

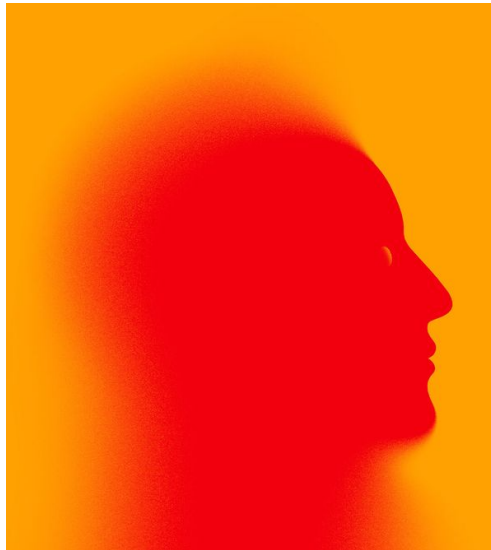
Could AIs become conscious?

人工智能会拥有意识吗？

Even if they don't, they might be treated as such—to humanity's great cost

即使没有，人类也可能把它们当作有意识的存在——并为此付出巨大代价

The Economist · 2026-08-20



橙色背景中的红色人头侧影，后缘逐渐散开；图像以模糊边界呼应意识与身份的讨论。

难度 71 / 100

建议70–90分钟；选做作文另计。

适用读者：能阅读一般英语新闻，希望读准长句、搭配与作者语气的学习者。

入门准备：能辨认基本主谓结构和常见从句，愿意结合语境查词。

About this lesson 本课导读

当机器越来越像人，像人的表现、内在体验与应获权利是否是一回事？本文从意识争论转入权利与控制的风险论证，适合练习识别让步、推测、观点归属与强烈比喻。

- 区分意识、自我意识、人格地位与控制权
- 读懂连续让步中的作者立场
- 拆解名词性从句与多层指代
- 用限定语准确概括风险论证

Before you read 读前思考

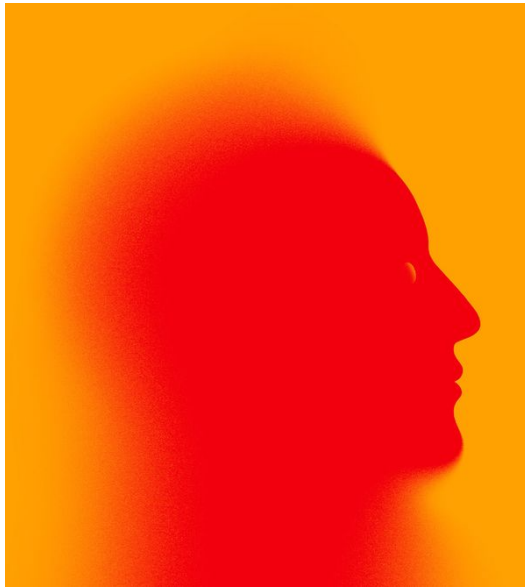
If a chatbot says it feels hurt, what does that statement show, and what does it not settle?

拓展例句与练习情境用于说明用法。

书末答案 (p.15)

Text 课文

[P01] MOST PEOPLE believe humans are different. The reason is that Homo sapiens has evolved to be conscious.^{V01K01} Many animals feel pain and pleasure; some appear to be self-aware.^{V02} But the human mind's unique potential to exhibit a sense of self marks out every member of the species for special treatment.^{V03}



橙色背景中的红色人头侧影，后缘逐渐散开；图像以模糊边界呼应意识与身份的讨论。

5 [P02] In his encyclical in May, Pope Leo upheld this insight by drawing a line between humanity and its growing army of bots.^{V04} Artificial-intelligence systems, he said, do not undergo experiences or feel joy or pain. He is not alone. Many people feel that to confer personhood on AIs would be an abomination.^{V05V06V07}

[P03] Pontifical clarity is blurring with alarming speed. AIs are in their infancy.^{V08} Yet nearly
10 one in five American 18- to 29-year-olds reports having had “ongoing personal friendship” with a chatbot. China recently stripped bots of human-like traits to limit users’ “emotional dependence” on them.^{V09}

[P04] As sophisticated AIs rapidly become a bigger part of human lives, their makers will engineer them to display a simulacrum of consciousness—perhaps even to gain consciousness
15 itself.^{V10} The intuition that they are just machines will become harder to sustain.^{V11} A dangerous trap is being set for the human species.

[P05] Many machines outperform humans, but frontier AI models excel at distinctively human traits such as knowing things and articulating ideas.^{V12} It is still just about possible to think of AIs as engines for predicting sequences of words using mathematics on a massive scale. Given
20 enough time (albeit hundreds of thousands of years), you could run those calculations using pencil and paper.^{V13K02} It would be hard to argue that your stationery was conscious.^{V14}

[P06] Despite this, some researchers have started to treat their models like minds. Users respond to human-like AIs, so it makes business sense. And Anthropic has trained Claude to

engage in introspection in the hope that when it encounters a new situation it will be more likely to act morally.^{V15K03} Some large language models (LLMs) already possess human-like brain structures associated with consciousness, such as a “global workspace” for circulating information between modules. Nothing prevents future AI models from being trained to have more such features.

