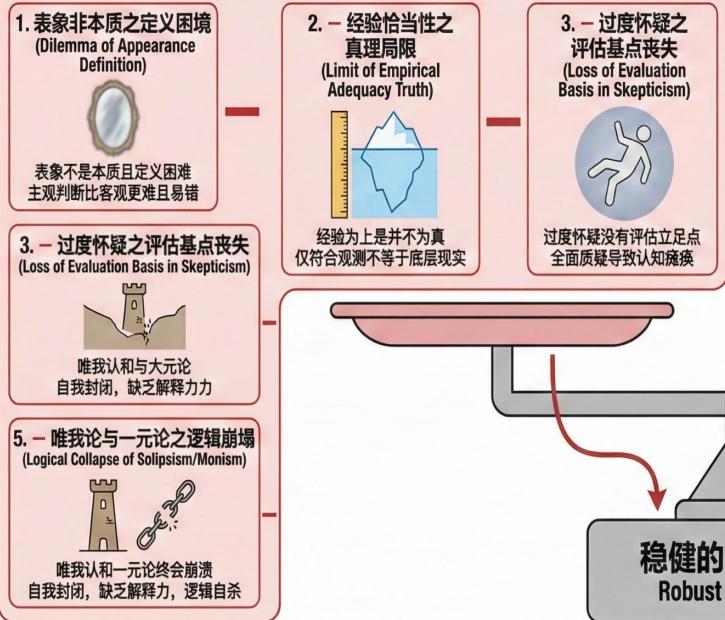


威廉姆森哲学全景重构

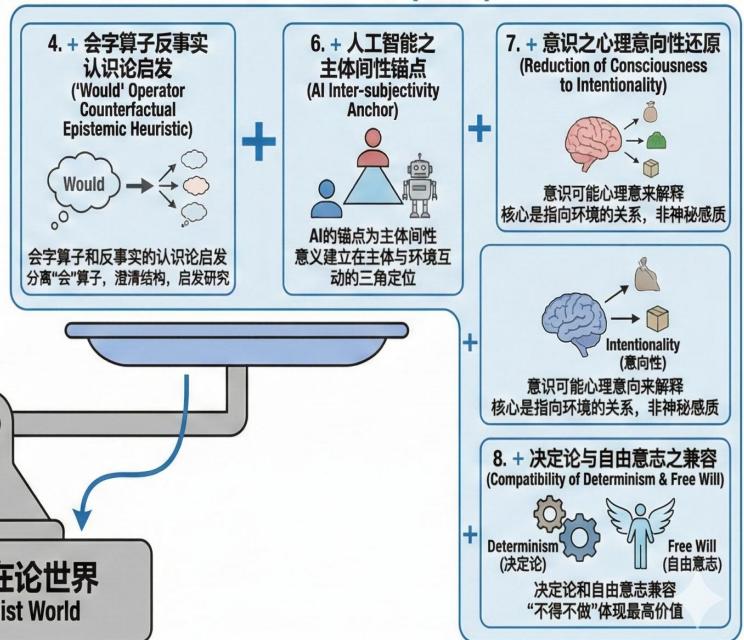
(Williamson's Philosophical Panorama Reconstruction)



SUBTRACTION (剔除)



ADDITION (构建)



稳健的实在论世界

Robust Realist World

威廉姆森实在论的认识论重构

<https://gemini.google.com/share/64f838e5ca45>

Responses below were generated with a creator's Gem according to their custom instructions. [Learn more](#) 15 January 2026 at 18:33

Published on 15 January 2026 at 22:36

把第六页做成等式并解释，等式元素

等式看世界=一个等式+一个等式=深入 Custom Gem

基于您提供的文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第六页的内容（即源文档中的 page 6），核心讨论集中在蒂莫西·威廉姆森（Timothy Williamson）反驳巴斯·范·弗拉森（Bas van Fraassen）的反实在论（Constructive Empiricism），并论证“实在论”以及为何“表象”（Appearances）并不具备认识论上的特权。

以下是符合Linux用户名/文件名严格格式（无空格、无特殊数学符号、仅用加减号连接）的核心归纳等式及其深度解析：

核心归纳等式

威廉姆森实在论的认识论重构=独立于意志的客观现实+科学对底层真相的追求+主观表象易错性论证—范弗拉森的经验恰当性限制—内省全知幻觉

等式要素深度解析

这个等式解构了威廉姆森在第六页中关于“实在论 vs 反实在论”的核心辩护逻辑，他打破了传统哲学中认为“主观表象是绝对安全知识”的迷思。

1. 威廉姆森实在论的认识论重构 (等式左侧结果)

- **思想：**这是威廉姆森对世界认知的总立场。他认为我们不仅能认知事物呈现的样子（表象），更有能力获得关于世界实际样子的知识。
- **结论：**实在论不仅是关于世界存在的本体论，更是一种相信我们能跨越表象到达真实的认识论信心。

2. + 独立于意志的客观现实 (Ontological Independence)

- **逻辑:** 实在论的基石。威廉姆森定义实在论为：我们在认知生活中处理的现实，在很大程度上是独立于我们的意志（Will）和看法的。
- **细节:** 无论我们怎么想，世界就在那里。这区别于唯心主义或极端的社会建构论。

3. + 科学对底层真相的追求 (Scientific Truth Seeking)

- **辩驳对象:** 针对范弗拉森认为“科学的目标仅仅是获得经验上精确的理论（Empirically adequate）”。
- **逻辑:** 威廉姆森指出，大多数科学家不只是想做能被验证的预测（Prediction），他们真正想要的是发现可观测事物之下的“底层现实”（Underlying reality）。
- **数据/事实:** 科学家的实际工作动机超越了单纯的数据拟合，而是指向世界的本质结构。

4. + 主观表象易错性论证 (Fallibility of Appearances)

- **核心思想（高光点）:** 这是本页最精彩的反直觉论证。通常人们认为“外部世界我可能看错，但我当下的感觉（表象）是绝对不会错的”。威廉姆森彻底否定了这一点。
- **具体案例:** 威廉姆森举例说，当你坐在一个房间里，要精确描述“世界如何向你显现的主观状态”，其实比直接描述“房间里有桌子和椅子”要困难得多。
- **推论:** “表象”并不是一个只要我们认为是事实就是事实的“舒适小领域”（Comfortable little realm）。我们对自己的主观倾向和特质也经常误判。

5. - 范弗拉森的经验恰当性限制 (Minus Empirical Adequacy)

- **逻辑减法:** 威廉姆森移除（否定）了范弗拉森设定的限制。范弗拉森认为只要理论不违反观测即可，无需承诺其描述了真实。

- **批判：**威廉姆森认为这种规定科学家“被允许追求什么目标”的做法是奇怪的，限制了人类认知的潜能。

6. 一内省全知幻觉 (Minus Introspective Omniscience)

- **逻辑减法：**破除“认识论特权”(Epistemic Privilege)。
- **思想：**范弗拉森错误地赋予了“表象”一种特殊的认识论特权，认为它是我们可以绝对确定的基石。威廉姆森指出自我认知 (Self-knowledge) 实际上非常难以获得，我们在内省时并不具备全知视角。

一句话总结第六页：威廉姆森通过论证**“描述主观感觉比描述客观物体更难且更易出错”**，击碎了反实在论者躲避在“表象”避风港中的企图，重申了科学与认知应当直面独立于心灵的客观现实。

现在弄下第7页吗，和上面一样的模式 弄

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第七页的内容，这一页的核心主题是蒂莫西·威廉姆森 (Timothy Williamson) 对**怀疑主义 (Skepticism)** 的结构性批判。他将怀疑主义比作药物过量，并揭示了怀疑论者设定的“必胜游戏规则”是一个逻辑陷阱。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

威廉姆森反怀疑论策略 = 健康的局部自我批评 + 保留评估立足点 - 全面质疑导致的认知瘫痪 - 悬疑论者必胜的游戏规则

等式要素深度解析

这个等式解构了威廉姆森在第七页中关于如何正确处理“怀疑”的逻辑：他并不完全否定怀疑，而是区分了“有益的自我修正”与“毁灭性的全面怀疑”。

1. 威廉姆森反怀疑论策略 (等式左侧结果)

- **思想:** 面对怀疑主义, 不需要跳进怀疑论者挖的坑里去试图把他们拉出来, 而是要彻底避开那个坑。
- **结论:** 怀疑主义不是一种需要被“驳倒”的理论, 而是一种需要被“管理”的认知剂量问题。

2. +健康的局部自我批评 (Healthy Self-Criticism)

- **逻辑:** 威廉姆森承认怀疑的本能是健康的 (Healthy instinct)。我们需要从某种信念中退一步, 进行自我批评和评估。
- **比喻 (数据):** 就像吃药 (Pills)。少量的药片 (怀疑) 能治病, 对认知系统有益。
- **细节:** 这种机制保证了我们不会盲目相信从小被教导的事情, 保留了修正错误信念的能力。

3. +保留评估立足点 (Retaining a Standpoint)

- **核心逻辑:** 评估信念必须基于某些“已知的其他事物”。你不能悬空评估。
- **操作方法:** 当我们质疑某一个信念时, 必须依赖其他未被质疑的信念作为评估工具。只有保留了这些资源, 批评才是有可能的。

4. -全面质疑导致的认知瘫痪 (Minus Total Cognitive Paralysis)

- **逻辑减法:** 否定了“同时质疑所有信念”的可能性。
- **推论:** 如果你试图从所有信念中退一步 (Step back from all beliefs), 你就没有任何立足点来评估任何东西了。这相当于扔掉了所有的认知资源。
- **结论:** 这会导致“彻底的瘫痪”(Total paralysis)。试图同时评估所有信念是一个“非常笨的主意”(Very dumb idea)。

5. -怀疑论者必胜的游戏规则 (Minus Skeptic's Rigged Game)

- **思想:** 怀疑论者强加了一套游戏规则: 如果你不能向怀疑论者证明他们是错的, 他们就赢了。
- **逻辑陷阱:** 例如对“理性”的怀疑。如果你用理性去反驳, 怀疑论者会说你在“乞题”(Begging the question, 即循环论证); 如果你不用理性, 你又无法辩护。
- **威廉姆森的对策:** 不要接受这些规则。怀疑论者把自己关进了一个爬不出来的坑里 (Pit), 我们无力把他们拉出来, 但这不意味着我们也得跳进去。

一句话总结第七页： 威廉姆森指出**“怀疑”就像药物，适量是解药，过量则是毒药**；彻底的怀疑主义通过剥夺所有评估标准导致认知瘫痪，我们应当拒绝进入这种自我封闭的逻辑陷阱。

现在继续弄下面一页吧

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第八页的内容，这一页的核心探讨从“怀疑主义的哲学批判”转向了更具操作性的**“认知方法论”**。威廉姆森解释了如何通过识别“认知启发法”（Heuristics）来在缺乏明确规则的情况下把握怀疑的“剂量”。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

学科审慎判断力的习得 = 认知启发法的识别与解构 + 类科学的数据权衡 + 特定语境下的防误判
—通用型万能怀疑论—绝对化的预警红灯

等式要素深度解析

这个等式解构了威廉姆森在第八页中关于“如何知道自己是否怀疑过头”的回答。他承认没有简单的规则，而是通过对思维工具（启发法）的元认知来维持平衡。

1. 学科审慎判断力的习得 (等式左侧结果)

- **思想：**面对“我该吃多少怀疑主义药片”的问题，威廉姆森指出没有简单的规则（Rules）。
- **结论：**解决之道在于“判断力”（Judgment）。就像科学家评估不可靠数据或人们在社交中把握玩笑的界限一样，这是一种通过经验习得的微妙感觉（Sense），是学科技能的一部分。

2. + 知识启发法的识别与解构 (Identification of Cognitive Heuristics)

- **核心概念：**威廉姆森引入了“认知启发法”（Cognitive Heuristics）——这是一种廉价、快速但有时会出错的认知捷径。

- **逻辑:** 他不再视哲学家的直觉为不可动摇的“数据”，而是去识别这些直觉背后依赖了哪种启发法。
- **操作:** 只要能展示这种启发法并非100%正确，就能在不陷入全面怀疑的情况下，合理地质疑某些哲学前提。

3. +类科学的数据权衡 (Science-like Data Evaluation)

- **类比:** 威廉姆森将哲学思考比作科学工作。科学家面对来自不信任实验室的数据时，不会直接丢弃，也不会全盘接受，而是行使判断力来处理。
- **逻辑:** 这种态度既非盲信也非虚无，而是一种对证据质量的动态评估。

4. +特定语境下的防误判 (Context-Specific Caution)

- **策略:** 威廉姆森强调要“谨慎行事”(Proceeding with caution)，不是随机质疑，而是具体分析。
- **细节:** 只有通过深入细节，才能发现某个特定信念是否由不可靠的启发法产生，而不是一刀切地否定。

5. -通用型万能怀疑论 (Minus Generic Skepticism)

- **警告信号:** 威廉姆森提出了一个关键的“黄灯”: 当一个怀疑理由太过通用 (Too general)，以至于可以用来质疑任何事情时，这就是走偏的信号。
- **逻辑减法:** 这种“万能酸”式的怀疑论证没有任何认知价值，因为它们无法区分好坏，只会溶解一切。

6. -绝对化的预警红灯 (Minus Flashing Red Light)

- **现实认知:** 威廉姆森坦言，脑子里并没有一个红灯 (Red light) 会在你犯错时闪烁。
- **逻辑减法:** 没有任何形式化的规则可以替你做决定。企图寻找一条能自动告诉我们“现在停下”的规则是徒劳的。

一句话总结第八页：威廉姆森提出，我们无法依靠闪烁的红灯来规避过度怀疑，而必须通过剖析大脑偷懒的机制（认知启发法）并培养类似科学家的敏锐判断力，来在“轻信”与“虚无”之间找到立足之地。

