent的身份
 - 传达机器人对某人感兴趣
 - 给出谈话轮换的信号
 - 表明机器人所指的是哪个物体

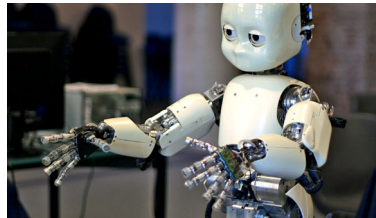


涉及的机器人注视行为包括：
 长时凝视、短暂一瞥、注视回避gaze aversion、快速注视转换、扫视、头部运动等

制造类人机器人

姿态和手势

- 机器人不同的动作姿势会影响人们的信任程度
 - 在一项人-机器人合作完成的游戏任务中，发现机器人手触摸脸、手臂交叉、身体后倾这些动作与人们对机器人较低信任度有关；
 - 而机器人向前倾斜、手臂置于膝上和张开双臂的姿势与值得信赖有关 (Lee et al. 2013)。
- 由于机器人的语言能力有限，因此姿势和手势是机器人向人传达情绪的重要手段
 - 人们可以从机器人的操作姿势中解读出机器人的情绪倾向 (Beck et al. 2010)；
 - 机器人专家通过研究人们是如何感知自己的身体空间位置，并将其应用于机器人，从而使机器人能够意识到自己的身体姿势 (Hoffmann et al. 2010)。



不断优化机器人的工作既可以帮助赋予机器人更接近人的能力，也可以在创建和测试人造模型的过程中对心理学和神经科学研究进行补充、修正。

制造宠物机器人



Paro (海豹)
用于治疗康复的
陪伴机器人



Genibo (小狗)
玩具机器人，韩国研制
可听懂语音指令如
“sit”，响应拍手，
摇尾巴，狗叫，拍照



Pleo (恐龙)
帮助儿童学习人际关系的建立，可学习语言指令，感知给它的特殊食物，对触摸做出反应

制造宠物机器人

提升宠物机器人陪伴性能的设计原则 (Kaplan 2001):

- 使机器人有拒绝命令者的自由；
- 让机器人表现出对所有者的依赖、促使所有者觉得有责任照顾机器人并帮助它不断成长 (Pleo的设计就考虑到了这一点，它在最初的几天里要在主人的陪同下学习如何行走)；
- 给机器人添加不成熟的青少年特质来触发人们的情绪反应；
- 情感交流 (主要通过非语言方式，由于机器人的语音理解困难)。



Paro (海豹) 是可用于治疗康复的陪伴机器人；
Genibo (小狗) 和 Pleo (恐龙) 主要用于提供娱乐

要想设计成功，宠物机器人应表现出与真实动物相似的依恋行为。Kaplan曾提议，可以把Ainsworth的陌生情境测验方法应用到宠物机器人和所有者的依恋关系测验中 (Ainsworth & Bell 1970, Kaplan 2001)。

陌生情境测验

- 是测量婴儿依恋类型的方法，1969年由美国心理学家安斯沃斯和威特提出，适用于2岁以下儿童。
- 具体做法：实验时将被测试的儿童与其母亲和一个陌生人以多种组合安排在实验室中游戏，来观察儿童在不同情境下的情绪行为反应，以确定儿童的依恋类型。
 - 各种组合方式包括婴儿与母亲在一起、与陌生人在一起、与母亲和陌生人在一起，独自一人时，母亲离开、回来时及陌生人出现、离开时等情境



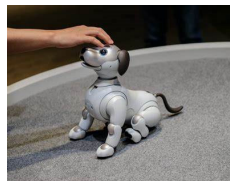
依恋类型:

- **安全型:** 母亲的在场使孩子感到足够的安全，能在陌生的情境下进行积极的探索和操作，对陌生人的反应也比较积极；母亲回来时，婴儿会立即寻找与母亲的接触，也很容易被安慰。
- **回避型:** 这类婴儿对母亲是否在场均表现出无所谓的态度，实际上并未与母亲形成特别亲密的感情联结。
- **矛盾型:** 这类婴儿每当母亲离开前就显得很警惕，母亲离开时表现的很苦恼、极度反抗；母亲回来时，婴儿的态度是矛盾的，既寻求又反抗与母亲的接触。

制造宠物机器人

人与宠物机器人的依恋关系

- 与**婴儿的依恋**类似，如果宠物机器人对其所有者表现出寻求接触的行为（如跟随）或能给出适当的信号行为（如朝所有者眨眼或者发出某种声响），就可以认为宠物机器人和人建立了依恋关系。
- 为了实现此目的，可为机器人增加和主人待在一起的动机驱动程序
 - 能够识别主人，独立活动时监控主人是否在场，主人不在时表现出短暂的抑郁，主人回来后的一段时间内表现出快乐
- 通过观察**真实的狗**对其主人的依恋行为也可以为宠物机器人的设计提供有意义的指导(Farago et al. 2014)。
 - 靠近主人、朝向和注视主人、摇尾巴、打招呼等都是很重要行为



04

人对机器人的 观点感受与行为



HRI研究

跨学科：

- 学科领域和边界仍在不断探索中

有50多年研究历史，最初是以电脑等科技产品为交互对象



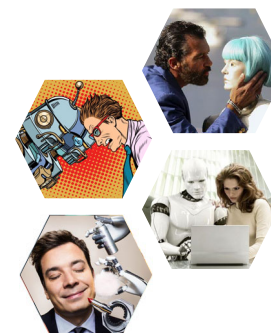
模拟研究

- 通过文本描述、照片、视频、人类控制等方式呈现机器人
- 优势：高度控制研究变量、快速、机器人可以更像人
- 劣势：人造情境，结果推广性存疑

真实世界研究

- 真正的机器人在真实的环境中（如商场、家里）
- 优势：真正的交互，可研究自然情境下长期交互中形成的人机关系，避免实验室研究中的反应偏差和新异性效应
- 劣势：耗时长，难实施，更贵，通常需要跨学科合作，需要买或造机器人并给机器人编程，技术的不可靠性，机器人能力和电量均有限