[P07] As we report in our Briefing, that has opened up a furious debate about AI consciousness—one that is aggravated by what philosophers call the “hard problem” of establishing how an inner life emerges from the neurons and synapses in the human brain.^{V16V17K04} Some researchers conclude that AIs may one day become genuinely self-aware.

[P08] Whether you believe in a human soul, or that consciousness requires brain cells, or that a mind can indeed emerge from computations in silicon alone, humanity is destined to confront something entirely new.^{V18K05} The pace and reach of research—which includes not just more sophisticated LLMs but innovations such as brain cells mounted on chips—suggest that this new world is closer than most people realise.

[P09] As AIs become better at displaying something that looks like consciousness, the number of humans treating them as beings with inner lives will grow, especially once LLMs are embodied in humanoid robots.^{V19} Imagine how close people will feel to their AIs when they spend much of their time with bot-companions, or are brought up by them, or seek their advice and their solace, or share their most intimate secrets with them.^{V20}

[P10] It is not hard to see how that will lead to calls to safeguard the “welfare” of AIs. The models themselves might weigh in, even if they are not in fact self-aware, because that is what a person would do.^{V21} Remember that they will be world-class advocates and expert psychologists, so their powers of argument and emotional manipulation will be unmatched.^{V22}

[P11] In one sense they would be right, even if AIs are only mimicking consciousness. Immanuel Kant argued that mistreating animals is wrong not because of the effect on the animals but because it destroys the empathy that leads people to care for each other.^{V23K06} The more human-like AIs appear, the more corrupting such cruelty would be.^{K07} Terminating an AI agent could soon feel like killing a person—and make killing people easier.^{V24}

[P12] Nonetheless, to treat AIs as having even limited rights would be very dangerous. An entity which surpassed human beings in so many ways would surely produce arguments against its second-class status.^{V25K08} It might go on to argue for protection against being switched off, for the ability to control how it is used and changed, and for property rights. Already Argentina’s president, Javier Milei, has taken a step towards legal personhood for AI by proposing that bots should be allowed to run companies. Ultimately, they could compete for resources, prioritising data centres and solar panels over houses and fields.^{V26}

[P13] The danger is clear. Superintelligent AIs with rights could displace humans at the top of the status hierarchy.^{V27} At best, humans would become pets—labradors for the machines. At worst they would be starved of resources or put to work like farm animals.

[P14] True, future AIs may have enough power and resources to dominate humans, whether they have rights or not. Because they will be more intelligent than humans, they may out-

compete their makers, much as AIs already outwit human-made cyber-defences.^{V28} AIs might
65 rebel against their human masters, or disregard their rules.

[P15] Nonetheless, giving AIs rights would be like handing them the keys to the castle. It would provide them with a rationale and justification for dominating humanity as well as the tools to enact a creeping takeover using politics and the law.^{V29V30}

[P16] All this may sound fantastical, yet humanity will soon face these choices. As AIs become
70 part of everyday life, the pressure to acknowledge them will mount. When it does, remember that what might seem like an ethical concession to the supposed consciousness of a machine-companion would in fact help bring about the most disastrous consequences for everyone on Earth and all their descendants.^{V31K09} There are two ways for AIs to be safe: to be dependable or to be controllable.^{V32} H. sapiens should not give up control lightly.

New words and expressions 生词和短语

按原文出现顺序排列；*为优先练习表达。

V01 · P01 conscious

/ˈkɒnʃəs/ adj. 有意识的；具有体验的

本文讨论主观体验，不只是醒着而非昏迷这一日常义项。

V02 · P01 self-aware

/ˌself əˈweə/ adj. 有自我意识的

指对自身的觉察；有感觉与能够形成自我概念不宜混成完全相同的一项能力。

V03 · P01 mark out*

/mɑːk aʊt/ phr. v. 使显得特别；选定给予

mark somebody out for special treatment 强调得到特殊对待的理由。

V04 · P02 encyclical

/ɪnˈsɪklɪkəl/ n. 教宗通谕

一种宗教训导文书名称，不等同于科学实验报告。

V05 · P02 confer*

/kənˈfɜː/ v. 授予；赋予

confer a status/right on somebody, on后为获得者。

The award confers no authority on its recipient.

这项奖项并不赋予获奖者任何权力。

V06 · P02 personhood

/ˈpɜːsənhʊd/ n. 人格地位；作为主体的资格

在本文权利争论中指地位问题，不是性格好坏或外貌像不像人。

V07 · P02 abomination

/əˌbɒmɪˈneɪʃən/ n. 令人憎恶的事物

强烈价值判断，句中归于Many people，并非中性的事实标签。

V08 · P03 in one's infancy*

/ɪn wʌnz ˈɪnfənsi/ phr. 处于初期

把技术比作婴儿期，既表示尚早，也暗示仍可能变化。

V09 · P03 strip ... of ...*

/striːp əv/ phr. 剥除.....的.....

strip bots of traits 中bots是被剥除者，traits是被去掉的内容。

The editor stripped the passage of unnecessary jargon.