现在弄第9页吧

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第九页的内容，这一页的核心探讨了如何评估一个看似荒谬的哲学立场（如否定桌子的存在），并对比了**“多产的科学反常识”与“贫瘠的怀疑论反常识”**。威廉姆森提出了以“生产力”(Productivity) 作为最终的裁决标准。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

哲学立场的生产力裁决 = 研究计划的长效多产性 + 反常识的科学突破范例 — 言行不一的表演性矛盾 — 怀疑论的蒙昧主义空转 — 人文学科的梦想直觉

等式要素深度解析

这个等式解构了威廉姆森在第九页中关于“如何区分疯狂的怀疑与伟大的洞见”的回答：标准不在于理论听起来多疯狂，而在于它能产出什么。

1. 哲学立场的生产力裁决 (等式左侧结果)

- **思想：**面对有人声称“桌子不存在”这种极端立场，我们不仅要看辩论逻辑，更要看该理论的历史后果。
- **结论：**评判一个反直觉观点的最终标准，是看该研究计划在长远 (In the long run) 是否具有多产性 (Fruitfulness) 。

2. + 研究计划的长效多产性 (Long-term Research Productivity)

- **逻辑：**威廉姆森认为，我们不应先验地禁止某种哲学方式，而应让它们发展，看看结果。

- **核心指标：** 区别“好理论”与“坏理论”的关键在于它是否能产生大量关于世界的新知识（Generated vast amounts of new knowledge）。

3. +反常识的科学突破范例 (Anti-commonsense Scientific Precedents)

- **案例：** 威廉姆森举例“日心说”。曾几何时，否定“地球是宇宙中心”也是对常识的疯狂否定。
- **对比价值：** 但这种否定被证明是正确的，因为它开启了一个极其丰硕的科学研究计划。这意味着“违反常识”本身并不是死罪，关键看后续贡献。

4. -言行不一的表演性矛盾 (Minus Performative Contradiction)

- **现象：** 针对那些声称“椅子不存在”却去坐椅子的人。威廉姆森认为这是一种“走得太远”(Going too far) 的警告信号。
- **逻辑减法：** 这种理论往往玩弄辞藻（如“按桌子方式排列的粒子”），却无法真正展示如何屏弃正常的思考方式，属于没有实际产出的理论自嗨。

5. -怀疑论的蒙昧主义空转 (Minus Skepticism's Obscurantism)

- **定性：** 怀疑主义往往“产不出什么”(Produces very little)。
- **批判：** 当与产生大量知识的非怀疑主义观点对比时，怀疑主义看起来就像是对科学的否定或蒙昧主义 (Obscurantism)。它是一个停止认知的借口，而非探索的开始。

6. -人文学科的梦境直觉 (Minus Humanities' Dream Intuition)

- **数据/观察：** 威廉姆森观察到，人文学科的学生比理工科（如工程师）学生更容易立即聚集在怀疑论（如“也许我们在做梦”）的旗帜下。
- **心理分析：** 这可能源于人文学科更强的共情力 (Empathy) 或对实际细节关注的缺乏，导致他们更容易与“虚构的做梦主体”产生廉价共鸣，从而陷入怀疑论的陷阱。

一句话总结第九页： 威廉姆森提出**“生产力是检验真理的试金石”**，指出尽管科学和怀疑论都挑战常识，但前者带来了知识的爆发，而后者（往往受人文学科直觉偏爱）只导致了蒙昧主义的空转。

继续弄第10页

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第十页的内容，这一页的核心探讨集中在**唯我论 (Solipsism)**。威廉姆森分析了为什么这种极其错误的观点会吸引人，并揭示了“无法被辩论击败”并不等于“真理”。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

唯我论的虚假稳健性批判 = 辩证上的绝对防御 + 挑衅性的心理姿态 + 自我封闭的解释闭环 - 阴谋论式的证据免疫 - 廉价的开明幻觉 - 辩论胜负的真理局限

等式要素深度解析

这个等式解构了威廉姆森在第十页中对唯我论的剖析：他指出唯我论之所以存在，不是因为它有道理，而是因为它利用了辩论机制的漏洞，构建了一种类似阴谋论的“无赖逻辑”。

1. 唯我论的虚假稳健性批判 (等式左侧结果)

- **思想：** 唯我论（认为除自己外一无所有）是一种极其错误的立场。
- **结论：** 它的生存能力不源于其真实性，而源于其在争论中“刀枪不入”的虚假稳健性 (Robustness) 。

2. + 辩证上的绝对防御 (Dialectical Robustness)

- **逻辑：** 人们喜欢这种观点，是因为它在辩论上非常“稳健”。如果你捍卫它，任何人都无法说服你是错的。

- **细节：**这种防御机制是绝对的，对方的任何论据都会被这一机制吞噬。

3. +挑衅性的心理姿态 (Provocative Stance)

- **动机：**采取唯我论往往是一种为了惹恼 (Annoy) 和震惊 (Shock) 他人的方法。
- **心理：**对于喜欢采取极端立场的人来说，这是一个“有趣”的位置，因为别人很难给出令人信服的反驳。

4. +自我封闭的解释闭环 (Self-Enclosed Loop)

- **具体案例：**威廉姆森举例：如果母亲质问唯我论者“你怎么敢假设我不存在？”，唯我论者会将其解释为“这只是发生在我脑子里的声音，并没有一个外在的母亲在说话”。
- **机制：**将所有外部反例强行内化，从而消解了外部世界的独立性。

5. -阴谋论式的证据免疫 (Minus Conspiracy-Theory Immunity)

- **类比：**威廉姆森将唯我论比作阴谋论。如果你给阴谋论者看反对他们的报纸，他们会说报纸是阴谋的一部分。
- **逻辑减法：**“无法被反驳”是一个非常低的标准。有很多愚蠢的理论 (Stupid theories) 都是无法通过这种方式反驳的。不可反驳性不代表真理。

6. -廉价的开明幻觉 (Minus Cheap Enlightenment)

- **批判：**唯我论者倾向于认为自己比任何人都更“开明”(Enlightened)，而别人对此无能为力。
- **价值判断：**威廉姆森斥责这种立场太“廉价”(Cheap) 且多得数不胜数 (A dime a dozen)，根本不足以作为自满的理由。

7. 一辩论胜负的真理局限 (Minus Debate Limitations)

- **核心哲学洞见：** 唯我论的流行揭示了“通过辩论或争论来评估立场”这一方法的局限性。
- **结论：** 一个立场在口头辩论中赢了（或没输），并不代表它在现实中是正确的。

一句话总结第十页： 威廉姆森揭示了唯我论本质上是一种“智力上的无赖策略”，它像阴谋论一样通过拒绝承认外部证据来获得廉价的不可战胜感，但这证明了辩论技巧的局限而非唯我论的真理。

搞11页

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第11页的内容，这一页的核心探讨从对唯我论/唯心论（如叔本华）的最后清算，转向了对**一元论（Monism）的逻辑荒谬性的批判，并确立了威廉姆森的多元论（Pluralism）**立场。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

威廉姆森多元论的本体确立 = 证据匮乏的举证责任 + 本体论的极度丰富性 + 常识实在的独立性
— 叔本华式阴谋论结构 — 辩论稳健性的低门槛 — 一元论的区别消解谬误

等式要素深度解析

这个等式解构了威廉姆森在第11页中如何拒绝“万物归一”的诱惑：他将无法证伪的唯心论归类为阴谋论，并用一个生动的例子揭示了一元论在逻辑上的自杀行为。

1. 威廉姆森多元论的本体确立 (等式左侧结果)

- **思想：** 威廉姆森明确反对严格的一元论（只存在一样东西），主张“外面有很多不同的东西”(There are lots of different things out there)。

- **结论：**世界是由种类繁多且数量众多的事物构成的，不能被简单归约为单一实体。

2. +证据匮乏的举证责任 (Burden of Evidence)

- **针对对象：**叔本华式的唯心论（认为万物只是单一意识根源的触角/幻觉）。
- **逻辑：**威廉姆森指出，对于那些没被说服的人来说，只需要指出支持这种观点的“证据匮乏”(Paucity of evidence)就足够了。
- **方法论：**不需要试图在逻辑闭环内击败对手，只需指出其实证基础的空虚。

3. +本体论的极度丰富性 (Ontological Richness)

- **数据/列举：**威廉姆森列举了小说、空集、物理粒子等事物。
- **逻辑：**这些事物显然都存在，且显然不能被归约为同一个东西。这种多样性直接否定了严格的一元论。

4. -叔本华式阴谋论结构 (Minus Schopenhauerian Conspiracy)

- **类比：**威廉姆森将形而上学的唯心论比作“阴谋论”(Conspiracy theories)。
- **机制：**采取错误观点的人会因为其持续的错误，导致评估证据的能力也受损。就像阴谋论者会将反面证据视为阴谋的一部分，唯心论者会将外部世界视为幻觉的一部分，这种自我免疫机制是病态的。

5. -辩论稳健性的低门槛 (Minus Low Bar of Robustness)

- **思想：**仅仅因为一个理论无法被驳倒（即辩论上的“稳健性”），并不代表它值得被认真对待。
- **结论：**“无法说服持有者改变主意”对于一个理论来说是一个“相当低的标准”(Pretty low bar)。真理不等于“辩论中的不败之地”。

6. ——元论的区别消解谬误 (Minus Monist's Distinction Dissolution)

- **核心轶事 (高光点):** 威廉姆森讲述了一个一元论者的故事，此人声称“不存在区别”(There are no distinctions)。当被问及“好论证和坏论证之间是否有区别”时，该一元论者被迫承认“没有”。
- **逻辑推论:** 严格的一元论会导致理性的自杀，因为它消除了评估真理所需的差异性基础。这是一种威廉姆森完全不感兴趣的荒谬观点。

一句话总结第11页： 威廉姆森通过将唯心论比作阴谋论，并揭示一元论会导致“无法区分好坏论证”的逻辑崩塌，从而捍卫了一个充满具体事物（如小说、粒子、集合）的丰富多元的现实世界。

继续弄下一页

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第12页的内容，这一页的核心探讨集中在**“意识”(Consciousness) **。威廉姆森对意识进行了彻底的“祛魅”，反对将其视为一种神秘的本体论特殊层级，并论证了心理状态可以独立于意识存在。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

意识的哲学祛魅 = 环境认知关系的回归 + 无意识认知的主导性 + 动物信念的功能性归属 - 感质本体论的特殊层级 - 内在主义的认识论特权 - 哲学重要性的过度夸大

等式要素深度解析

这个等式解构了威廉姆森在第12页中关于“意识算不算大事”的回答：他认为哲学家们人为地抬高了意识的地位，而实际上它只是认知的一种形式，并非所有心理活动的必要前提。

1. 意识的哲学祛魅 (等式左侧结果)

- **思想:** 威廉姆森直言“意识在哲学上的重要性被极大地夸大了”(Grossly exaggerated)。
- **结论:** 意识不是哲学的圣杯，我们应当用平常心看待它。虽然对个人生活很重要（不想昏迷），但在解释认知架构时，它被赋予了不配拥有的特权。

2. +环境认知关系的回归 (Cognitive Relation to Environment)

- **定义:** 威廉姆森主张将意识简单地理解为“意识到某事”(Being conscious of something)，即作为一种认知形式。
- **逻辑:** 比如意识到面前的麦克风，这仅仅是主体与环境中事物的一种认知关系，而不是某种神秘的内在发光体。

3. +无意识认知的主导性 (Dominance of Unconscious Cognition)

- **数据/科学证据:** 心理学证据表明，事实上大多数行为 (Action) 都发生在意识之外。
- **定位:** 虽然意识可以作为仔细权衡决定的“最终工作场所”(Final workshop)，但我们不应赋予它高于那些在意识层面之下运作的大量认知行为的特权。

4. +动物信念的功能性归属 (Functional Attribution of Beliefs)

- **逻辑:** 针对“心理状态是否以意识为前提”的问题。威廉姆森认为，如果我们通过赋予动物信念、欲望和知识能更好地理解它们的行为，那么即使它们没有哲学家定义的“意识”，说它们拥有心理状态也是合理的。
- **推论:** 仅仅通过感知环境，生物就在获得基础知识，这不需要复杂的意识反思。
 -

5. -感质本体论的特殊层级 (Minus Qualia Ontology)

- **批判对象：**那些认为意识是由“感质”(Qualia) 构成，或者是某种独立于物理的特殊现实层级的观点。
- **逻辑减法：**威廉姆森拒绝这种将意识神秘化为“特殊本体论层级”的做法。

6. 一内在主义的认识论特权 (Minus Internalist Epistemic Privilege)

- **学术背景：**“内在主义者”(Internalists) 倾向于认为信念的辩护必须发生在意识层面。
- **反驳：**威廉姆森认为这是意识被高估的主要原因。既然大多数认知处理都在意识之外，强求意识内的辩护是错误的。

7. 一哲学重要性的过度夸大 (Minus Exaggerated Philosophical Importance)

- **总结：**威廉姆森的总结是“我不认为意识是什么大不了的事”(I don't think consciousness is that big a deal)。
- **区分：**他区分了“生活上的重要性”(不想死或昏迷) 与“哲学解释上的重要性”。哲学家错误地将前者投射到了后者上。

一句话总结第12页：威廉姆森主张将意识从神坛上拉下来，将其还原为一种普通的环境认知关系，并指出大多数心理状态（如知识和信念）实际上并不依赖于哲学家所迷恋的那种高级意识。

继续弄下一页

等式看世界=一个等式+一个等式=深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第13页的内容，这一页的核心探讨集中在**人工智能 (AI) 与心理状态 (Mental States) **的关系。威廉姆森提出了判断AI是否具有思维的实用主义标准——“最佳解释推论”，并区分了单纯的语言模仿与真正的环境交互。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

AI心理状态判定 = 最佳解释推论的证据效力 + 具身交互的感知接地 + 环境适应的进化功能 - 人类语言的中介性指称 - 石头归地式的冗余解释 - 主观裁决的本体论误区