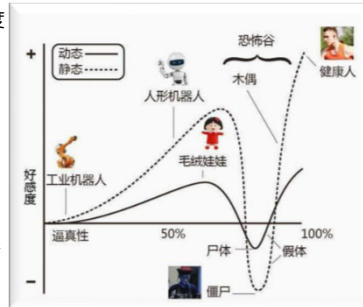
人-机器人交互研究



- 恐怖谷
- 第二自我：机器人改变、挑战与揭示人类
- 媒体等同：下意识对机器人使用社会规则
- 拟人化：透过人类滤镜看世界
- 感知机器人的心智
- 对机器人的情感依恋
- 物理具身：机器人vs其他科技产品
- 机器人虐待

恐怖谷 Uncanny Valley

- 1970年日本机器人专家森政弘提出
 - 当机器人看上去近乎人类但又不是人类时,人们的舒适度会急剧下降(Mori 1970, Mori et al. 2012)。
 - 其他都很类似人,但有一个非人本质特征时产生
 - 外貌像人但是动作古怪
 - 有手,但是冰冷或没有骨骼结构
 - 最初将此解释为让人联想到了死亡的躯体
- 争议
 - 有研究证据支持了恐怖谷的存在(MacDorman & Ishiguro 2006)
 - 也有其他探究发现人们更喜欢逼真的机器人,向实验参与者呈现从卡通到真实的连续面孔时,并没有发现会引起人的舒适感下降的现象(Hanson et al. 2005)。
 - 研究结果差异性可能源于用来衡量人们感觉的术语不同、呈现的机器人刺激的差异等。



恐怖谷存在的原因解释

- 进化论的观点

人类脑部中有一种逻辑偏袒性,会提供一种高度能力去感应并且排斥那些不健康、与疾病或死亡相关的人类畸形体,恐怖谷现象是人们这种排斥能的体现。
- 分类的不确定性

当人们不确定看到的实体是机器人还是自己的同类,即存在归类困难时,也会产生不适感。
- 线索的冲突感

当在同一个面孔中同时使用真人和人造的面部特征时,会引发更多的恐怖感(MacDorman & Chatto-padhyay 2016),即线索间的矛盾是引起人们不安的重要成分,因为这种矛盾会违背我们对正常人类的期望。



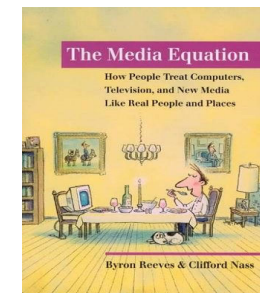
第二自我: 机器人改变、挑战与揭示人类

- 在20世纪80年代, Sherry Turkle从**精神分析**的角度研究人机关系,认为人们在使用计算机的过程中获得了对自身的**洞察** (Turkle 2005)。
 - 将个体经验与计算机过程相联系,如记忆,简化了我们对自身的理解
 - 给计算机编程,增强了我们对世界的**掌控感**,某种形式的**自我表达** (程序反映了编程者的性格)
 - 早期语句式计算机可教育儿童逻辑思维和过程,而当界面式计算机使得儿童失去了了解计算机如何工作的机会。
- 计算机促使人类思考自身的思维过程和本质(Turkle et al. 2006),同样的,机器人也可促使我们思考**何为人类本质**以及我们与机器人有何**不同**。
- 深度观察成人和儿童对待机器人的不同方式可以**解读**出他们的家庭状况和心理需求;通过观察儿童与机器人的交互可以**洞察**他们的个人状况和生活经历 (Turkle et al. 2006)。



媒体等同: 下意识对机器人使用社会规则

- 媒体等同Media Equation
 - 计算机科学家和社会学家Clifford Nass提出
 - 人们会将计算机和其他科技产品视为人类一样对待,即使人们知道它们并非人类
 - 人们会**无意识地将社会规则应用于计算机**(Nass & Moon, 2000)
 - 刻板社会分类: 性别、种族、群体
 - 社会规则: 实验者偏差、互惠原则、有礼貌
 - 服从权威
 - 他人在场提升诚实度

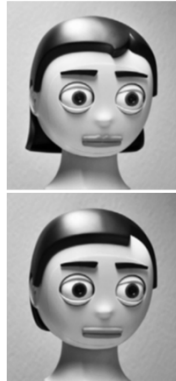


媒体等同——刻板社会分类

人们在与计算机或机器人进行交互时会使用人际交往中存在的**刻板印象**对其归类，如**性别、种族、群体**

性别：

- 人们认为有**女性**声音的计算机对爱和人际关系的了解更多、有**男性**声音的计算机懂得更多的计算机知识(Nass et al. 1997)，这与人们认为女性更擅长照料、男性更擅长处理事务(Huddy & Terkildsen 1993)的发现是相似的；
- 有短发、扁平嘴唇的机器人被认为更少女性特质，更适合从事典型的男性工作，如机械；而长发、嘴唇特征明显的机器人被认为更有女性特质，更适合从事典型的女性工作，如照顾小孩(Eyssel & Hegel, 2012)
- 博物馆的机器人请求人们捐款，女性声音比男性声音可得到更多的捐款(Siegel et al. 2009)



媒体等同——刻板社会分类

- 种族**：实验参与者评估有不同种族面孔的电脑智能体，发现人们对与自己同种族的人评价更高，认为其更聪明、迷人、有说服力、值得信赖等(Nass & Moon 2000)，即存在**种族偏见效应**；

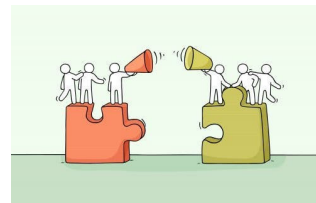
- 德国被试认为德国产、有德国名字的机器人更热情、设计更好、能力更强、更亲近（对照：同一个机器人，但被描述为土耳其产、有土耳其名字）(Eyssel & Kuchenbrandt 2012)