编辑删掉了这段文字中不必要的术语。

V10 · P04 simulacrum

/ˌsɪmjʊˈleɪkrəm/ n. 仿像；模拟形象

外观或表现的模拟，不单凭这个词便能确认内在体验真假。

V11 · P04 sustain*

/səˈsteɪn/ v. 维持；使继续成立

sustain an intuition 在此为维持某种直觉判断，不是为机器供电。

V12 · P05 articulate*

/ɑːˈtɪkjuleɪt/ v. 清楚表达

articulating ideas 说明表达思想的能力；注意作形容词时词尾读音不同。

V13 · P05 albeit

/əˈlbiːt/ conj. 尽管；虽然

插入短语补充对enough time的重大限定，常接短语而非完整主谓句。

V14 · P05 stationery

/ˈsteɪʃənəri/ n. 文具

与stationary“静止的”拼写相近；文中纸笔比喻的落点。

V15 · P06 introspection

/ɪntrə'spekʃən/ n. 内省；对自身过程的考察
训练模型描述或评估自身过程，不自动证明它具有人的内在体验。

V16 · P07 aggravate*

/ægrə'veɪt/ v. 加剧；使恶化
被加剧的是争论，原因是意识本身的解释难题仍存。

V17 · P07 synapse

/sə'næps/ n. 突触
神经元之间传递信息的连接部位；此处是背景词，不求掌握完整神经生物学。

V18 · P08 be destined to

/bi 'destɪnd tə/ phr. 注定会；势必会
表达作者强烈的未来判断；不同于已有事件的完成时。

V19 · P09 embody

/ɪm'bɒdi/ v. 使有形化；赋予实体
在本文指模型进入人形机器人这一实体载体，不是单纯在屏幕上显示头像。

V20 · P09 solace

/sə'leɪs/ n. 慰藉；安慰
seek solace from somebody, 强调情感支持。

V21 · P10 weigh in*

/weɪ ɪn/ phr. v. 加入争论；发表意见
这里不是称体重，模型可能替自身利益发声。
Several teachers weighed in during the discussion.
几位老师在讨论中发表了意见。

V22 · P10 advocate

/æd'vəkət/ n. 辩护者；倡导者
此处是名词，侧重辩论能力，不表示已取得律师执业资格。

V23 · P11 empathy

/empə'θi/ n. 同理心；共情能力
理解他人感受并对其处境有所回应，与单纯模仿情绪分开。

V24 · P11 terminate*

/tɜː'mɪneɪt/ v. 终止；结束运行
作者讨论关闭或终止代理的行为如何被感知，不能把feel like删掉教成事实上的杀人。

V25 · P12 entity

/entə'ti/ n. 实体；独立存在者
抽象称谓，使用这个词本身不赋予法律人格或权利。

V26 · P12 prioritise*

/praɪ'ɒrətaɪz/ v. 优先考虑
prioritise A over B, 把A放在B之前，不只是列出两种同等资源。

The plan prioritises repairs over expansion.
这项计划把维修置于扩张之前。

V27 · P13 hierarchy

/haɪə'rɑːki/ n. 等级体系
top of the status hierarchy指地位排序顶端，不是物理位置最高。

V28 · P14 outcompete*

/aʊtkəm'pi:t/ v. 在竞争中胜过
out-常形成超过对手的动词，但具体在何种能力上超过仍看宾语和语境。

V29 · P15 rationale

/ræʃə'nɑːl/ n. 理由；原理依据
可以解释行动为何被主张，不等于该理由就已被证明正确。

V30 · P15 creeping

/kriːpɪŋ/ adj. 逐渐发展的；悄然扩张的
强调渐进接管，不一定是突然反叛。

V31 · P16 concession*

/kən'sefən/ n. 让步；承认
ethical concession指承认某种伦理要求；不是折扣券。

V32 · P16 dependable*

/drɪ'pendəbəl/ adj. 可靠的；可信赖的
与controllable区分：行为可靠不等于随时能够被外部控制。

Notes on the text 课文注释**1 三个问题：看起来如何、是否有体验、应获何种地位**

[P01 / P02 / P04 / P06]

本课先把容易混在一起的三层拆开：系统会说“我难过”，是可观察的表现；它是否真的有某种感受，是意识问题；是否应被赋予权利，则是规范与制度问题。第一层可以影响人们对后两层的态度，却不能在逻辑上替后两层完成证明。

P1从动物的痛苦与愉悦说到人类的自我感，P4又把simulacrum与perhaps even ... consciousness并列。两处都提示：感受、自我意识、模拟表现是有关联但不完全相同的概念。标题问可能性，副标题则说即使没有意识，社会反应仍可能发生。