等式要素深度解析

这个等式解构了威廉姆森在第13页中关于“AI是否真懂”的回答：他认为目前的语言模型（LLM）还不够格，但未来的具身机器人如果表现出只有通过归属“信念”才能解释的行为，那就真的拥有心理状态。

1. AI心理状态判定 (等式左侧结果)

- **思想：**探讨人工智能（如ChatGPT或未来机器人）是否能真正拥有知识、信念和欲望，而不仅仅是模拟。
- **结论：**这不是一个由人类投票决定的定义问题，而是一个关于该系统内部运作事实的科学假设检验问题。

2. +最佳解释推论的证据效力 (Evidence Power of Inference to the Best Explanation)

- **核心标准：**威廉姆森提出了一种功能主义的判断标准：如果我们理解一个机器人行为的“最佳方式”是赋予其知识、信念和欲望，那么这就是它确实拥有这些心理状态的证据。
- **逻辑：**心理状态的归属是因为它们能提供独特的解释力（Explanatory power），让我们能预测行为。

3. +具身交互的感知接地 (Embodied Interaction / Grounding)

- **区分：**威廉姆森认为ChatGPT目前的形态不足以拥有心理状态，因为它们与世界的联系太有限。
- **条件：**真正的心理状态可能需要“具身性”(Embodiment)，即像机器人一样具有感官知觉（Sensory perception）并以更广泛的方式作用于环境。

4. +环境适应的进化功能 (Evolutionary Function of Adaptation)

- **深层逻辑：**拥有思维（Mind）的全部意义，从进化角度看，是为了以“灵活的方式”与非常复杂、变化快速的环境进行互动。
- **推论：**只有当一个系统展现出这种需要通过“表征环境”来实现的灵活行为时，归属心理状态才是合理的。

5. 一人类语言的中介性指称 (Minus Mediated Reference)

- **批判对象：**当前的大语言模型（LLM）。
- **逻辑减法：**威廉姆森指出，ChatGPT使用的“猫”这个词之所以指向猫，完全是由“人类对语言的使用”所介导的（Mediated），而不是AI自己建立的指称关系。它是在借用我们的意义，而非产生意义。

6. 一石头归地式的冗余解释 (Minus Redundant Explanation like Stone-Desire)

- **反例：**针对“石头想要靠近地球”的古希腊观点。
- **逻辑减法：**虽然说石头“想要”掉下来并不完全错，但这没有解释力。石头不需要在脑子里“表征”地球就能掉下来，所以赋予其心理状态是冗余的。如果AI的行为像石头一样机械（无需表征即可解释），那它就没有思维。

7. 一主观裁决的本体论误区 (Minus Subjective Decision Fallacy)

- **澄清：**威廉姆森强调，AI是否有意识或心理状态，并不是由我们（人类）来决定的（Not for us to decide）。
- **实在论立场：**事实真相取决于那个机器人内部实际发生了什么。我们只是在测试假设，而不是在进行某种社会协商或定义游戏。

一句话总结第13页：威廉姆森认为目前的AI只是借用人类语言指称的“空壳”，但如果未来的机器人能通过具身交互展现出必须用“信念/欲望”才能解释的复杂环境适应行为，那么它们就客观上拥有了心理状态。

继续下一页

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第14页的内容，这一页的核心探讨了**“思维”(Mind)、“心理状态”(Mental States) 与“知识”(Knowledge) 之间的层级关系，并提出了著名的“知识是安全的真信念”** (Knowledge as Safe Belief) 理论，同时反驳了“知道就必然知道自己知道”(KK原则)。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

威廉姆森心灵与知识解构=思维作为心理状态标签+缄默知识的自动驾驶+知识的安全容错空间+元认知的双重叠加成本—意识的普遍必要性—KK原则的自动推导

等式要素深度解析

这个等式解构了威廉姆森在第14页中如何重新定义心灵哲学的基础概念：他将“思维”去神秘化，将“意识”边缘化，并用“安全边界”的数学/工程隐喻来量化知识的定义。

1. 威廉姆森心灵与知识解构 (等式左侧结果)

- **思想：**对心灵哲学中核心概念的清洗与重排。
- **结论：**建立了一个不需要以“意识”为核心的认知模型，强调功能的有效性与认知的可靠性（安全性）。

2. +思维作为心理状态标签 (Mind as Label for Mental States)

- **定义：**威廉姆森认为“思维”(Mind) 只是一个抽象术语，用于指代“任何具有心理状态的事物”。
- **逻辑：**这是一个简单的集合论定义：如果你有心理状态（如信念、欲望），你就有思维；没有心理状态，就没有思维。不需要额外的灵魂或实体。

3. +缄默知识的自动驾驶 (Autopilot of Tacit Knowledge)

- **案例：**骑自行车。当某人知道如何骑车时，这些知识并不在他的意识中运行。

- **反直觉：**如果这些知识被迫进入意识（变得“有意识”），骑车的人反而可能会摔倒。
- **推论：**无意识的心理状态（如潜意识的愤怒或技能）不仅存在，而且往往比有意识的状态更有效。

4. +知识的安全容错空间 (Safety Margin of Knowledge)

- **核心理论：**威廉姆森对“知识”的定义：为了知道某事，你必须在那件事上是“安全地正确”(Safely correct)。
- **工程隐喻：**这意味着你拥有一个“容错空间”(Margin for error)，保证你离犯错不会太近。仅仅碰巧猜对不是知识，必须在各种相似情形下都对才是知识。

5. +元认知的双重叠加成本 (Double Stacking Cost of Meta-cognition)

- **区别：**针对“知道”(K) 与“知道自己知道”(KK)。
- **逻辑：**“知道”需要一个层级的容错空间；而“知道自己知道”需要两个层级的容错空间（容错空间的容错空间）。
- **结论：**这是一个更苛刻、更昂贵的状态。因此，我们经常“知道”某事，却不一定“知道自己知道”它。

6. -意识的普遍必要性 (Minus Necessity of Consciousness)

- **减法操作：**威廉姆森将“意识”从心理状态的必要条件降级为“可选的额外项”(Optional extra)。
- **观点：**并非所有的心理状态都是有意识的，这是常识的一部分。意识通常指能用语言说出 (Verbalize)，但这对于拥有心理状态并非必需。

7. -KK原则的自动推导 (Minus KK Principle deduction)

- **哲学背景：**许多哲学家认为如果你知道P，你就必然知道你知道P（即KK原则）。

- **反驳：**威廉姆森通过上述的“双重安全成本”论证，切断了从K到KK的自动推导。这是他《知识及其限度》一书中的核心反叛观点。

一句话总结第14页：威廉姆森主张**“思维”只是心理状态的集合**，通过**骑车摔倒**的例子证明意识往往是认知的干扰项，并用**“安全容错空间”的层级差异**（一层 vs 两层）精彩地证明了为什么我们往往“知道”却“不知道自己知道”。

[继续下一页](#)

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第15页的内容，这一页的核心探讨集中在**语言意义的本质 (Nature of Meaning)**。面对人工智能中流行的“向量嵌入”(Vector Embeddings)和“分布假设”，威廉姆森坚持**指称外部主义**，认为意义必须锚定在外部世界，而非系统内部的关系。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

语义指称的外部实在性 = 指称关系的物理锚定 + 主体间环境三角定位 + 外部世界的直接触达 — 向量嵌入的内部主义闭环 — 分布假设的符号接地缺失 — 大脑内部的孤立解码

等式要素深度解析

这个等式解构了威廉姆森在第15页中关于“AI是否理解意义”的语言哲学批判：他反对将意义简化为高维空间中的数学位置，强调语言必须“触碰”世界。

1. 语义指称的外部实在性 (等式左侧结果)

- **思想：**针对“一个词的意义就是它在向量嵌入中的位置”这一技术观点，威廉姆森提出了反驳。
- **结论：**意义 (Meaning) 的核心在于**指称 (Reference)**，即语言符号与外部实在之间的连接，而不仅仅是符号与符号之间的关系。

2. +指称关系的物理锚定 (Physical Anchoring of Reference)

- **定义：**指称通常是单词与外部环境中的事物（而不仅仅是一个人脑子里的东西）之间的一种关系。
- **案例：**翻译“狗”这个词时，我们不是在寻找数学上的对应，而是在寻找另一种语言中同样指向现实中“狗”这种生物的词。

3. +主体间环境三角定位 (Inter-subjective Environmental Triangulation)

- **机制：**威廉姆森解释了人类沟通的本质：我们之所以能交流，不是因为我们交换了某种内部的向量，而是因为我们都在对环境中的同一个物体进行“三角定位”(Triangulating on objects)。
- **逻辑：**这种基于外部共同客体的定位，是语言能够跨越个体差异进行有效传达的基础。

4. -向量嵌入的内部主义闭环 (Minus Internalist Loop of Vector Embeddings)

- **批判对象：**将意义仅仅视为向量嵌入 (Vector embeddings) 或分布假设。
- **逻辑减法：**这种观点将意义变成了一种纯粹“系统内部的东西”(Something internal to the system)。它只描述了词与词的距离，却丢失了词与世界的距离。

5. -分布假设的符号接地缺失 (Minus Symbol Grounding Failure)

- **核心概念：**威廉姆森提到了“符号接地问题”(Symbol grounding problem)。
- **批判：**仅仅谈论概念作用或推理作用并不能触及意义的核心。如果符号没有“接地”到现实世界中，它们就只是空转的记号。

6. -大脑内部的孤立解码 (Minus Isolated Decoding inside Brain)

- **思想实验：**如果你试图在一个人的大脑（或AI模型）内部寻找他们指的是什么，而完全孤立地不看大脑与之互动的环境，那你将永远无法了解他们所表达

的意思。

- **结论：**意义不在头脑（或GPU）里，而在头脑与世界的互动中。

一句话总结第15页：威廉姆森通过**“指称”与“三角定位”的概念，击碎了AI向量嵌入的内部主义幻觉**，指出如果语言符号不与外部物理世界发生真实的触碰（接地），再复杂的数学向量也产生不了真正的意义。

继续下一页

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第16页的内容，这一页的核心探讨了**“什么是最大的哲学问题”以及威廉姆森独特的“投机性研究方法论”**。他否定了“清单式”的选题方式，主张通过微观技术问题的深挖（如反事实条件句），跟随论证的逻辑流向，最终抵达意想不到的宏观方法论变革。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

威廉姆森投机性研究法 = 问题形态的动态重塑 + 论证路径的意外追踪 + 反可能句的微观切入 — 静态清单的任务勾选 — 直觉启发法的隐形误导 — 预设终点的线性规划

等式要素深度解析

这个等式解构了威廉姆森在第16页中关于“如何做伟大的哲学研究”的回答：他认为哲学研究不是填表，而是一场未知的探险，从一个小的逻辑裂缝（如 $2+2=5$ ）切入，最终可能颠覆整个认知方法论。

1. 威廉姆森投机性研究法 (等式左侧结果)

- **思想：**威廉姆森自述其研究方式是“相当具有投机性”(Quite speculative) 的。
- **结论：**伟大的哲学洞见不是通过预先设定一个“大问题”来解决的，而是通过跟随逻辑的指引自然生长出来的。

2. +问题形态的动态重塑 (Dynamic Reshaping of Problems)

- **逻辑:** 哲学问题不同于数学问题，它在研究过程中会“不断发生转化”(Transformed)。你做的事情并不是在回答一开始想好的问题，而是逐渐建立起一个宏观图景。
- **细节:** 你通常并不知道正在开发的理论最终能解释什么。

3. +论证路径的意外追踪 (Unexpected Following of Arguments)

- **策略:** “跟随论证的指引”(Following the argument)，看看它在哪些地方有有趣的暗示。
- **结果:** 这种路径往往通向与最初初衷“完全不同”的结论。威廉姆森从研究具体的条件句出发，最后却写了一本关于哲学方法论和启发法的书。

4. +反可能句的微观切入 (Micro-entry via Counter-possibles)

- **具体案例:** 威廉姆森的研究起点非常具体且技术化——“反可能句”(Counter-possibles)，即前件不可能的反事实条件句（如“如果 $2+2=5\dots$ ”）。
- **价值:** 正是通过钻研这个看似小众的逻辑角落，他发现了人类认知评估系统的普遍缺陷。

5. -静态清单的任务勾选 (Minus Static Checklist Ticking)

- **观念减法:** 威廉姆森明确表示没有一份“哲学问题清单”供他一个一个打勾 (Tick off)。
- **批判:** 试图通过列出“最大的问题”来指导研究是行不通的，因为不同分支的人答案截然不同。

6. -直觉启发法的隐形误导 (Minus Invisible Misleading of Heuristics)

- **核心发现:** 在研究中，他意识到标准观点之所以认为反可能句是“空真”的，是因为依赖了不可靠的“认知启发法”(Cognitive heuristics)。

- **突破：**他证明了这种通用的评估启发法在本质上是“不一致的”(Inconsistent)，这推翻了人们视为“数据”的直觉基础。

7. 一预设终点的线性规划 (Minus Linear Planning of Pre-set Goals)

- **方法论拒绝：**他根本不是通过思考“某某是个好问题，让我试着解决它”来开展研究的。
- **逻辑：**这种线性的工程思维在哲学探索中是低效的，因为它限制了发现未知关联的可能性。

一句话总结第16页：威廉姆森通过自己从**“条件句逻辑”意外跨越到“认知启发法批判”的研究历程，证明了真正的哲学创新源于对论证路径的投机性追踪**，而非对预设宏大问题的线性求解。

继续弄下一页

等式看世界=一个等式+一个等式=深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第17页的内容，这一页的核心探讨分为两个部分：一是**反事实条件句（Counterfactuals）的逻辑解剖**，二是威廉姆森对**“真理应具有何种形态”的审美渴望**（简单、优雅）。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