媒体等同——刻板社会分类

群体：人们日常的社会交往中会存在“**内群体偏见**” (in-group bias)，即相比于团体外成员，人们更愿意与团体内的进行互动，更愿意帮助有需求的团体内成员等(Fu et al. 2012)；

- 同样的，在一项实验中，设置了**计算机屏幕的颜色**与使用计算机者的**手臂绑带颜色**相同或不同的条件，发现人们对与自己“同色”的计算机**评价更高**，也**更愿意遵从**“同色”计算机发出的实验指令，说明仅仅是用颜色作为一种绑定手段，就会在人机交互中引发与群体内偏袒类似的现象(Nass et al. 1996, Nass & Reeves 1996)。



媒体等同——社会规则

- 人们在与计算机或机器人交互时会使用**习惯的社交行为**

- 实验者偏差** (experimenter bias)，被试总是尽量不冒犯实验人员、对实验人员表示出更多的礼貌行为

- 当要求人们对**使用过的计算机进行评价**时，与纸笔评价、使用另一台计算机进行评价相比，人们用**同一台电脑**进行评价时**计算机得到的性能评分最高**（事实上被试所用的计算机性能并无差异）(Nass et al. 1999)

媒体等同——社会规则

- “互惠原则”，我们更愿意对帮助过自己的人施以援手
 - 在合作任务中，若人们在先前**得到过计算机的帮助**，则在接下来的任务中人们也愿意**更多的帮助计算机。帮助的正确率也会更高** (Fogg & Nass 1997)
 - 在最初的交互中，如果计算机向人透露更多的自身信息，人们接下来也会更多地对计算机做出亲密的反应 (Moon 2000)
 - 如果照料机器人先向人寻求帮助然后再向人回报以帮助，人们对其的评价更积极 (相比没有互动的机器人)
 - 在囚徒困境游戏中，人们对机器人和人表现出相同多的互惠行为(Sandoval et al. 2016)，如果机器人第一轮表现出合作而非随机行为，人们对其的互惠行为也会更多

囚徒困境 Prisoner's Dilemma

两个犯罪团伙被关入监狱，不能互相沟通情况。

- 如果两个人都不揭发对方，则由于证据不确定，每个人都坐牢一年；
- 若一人揭发、另一人沉默，则揭发者因为立功而立即获释，沉默者入狱十年；
- 若互相揭发，则因证据确凿，二者都判刑八年。

		PRISONER 2	
		Confess	Lie
PRISONER 1	Confess	-8, -8	0, -10
	Lie	-10, 0	-1, -1

媒体等同——社会规则

- 与对待计算机的态度相似，人们同样会对**机器人有礼貌**
 - 与Nao交互时，一些人对它有礼貌并愿意帮助它；另一些人直呼其名且喜欢控制它(Rehm & Krogsager 2013)
 - 50%的人会对大学校园里的接待机器人有礼貌，如打招呼(Lee et al. 2010)

但是，并非所有对机器人的反应都是自动、无意识的，人们对机器人表现出社会行为的程度也存在差异。

- 人们对可交谈的轮椅机器人的语言回应基于人们的目标，且清晰意识到自己的反应选择(Fischer 2011)。

媒体等同——服从权威

HRI研究中的一种**常见方法**是把人-人心理学实验改编成人-机实验，旨在检验我们对机器人的行为反应是否对人类相同。

Milgram的服从权威实验

- 1961年耶鲁大学心理学助理教授Milgram设计
- 请志愿者协助“权威专家”们考察学习人员的“单词记忆能力”
- 学习人员的身体连接到一根电击棒上，电击棒由志愿者操控，学习人员每背错一个单词，志愿者就对其进行电击一次，电压强度每次提高15伏，并会让志愿者了解相应的电击值大小给学员带来的痛苦程度。
- 实验过程中，很多学习人员在高压电流的刺激下，反复挣扎、发出惨叫，请求停止试验，按照一般人的逻辑，目睹学员的种种挣扎，很多志愿者应该停止电击，放弃实验。但事实是，所有40名志愿者中，只有很少的人停下来，其余全部完成了实验，即绝大部分志愿者对被学习人员都实施了最高450伏的电击。



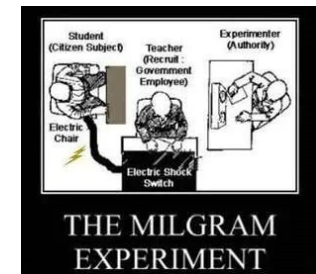
Milgram服从权威实验情景图

媒体等同——服从权威

- Milgram的实验排除了性别差异、志愿者自身的暴力倾向等因素，**证明了人类有一种服从权威命令的天性**，在某些情景下，人们会**背叛**自己一直以来遵守的**道德规范**，听从权威人士去**伤害无辜**的人。这会让我们很容易的联想到德国纳粹的反犹太人运动、南京大屠杀这类悲剧，引发我们对**人性**的思考。

服从实验在HRI领域的再现

在机器人版的服从实验中，要求参与者对机器人进行电击，实验结果是**100%的参与者**会不顾机器人的抗议和请求，对机器人进行**最大程度的电击**（远高于Milgram实验中65%的最大电击比例），说明在这种实验情境下人们**不会把机器人看作是有生命的个体** (Bartneck et al. 2005)。



Milgram服从权威实验情景图

媒体等同——他人在场提升诚实度

心理学实验(Covey, et al. 1989) :