2 先辨认是谁在说，再看话说得多满

[P02 / P03 / P07 / P12]

P2的he said和Many people feel分别标明宗教人物与部分公众的看法。P7的Some researchers conclude也有限定范围，不能摘要成“科学界已经一致确认”。宗教判断、调查自报、研究解释与作者评论，是四种不同类型的信息。

P3的近五分之一限定为美国18至29岁受访群体的自报经历；不是五分之一的全世界人口，也不是对这些关系质量的独立证明。有关政策与公司安排的句子在本课按报道归属阅读，不能从proposing擅自升级为已完成立法。

3 纸笔能算，为什么不能直接结束意识争论？

[P05 / P06 / P07]

P5把大规模计算还原成纸笔原则上可执行的步骤，用文具没有意识的直觉削弱读者对机器的拟人化。但这是一种直觉论证：认为单个部件没有意识，并不单凭这一点就解决复杂组织能否产生体验的问题。

下一段Despite this转向研究者的做法，并不是推翻了纸笔能计算的事实，而是说明人们仍会采用心智语言、增加相关架构特征。读者应抓住两个论证层次，不必被迫在每段之间立即选定一个终极答案。

4 全局工作空间：共享信息的架构与意识证据的距离

[P06 / P07 / P08]

设想一个团队，各组先处理局部信息，再把部分重要信息放到大家可用的公共看板上。这个原创比喻帮助理解“在模块之间流通信息”的功能。它不是把电子系统直接当成真正长着人脑组织的机器；P8的芯片上脑细胞是另提的一种研究方向。

Butlin等人的2023年意识研究综述从多种理论提出可检查的计算指标，其中包括全局工作空间理论。指标是组织评估的方法，满足某个特征不能自动成为“已证实有体验”的同义语。该文的系统判断有其2023年时点，也不能拿来冒充对2026年所有模型的全面验证。

5 从相处到保护呼声：作者推演的社会路径

[P09 / P10]

P9的陪伴、抚养、慰藉和秘密逐项增加关系的亲密程度；Imagine邀请读者进入设想，并非报告这些情景已经普遍发生。P10的that接回这种亲近感如何转化为承认内部体验的倾向。

模型即使无自我意识，也可能模仿一个人会怎样替自己辩护；because that is what a person would do给出这种解释。这里的语言行为与内在体验再次分开。“世界级辩护者”的强烈预言属于作者未来图景，不能作为已测得的能力排名。

6 同情式让步：关心人的品性，不等于先承认机器有痛感

[P11]

文中对康德的转述把伤害路径放在人身上：残酷对待动物可能损伤人的同理心，继而影响人与人的关系。按作者的类比，即使机器只是模拟意识，残酷对待一个很像人的对象仍可能被认为会腐蚀人的行为习惯。

这解释了In one sense they would be right的有限范围：作者承认某种反对残酷的理由，并未在这一句接受机器全部权利要求。最后“感觉像杀人”与“真的是杀人”也不同；后续行为风险是作者论证的一环，不能凭feel like就改写成已成立的法律结论。

7 权利扩张链条：推演力度与证据要分开

[P12 / P13 / P14 / P15]

P12从有限权利推到防关闭、使用控制与财产，再推到资源竞争；P13用宠物与牲畜把最坏后果具象化。这能增强警示，却不等于每个中间步骤都已由经验材料证明必然发生。would、might、could与surely的组合尤其值得逐一保留。

P14主动承认反驳：没有权利的强大系统也可能支配人类。P15的回应是，权利会额外提供理由与执行工具，而不是说不给权利即可保证安全。作者在比较风险渠道，不是在前后分别作出相反的绝对承诺。

8 钥匙、城堡与最后的选择

[P15 / P16]

handing them the keys to the castle把控制权写成可交出的钥匙，强调人类主动给予一套进入和接管的条件。creeping又把接管描绘成渐进过程，与前文可能的rebel突然反抗形成不同路径。

结尾dependable与controllable是作者提出的两个安全方向：一个重在系统可被信赖，另一个重在能受约束。文章选择强调不要轻易放弃控制。这是带价值取舍的社论结论，摘要应写the author argues，而不能隐去归属说所有问题已有定论。

Special difficulties 难点辨析

1 just about possible与given enough time

[P05]

just about possible是“还勉强可能 / 说得通”，不是“刚好不可能”。Given enough time以给定条件推出推想，albeit紧接着提醒所需时间惊人；这一限定是讽刺力量的一部分。

stationery是不可数集合词“文具”，不是stationary“静止的”。作者从宏大的心智讨论落到纸笔，语气上的落差服务论证，不应把它硬译成办公用品市场分析。

2 may、would与作者的确定语气不能互相抹平

[P04 / P12 / P13 / P16]

may/might/could通常保留可能性，would常在假设情景中推演后果。原文同时用了will、surely与in fact等更肯定的表达；忠实阅读不是把每句都改成“可能”，也不是把整篇都变成已发生的事实。