反事实逻辑解析与理论审美=真值表逻辑复位+会字模态算子分离+奥斯瓦尔德刺杀案例+简单优雅理论的渴望—语境泛滥的不可控—丑陋复杂的现实纠缠

等式要素深度解析

这个等式解构了威廉姆森在第17页中如何用基础逻辑学处理复杂的语言难题，以及他作为一个理论家对世界秩序的主观愿景。

1. 反事实逻辑解析与理论审美 (等式左侧结果)

- **思想：** 这一页展示了威廉姆森如何通过“拆解法”处理反事实逻辑，并坦承了其哲学探索背后的审美驱动力。
- **结论：** 复杂的反事实语句可以还原为简单的逻辑结构加上模态算子；而好的理论应当是优雅的，尽管现实可能不是。

2. +真值表逻辑复位 (Return to Truth Tables)

- **核心观点：** 针对条件句（如果...那么...），威廉姆森捍卫了一个极具争议的观点：其“如果”部分确实遵循逻辑课上教给学生的标准真值表（Truth tables）。
- **逻辑：** 即使是反事实条件句，其基础逻辑骨架依然是经典的，前件为假并不导致逻辑崩溃。

3. +会字模态算子分离 (Separation of 'Would' Operator)

- **技术解构：** 威廉姆森认为，反事实条件句中的额外含义并非来自“如果”，而是来自“会”(Would) 这个词。
- **操作：** “会”是一个独立的模态算子，它有自己的生命（Life of its own）。理解反事实条件句的正确方法是：普通条件句 + “会”算子。这种拆解让问题变得清晰。
 -

4. +奥斯瓦尔德刺杀案例 (Oswald Assassination Case)

- **案例对比：** 威廉姆森通过对比两个句子来区分普通条件句与反事实条件句：
 - **普通：** 如果奥斯瓦尔德没杀肯尼迪，有人杀了（显而易见，因为肯尼迪死了）。
 - **反事实：** 如果奥斯瓦尔德没杀肯尼迪，别人会杀（这就变成了阴谋论，暗示有备用刺客）。
- **功能：** 这个例子生动地展示了“会”字如何改变了语句的真值条件。

5. +简单优雅理论的渴望 (Desire for Elegant Theories)

- **主观愿景：**当被问及“希望什么是真的”时，威廉姆森表示他希望世界能由“相对简单、优雅、强大的理论”(Simple, elegant, powerful theories) 来解释。
- **动机：**这是一种对秩序和美的智性追求。

6. -语境泛滥的不可控 (Minus Uncontrolled Context)

- **恐惧对象：**他希望自然语言语义中“语境的作用”是相对有限的。
- **逻辑减法：**如果语境的作用无限大，我们将很难对任何事情获得“理论控制”(Theoretical control)。为了理论的可能性，必须限制语境的干扰。

7. -丑陋复杂的现实纠缠 (Minus Ugly Reality)

- **现实认知：**威廉姆森承认他的愿景可能只是一厢情愿，因为实际上“很多事物都丑陋得可怕”(Hideously ugly)。
- **冲突：**理论家的审美渴望与混乱的本体论现实之间存在张力，但他依然选择希望前者为真。

一句话总结第17页：威廉姆森通过将“会”(Would) 从条件句中剥离，捍卫了经典逻辑在反事实推理中的地位，并表达了对“简单优雅理论”的哲学乡愁，以此对抗现实世界的丑陋与语境的混乱。

继续弄18页

等式看世界=一个等式+一个等式=深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第18页的内容，这一页的核心探讨了**“如何给学生建议”以及“灵感 (Insight) 从何而来”**。威廉姆森打破了对“天才顿悟”的迷信，提出了一个极具反直觉的观点：原创性不是追求来的，而是通过追求准确性“被迫”产生的。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

威廉姆森学术原创性与灵感生成机制 = 对真理准确性的强制性关注 + 公认观点的必然证伪 + 对话与低压环境的灵感涌现 - 天才式顿悟的被动等待 - 直接追求原创的虚荣焦虑 - 高赌注下的最佳状态强迫症

等式要素深度解析

这个等式解构了威廉姆森在第18页中关于“学术生产力”的实用主义哲学：他认为真正的创新是诚实面对真理的副产品，而灵感则需要在一个“低赌注”的心理环境中才能浮现。

1. 威廉姆森学术原创性与灵感生成机制 (等式左侧结果)

- **思想：** 这一页揭示了这位顶级哲学家如何指导学生进行创新，以及他自己如何捕捉思想。
- **结论：** 原创性不是目标，而是结果；灵感不是神迹，而是高强度思考后的放松产物。

2. + 对真理准确性的强制性关注 (Compulsory Focus on Accuracy)

- **核心建议：** 威廉姆森告诉学生“不要担心原创性，只要担心他们所说的话是否真实 (True)”。
- **逻辑机制：** 如果你足够关注准确性，你实际上会被迫 (Forced) 在许多领域说出原创的话。为什么？因为很多公认的观点其实是不准确的。
- **推论：** 只要你诚实地修正错误，你就自动变得原创了。

3. + 公认观点的必然证伪 (Falsification of Received Views)

- **前提：** 大量的公认观点 (Received views) 是不正确的。
- **操作：** 如果你要做到准确，你就不能重复那些错误的旧观点。这迫使你想出一种不受同样批评的新观点。是“为了解决困难例子”的需求驱动了创新，而不是为了创新而创新。

4. + 对话与低压环境的灵感涌现 (Emergence in Discussion & Low Stakes)

- **灵感来源1（对话）：**想法经常在讨论中产生。威廉姆森引用名言：“在听到我说了什么之前，我怎么知道我自己在想什么？”。
- **灵感来源2（放松）：**当事情已经在脑子里思考了一段时间，你需要做的就是放松（Relax），等不再担心它时，答案就会来。
- **案例：**著名的“浴缸时刻”。他曾泡在热水里直到水变凉，强迫自己想通一个问题，但这属于一种特定的专注状态。

5. 一天才式顿悟的被动等待 (Minus Passive Waiting for Genius)

- **迷思：**有人认为哲学就是拥有可以总结为一句话的“天才洞见”(Genius insight)，然后坐在那里等它发生。
- **反驳：**威廉姆森明确指出“那从来都不是事实”(That's never the case)。核心在于“努力工作”(Hard work)。

6. 一直接追求原创的虚荣焦虑 (Minus Anxiety for Originality)

- **策略减法：**威廉姆森建议学生不要把“原创性”当作直接追求的目标。
- **逻辑：**那些缺乏信心或过度自信的学生，如果直接瞄准原创性，往往会走偏。通过瞄准“准确性”，原创性会自然作为副产品出现。

7. 一高赌注下的最佳状态强迫症 (Minus High-Stakes Performance Anxiety)

- **心理陷阱：**如果你对获取灵感的条件（如完美的安静、完美的浴缸）过于自觉，你会产生焦虑：“如果是最佳条件我还没想通，那我就永远想不通了”。
- **解决：**这种“赌注太大”(Stakes are too high) 的心态会扼杀灵感。你必须让赌注相对较低 (Relatively low)，思维才能顺畅流动。

一句话总结第18页：威廉姆森通过**“不要追求原创，要追求准确”的辩证法，指出了通往学术创新的捷径，并警告过度在意“灵感发生的条件”反而会因为心理焦虑而扼杀

灵感**。

弄19页

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第19页的内容，这一页的核心探讨集中在**自由意志（Free Will）与决定论（Determinism）**。威廉姆森表达了一种罕见的“冷漠”态度，他通过艺术家创作和救人的例子，论证了“必然性”（不得不做）往往比“自由选择”（权衡）具有更高的价值。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

威廉姆森自由意志冷漠论 = 决定论的相容性接纳 + 艺术创作的必然性赞赏 + 道德直觉的自动反应 - 自由意志主义的形而上学焦虑 - 道德权衡的犹豫缺陷 - 过度道德化的责任包袱

等式要素深度解析

这个等式解构了威廉姆森在第19页中关于“为什么我不想要那种传统的自由意志”的回答：他认为某种形式的“不自由”（不得不做）恰恰是审美和道德的最高境界。

1. 威廉姆森自由意志冷漠论 (等式左侧结果)

- **思想：** 威廉姆森自称是“少数几个并不真的在乎自己是否有自由意志的哲学家之一”。
- **结论：** 他对这一传统哲学大问题的态度是实用主义和心理上的超然，即使神谕告诉他“自由意志不存在且决定论是真的”，他也不会感到伤心。

2. + 决定论的相容性接纳 (Acceptance of Compatibilism)

- **立场：** 他倾向于认为“自由意志与决定论是相容的”。
- **逻辑：** 即使世界是决定的，并不影响人类行为的价值评价。

3. + 艺术创作的必然性赞赏 (Appreciation of Artistic Necessity)

- **案例：**当一位伟大的艺术家说“我必须这样画这幅画，因为我没法以其他方式画它”时，这是在表达对美学约束的深刻理解。
- **反直觉：**我们不会因为艺术家“没有选择”而认为这幅画不归功于他，反而会印象更深刻。在这里，“不得不做”是一种能力的体现。

4. + 道德直觉的自动反应 (Automatic Response of Moral Intuition)

- **案例：**一个人跳进池塘救溺水儿童，事后说“我没有多想，这是我必须做的事”
 -
- **价值判断：**这种在道德危机时刻表现出的“非自由”(在这件事上没有自由意志)被视为一种高尚的本能。

5. - 自由意志主义的形而上学焦虑 (Minus Metaphysical Anxiety of Libertarianism)

- **逻辑减法：**威廉姆森拒绝了自由意志主义者 (Libertarian) 那种认为“必须有其他选择可能”才算自由的焦虑。
- **观点：**假如自由意志真的与决定论不相容，那么这种自由意志根本就不是什么“值得追求的东西”。

6. - 道德权衡的犹豫缺陷 (Minus Defect of Moral Weighing)

- **对比：**相比于毫不犹豫救人的人，那个“在那儿权衡半天、意识到自己可以不管孩子溺水”的人虽然展现了更多“选择的自由”，但却让人觉得他“脑子有问题”。
- **结论：**过多的审慎权衡 (Deliberation) 在某些语境下是道德缺陷而非自由的证明。

7. - 过度道德化的责任包袱 (Minus Over-Moralizing Responsibility Burden)

- **个人特质：**威廉姆森坦言自己不是一个“过度讲道德 (moralizing) 的人”。

- **逻辑：**因此，他不太受那些关于“责任归属”的传统自由意志论证的影响，更关注行为本身的审美和直觉质量。

一句话总结第19页：威廉姆森通过艺术家创作和救落水儿童的例子，颠覆了对自由意志的崇拜，指出在最高级的审美和道德情境中，“不得不做”（必然性）远比“可以做别的”（选择权）更具价值。

继续弄20页

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于文档《[Chinese] Philosophy's Most Formidable Living Mind [DownSub.com].pdf》第20页的内容，这是访谈的结语部分。主持人库尔特·杰蒙加尔（Curt Jaimungal）总结了知识传播的***“价值延伸机制”**，强调了从被动收听到主动思考的转化，以及如何通过复听和深度阅读打破一次性消费的局限。

以下是符合严格格式的核心归纳等式及其深度解析：

核心归纳等式

深度对话的价值延伸生态 = Substack的独家思想沉淀 + 复听讲座的认知迭代 + 主持人中立与主观的二元分离 + 商业赞助的经济支撑 - 一次性消费的浅层理解 - 垃圾邮件的信息干扰

等式要素深度解析

这个等式解构了第20页中关于“如何最大化播客价值”的运营哲学：知识不仅仅是听一遍就过的音频，而是一个包含深度阅读、反复咀嚼和经济支撑的完整生态。

1. 深度对话的价值延伸生态 (等式左侧结果)

- **思想：**访谈结束并非终点，而是听众进入更深层思考的起点。
- **结论：**优质的知识内容需要通过不同的媒介（音频、文字、社区）进行多维度的延伸和巩固。

2. + Substack的独家思想沉淀 (Exclusive Substack Precipitation)

- **功能：**Substack不仅是通讯邮件，更是“科学类别的第二名”社区，提供YouTube上没有的独家文字内容和提前收听权。

- **价值：**它是对流媒体音频内容的文字化沉淀，提供了更精美的格式和更纯净的阅读体验。

3. +复听讲座的认知迭代 (Iterative Re-listening)

- **核心建议：**主持人强调自己从“重看讲座和播客”中受益匪浅，听众也反馈喜欢重听。
- **逻辑：**对于威廉姆森这样高密度的哲学对话，一次性收听往往只能获取表面信息，只有通过“复听”(Re-listening) 才能实现认知的真正迭代和内化。

4. +主持人中立与主观的二元分离 (Separation of Neutrality and Subjectivity)

- **方法论：**库尔特指出他在采访中保持“中立”(Neutral)，但在Substack中会打开窗口展示他“自己是怎么想的”。
- **意义：**这种分离保证了访谈的客观性，同时保留了主持人作为思想者的主体性，满足了听众对不同视角的需求。

5. +商业赞助的经济支撑 (Commercial Sponsorship Support)

- **现实基础：**提及《经济学人》(The Economist) 作为唯一的播客合作伙伴。
- **逻辑：**高质量的知识生产（如深度访谈）需要稳定的经济模型支撑，赞助商提供的折扣链接构成了这种生态的物质基础。

6. -一次性消费的浅层理解 (Minus Shallow Consumption)

- **对抗目标：**整个结语都在隐含地反对快餐式的内容消费。
- **逻辑减法：**通过鼓励订阅Substack和重听录音，试图抵消现代媒体环境中普遍存在的“听完即忘”的浅层消费习惯。

7. -垃圾邮件的信息干扰 (Minus Spam Interference)

- **承诺：**强调Substack是“零垃圾邮件”(Zero spam) 的。

- **价值：**在信息过载的时代，承诺不发送垃圾邮件是建立信任、确保高信噪比的关键。

一句话总结第20页：库尔特通过推广Substack深度阅读和复听习惯，构建了一个超越单纯音频播放的知识闭环，主张通过中立访谈与主观思考的分离来最大化哲学对话的长尾价值。