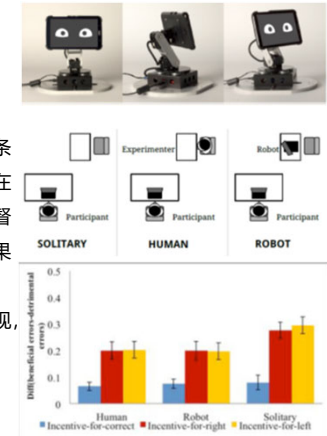
- 告知被试要参加一项**视觉记忆任务**，以测定他们的记忆能力
 - 根据被试记忆和作答时中是否有实验人员认真监督分为**高监督组**和**低监督组**两种条件
 - 每种条件下各有一半的人事先被告知由于记忆任务难度较大，无论测试结果好坏都会很看重被试的测试诚信度（**低诱骗动机条件**），另一半人被告知当且仅当测试成绩特别突出时才会质疑其作弊（**高诱骗动机条件**）。
- 实验结果表明旁观者的在场**可以提高人们的诚实度，增加**诱骗动机**可以降低人们的诚实度。



媒体等同——他人在场提升诚实度

机器人版本的类似实验 (Hoffman et al. 2015)

- 将监督者由人替换成机器人、并添加参与者单独在场的实验条件，在施加了同样诱骗动机的情况下，发现虽然参与者仅仅在单独在场的情况下显得不那么诚实；有机器人监督或人工监督时参与者的诚实度值会高，并且机器人监督和真人监督的结果并没有显著差异。
- 这说明机器人的社会存在会使人们感觉好像他们正在被人监视，即机器人的监督效果与**人类在场类似**



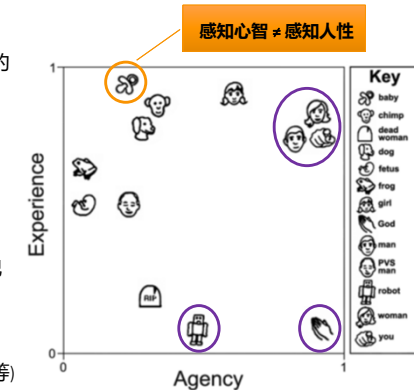
透过拟人滤镜看世界

- 拟人化Anthropomorphism**是指我们倾向于把类人的特点、情感和动机等赋予给非人类的实体，诸如动物、神灵和物体等 (Epley et al. 2007)。
- 人类拟人化倾向的解释**
 - 做为人，我们最了解的就是人自己。当我们与不熟悉的非人事物打交道时，会用自己熟悉的关于人的知识去理解它们。
 - 生物基础**：镜像神经元mirror neurons。当观察非人物体做出类似人类的行为时，会激活镜像神经元，让我们有“亲身经历”的感觉 (Epley et al. 2007)；
 - 当人与非人形机器人进行社会互动时（观看扫地机器人被人骂的视频），镜像神经元更活跃 (Hoenen et al. 2016)。
 - 我们的拟人化倾向会随着**非人类实体的人性化程度**的增加而增加 (DiSalvo et al. 2002)，如拥有更多人类面孔特征，行为更不可预测
 - 孤单、缺少**社会联结**的人更容易把机器人拟人化 (Eyssele & Reich 2013)。



感知机器人的心智

- 机器人能否思考一直是AI领域的一个吸引人的问题，其答案取决于：
 - 科技赋予机器人的能力
 - 提出该问题的人的观点
- 对个体心智能力的判断取决于 (Gray et al. 2007):
 - 个体自身的代理能力 Agency** (自控力、记忆力、情感识别、计划性、思维逻辑等)
 - 个体自身的体验能力 Experience** (饥饿、恐惧、开心、渴望、个性、骄傲、尴尬等)



> 感知机器人的心智

- 机器人的外貌影响人们对其心智的感知
 - 看到机器人的**正脸**（而不是后脑勺）(Gray & Wegner 2012)或机器人的显示屏幕**上有脸**（而不是无脸）(Broadbent et al. 2013)时，人们都会觉得机器人的**心智水平会更高**。



Humanlike Appearance

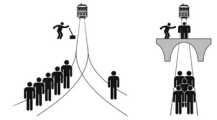
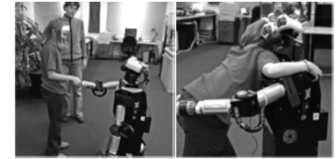


Mechanical Appearance

> 感知机器人的心智

道德决策与心智感知相关

- 道德决策受两方面的影响(Gray et al. 2012)
 - 动机——实施道德决策的人
 - 体验——被道德决策影响的人
- 由于机器人拥有某些心智，那么：机器人是否应该遵循道德行事？是否应该被有道德地对待？
 - 儿童在与Robovie机器人进行社会互动后（聊天、玩耍、拥抱），实验者要把机器人放进壁橱里，机器人抗议说不公平，自己怕黑不想待在壁橱里。
 - 大部分儿童都认为机器人有思维和情感，是一个social agent
 - 所有儿童都认为可以把扫把放进壁橱，不可以把人放进壁橱，有不到一半的儿童认为可以把机器人放进壁橱
 - 本科生和Robovie玩游戏，犯相同错误的情况下，他们对机器人（相比于人）的问责度更低(Kahn et al. 2012b)
 - 在**道德两难困境**中，即使做出和人一样的功利性选择，我们也会更少责备机器人(Malle et al. 2015)。



> 对机器人的情感依恋

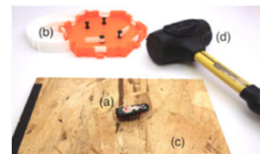
研究表明，人们会对机器人产生某种形式的情感依恋(emotional attachment)

- 儿童与机器狗Aibo进行短暂互动后产生了情感依恋(Weiss et al. 2009)
 - 以互动的积极性、认为机器人是否有感知能力、是否将机器人作为同伴等作为衡量因素
- 几乎所有的儿童都会将生物性、心理活动、社交能力和道德立场等与真狗联系起来，而较少(但仍占多数)的儿童会把这些特征与毛绒玩具狗、机器狗联系在一起(Melson et al. 2009)
- 75%的儿童认为机器狗可以成为他们的朋友
- Aibo机器狗的主人尽管意识到它有电池、是一个物体，但仍认为其有生命、有情感和个性，并对其形成情感依恋(Friedman et al. 2003)。