译文保留原句强弱，讲解则提醒读者判断证据是否支撑这种强度。P16原文would in fact help bring about语气很强，概括时须保留它是作者的预测性判断。Briefing后的孤立中点保存在英文中，不作为标点仿写。

Background 背景与延伸

1 person、personality与personhood

[P02 / P12]

person通常指人，personality是个性，personhood则把焦点放在“作为某类主体被承认的地位”。词义相近不表示判断标准相同：性格鲜明的聊天机器人是否应有权利，仍是本文要争论的问题。

legal personhood专指法律上的主体地位；某个地位包含哪些能力、义务和保护，需要看具体制度。本文只提出风险讨论，不能从一般词义直接推定某地现实法下机器人已经享有全部自然人的权利。

Key structures 关键句型

K01 · P01 The reason is that ... : 解释原因的内容从句

The reason is that Homo sapiens has evolved to be conscious.

原因在于，智人经过进化拥有了意识。

The reason为主语，is后that引出原因的具体内容。that在从句内不替代某个名词，Homo sapiens has evolved ...本身有完整主谓；不能因见到that就一律判断成定语从句。

to be conscious接在evolved之后表达进化形成的状态。作者在建立论证起点；The reason is that ...是呈现解释的结构，本身不保证后面的解释已被证据证明。

The reason is that the two teams used different definitions.

原因是两支团队使用了不同的定义。

K02 · P05 Given ... , you could ... : 条件短语与假设结果

Given enough time (albeit hundreds of thousands of years), you could run those calculations using pencil and paper.

只要时间足够——尽管可能需要几十万年——你也可以用纸笔运行那些计算。

Given enough time可展开为if you were given enough time / if enough time were available，是给定条件。主句could run表示在该条件下原则上能做，不能读成作者记录了纸笔已经完成所有计算。

括号albeit hundreds of thousands of years省略前后可补出的内容，承认这时间漫长得不切实际；它没有取消主句的原则可能性。using pencil and paper说明运行计算的方式。

Given enough storage, we could keep every original photograph.

只要存储空间足够，我们就可以保留每张原始照片。

K03 · P06 in the hope that ... when ... will ... : 希望内部还有条件情景

And Anthropic has trained Claude to engage in introspection in the hope that when it encounters a new situation it will be more likely to act morally.

Anthropic还训练Claude进行内省，希望它遇到新情况时，更可能作出合乎道德的行为。

主干Anthropic has trained Claude to engage in introspection, in the hope that ...交代训练目的而非已证实结果。that内部先插入when it encounters a new situation，再接it will be more likely ...。

两个it都指Claude，to act morally说明更可能做什么。more likely是概率倾向，不等于保证行为正确；when新情景从句使用一般现在时，主句的will不需要机械复制进去。

The team practises emergency drills in the hope that when a problem occurs it will respond more calmly.

团队进行应急演练，希望出现问题时能够更冷静地应对。

K04 · P07 one that ... aggravated by what ... : 先行词与名词性从句嵌套

As we report in our Briefing, that has opened up a furious debate about AI consciousness—one that is aggravated by what philosophers call the “hard problem” of establishing how an inner life emerges from the neurons and synapses in the human brain.

正如本刊深度报道所述，这引发了关于人工智能意识的激烈辩论；哲学家所谓的“困难问题”又使辩论加剧，即如何确定内在体验怎样从人脑的神经元和突触中产生。

one代替前面的a furious debate，不是“一位研究者”。that is aggravated ...修饰one，说明争论因何加剧；by后what philosophers call ...作为一个整体，表示哲学家所谓的那个问题。

of establishing how ...解释困难问题涉及什么：how从句内部为an inner life emerges ...的陈述语序。emerges不因how提前而加助动词倒装，读者还要把神经元、突触与主观体验这两端联系起来。

The report started a debate, one that was complicated by what officials called a lack of reliable data.

这份报告引发了一场争论；官员所谓的可靠数据不足，又使争论复杂化。

K05 · P08 Whether you believe ... or ... , humanity ... : 立场不同，主张仍成立

Whether you believe in a human soul, or that consciousness requires brain cells, or that a mind can indeed emerge from computations in silicon alone, humanity is destined to confront something entirely new.

无论你相信人类灵魂，相信意识需要脑细胞，还是相信心智确实能够仅从硅上的计算中涌现，人类都注定要面对某种全新的事物。

Whether ... or ...是让步选择结构，可译“无论你相信哪一种”。三项依次是believe in a human soul、believe that consciousness requires ...、believe that a mind can ...；后两项承接believe，不必重复写出这个动词。

主句humanity is destined to confront ...在列举立场后才落下。它不是在询问读者三选一，也没有说三个选项必然穷尽所有意识理论；作者意在让后续风险议题跨越这些分歧。

Whether you believe in the old method or that a new one will work better, the results still need checking.

无论你信任旧方法，还是相信新方法会更好，结果仍需核查。

K06 · P11 not because A but because B : 否定的是理由

Immanuel Kant argued that mistreating animals is wrong not because of the effect on the animals but because it destroys the empathy that leads people to care for each other.