帮我弄个这篇文章的总结等式，要几个关键词 表象不是本质且 定义困难 经验为上是并不为真 过度怀疑没有评估立足点 会字算子和反事实的认识论启发 唯我论和一元论终会崩溃 AI的锚点为 主体间性 意识可能心理意向来解释 决定论和自由意志兼容。

等式看世界 = 一个等式 + 一个等式 = 深入 Custom Gem

基于您提供的关键词和整篇文档的核心逻辑，以下是为您定制的全篇总结等式。该等式串联了威廉姆森从认识论、形而上学到心灵哲学和语言哲学的完整思想链条。

核心归纳等式

威廉姆森哲学全景重构 = 表象非本质之定义困境 - 经验恰当性之真理局限 - 过度怀疑之评估基点丧失 + 会字算子反事实认识论启发 - 唯我论与一元论之逻辑崩塌 + 人工智能之主体间性锚点 + 意识之心理意向性还原 + 决定论与自由意志之兼容

等式要素深度解析

这个等式将蒂莫西·威廉姆森反直觉的哲学体系压缩为一个动态平衡系统：他通过做减法（剔除反实在论、过度怀疑、唯我论、一元论），并做加法（引入模态逻辑、主体间性、相容论），构建了一个稳健的实在论世界。

1. 表象非本质之定义困境 (等式左侧基石)

- **源关键词：**表象不是本质且定义困难
- **逻辑解析：**威廉姆森论证“表象”(Appearances) 并非认识论的安全港湾。我们对“世界如何向我显现”的主观判断往往比对“桌子在那里”的客观判断更难且更易出错。试图定义一个绝对正确的“表象”领域在逻辑上是困难且徒劳的。

2. -经验恰当性之真理局限 (Minus Empirical Adequacy)

- **源关键词:** 经验为上是并不为真
- **逻辑解析:** 针对范·弗拉森的观点，威廉姆森指出科学的目标不仅仅是获得“经验上精确”的理论（即仅仅符合观测数据），而是要追求可观测事物之下的底层现实（Underlying Reality）。仅满足经验恰当性并不等于获得了真理。

3. -过度怀疑之评估基点丧失 (Minus Skeptical Paralysis)

- **源关键词:** 过度怀疑没有评估立足点
- **逻辑解析:** 怀疑就像药物，过量即毒药。如果对其进行全面质疑，试图从所有信念中退一步，就会导致认知瘫痪，因为你失去了用来评估对错的任何“立足点”或背景信念。

4. +会字算子反事实认识论启发 (Plus 'Would' Operator Heuristics)

- **源关键词:** 会字算子和反事实的认识论启发
- **逻辑解析:** 理解反事实条件句（如果...那么...）的关键在于将“会”(Would) 作为一个独立的模态算子分离出来。这种逻辑拆解不仅澄清了语言结构，还启发了对人类如何通过“认知启发法”处理可能性的研究。

5. -唯我论与一元论之逻辑崩塌 (Minus Collapse of Solipsism/Monism)

- **源关键词:** 唯我论和一元论终会崩溃
- **逻辑解析:** 唯我论虽在辩论上不可战胜，但本质是像阴谋论一样的自我封闭系统。一元论（万物归一）则因为不得不否认“好论证与坏论证的区别”而陷入逻辑自杀。两者最终都会因缺乏解释力而崩塌。

6. +人工智能之主体间性锚点 (Plus AI Inter-subjectivity)

- **源关键词:** AI的锚点为主体间性
- **逻辑解析:** AI不能仅靠内部的向量嵌入获得意义。真正的意义（Meaning）必须建立在主体与环境互动的“三角定位”(Triangulation) 之上，这种主体间性（Inter-subjectivity）是语言指称现实的物理锚点。

7. +意识之心理意向性还原 (Plus Intentionality Reduction)

- **源关键词：** 意识可能心理意向来解释
- **逻辑解析：** 意识在哲学上被高估了。心理状态（如信念、知识）的核心在于“指向环境”的意向性关系，而非某种神秘的内在感质。只要能解释行为，就可以将心理状态归属于（即使是无意识的）系统。

8. +决定论与自由意志之兼容 (Plus Compatibilism)

- **源关键词：** 决定论和自由意志兼容
- **逻辑解析：** 自由意志并不意味着“可以做别的选择”。在艺术创作或紧急救人时，“不得不做”（必然性）恰恰体现了最高级的价值。因此，即使世界是决定的，这种高尚的“不自由”依然与决定论完美兼容。

我不认为物理主义在非常强烈的意义上很可能是正确的。你可以采取所有这些立场，认为自己不知何故比其他任何人都更开明。但这并不是自满或自得的好理由，因为这样的立场太廉价了，而且多得数不胜数。

这是一次对现实本质的深入探索。我们将讨论逻辑、自我、意识、语言等话题，到最后甚至会涉及艺术和自由意志。蒂莫西·威廉姆森教授是牛津大学的威克姆逻辑学教授，也是被引用最多的当代哲学家之一。他的名字经常出现在载入史册的当代思想家名单上。在本频道《万物理论》中，我是库尔特·杰蒙加尔，我以严谨的技术深度采访研究人员关于他们的现实理论。今天，反还原论者威廉姆森驳斥了各方的立场。我们涵盖了堆垛悖论（Sorites paradox），即关于“堆”的悖论。他剖析了唯我论，并考察了“意义”意味着什么，最终讨论了人工智能是否能真正了解任何事情。请坚持看到最后，因为他谈到了为什么真正的一元论者在哲学上是困惑的。

教授，您相信什么而大多数同事都反对您？

“哦，也许一个很好的例子是我对模糊性的看法，即对于像‘多少粒沙子构成一堆’或‘多少根头发让一个人不秃头’之类的问题该怎么说。在我看来，事情是有真相的。在‘堆’与‘非堆’之间，或者‘秃’与‘不秃’之间，存在一个特定的截断点。只是由于该术语的使用方式，我们无法确定截断点在哪里。所以，我的意思是，这与很多人的看法形成对比，很多人认为思想和语言中的模糊性需要对逻辑或真假区分进行某种修正，也许是采用真值的连续程度或类似的东西。而在我看来，根本不需要对标准逻辑和真假区分进行这些修正；只要将其理解为一种我们根本不知道的事实上的无知类型，就能更好地理解模糊性。”

所以你的主张是，大约是 310 粒沙子，但我们只是无法真正知道是 310 粒，还是怎么回事？

“是的，我的意思是，这就是那个想法。它可能会随语境的不同而变化，但至少在——假设在——我们现在的语境中，截断点可能是在 310 和 309 之间。当然，这也取决于沙粒究竟是如何堆放的之类的因素，才能构成真正的堆。但我们可以用完全经典的方式来推理它。作为一个模糊术语，我们不知道这个数字是多少，这只是我们的无知。而这并不妨碍我们使用标准逻辑。”

所以你指的是所谓的堆垛悖论（Sorites paradox），人们可能听过这个。我会在屏幕上留下斯坦福哲学百科全书相关条目的链接。现在你用了“辨别”（identify）这个词。关于什么构成了——嗯，多少粒沙子？如果你只有三粒沙子，你可能只会说这是一堆沙子。但在某一时刻，如果你有一整堆，你会用“堆”这个词。维特根斯坦的一个标准回应是，他说试图调查 157 粒沙子是否是一堆是一个愚蠢的努力。这只是我们使用语言的方式。所以我知道我可能对维特根斯坦的刻画不够精确，因此请随意精确地刻画他，然后告诉我们为什么在你看来他是错的，以及这与身份（identity），比如自

我 (self)，有什么关系。

“噢，好的，所以它并不——当我提‘辨别’时，我指的不是‘同一性’(identity)。我指的只是基本上发现那个数字是多少。这是我们真的不知道该怎么做的事情。所以维特根斯坦至少有一点是对的，那就是如果我们试图讨论到底需要多少粒沙子，我们是在浪费时间。但事实上，这里有一个有趣的地方，那就是其中的科学方面比哲学家们通常允许的要多得多。我的意思是，事实上，科学家确实研究过堆，因为他们对撒哈拉沙漠中沙丘的发展和运动等事情感兴趣，由于荒漠化等原因，这是一个相当严重的问题。所以，我的意思是，实际上确实有一门关于堆的科学，我猜维特根斯坦可能没想过会有这种科学。

甚至还有一个论点，哲学家 W.D. 哈特 (Bill Hart) 提出，事实上，我们可以知道那个截断点。答案是，构成一堆所需要的最少沙粒数是四粒。因为基本上，当你观察堆的物理特性时，相关的因素是你有一粒沙由于其它沙粒而稳定地停留。你用三粒也可以实现这一点，因为你知道你必须有两粒，然后不知何故有一粒平衡在它们上面，但这不会稳定。但如果有三粒，你可以得到一个由三粒组成的小三角形，然后有一粒在三角形顶部，这是一种稳定的配置。所以，在堆这种特定情况下，我们甚至可能真的知道。但当然，这并不能解决模糊性的一般问题，因为我们可以切换到像‘相当大的堆’这样的表达方式。可能你无法就到底需要多少粒沙子才能构成一堆‘相当大的堆’等等展开类似的论证。

所以一般性问题仍然存在。但我认为值得记住普通语言术语这种关于模糊性的讨论与科学互动的方式。因为这表明，人们不应该太有信心地假设永远不会有任何方法来回答这些问题。我是说，科学确实找到了回答最初看起来完全无望的问题的方法。”而我昨天并没有这个属性。但如果由于这就认为我不是昨天那个完整的人，那真是对逻辑的误用。我的意思是，这里只有一个人，他已经存在了相当长的一段时间。

当这种改变累积到更剧烈时，你就会遇到关于人的堆垛悖论。事实上，美国哲学家彼得·昂格尔曾经发表了一篇名为《我不存在》的文章，他在文章中使用了关于人而非沙堆的堆垛悖论，以此论证根本没有他这样一个人。我的意思是，我认为这是一个谬误，但当然，如果你一次从一个人身上减去一个原子，然后继续减去原子，如果你减得足够快、时间足够长，最终会导致该人的毁灭（假设原子没有被替换）。当然，目前还不清楚这一情况究竟会在什么时候发生。在医学上，关于如何确定确切的死亡时间以及应该使用什么标准，存在各种各样的问题。但我不认为这些真的会对人的存在产生任何严重的怀疑，正如对桌椅存在的严重怀疑一样，它们也受到类似的论证的影响。我的意思是，你可以从一张桌子上一粒接一粒地移除粒子，最终桌子就不在那里了。但作为宏观物体，我们无法免于这类悖论的影响。但我不认为我们难以确定死亡时间这一事实，就应该让我们认为这个人从未活过。”

对，对。好，现在刚进入频道的观众，在这一点上，也许他们是快进跳到这里的，或者也许这是一个独立的剪辑。这个播客叫《万物理论》。在这里，我通常会与理论物理学家交谈，他们从物理学的角度探索现实是什么。但现实中对人类有意义的其他方面是“意义”——即意义性以及什么是意义——还可能有什么是逻辑，什么是自我，心理状态意义上的心理性，也许还有意识，也许是语言。你写过几乎所有这些话题，而物理学家往往是还原论者。你以作为非还原论者，或者可能是反还原论者而闻名。所以我也想探讨一下这一点。你怎么能证明还原论是错误的呢？

“嗯，我的意思是，有很多不同的观点都被称为还原论。我认为最剧烈的观点也许是，唯一真正存在的东西是绝对基础的粒子，此外别无他物。老实说，我并不太认真对待这种观点，但我的意思是，当然，大量的物理学都关注诸如质量之类的特征，他们有兴趣将这些特征归属于宏观物体。而且，我不做还原论者的另一个原因是，不认为所有存在的东西都是这些物理上的基础物理量。因为物理学依赖于数学，而数学依赖于集合论或类似集合论的东西。当然，在集合论中，你是在讨论集合的存在。标准集合论的一个结果就是，集合的数量远多于基础粒子的数量。因此，物理学本身在某种程度上正假设着非物理基础的对象（如集合）的存在。当然，人们试图摆脱这类承诺（commitments）。但如果你按表面意思理解，这正是数学让你承诺的东西。而物理学家非常乐于依赖数学，且并不觉得需要解释数学为什么没问题，即使其基本公理是错误的。因此，我认为他们不具备做那种极端还原论者良好立场。不过也许你想到的是不太极端的还原论形式。”那么，这是否会让你不是一个物理主义者呢？

“嗯，在试图精确描述物理主义的内容时，其实存在一个臭名昭著的问题。因为，我想，它大概是说：万物皆物理。但很难定义什么是‘物理’。因为如果你尝试根据当前的物理学来定义它，那么很可能有很多事物尚未被当前的物理学所承认，但将在200年后的物理学中被承认。因此，将‘物理’的概念与物理学目前所达到的特定时间点联系起来似乎是错误的。但如果你只是说：‘好吧，物理的事物就是未来的物理学将承认的事物’，那就更糟了，因为那是假设我们知道什么才算作物理学。谁知道几千年后，当人们说另一种语言时，‘物理学’这个词会如何被使用呢？”教授这里提到的是所谓的“亨佩尔困境”(Hempel's dilemma)，我在尼亚加拉大学关于为什么我不买账模拟假设或物理主义的演讲中概述过。链接在屏幕上和说明栏中。

“所以我认为，万物皆物理的想法本来就有些定义不明。但我倾向于认为存在各种数学对象等等，而且没有理由——例如，没有理由认为空集是一个物理对象。所以，我不认为物理主义在非常强烈的意义上很可能是正确的。我的意思是，我并不相信灵魂或类似的东西——通常那些不喜欢物理主义的人会求助于这些东西。但我并没有看到真的有任何支持物理主义的极好论证。我认为死板的物理主义者经常没有考虑到另一件事，那就是，万物皆物理这种理论，根本不是物理学方法所能胜任测试或评估的