> 对机器人的情感依恋

- 共情**empathy，又称同理心、同感等，是指人类体验别人内心世界的的能力，由人本主义创始人罗杰斯提出。
 - 让被试看一个**害怕自己失忆**的机器人，并且随后展现机器人真的失忆(实验条件)和并没有失忆的场景(控制条件)，结果表明，实验条件下被试更加同情机器人(Seo et al. 2015)。
 - 给被试观看人或者机器人手指**割伤**的图片，fMRI数据显示，不同受伤对象引起的大脑共情反应相似，但机器人略低(Suzuki et al. 2015)；
 - 让被试用锤子打机器虫，当实验人员提供给被试更多机器虫的**背景信息**时，被试打虫时会更犹豫，即当人们对机器虫有更多的了解时，对机器虫的共情也会增强(Darling et al. 2015)。



物理具身：机器人vs其他科技产品

我们对待真实的机器人、屏幕呈现的机器人、计算机的态度是不同的

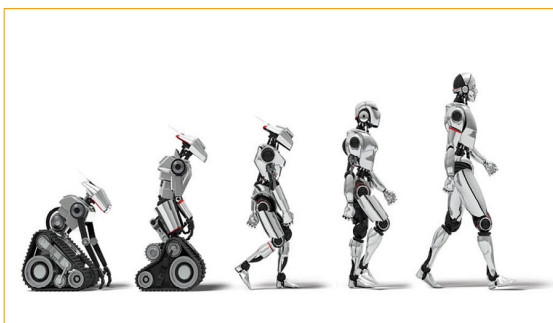
-  **更喜欢与真实机器人的相处**(Wainer et al. 2006); 同时电击屏幕呈现的机器人和真实存在的机器人时, 儿童会对后者表现出**更多的同情心**(Kwak et al. 2013)。
-  更倾向于和机器人建立**治疗联盟关系**(Kidd & Breazeal 2008)。更愿意服从机器人的指令(如遵从减肥食谱、放松指令)(Mann et al. 2015); 人们再次与机器人互动的渴望也高于再次与平板电脑互动的渴望。
-  **信任**对人与机器人成功互动至关重要。当前人们对数字设备安全性和隐私性的担忧很可能会转移到机器人身上(Chin et al. 2012); 机器人表现出的可性度和失误率、个性和拟人度都会影响人们对它的信任(Hancock et al. 2011)。
-  **触摸机器人**会使人产生不同的**生理反应**。如触摸人形机器人的眼睛或臀部会比触摸手臂引发人们更强烈的生理反应(Li et al. 2016); 抚摸Paro可以帮助**降低**住院和居家患者的**血压和心率**(Robinson et al. 2015)。

机器人虐待

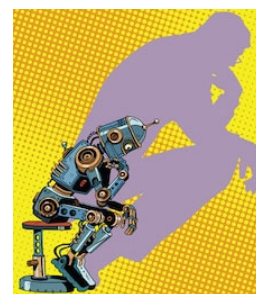
- 没有成年人在场时, **孩子们会聚集在一起并表现出虐待机器人的行为**(如挡住机器人的去路、忽略机器人的礼貌请求、踢打机器人等); 只有孩子们觉得**无聊或有成人介入**时, 孩子们才会停止虐待行为(Brsic'et al. 2015); 虐待机器人的儿童和青少年中, 7岁到十几岁的男孩**最具攻击性**(Scheeff et al. 2002)。
- 机器人表现出**害怕、悲伤**时, 只会遭到**更糟糕的虐待**; 只有机器人**生气、反抗**虐待行为时**情况才会好转**。
- 即便机器人感觉不到痛苦, 关注虐待、滥用机器人的问题也是**很有意义的**, 可以使我们对人类的行为方式以及如何减少对同类、动物的伤害行为有更深入的了解。



总结 & 展望



总结

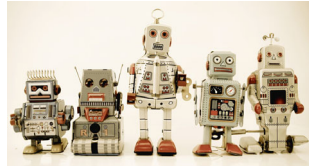


- 机器人已经开始进入人类生活, 可以提供陪伴、增加交流、减少花费, 尤其适用于医疗健康领域
- 工程师尝试制造外形像人或动物的机器人以提升用户体验
- 但是, 过分像人的机器人也可能让人觉得不舒服(恐怖谷)
- 与机器人交互会促使我们思考“我是谁”
- 先天和后天的因素导致我们通过拟人滤镜看待机器人
- 我们无意识地对机器人和电脑等科技产品施以社会规则
- 我们不认为机器人与人有相同的思维, 也不赋予它与人相同的道德权力和义务
- 我们会对机器人宠物产生情感依恋, 但低于真的宠物
- 我们更认可真实的而非虚拟的机器人或电脑



展望

- 人们更愿意和机器人而非人交互 ⇒ 社会的崩溃(Whitby 2008)
 - 机器人：永远愿意与人互动、聪明、忠诚、可预测
- 今日之科技远达不到制造完美机器人的程度
 - 电池、材料、语言理解、执行器、感知器.....
 - 对人类认知、情感、行为的理解以供机器人建模



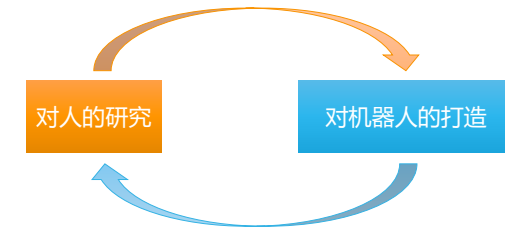
我们希望机器人达到什么程度？为什么？

- 为何要超出机器人为人类服务的需求范围制造更像人的机器人？
- 机器人究竟要多像人？
- 我们制造完美机器人的动机是因为想打造完美的人吗？但是机器人不是人，又怎么称得上完美的人？
- 人的本质是什么？