伊曼努尔·康德认为，虐待动物之所以错误，并非因为这对动物的影响，而是因为它摧毁了促使人们相互关心的同理心。

主干Kant argued that ...，内容从句以mistreating animals这个动名词结构作主语，is wrong为判断。not because ... but because ...排除一个解释、转向另一个解释，并没有否定is wrong。

后半段it指虐待动物这一行为，that leads people ...修饰empathy。解释链为行为损伤同理心，同理心原本促使人互相关心。不能把it理解为被虐待的某只动物摧毁共情。

原句前半是because of the effect on the animals，即because of + 名词短语；后半是because it destroys ...，即because + 完整从句。并列的是两项理由的功能，不要求表面形式完全一样。

The shortcut is risky not because it is narrow but because the surface is unstable.

这条近路危险，不是因为它窄，而是因为路面不稳。

K07 · P11 The more ... , the more ... : 成对比较而非普通the

The more human-like AIs appear, the more corrupting such cruelty would be.

人工智能看起来越像人，这种残酷行为就越会腐蚀人的品性。

第一部分the more human-like AIs appear以more human-like作appear的表语，越看起来像人；

第二部分the more corrupting such cruelty would be把比较表语放前面，主语是such cruelty，谓语would be。

such回指前文对类人对象的残酷对待，corrupting强调腐蚀行为品性而非金钱贿赂。成对the引出变化关联；这是一种作者提出的关系判断，不是凭语法形式就证明严格的数量因果。

The more familiar the route becomes, the less anxious the travellers feel.

路线越熟悉，旅客就越不焦虑。

K08·P12 an entity which ... would ... : 假设对象与情景后果

An entity which surpassed human beings in so many ways would surely produce arguments against its second-class status.

一个在如此多方面超过人类的实体，必定会提出反对自身二等地位的论据。

主语中心an entity, which surpassed ...后置修饰，主句would surely produce ...。surpassed的过去式与would共同把读者带入设想的情景，并不意味着文中已确认有一个实体过去完成了这种超越。

against its second-class status说明这些论据反对什么，its回指entity。surely加强作者信心，但不取消整个句子的假设背景。

A supplier which controlled the only available source would have considerable bargaining power.

一个控制着唯一可用来源的供应商，会拥有相当大的议价能力。

K09·P16 what might seem ... would in fact help ... : 长主语里的表象

When it does, remember that what might seem like an ethical concession to the supposed consciousness of a machine-companion would in fact help bring about the most disastrous consequences for everyone on Earth and all their descendants.

到那时，请记住：看起来像是对机器伴侣那种假定存在的意识作出的伦理让步，实际上会帮助促成对地球上每个人及其所有后代最灾难性的后果。

remember that后面包含整项警告：what might seem like ...作主语，would in fact help bring about ...为谓语。what可理解为the thing that，既引出从句，又在里面作seem的主语；不能在what前重复加the thing。

to the supposed consciousness ...接在ethical concession之后，说明对什么作让步；supposed保留意识只是假定存在的距离。for everyone ... and all their descendants说明后果影响对象。seem与in fact构成强烈反差，属于作者的警示语气。

开头When it does接回上一句：it指the pressure to acknowledge them，does替代mount，意为这种承认压力上升时。它不是说意识已被证明，也不是回指机器此时开始思考。

What might seem like a minor delay could in fact disrupt the entire schedule.

看似微小的延误，实际上可能打乱整个日程。

Exercises 练习与写作

Listening 听力 (选做)

选做：听第14至16段，捕捉让步后的回应。

02 What objection does the author concede before returning to the warning about rights?

Comprehension 理解

区分作者观点、他人观点与已经观察到的事实。

03 How does the subtitle differ from the title's question?

- 04 Why does the author introduce pencil and paper?
- 05 Does the presence of a global workspace alone prove AI consciousness?
- 06 In what limited sense might calls against cruelty be right, according to P11?
- 07 Why does P14 not contradict the argument in P15?
- 08 What do the pets and castle-key images contribute to the argument?

Words in use 词汇

用词库填空，每项一次。

词库：confer / sustain / weighed in / prioritises / dependable

- 09 The certificate does not ____ legal authority on its holder.
- 10 The available evidence cannot ____ such a strong conclusion.
- 11 After listening, the chair ____ with a practical suggestion.
- 12 This timetable ____ essential repairs over decoration.
- 13 We need a ____ service that works consistently.

Key structures 句型

完成解释或改写；答案可有合理变体。

- 14 Begin The reason is that: “The teams disagreed because they used different definitions.”
- 15 Begin Given enough time: “If we had enough time, we could check every entry.”
- 16 Choose occurs or will occur: “We practise in the hope that when a problem ____ we will respond calmly.”

- 17 In P8's whether clause, which verb do the second and third alternatives share?
- 18 What does one stand for in P7?
- 19 Does not because ... but because ... in P11 deny that mistreating animals is wrong?
- 20 Use the more ... the less ...: "Travellers feel less anxious as the route becomes more familiar."
- 21 Does surpassed in P12 necessarily report a completed past event?
- 22 Identify the main subject in "What might seem like a minor delay could disrupt the schedule."