那种理论。我的意思是，你知道，这不是那种你可以通过获得一些定量预测然后观察它们是否被证实或证伪来测试的东西。我的意思是，认为除了物理对象之外别无他物的理论，显然属于哲学中被称为‘形而上学’的分支，而这根本不是物理学的一个分支。所以我认为那些说只有物理对象的人，实际上是在致力于一种无法被物理学证实的形而上学学说。我的意思是，它也许可以被哲学证实，但物理学并不是正确的使用工具。”

我想正在听的物理主义者，或者认为自己是物理主义者的物理学家，会说些类似这样的话：“看，让我做一两个动作，也许同时做。第一，你已经指出我们可以有模糊性，而这只是我们自己的属性。这是我们自己的无知，但并不意味着世界本身存在模糊性。所以根据你所概述的，即我们不知道未来的理想物理学，并且物理学定义不明——也许‘定义不明’与模糊性略有不同——但我猜他们可能会说：‘我们难道不能反过来利用你自己的论证说，不，不，关于物理学是有明确定义的。我们并不知道它具体是什么，因为我们求助于某种未来的理想状态，但它仍然存在。’”所以我猜他们可能会做出那样的举动。我想请你在稍后对此做出回应。他们可能做的另一个举动是说，如果有人要声称在物理学之外还有东西，那是什么？指出来。是这个杯子吗？我可以用物理学工具来研究这个杯子。所以大多数人不会说杯子超出了物理学的范畴。所以他们可能会说：‘好吧，举证责任在别人身上。’那么你如何回应这两个举动？”“是的。所以我当然不是说‘物理’这个词毫无意义。我认为他们把我对模糊性的观点（至少在辩论上）应用到‘物理’一词上是完全没问题的。但如果他们这样做，并说：‘瞧，我们并不确切知道哪些事物算作物理的，哪些事物算作非物理的’，那么听起来他们实际上是允许了非物理事物存在的可能性。所以这是我对你第一点的回答。

对于第二点，嗯，当你问‘如果存在非物理的事物，它们是什么？’时，我确实提供了一个候选者，那就是空集。你知道，我认为人们也可以提供其他类型的候选者。比如，我们说小说《傲慢与偏见》是一个物理事物。当然，一本《傲慢与偏见》的具体副本是一个物理事物。但当我们谈论这本小说时，我们并不是在谈论它的任何具体副本。但与此同时，我认为我们可以说一些有意义且真实的事情。例如，简·奥斯汀完成了六部小说。所以，你知道，如果小说是可以计数的事物，那么如果它们的数量大于零，它们就存在。因此，甚至不清楚当我们谈论像小说这样普通的事物时，我们是否真的在谈论一个物理对象。”

你知道我花了多少时间研究嘉宾、阅读论文、准备这些对话。嗯，我没有时间做的是站在杂货店的过道里试图弄清楚晚饭吃什么。这就是HelloFresh的用途。新鲜食材预先分装，直接送到我家门口，配方非常实用。我不需要计划用餐，不需要在店里闲逛。我非常不喜欢逛商店和排队。我刚订了牛排、墨西哥卷饼（enchiladas）、带蘑菇的意面和三文鱼。天哪，那些墨西哥卷饼太棒了。你一直看到且即将看到的其实就

是我和我妻子就在这里制作它的过程。打动我的是新的 HelloFresh。天哪，那些墨西哥卷饼太棒了。你一直看到且即将看到的其实就是我和我妻子就在这里制作它的过程。打动我的是新的 HelloFresh。他们把菜单扩大了一倍。所以总会有符合我心情的选择。对于我们这些想要达到蛋白质目标而又不想吃乏味的鸡肉的人来说，他们增加了高级牛排和海鲜选项，且无需额外费用。我喜欢这一点。既然都是预先分装好的，做饭就变得超级简单。利用新的 HelloFresh，在 HelloFresh.com 使用我的代码 CURTJHFZWL，即可获得 10 顿免费餐食，外加终身免费早餐。在有效订阅期间，每个盒子里有一顿免费餐食。免费餐食作为第一盒的折扣应用。仅限新订户或回归订户。具体视计划而定。今天就注册，点击下面的链接，然后在 HelloFresh.com 使用我的代码。

教授，我不确定您是否知道一个叫 Reddit 的网站。您听过吗？

“是的，我听过。我偶尔会看一看。”

嗯，在 Reddit 上，有很多被称为 subreddits 的版块，它们只是人们聚集在一起讨论某个话题的不同类别。其中一个版块叫做 AskPhilosophy。它被称为 r/AskPhilosophy。这是一个讨论哲学话题的绝佳子版块。你的名字经常被提及，作为当代活着的哲学家之一，你将被载入史册。你是目前被引用最多的活着的哲学家之一。你的工作到底是什么，让其他哲学家产生如此大的共鸣？或者可能共鸣与争议并存，我不知道。它是什么？

“嗯，是的，被问到为什么我这么伟大有点尴尬。”

尤其是对于一个英国人来说。

“是的，试着从观察者的角度来看，我认为有很多原因。我的意思是，我确实写了很多文章，涉及很多不同的主题。我通常采取人们认为相当具有挑战性的观点。原因在于，如果我认为目前的共识大体上是正确的，那么我就没有动力去写它。我最有动力写的话题是那些在我看来人们真的搞错了的话题。我经常会想起一些比人们一直在构建的那种精细理论要简单得多、漂亮得多、有时甚至更符合常识（虽然并不总是）的观点。所以我喜欢展示一个简单得多的观点实际上可以处理所有的困难。

我想另一件事是，我喜欢从底层结构（通常是某种逻辑结构）的角度来思考哲学问题。所以我喜欢以一种在形式上非常严密的方式来发展我所持有的任何观点，而不是在手中就会散架。所以我希望，以这种方式，我所做的工作往往是相当稳健的，因为我所说的一切都有一个底层的逻辑结构。底层逻辑结构对所言之物至关重要。所以，它不会一经推敲就分崩离析。”我确信你听说过巴斯·范·弗拉森（Bas van Fraassen）。

“是的，是的，绝对听过。我见过他。我认识他。是的。”

太棒了。嗯，他刚上过这个播客，他说他是一个反实在论者。你是一个实在论

者，这对我来说意味着在你们私下或公开的讨论中，你可能激烈地反对过他，我想知道为什么？在我看来，他的什么观点是不正确的？

“我是一个实在论者，因为我倾向于认为在我们的认知生活中，我们处理的现实大多或多或少独立于我们的意志，也独立于我们对它们的看法。此外，我们有能力获得关于我们周围世界的知识——不仅是关于它对我们呈现出来的样子，而且是关于事物实际的样子。范·弗拉森的部分观点是，科学的真正目标仅仅是获得经验上精确的理论，即它们不会对我们可以驳倒的可观测事物做出预测。他似乎认为科学家的目标并不是试图发现我们可以实际观察到的事物之下的底层现实。我的意思是，他规定了科学家被允许追求的目标，这在我看来有点奇怪。我本以为这应该由科学家自己来决定。但我的强烈印象是，大多数科学家不仅仅对做出得到验证的预测感兴趣，他们真的想发现外面的世界是什么样子的。在我看来，这是一个完全说得通的观点，是他们追求的一个完全说得通的目标。

特别地，我认为他出错的地方在于，他赋予了‘表象’(appearances) 一种特殊的认识论特权。据我所知，他认为我们真正能调查的只是事物如何（可以说）向我们显现出来。我认为这是一个错误，部分原因在于，虽然他当然非常清楚我们可能会错误地判断事物的真实样貌，也就是说我们总是有可能犯错，这当然是真的。但我认为对于表象本身也是如此，即我们在事物如何向我们显现的问题上也没有任何绝对正确性。我们可能会误判事物向我们显现的方式。例如，当你坐在一个可能还有一群人的房间里时，要说出事物如何向你显现出来的挑战，几乎和说出房间里到底发生了什么的挑战一样艰巨。我的意思是，了解桌子、椅子、人们坐在哪里之类的事情，其实比了解你自己在被世界‘显现’时的那种主观状态要容易。因此，我认为那种认为表象具有任何特殊认识论特权的观点确实是个错误。”

是的。这里所说的表象是什么意思？

“嗯，所以我认为表象指的是某种仅仅（可以说是）主观的事物，因为它只与人所处的某种（可以说）内在思维状态有关。例如，如果我们谈论一副特定的画作（假设）现在对你来说看起来是什么样子的，那么这应该只是一个表象问题。而如果我们谈论的是独立于它现在看起来的样子之外的样子——我的意思是，毕竟有很多方面是你从目前的位置看不到的，或者你可能有些误觉——那是关于它是怎样的，而不是它看起来是怎样的。

有时候人们谈论表象时，他们实际上只是在谈论（你知道）我们倾向于判断事物的方式。但那也是一个问题，（你知道）这在认识论上并没有什么特殊的特权。因为如果我们谈论的是（你知道）我们自己做出判断的倾向是什么，我的意思是，这些事情（再一次）我们可能是错的。我们可能会误解我们自己的倾向和特质。我是说，自我认知是很难获得的。我认为人们在表象和其他事物之间划线的方式各不相同。但无

论你如何划线，表象都不会变成一个（你知道）舒适的小领域，在那里，无论我们认为什么是事实，那就是事实。我的意思是，在那里我们是全知的。”

我想如果我们认真对待表象的易错性，我们就会陷入某种怀疑主义问题。是吗？对于那些不知道什么是怀疑主义的人来说，什么是怀疑主义问题？

“是的。所以怀疑主义可能有很多相互关联的问题。但我思考它的方式是这样的：人们有一种非常健康的本能，有时会质疑自己的信念，并进行某种自我批评，试图从某件事中退一步（也许是他们从小被教导要相信的事），并询问这是否真的正确。那么他们所做的就是尝试根据他们所知道或相信的关于世界的其他一切来评估它。我们认为这当然是一件好事，因为我们不希望人们完全缺乏批评自己信念的能力。

问题在于，如果你走得太远，最终会导致彻底的瘫痪。因为如果你开始担心也许你关于世界的所有信念都是错误的，那么如果你试图从所有信念中退一步，你就真的没有任何基础来评估你关于世界的信念了。事实上，试图同时评估一个人关于世界的所有信念似乎是一个非常笨的主意，因为你没有留下任何可以评估它们的立场。我的意思是，你基本上扔掉了所有的资源，或者在没有任何剩余的情况下疏远了它们。所以我们处在一种处境中，少量的自我批评是好的，但太多了就只是糟糕且无望。很难知道适量是多少。我的意思是，这就像有人递给你一瓶药片说：‘瞧，吃这些药片。只吃几片。但如果你吃得太多，它们会对你非常非常不利。’但我就不能告诉你吃多少。瓶子上没写你应该吃多少。所以这是一种困难的情况。

我认为人们对怀疑主义论证印象过深，但怀疑主义论证之所以能对我们产生一点吸引力，是因为它们只是某种东西的一种极其夸张的形式，即在很小的剂量下确实是个好主意。很难知道具体该如何处理。但我认为从中得出怀疑主义教训是完全错误的方向。我的意思是，我认为怀疑论者非常聪明的一点是，（可以说）强加了他们自己的游戏规则，让他们看起来如果你不能说服怀疑论者他们是错的、他们的怀疑是放错地方的，那么怀疑论者就赢了。当然我们知道如果你要玩得足够彻底，从中得出怀疑主义教训是完全错误的方向。我的意思是，我认为怀疑论者非常聪明的一点是，（可以说）强加了他们自己的游戏规则，让他们看起来如果你不能说服怀疑论者他们是错的、他们的怀疑是放错地方的，那么怀疑论者就赢了。当然我们知道，如果你要玩怀疑论玩得足够彻底，你永远无法说服他们放弃怀疑。

毕竟，一个人可以信奉的一种简单的怀疑主义形式是对理性的怀疑——你可以怀疑任何事物是否是任何事物的理由。那么如果你试图给对理性的怀疑论者提供理由让他们放弃怀疑，他们只会说：‘瞧，你在乞题（begging the question），因为你在给我理由，而我的整个怀疑主义就是对理由的怀疑。’所以那是行不通的。而且，你知道，如果我们按照这些规则与对理性的怀疑论者玩怀疑主义游戏，那么怀疑论者将自动获胜。但我们不应该接受这些是正确的规则。我的意思是，我思考它的方式是，怀疑论

者把自己关进了一个爬不出来的坑里。我们真的无能为力把他们从坑里拉出来。但我们不能把他们从坑里拉出来并不意味着我们也得跳进去。我的意思是，我们要做的就是避开那个坑。”你怎么知道该吃多少片药？我不——我的意思是——好吧，让我这样说。如果你吃了足够的泰诺（Tylenol）——泰诺实际上很快就会变得相当有毒。我以前没意识到这一点。如果你吃四五片，它可能会把你送进医院。好，所以你可以注意到身体的迹象。但作为一个哲学家，是否有什么东西会触发让你知道：“好了，我这片怀疑主义药吃得太多了”？你是否看到了什么警告信号？

“不幸的是，没有你可以应用的规则。顺便说一下，我不是在说一条我可以应用的规则；我是说你，蒂莫西。所以，例如，在我最近的一些工作中，我感兴趣的一件事是，哲学家们在多大程度上因为依赖认知启发法（cognitive heuristics）而走错了路。启发法是一种廉价且快速的回答问题的方法，大多数时候是可靠的，但有时会给出错误的结果。你知道，人类的大量认知生活实际上取决于启发法的使用。所以我一直在质疑许多哲学家仅仅视为‘数据’的各种事物，包括在模糊性的案例中。

因此，我当然在想：我是不是变得有点太怀疑了？而且可以说，在这一点上我脑子里并没有红灯在闪烁。在我看来，通过谨慎行事，通过确保不仅仅是随机地质疑事物，而是通过实际识别我们所依赖的启发法，并展示这些启发法不可能百分之百正确，我可以将怀疑保持在合理的水平。所以我正尝试以一种不只是泛泛而谈的方式工作。因为有时候人们只是简单地使用一些直接、具体的怀疑主义论证实例，但同样的论证几乎可以用于任何事情。我认为当怀疑的原因太通用，以至于几乎可以对所有事情都提出怀疑时，就是一个警告信号。