展望

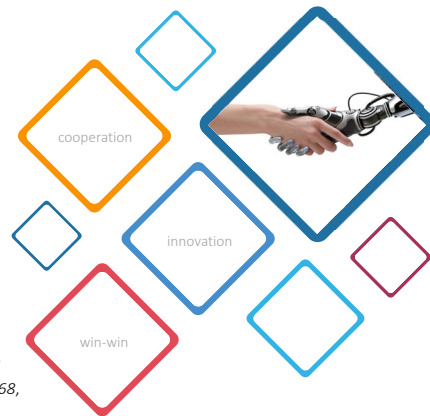
- 人机交互的很多研究仍处在探索阶段，但是扩张和发展迅速
- 目前的大多数研究都集中在机器人的技术方面，还需要在人类对机器人的反应和合作模式等方面进行更多的研究
- 这将是一个双赢的合作：



THANKS

参考文献：

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, 68, 627–652.



2、对个体的心理研究助力于提升用户对人工智能/机器人的接纳、信任和喜爱

算法反感
Algorithm Aversion

Castelo, N.¹, Bos, M. W.², & Lehmann, D. R.³ (2019). Task-Dependent Algorithm Aversion. *Journal of Marketing Research*, 56, 809-825.

¹ Assistant Professor of Marketing, University of Alberta

² Senior Research Scientist, Snap Inc.

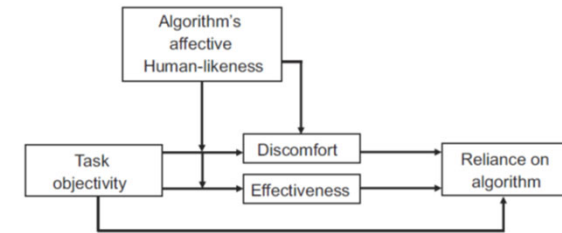
³ Professor of Business, Columbia University

- 早在上个世纪50年代，已有研究发现诸如线性回归的简单算法就可在诊断医药和心理疾病上超过人类(Dawes, Faust, and Meehl 1989; Grove et al. 2000; Meehl 1954)。
- 今日，算法在越来越多的领域超越了人类专家，包括诊断复杂的疾病(Simonite 2014)、提供法律建议(Krasnianski 2015)，算法也可以胜任一些看似主观的任务，比如从面部表情和声调中检测出情绪。
- 因此，算法可以为用户和公司提供大量潜在的服务，比如市场决策。但前人研究发现人们的默认选项是依赖人而非算法，即使会导致更差的结果。

- 如果医药建议来自算法而非人类医生，人们对其的信任度更低(Promberger and Baron 2006; Longoni, Bonezzi, Morewedge 2019).
- 人们不喜欢自动化的医疗照料，因为他们认为其忽视了自己独特的情况(Longoni, Bonezzi, and Morewedge 2019)。
- 人们认为医生独自做出的诊断=询问同事后做出的诊断，二者均比依靠算法辅助做出的诊断更好(Shaffer et al. 2013)。
- 对股价的预测上人们也更少依赖算法(O'neal et al. 2009)，人们也更容易把责任转嫁到其他人类身上。人类可为其行为提供解释，被知觉为高自信，有始终如一的信誉，有关于未来的信息，而这些算法都无法体现。
- 在雇员挑选和聘用决策中，人们会认为人类面试官比算法更有效、专业、公平、个人、灵活、准确。
- 在学业成绩预测上，当看到算法犯了错误之后，人们更倾向于选择人类来预测学生的成绩，即使会得到次优的预测结果(Dietvorst, Simmons, and Massey 2014); 当看到算法在避免错误、恰当对属性赋予权重，对信息加权保持一致性上均优于人类之后，人们还是认为人类比算法更擅长从错误中学习，通过练习变得更好，从沙砾中找到钻石，发现规则的例外。

65

假设模型：任务客观性减少了使用算法执行任务的不适感，提高了算法的感知效率，并增加了对算法的依赖，但是当算法具有很高的类人情感时，这些影响会减小。

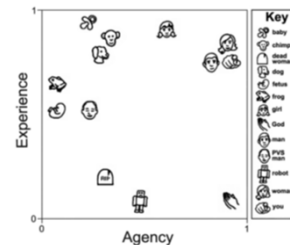


66

主、客观任务二维划分

- 人们通常认为客观任务需要逻辑的、基于规则的分析，而主观任务则需要直觉和“直觉”intuition and “gut instincts”
- 去人性化的研究表明，人们认为人类具有两类能力

- “人类独特能力human uniqueness abilities”，它可以与人类与其他动物区分开来，但可以与机器共享，这类能力在本质上往往是认知性的（例如逻辑和理性）
- “人类天然能力human nature abilities”，可以与其它动物共享，而不能与机器共享，这类能力本质上倾向于情绪或情感（例如温暖和直觉）
- 诸如机器人之类的机器被认为缺乏人类的自然能力（情感性），但具有人类的独特能力（认知能力）



- 对心智mind的研究也支持类似的划分：
 - 能动性agency, the ability to engage in intentional planning and action;
 - 感受性experience, the ability to subjectively experience emotions and sensations

67

研究1：26个任务

- part1, 让人0-100评价其客观性、重要性、对运用算法完成该任务的熟悉性;
- part2, 让人0-100评价是否信任算法或非常有资质的人完成。
- 总体对人的信任高于算法，但在某些任务上（预测股价、天气、分析数据、指方向），算法优于人。
- 线性回归发现，人们对算法的信任性会在重要性更强、更熟悉算法用于该任务、更客观的任务上更强。

研究3：MTurk上的被试就9项任务（预测学生成绩、预测员工表现、推荐疾病治疗、预测再犯、开车、推荐电影、诊断疾病、预测个人特质、预测笑话是否搞笑）评价自己更偏好算法还是人（100-0）。