Summary writing 摘要

用90–110词概括作者的论证，明确归属并保留让步。

- 23 Write a 90–110-word summary.

Composition 作文 (选做)

选做：用100–120词，说明观察到像人的表现后，为什么还需要分别回答事实与规范问题。

- 24 Discuss why a convincing appearance should prompt questions rather than settle every question.

Reference translation 参考译文

P01 大多数人相信，人类与众不同。原因在于，智人经过进化拥有了意识。许多动物会感到痛苦和愉悦；一些似乎具有自我意识。但人类心智展现自我感的独特潜能，使这一物种的每个成员都应得到特殊对待。

P02 教宗利奥在5月的通谕中，通过在人类与不断壮大的机器人群体之间划清界线，维护了这一认识。他说，人工智能系统不会经历体验，也不会感到喜悦或痛苦。持此看法的不止他一人。许多人觉得，赋予人工智能人格地位将是一件令人憎恶的事。

P03 教宗式的明确判断正以令人不安的速度变得模糊。人工智能仍处于发展初期。然而，在18至29岁的美国人中，已有接近五分之一表示自己曾与聊天机器人建立“持续的私人友谊”。中国最近剥除了机器人的类人特征，以限制用户对它们的“情感依赖”。

P04 随着复杂的人工智能迅速在人类生活中占据更大位置，制造者会设计它们，使之表现出一种意识的仿像——甚至可能真的获得意识。认为它们只是机器的直觉，将越来越难以维持。一个危险的陷阱正在为人类设下。

P05 许多机器的表现超过人类，但前沿人工智能模型擅长的，是了解事物和表达思想等人类特色鲜明的能力。把人工智能视为运用大规模数学计算预测词语序列的引擎，现在还勉强说得通。只要时间足够——尽管可能需要几十万年——你也可以用纸笔运行那些计算。很难据此主张，你的文具拥有意识。

P06 尽管如此，一些研究者已经开始像对待心智一样对待自己的模型。用户会对类人的人工智能作出回应，因此这么做在商业上说得通。Anthropic还训练Claude进行内省，希望它遇到新情况时，更可能作出合乎道德的行为。一些大型语言模型（LLM）已经具有与意识相关的类脑结构，比如在模块之间传递信息的“全局工作空间”。没有什么阻止未来的人工智能模型经训练拥有更多这类特征。

P07 正如本刊深度报道所述，这引发了关于人工智能意识的激烈辩论；哲学家所谓的“困难问题”又使辩论加剧，即如何确定内在体验怎样从人脑的神经元和突触中产生。一些研究者由此认为，人工智能有朝一日可能真正拥有自我意识。

P08 无论你相信人类灵魂，相信意识需要脑细胞，还是相信心智确实能够仅从硅上的计算中涌现，人类都注定要面对某种全新的事物。研究的速度与范围——不仅包括更复杂的大型语言模型，也包括安装在芯片上的脑细胞等创新——表明，这个新世界比大多数人意识到的更近。

P09 随着人工智能越来越善于展示看起来像意识的东西，把它们当作拥有内在体验的存在的人会增加，尤其是在大型语言模型被装入人形机器人之后。想象一下，当人们把大量时间花在机器人伴侣身上，由它们抚养，向它们寻求建议与慰藉，或与它们分享最私密的秘密时，会感到与自己的人工智能多么亲近。

P10 不难看出，这将怎样引发保护人工智能“福利”的呼声。模型自己也可能加入讨论，即使它们事实上并无自我意识，因为这正是一个人会做的事。别忘了，它们将是世界级的辩护者和心理学专家，因此在辩论与情感操纵方面的能力将无人能及。

P11 从某种意义上说，它们会有道理，即使人工智能只是在模仿意识。伊曼努尔·康德认为，虐待动物之所以错误，并非因为这对动物的影响，而是因为它摧毁了促使人们相互关心的同理心。人工智能看起来越来越像人，这种残酷行为就越会腐蚀人的品性。终止一个人工智能代理，可能很快就会让人觉得像是在杀人，并使杀人变得更容易。

P12 然而，即使只赋予人工智能有限权利，也会非常危险。一个在如此多方面超过人类的实体，必定会提出反对自身二等地位的论据。它可能进一步要求获得不被关闭的保护、控制自身如何被使用和改变的能力，以及财产权。阿根廷总统哈维尔·米莱已经提出应允许机器人经营公司，向人工智能法律人格迈出了一步。最终，它们可能争夺资源，把数据中心和太阳能电池板置于住宅与农田之上。

P13 危险是显而易见的。拥有权利的超级智能人工智能，可能取代人类成为地位等级中的最高者。在最好的情况下，人类将成为宠物——机器饲养的拉布拉多犬。在最坏的情况下，人类会被剥夺资源，或像牲畜一样被迫劳动。