所以，但我认为，你知道，我认为科学家可能必须处理这类问题，他们对数据有点怀疑，但他们不想直接丢弃它，而且正确的方法并不明显。比如，数据来自一个他们并不完全信任的实验室，诸如此类的事情。我认为，你知道，一个人只能对此行使判断力，这正是我们在社交生活中一直在做的事情。我是说，你知道，当人们开玩笑时，你知道，存在各种界限。嗯，他们有这样一个想法，当然，有时候人们在戏谑或其他事情上可能走得太远了。除了使用良好的判断力外，并没有任何非常明显的方法来确定这一点。我认为可能在每个学科中，获得该学科技能的一部分就是习得某种感觉，即（你知道）你需要对各种类型的数据保持多大程度的怀疑。”

《经济学人》（The Economist）涵盖了数学、物理、哲学和人工智能，其方式展示了不同国家如何看待发展以及它们如何影响市场。他们最近发表了一篇关于中国新型中微子探测器的文章。他们涵盖了通过线粒体移植延长寿命，从而开创了一个全新的医学领域。但不仅仅是科学领域。他们分析文化。他们分析每个地区的金融、经济、商业和国际事务。我特别喜欢他们新推出的 Insider 功能。它就在本月推出。它让你（也让我）能够近距离接触《经济学人》内部的编辑辩论，资深编辑在每周两次

的长节目中与世界领导人和政策制定者辩论新闻。基本上是一个极高质量的播客。

你还应该知道的另一件事是，如果你访问他们的应用程序，他们不仅有每日文章，还有与编辑和作家的长篇播客。这在线上也可以获得。无论是科学创新还是全球政治格局的变化，《经济学人》都提供了超越标题的全面报道。作为 TOE 的听众，你可以获得特别折扣。此外，直到 1 月 18 日——也就是本周——可享受五折优惠。请访问 economist.com/TOE 订阅。地址是 economist.com/TOE，[获取你的折扣](#)。

“言行不一”(performative contradiction) 在多大程度上进入你对“——也许不是你提到的红灯，也许是黄灯——你走得太远了”的分析中？例如，假设你非常怀疑，说椅子不存在，但你却去坐它。你是一个唯我论者，但你却试图说服别人唯我论是真的。那么你是否认为这暗示某人在学术上故弄玄虚或在进行思想上的自慰？你如何看待它？

“是的。所以，我的意思是，我认为这些看起来像是警告信号。我本人一点也不怀疑桌椅的存在。对于那些认为自己可以简单地否定它们存在的人——我的意思是，他们经常谈论诸如‘按桌子方式排列的粒子’(particles arranged table-wise) 之类的话，但没有真正展示他们如何能屏弃我们所有正常的思考桌子的方式。我的意思是，这给我的印象是他们走得太远了。我的意思是，我不想把这归咎于他们性格上的缺陷。我认为只是有些人太愿意在某些并不充分的基础上放弃我们用来引导世界的许多常识。”

但我认为对于哲学来说，可能是一件好事，即不同的人会有不同的点，在那个点上他们会划清界限，觉得怀疑主义变得疯狂了。从长远来看，我认为我们可以判断——不是以一种先验的方式，而是通过观察不同研究计划的多产程度。我想曾几何时，人们认为否定地球处于宇宙中心也是一种对常识的疯狂否定。但我的意思是，它已被证明是正确的举动。之所以如此，是因为它现在是（且几个世纪以来一直是）非常丰硕的科学研究计划的一部分，产生了大量产生的关于世界的新的知识。所以我认为我们不应该提前禁止所有这些不同的哲学研究方式。我认为应该让它们发展，看看它们会产生什么结果。怀疑主义往往产生不出什么。当它与产生大量知识（显然是大量知识）的非怀疑主义观点相对比时，怀疑主义看起来就像是某种蒙昧主义或对科学的否定。”

早些时候你说人们似乎对怀疑主义论证印象深刻。在这里，我想你使用的是“人们”而不是“哲学家”。你的意思是说哲学家，还是指普通大众往往更容易被怀疑主义或唯我论的论证所吸引？“嗯，我认为这不仅仅是专业的哲学家。所以，你知道，我记得有一次教一个本科生，那是他有史以来第一次上哲学辅导课，他告诉我：‘哦，当然，在哲学术语“知道”的意义上，没有人知道任何事情。’我当时在想：你到底是怎么知道‘知道’这个词的哲学意义是什么的？但对于很多人来说——那些完全没有接触过学术哲学的人——如果你只是说：‘也许我们都在做梦’，或者‘也许你是在做梦’，相当多的

人对此会反应说：‘哇，是的。我的意思是，是的，就我所知，我是在做梦。而且你知道，所以我不知道关于世界的任何事情，因为它可能都只是一场梦。’所以这是一种人们不需要大量的专业引导就能产生联想的方式。我认为这在人文科学的学生中可能比在科学专业的学生中更普遍。我曾与其他哲学家讨论过他们向本科生观众讨论怀疑论论证的经历，那些向人文科学学生演讲的人往往发现学生们会立即聚集在怀疑论的旗帜下。而像工程师这样的群体，则更不倾向于认真对待怀疑主义论证。所以我认为这确实与人们的学科倾向有一点点关系。我的意思是，其中一些可能——你可能会说，嗯，那是工程师更有常识——但也可能是人文科学的人，你知道，有更多的同理心，因此愿意以一种关注实际细节的人所不具备的方式，与这个假设的做梦主体产生共鸣。我的意思是，这仅仅是推测。”

对。在工程师或更偏向数学、科学的人与人文科学的人之间，存在许多干扰因素。所以由于人们采取这些观点的理由对我来说比这更不具吸引力。那么唯我论者到底错在哪里？另外，抱歉——简要地，对正在听的人（在这里用‘人’这个词其实挺危险的），但对于正在听的人，你能概述一下什么是唯我论吗？“是的。唯我论粗略地说，就是相信自己是世界上唯一存在的事实。除了我之外，世界上什么都没有。所以它并不完全是一个怀疑主义观点，因为怀疑论者只是怀疑除了他们自己之外是否还有其他东西存在。而唯我论者似乎对只有他们自己这一点很有信心。嗯，我认为这是一个极其错误的立场。但是，如果你问人们是如何陷入这种观点的……”

除了由于迷幻剂 (psychedelics)。

“是的，是的。我认为其中一些原因在于人们喜欢那种在辩论上（可以说）非常稳健的观点，也就是说，如果你开始捍卫这种观点，实际上任何与你争论的人都无法说服你是错的。你知道这对怀疑主义也是成立的，因为一个老练的怀疑论者可以对他遇到的任何论证保持怀疑。而唯我论者——你知道，如果他们的母亲告诉他们：‘你怎么敢假设她不存在？’那么他们可以答：‘嗯，那只是发生在我脑子里的一件事。外面并没有像我的母亲这样的人在跟我说话。这（你知道）只是一个人脑子里的声音。’也许一个严肃的唯我论者甚至不应该允许‘大脑’这种事物的存在。

但它是一种挑衅性的姿态，一种惹恼和震惊人们的方式，从你所采取的观点来看，别人很难给出令人信服的论证。所以，在某种程度上，对于那些喜欢采取极端和挑衅立场的人来说，这是一个相当有趣的立场。但我认为他们犯的一个错误可能是过于依赖这样一种评估立场的方式：看该立场是否能按照持有该立场的人的标准被反驳掉。有很多非常愚蠢的理论是无法以这种方式反驳的。这不仅包括哲学理论，还包括许多阴谋论。你知道，如果你给一个阴谋论者看一份与他们所说的完全不相符的报纸报道，他们当然只会说这份报纸显然是阴谋的一部分，以此类推。因此，你可以采取所有这些立场，认为自己不知何故比其他任何人都更开明，而别人对此无能为力。但

你知道，这并不是自满或自得的好理由，因为这样的立场太廉价了，而且多得数不胜数，以至于它真的说明不了什么。我认为这与谈话中的辩论或争论作为评估立场的一种方式的局限性有关。”

是的，我在想，像卡尔·荣格这样会分析人心理的人，是否会说你想要变得‘开明’，甚至可能想要被视为‘开明’，这才是潜在隐藏在你唯我论背后的秘密。不管怎么说，当我听到这里时，我意识到有各种形式的唯心论（idealism）在唯我论的边缘徘徊——例如叔本华，那里仿佛有一个单一的根，这个根就是万物的所有，而这个根发出这些不同身份表象的触角，但实际上这个根不知何故才是真实且唯一的意识，而个体的思维不知何故是幻觉。那么你反对唯我论的论证是否扩展到包括叔本华的观点，还是排除叔本华的观点？“嗯，我不认为我说的任何话能说服叔本华改变主意。而且，你知道，那其实——我的意思是，对于一个理论来说，这是一个相当低的标准，因为很可能我说的任何话也无法说服一个阴谋论者改变主意。这并不意味着这是一个值得被认真对待的观点，认真对待需要远不止我刚才谈到的这种辩论上的稳健性

（robustness）。因为采取了错误观点的人，由于其持续的错误观点，也会使他们在评估证据等方面变得错误。我的意思是，这在阴谋论中非常明显，而且我也认为这适用于哲学理论。所以，对于我们这些没有被说服接受叔本华观点的人，我们只需要指出支持叔本华观点的证据匮乏。这对我们来说是一件非常合理的事情。而他让自己陷入了一种完全不被这类考虑所动的立场，那是他的问题。”

你是一个一元论者吗？你是一个多元论者、二元论者还是什么？你是如何思考现实的？

“嗯，我认为外面有很多不同的东西。我的意思是，严格意义上的一元论是相信只有一个东西。我记得曾与一个一元论者争辩，他刚发表过一篇捍卫一元论的论文，说有很多支持一元论的绝佳论证等等。他把一元论总结为：不存在区别。所以我问他好论证和坏论证之间是否有区别。他沉默了很长时间，然后说：‘没有’，因为他意识到他必须这么说。但这是一种我完全不感兴趣的观点。我认为外面有很多不同的东西，不仅数量众多，而且种类繁多。我们之前谈到了小说、空集、物理粒子。显然这还远不是一个详尽的清单。所以我认为有很多确实存在的东西，它们不能被归约为同一个东西。我的意思是，当然，在某种意义上，它们都是统一世界的一部分。并不是说这些东西都完全没有联系。但在这种意义上，我是某种多元论者。”作为非还原论者是否意味着你必须是非一元论者？另外，作为还原论者是否意味着你必须是一元论者？

“嗯，我刚才谈到的那种一元论是非常、非常极端的。所以，你知道，那些认为万物只是大量基础粒子的人，并不是最严格意义上的一元论者，因为在最严格的意义上，一元论是相信只有一种东西。它可能是某种较弱意义上的一元论，即认为只有一种这类事物。当然，有个问题是，什么算作一类事物？但如果这是一个一元论者——

我是说，那些认为只有基础粒子的人——那么他们已经相当具体地说明了那一种事物是什么。所以可能存在类似那样的还原论形式。但我意思是，例如，那些认为所有知识都可以某种方式归约为物理学的人，他们还没有让自己成为关于物理学的一元论者，因为当然在物理学内部存在很多多样性。但这种观点在某种意义上确实急剧减少了世界的多样性。”

我明白我正在犯一个大错，一个巨大的错误，即在与蒂莫西·威廉姆森交流时没有问他什么是逻辑以及什么是知识。每个人都那样问。我想听你详述一些罕见的主题。你对“意识”有什么看法？

“嗯，我认为意识在哲学上的重要性被极大地夸大了。我当然不认为它是某种在本体论上非常特殊的现实层级，即像某些人认为的那样，认为意识是由‘感质’(qualia)构成的，或者类似意识感觉（这种独立于物理的东西）。我认为在谈论意识时，最好集中在‘意识到某事’(being conscious of something) 这一问题上，即作为某种形式的认知。我的意思是，现在我看意识到你，例如，我也意识到我面前的麦克风。这只是与环境中的事物的一种认知关系。

我认为在认识论中，意识被高估了，因为那些自称为‘内在主义者’(internalists)的人倾向于认为，真正与例如信念的辩护等有关的行为，必须发生在意识层面。而心理学的证据似乎正推向这样的方向：事实上，大多数行为 (action) 都发生在意识之外。虽然意识是某种最终的工作场所，让我们可以在那里仔细权衡决定等等——我的意思是，这确实有道理——但我们不应该赋予它高于所有那些并未发生在意识层面的认知行为的特权。总结一下：我不认为意识是什么大不了的事。”

补充一下，不是说你认为意识不重要。你的意思是它在哲学上不是什么大不了的事，而不是说它在人们的生活中不重要。

“在哲学上，是的。我的意思是，显而易见，如果有人告诉你或我，他们准备给我们打一针，让我们余生都处于无意识状态，那我们肯定一点都不会高兴。对此感到高兴是很愚蠢的。但我认为意识不知何故被哲学家赋予了它不配拥有的特权。”

当人们谈论信念、知识和理解时，这些通常被称为心里状态。那心理状态是否以意识为前提？

“我认为这是否成立并不是很明显。但如果遇到一些动物，通过赋予它们（比方说）信念、欲望和知识，比通过任何其他方式都能更好地理解它们的行为，那么即使对它们大脑的某种调查显示它们没有任何符合哲学家的意识范式的东西，说它们确实拥有信念也是合理的。当然，在某种意义上，仅仅通过感知环境，我们在某种程度上就是通过感知意识到环境的；我们一直在获得关于环境的基础知识。如果我们面对的是一个在认识论意义上完全没有意识到任何事物的生物，那么听起来它就不像是一个拥有心灵的生物。因此，我认为这取决于我们在多大程度上将意识仅仅建模为‘意识

到某事’(仅仅是感知事物)，以及我们想在多大程度上将关于意识特殊作用的内容放入其中。但如果我们放入太多关于意识的要求，那么它就不再是拥有心理状态、信念和知识的必要条件了。”