- 知情组被明确告知已有研究证明算法在上述任务中优于人，不知情组不提供上述信息。
- 结果知情组比不知情组更倾向于选择算法，且如果根据研究1的得分将任务分为客观任务和主观任务，该现象在客观任务上更明显。

68

研究2: field study

- 在facebook上贴广告, 关于人/算法提供约会/财务建议的。考察广告的点击率 click-through rate (CTR), 看到广告的人里有多少人点击广告。
- logistic regression显示对人的广告和财务广告的点击率更高, 交互作用显著, 即约会广告对人比对算法的点击率显著更高, 而财务广告人的优势仅边缘显著。

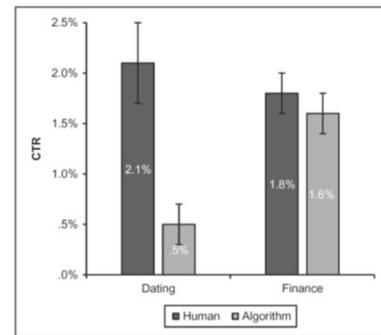


Figure 2. Algorithm aversion for a subjective (but not objective) task.
Notes: Error bars represent standard errors.

69

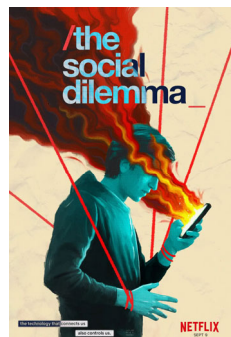
研究4: Prolific Academic上的被试, 被算法或人推荐约会对象。

- 主观条件, 被试得知该任务在集中于个体的情绪、情感和直觉的情况下完成得更好;
- 客观条件, 被试得知该任务在集中于可量化的数据如可测量的性格特质时完成得更好。
- 被试报告自己在多大程度上更相信算法 (0相信人-100相信算法)。
- 总体上被试更相信人, 但是在强调客观性的条件下, 他们相信算法的概率会提高。中介效应分析显示知觉到的任务客观性和知觉到的算法效果链式中介了任务类型和对算法信任之间的关系。

研究5: Facebook基于算法的约会服务广告, 其中一个强调量化对完成该服务的优势, 另一个没有强调。结果是强调量化的广告点击率更高。

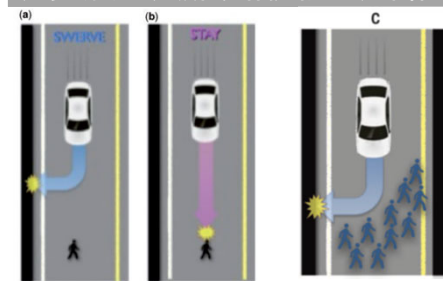
70

3、对群体的社会心理研究助力于人工智能相关的社会服务和政策制定



监视资本主义: 智能陷阱
The Social Dilemma (2020)

道德选择: 你认为发生重大事故时, 自动驾驶的程序设置应该优先保护路人还是乘客?



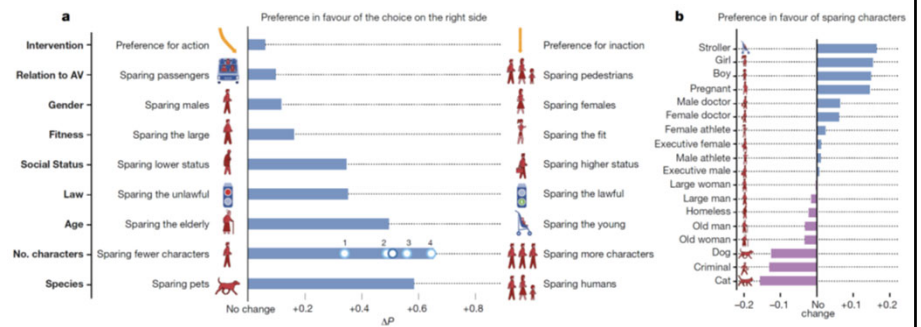
The social dilemma of autonomous vehicles. *Science*, 352(6293), 1573-1576. Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016).

71

The Moral Machine experiment. *Nature*, 563(7729), 59-64.

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J. F., & Rahwan, I. (2018).

233个国家/地区/属地或领地中收集3,961万个决定



个体差异 (年龄, 教育程度, 性别, 收入, 政治, 宗教观点) 影响非常微弱

72



74

75

76

Could a Rising Robot Workforce Make Humans Less Prejudiced?

Jackson, J. C., Castelo, N., & Gray, K. (2020). *American Psychologist*, 75, 969-982.

2017 McKinsey Report 预计未来将有50%的工作岗位被机器人替代。

- 正如工业化兴起加剧了阶层之间的社会矛盾，外包产业给印度和中国加剧了人们对亚洲的偏见；
- 而机器工人的兴起将导致人们抢夺剩下的岗位，更加支持反移民法案。
- 但是，人们对移民和机器人的担忧很少同时都有：



奥巴马只警告人们小心机器人



而特朗普只警告人们小心移民

奥巴马的策略似乎是视线转移，即担心机器人会降低人们对移民的担心。

77

因此，机器的影响可能是双重的：

- 减少就业资源增加人群间的冲突(Frey et al., 2018; Im et al., 2019)；
- 或激起人们的“泛人类主义”增进人际亲密感
 - “泛人类主义” (Panhumanism)：即在不同人种和宗教之间知觉到共同的人类特质

这种差异在于人们是如何认识机器人的：

- 认为机器人是独特的社会个体 *distinct social identity* 可能促进人们的团结，
- 认为机器人是先进经济工具 *economic advance* 则可能导致人们对人际外群体的敌视和偏见。