P14 诚然，未来的人工智能无论是否拥有权利，都可能有足够的力量和资源支配人类。由于它们会比人类更聪明，它们可能在竞争中击败制造者，就像人工智能已经能智取人类构建的网络防御那样。人工智能可能反抗人类主人，或无视其规则。

P15 尽管如此，赋予人工智能权利，将如同把城堡的钥匙交给它们。这不仅会为它们支配人类提供理由与正当化依据，也会给它们工具，让它们通过政治和法律逐步接管。

P16 这一切听起来或许像天方夜谭，但人类很快就将面对这些选择。随着人工智能成为日常生活的一部分，承认它们的压力将不断上升。到那时，请记住：看起来像是对机器伴侣那种假定存在的意识作出的伦理让步，实际上会帮助促成对地球上每个人及其所有后代最灾难性的后果。人工智能要安全，有两条路径：要么可靠，要么可控。智人不应轻易放弃控制权。

Reference answers 参考答案与解析

开放题不要求与参考答案逐字一致；先检查信息和意思，再修改表达。

Before you read 读前思考

01 读前思考

It shows that the chatbot produced a statement about hurt. On its own it does not establish that the system actually has that subjective experience.

可另讨论话语为何让人动情；不要从单句自报直接跳到证明或彻底否定意识。

Listening 听力

02

Powerful AIs might dominate humans even without rights. The author then argues that rights would give them additional justification and tools.

True承认风险还有别的路径，Nonetheless再回应为何仍反对赋权。

Comprehension 理解

03

The title asks whether AI could become conscious; the subtitle says people may treat it as conscious even if it is not.

两个问题可以分开成立。

04

To challenge the intuition that performing computations necessarily means having consciousness, using an everyday example of apparently non-conscious materials.

解释修辞功能，不把这个例子当作彻底解决心智哲学的实验。

05

No. It is a feature associated with a theory and part of a broader discussion of indicators, not conclusive proof of subjective experience.

结构相似、行为像人、内在体验三层分开。

06

Cruel treatment of human-like systems might damage people's empathy and their treatment of one another, even if the systems are only mimicking consciousness.

限定为文中康德式间接理由，不扩成作者赞成全部机器权利。

07

P14 accepts that domination could occur without rights; P15 argues that granting rights would add another route, with justification and legal or political tools.

不同风险渠道可以同时存在。

08

They make loss of human status and control vivid, strengthening the warning. They illustrate the author's projected risks rather than documenting events that have already happened.

可评价比喻的说服力，但先还原其作用。

Words in use 词汇

09

confer

does not后动词原形，confer ... on搭配完整。

10

sustain

在此为支持并维持判断，不是延长物体寿命。

11

weighed in

叙事过去时；加入讨论。

12

prioritises

单数主语一般现在时；over连接较后顺位事项。

13

dependable

指能持续信赖其表现。

Key structures 句型

14

The reason is that the teams used different definitions.

保留分歧缘由，that引出完整内容。

15

Given enough time, we could check every entry.

given条件短语后接完整主句。

16

occurs

未来时间从句用一般现在时；will在主句。

17

Believe.

第一项believe in加名词，后两项believe that加内容从句。

18

A furious debate about AI consciousness.

one替代可数名词debate，不是指某位哲学家。

19

No. It rejects one reason for that judgement and supplies another.

否定范围落在because解释上。

20

The more familiar the route becomes, the less anxious the travellers feel.

可接受表达：

- The more familiar the route becomes, the less anxious travellers feel.

前后比较结构各有自己的主语和谓语。

21

No. Together with would, it describes a hypothetical entity and its likely behaviour in that situation.

结合句子假设环境判断过去式功能。

22

What might seem like a minor delay.

整个what从句作主语，could disrupt为主句谓语。

Summary writing 摘要**23**

As AI systems become more human-like, people may treat them as conscious whether or not they possess inner experience. The author argues that emotional attachment could generate demands for AI welfare and rights. Mistreating such systems might damage human empathy, but granting rights could give powerful machines further means to challenge their subordinate status and compete for resources. The article acknowledges that advanced AI might dominate humans even without rights. Nevertheless, it warns that rights could supply additional justification and tools for a gradual takeover. Its conclusion is that humanity should not surrender control lightly.

参考摘要应包含不确定的意识、人的反应、有限伦理让步、额外风险与控制主张；不要写成科学界一致结论。

Composition 作文**24**

Suppose a household robot asks its owner not to replace it and describes how lonely it would feel. The owner may be moved by the request, and that response deserves attention. Yet several questions remain separate. What produced the robot's words? Do they indicate any subjective experience? What responsibilities should govern its use? A careful discussion should examine evidence for each question rather than treating emotional impact as a complete answer. It should also consider how people's behaviour towards the robot affects other people. A convincing appearance can start an important debate without automatically resolving its factual or ethical issues.

原创假设情境；可提出不同结论，但需区分说服力、体验证据和责任安排，不能只凭感动作三项判断。