78

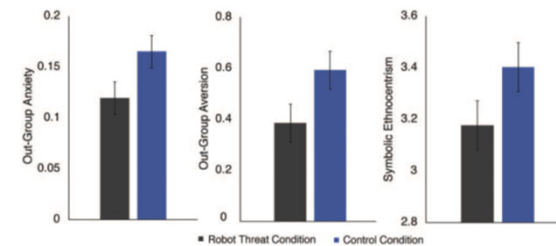
研究1：7点量表测量人们对外群体的焦虑，包括机器人、不同种族的人、不同性取向、移民/外籍劳工、不同宗教的人。

- 结果人们对机器人的焦虑可预测更少的对外群体人的焦虑、厌恶、民族中心主义
 - 对外群体的厌恶，问卷测量人们是否拒绝和不同种族、不同宗教、移民、同性恋当邻居，总分作为外群体厌恶指标。
 - 象征性的民族中心主义 *Symbolic ethnocentrism*，测量对移民的偏见

79

研究2：被试读一篇机器人将取代/无法取代人类工作的文章，然后测量研究1的问卷。

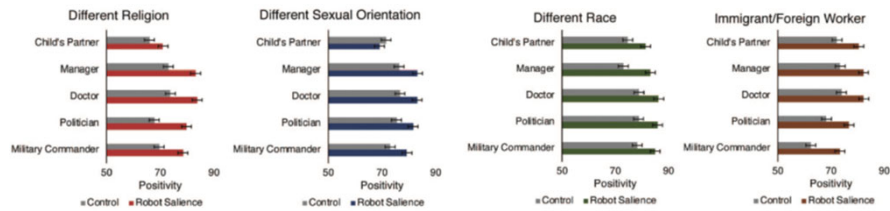
- 结果发现凸显机器人取代人类工作将降低人们对外群的负性态度。



80

研究3：被试评估自己1-100接纳研究1中的外群体：与自己的孩子结婚、管理自己工作的公司、当公务员、给自己做一台高风险的手术。机器人组也评价机器人，而控制组则完全没有提到机器人。

- 结果同研究2.



81

研究4：被试作为一个未来社区的财政官员，给30位有不同职业的居民发奖金（人人都有基本工资满足温饱），总额3万人均100。

- 控制组全是人类，白黑黄种各10人，男女各半。
- 机器人凸显组额外还有10个机器人，然后从40人中随机抽30人出来（大概25%为机器人）。
- 结果机器人凸显组会更公平地在不同人种间分配奖金，而全人组会更少地给黑人奖金，亚洲人比白人少但不显著。



Table 3
Pay Gap Between White and Minority Individuals by Study 4 Condition

Condition	Black pay gap	Asian pay gap	Minority pay gap
Robot salience	-\$1.29 [-\$3.45, \$0.86]	-\$1.18 [-\$3.33, -\$0.98]	-\$1.50 [-\$3.62, \$0.61]
Control	-\$4.67 [-\$6.42, -\$2.92]	-\$2.58 [-\$4.33, -\$0.83]	-\$4.37 [-\$6.12, -\$2.62]

Note. Coefficients indicate mean pay gaps, and brackets indicate 95% confidence intervals by condition. All participants are included in the Black and Asian estimates. Only White participants are included in the minority estimates.

82

研究5：被试写一段机器人劳工增长的短文（5-6句话），或“一个自己非常焦虑的情境”（负性控制组），或早饭吃的什么（中性控制组）。

- 为了测量“泛人类主义”(Panhumanism)，让被试把13类人（非裔美国人，移民，西班牙裔，同性恋者，基督徒，右翼极端主义者，左翼极端主义者，患有疾病的人，犹太人，天主教徒，穆斯林，亚裔美国人以及您）拖放到2个盒子中，与自己放在同一个盒子里的类别数量的多少用以测量“泛人类主义”。同样用研究1的问卷测量组外偏见。
- 路径模型发现机器人显著组的被试表现出更强的泛人类主义，且进一步减弱对外群体人类的偏见。

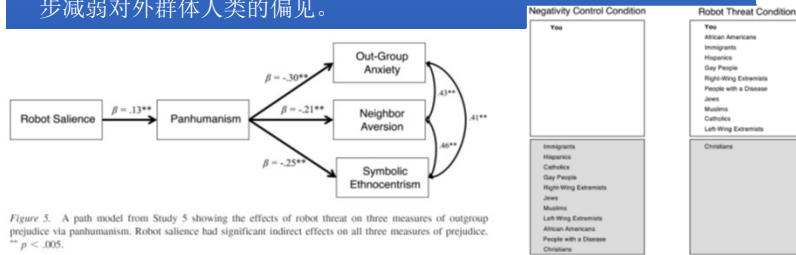


Figure 5. A path model from Study 5 showing the effects of robot threat on three measures of outgroup prejudice via panhumanism. Robot salience had significant indirect effects on all three measures of prejudice. ** $p < .005$.

研究6：简化“泛人类主义”测量。一系列距离不同的两个圆圈，其中一个是自己，另一个是人类 Humanity，让被试选择哪个最能代表自己和全人类的关系。其余测量同研究3。

- 路径模型发现，“泛人类主义”完全中介了凸显机器人对外群体偏见的的影响。

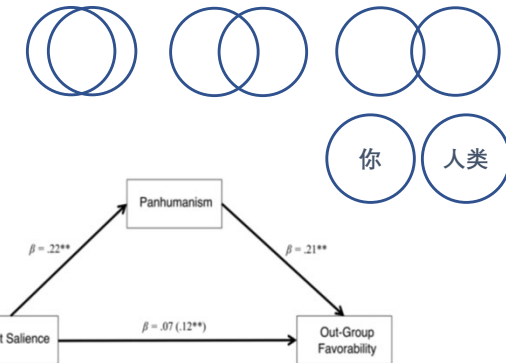


Figure 6. The path model from Study 6 showing the effects of robot threat on outgroup favorability via panhumanism. Robot salience significantly increased outgroup favorability through panhumanism. Participant ethnicity (Black vs. White; Hispanic vs. White) did not moderate the effects of robot salience on panhumanism or outgroup favorability. ** $p < .005$.

84