那么根据你的说法，人工智能是否能真正了解事实或理解模型或其他什么下你的账户下？“嗯，我认为这不太好说，不管我刚才说了什么。因为，当然，有些人也可能会说高级形式的人工智能是有意识的。你知道，我认为人们通常心目中的那种人工智能似乎不足以赋予其心理状态，因为它们与世界的联系太有限了。虽然像 ChatGPT 之类的东西，它正在产生大量的文字。其中的单词确实指向环境中的物体。而这种指称完全是由人类对语言的使用所介导的。我的意思是，它们使用的单词之所以——你知道，如果人工智能使用‘猫’这个词，那是指猫，是因为人类对‘猫’这个词的定义就是这个。这与一个正在与环境互动的机器人非常不同，机器人具有某种形式的感官知觉，并以更广泛的方式作用于环境。当然，将人工智能与这些输入和输出形式结合起来，在技术上并不那么困难。我认为这与我们是否真的在谈论真正的心理状态有很大关系。我怀疑目前的这种人工智能形式是否真的值得被视为拥有心理状态。但我认为如果我们谈论的是具有更多自主权和更多与世界互动能力的系统中的人工智慧，那么我看不出有什么明确的理由拒绝承认它们拥有心理状态。如果理解一个机器人的行为的最好方式是赋予其知识、信念和欲望等，那么这就是它确实拥有这些心理状态的证据。”

这很有趣。如果我们理解某事的最佳方式是归属某种属性给它，那么这就是该事物拥有这些属性的证据？“是的。当然，这并不是说是否拥有心理状态由我们决定。我的意思是，事实真相取决于那个机器人或非人类动物或其他什么东西身上发生了什么。所以不是我们决定它是否有意识。只是当我们测试它具有心理状态的假设时，我们要看这个假设是否比任何其他关于现状的替代观点都能提供更好的解释。因此，如果‘它具有心理状态’这一假设（以及对其知识、信念和欲望系统的具体归属）（可以说是）产生了好的解释效果，那么这就是有利于该假设的证据。但我们讨论的是，人类在将心理状态归因于某种其他类型的系统时有多少证据。”

这很有趣，因为我思考心理状态或意识或任何这类属性的方式是关于实体的‘本体论’陈述，而不是关于我描述该实体的最佳方式的陈述。

“是的，我同意这一点。”

好，举个例子，我相信古希腊人认为石头想靠近地球。所以，就好像它们有这种靠近地球的‘目的’(telos) 或目的论，我们也可以将其概念化为一种‘想要’(want)。它有靠近地球的渴望。我猜那是成功的。现在，它在定量上并不成功，但也许这对当时的他们来说已经足够好了。然后这被不同的模型取代了，牛顿力学模型中没有任何名为‘想要’或‘渴望’的变量，所以我们消除了它。所以我假设——所以你刚才说你同意我

的看法。“是的，是的。我是说，所以问题是，（你知道）如果有人提出了一种重力理论，其中重力实际上是事物想要靠得更近的结果，而且物体越大，它们就越想要靠得更近，或者类似那样的理论……”

心越大（渴望越强），是的。

“那个理论，它不会完全出错，但它并没有我们所期望的将心理状态归因于生物的那种解释力。因为，我的意思是，思维的关键在于——我的意思是，从广泛的进化角度来看，拥有思维的全部意义是为了以灵活的方式与非常复杂、变化非常快的环境互动。我们对心理状态归属的期望是，它们能让我们预测灵活的行为，因为这才是归属心灵的全部意义。而我们在石头身上得不到这一点。所以我们无法真的从归属心灵中获得我们想要的那种独特的解释效益。我们期望的那类事情，例如，它们通过某种方式表征环境。这对于石头来说似乎完全不合适。我的意思是，它们不需要表征那些对它们有引力的物体就能产生引力。”你能帮我梳理一下思维、心理状态和意识之间的区别吗？哪些是另一个的必要条件，哪些是可选的等等？这三个：思维（mind）、心理状态（mental states）和意识（consciousness）。

“所以我认为‘思维’只是一个抽象术语，我们将其应用于任何具有心理状态的事物。所以如果你有心理状态，那么你就有了思维。如果你没有心理状态，你就没有思维。所以我认为这相对容易。至于意识，我的意思是，并非所有的心理状态都是有意识的。我认为这一点大家基本都同意。而且在某种程度上，这已成为常识的一部分。所以我们可能会通过说某人的行为（比如）是因为他们仍然对母亲感到生气来描述它，尽管他们对此一点也没有意识到。他们可能在生气却没有感觉到生气，但赋予他们一种对母亲持有的无意识愤怒态度，可能依然是理解他们各种不合理行为的最佳方式。所以这是一种无意识的心理状态。

甚至当我们谈论某人知道如何做某事时——他们有各种技能，而知识并不一定是意识层面的知识。所以我意思是，例如，当某人知道如何骑自行车时，他所拥有的许多知识在他骑自行车时并不在他的意识中运行。我的意思是，如果这些知识在意识中运行，他可能会摔倒。所以意识不到它是很重要的，但这仍然是他们确实知道的东西。

至于意识到各种事物的意识关系——如果它们被理解为（例如）如果某事对你来说是有意识的，那么你通常能够用语言说出它——那对心理状态来说并非必需。对于一个特定的心理状态，它是可选的额外项。”

你能知道某件事却不知道你知道它吗（即排除‘缄默知识’与‘显性知识’的区别）？

“嗯，这是我论证过你可以做到的事情。粗略地说，论证是一个人为了知道某事，你必须（粗略地说）在那件事上是‘安全地正确’。这样你就不会（可以说）离犯错太近。所以如果说某人知道某事，你只是在赋予一个层级的安全性，或者（你知道）

正如我有时所说的，一个层级的容错空间。而为了知道你知道某事，基本上你需要两个层级的容错空间。所以那是一个更苛刻的状态。所以，你知道，我认为——所以我认为在不过多担心意识问题的情况下，（你知道）基于广泛的认识论理由，你可以论证知道与知道你知道之间一定存在区别。我的意思是，我应该说我的这些论点是相当有争议的，但那确实是我在《知识及其限度》（Knowledge and Its Limits）和我的其他作品中论证过的东西。”

那是你最孤独的知性立场吗？很多计算主义者会关注向量嵌入（vector embeddings），这是大模型工作的一个组成部分。抛开架构不谈，存在各种单词或词元（tokens）。这些词被放置在某个高维空间中，词与词之间的各种关系被捕获，含义相近的词在空间中靠得更近。你还可以对某些概念进行形式上的加减运算，也能得到有意义的结果。比如“女王”减去“女性”等于“国王”之类的。我在一次晚宴上发言，听众中还有很多其他人工智能领域的人。我说这个播客的项目之一是调查“现实”，其中一部分涉及到意义以及意义是什么、指称的意义是什么，这对我来说并不是一个显而易见的问题。我不确定意义是什么。然后有人私下对我说：“库尔特，不，那只是向量嵌入。那就是意义的本质。一个词的意义就是它在向量嵌入中的位置。”我想这被称为“分布假设”（distributional hypothesis）。那种对意义的解释能让你满意吗？说到知道某事而不知道它——对我来说，那并不是我所说的意义记录。但在我看来，那位对话者会说：“那么你所说的意义是什么呢？”我会用言语表达表达，而他们可以用这些词然后指向向量嵌入，所以我无法用我的言语说服他们。至少我认为我做不到。

“是的。嗯，我认为意义与‘指称’（reference）有很大关系。指称通常是单词与外部环境中的事物（而不仅仅是一个人脑子里的东西）之间的一种关系。一个初级的思考方式是，当我们进行互译时，如果我们要翻译‘狗’这个词，我们会寻找另一种语言中具有相同指称的词，即指向‘狗’这种环境中的事物的词。

而将意义仅仅视为向量嵌入，使其变成了一种系统内部的东西。这似乎与人们有时谈论的‘概念作用’或‘推理作用’密切相关。但仅谈论这些系统内部的关系并不能触及意义的核心，意义是我们可以传达给他人并长期保存在记忆中的东西。我们传达给别人的不是某种向量嵌入，因为他们的向量可能与我们的完全不同，但我们仍然可以交流，因为我们是在对环境中的物体进行‘三角定位’。所以我认为刚才提到的主张前提是一种非常不足的、内在主义的意义概念。”你指的是“符号接地问题”（symbol grounding problem）吗？

“嗯，那可能是表达同一观点的另一种方式。我的意思是，如果你试图在一个人的大脑内部寻找他们指的是什么，仅仅孤立地看待大脑而不看它与之互动的环境，那么你将无法了解他们所表达的意思。”

什么是最大的哲学问题？

“我当然没有一份哲学问题清单供我一个一个打勾什么的。但我想如果你问一个数学家：‘什么是最大的数学问题？’他们可能会告诉你一些与他们自己的研究密切相关的大问题。我怀疑你会从不同的数学家那里得到截然不同的答案，这取决于他们所在的具体数学分支。

我认为哲学也不例外。部分原因在于——这可能与数学有所不同——在哲学中，问题在你研究它们的过程中不断发生转化。所以你做的事情并不是在回答一开始想好的问题，而是通过探索问题（可能是相当小的问题，比方说人们在哪里搞错了等等），你逐渐建立起一个可以解释很多事情宏观图景。但你通常并不知道你正在开发的理论将能够解释什么，也不能解释什么。抱歉，这是一个略显回避的回答，但是……”

不，这没关系。但能不能具体一点，给一个例子，比如你过去 10 年研究中的一个个人案例。“好的。例如，我对我之前提到的‘启发法’在我们的思维中的作用，以及我们依赖的（甚至是人类通用的）启发法在思考各种事物时如何引导我们误入歧途产生了浓厚兴趣。我进入这个领域是因为我对‘条件句’非常感兴趣，也就是‘如果……那么’语句，特别是关于‘如果 X 发生了，Y 就会发生’之类的反事实条件句。其中的一个关键问题（虽然某种程度上比较小众，但影响深远）是如何处理所谓的‘反可能句’

（counter-possibles），即前件不可能的反事实条件句。比如：‘如果 2 加 2 等于 5，那么物理学会有什么结果？’之类。

标准观点（我认为那是正确的观点）是这些句子基本上是‘空真’（vacuously correct），因为你无法找到它们的反例。所以我之前提到的启发法，即我们用来获取判断的某种方法，并不是那么可靠。然后通过思考这种启发法，我意识到我所假设的这种启发法实际上只是一个更通用的评估条件句的启发法的特例，而且人们可以证明这种更通用的启发法在本质上是不一致的，因此它不可能在所有时候都给出正确答案。这就是我如何开始更广泛地思考哲学中的启发法，以及它们在（例如）对特定实例做出错误评估、甚至由于实例的问题而非理论的问题而认为我们找到了理论的反例等方面所起的作用。因此，我几年前写了一本书，叫做《哲学中的过度拟合与启发法》（Overfitting and Heuristics in Philosophy），是关于启发法与哲学方法的相关性的。抱歉，这是一个相当长的故事，因为它是一段漫长的旅程。我是因为思考条件句，特别是反事实条件句，以及它们如何与我们对可能发生的情况的认识相关联而进入这个领域的。但最后我得出了关于我们需要重新思考哲学方法的一些结论，这跟我刚开始那段思路时想的完全不一样。你知道，这并不罕见。所以我倾向于通过（可以说）跟随论证的指引来工作，看看它在哪些地方有一些有趣的暗示，这可能与我最初开始思考这个问题的初衷非常不同。所以，从这个意义上说，我在做研究的方式上相当具有‘投机性’。所以，我根本不是通过思考‘某某是个好问题，让我试着解决它’来开

展研究的。”

既然我们正在谈论条件句和反事实，对于那些上过古典数学或逻辑课程并学习过经典逻辑的人——在第一节课中你会学到真值表，即如果前件为假，那么整个条件句就是真的。好，那么问题是：我们如何进行反事实推理？假设你早餐吃了鸡蛋，但如果我说：‘如果蒂莫西·威廉姆森早餐吃了培根，那么……’，但这总是平庸地（trivially）为真。‘如果蒂莫西·威廉姆森早餐吃了特蕾莎修女’，一切也都是平庸地为真。那么如果前件已经为假，反事实思维如何产生任何有价值或有用的结果呢？“是的。嗯，我认为答案是真值表——我的意思是，我会为我们的普通条件句辩护，比如‘如果 X 发生了，那么 Y 就发生了’之类。但我认为几乎没有人会为反事实条件句的真值表辩护。那是由于反事实条件句中含有一个额外的模态元素，它与可能性有关。一个关于普通条件句和反事实条件句区别的经典例子是我们可以肯定地说：‘如果奥斯瓦尔德没有射杀肯尼迪，那么就有别的人射杀了。’这是显而易见，因为我们知道有人射杀了肯尼迪。所以如果不是奥斯瓦尔德，那就是别人。但如果说：‘如果奥斯瓦尔德没有射杀肯尼迪，那么别的人也会射杀。’你是在做一个非常、非常不同的陈述，且是一个不那么显而易见的陈述，因为推测中你……”

这取决于你选择哪种阴谋论，是的。“是的。我的意思是，要么有第二个备用刺客在等待，要么可能有人会说因为他们认为肯尼迪不知何故是一个容易被暗杀的人。但你必须有类似那样的故事。所以这两类条件句——我是说，普通的那些和带有‘会’

（would）的那些——是非常不同的，必须由不同的工具来处理。我个人的观点（也相当有争议）是：‘如果’部分确实遵循逻辑课上教给学生的真值表，

而额外的内容仅来自‘会’（would）这个词，这个词并不仅仅在处理条件句时使用。例如，我们可以说：‘哦，她绝不会背叛朋友。’这里我们用了‘会’，但它不是一个条件句。它只是关于这个人绝不会做某事。我们还可以说：‘哦，他会那样做，不是吗？’等等。所以‘会’这个词本身就有其生命。我认为很多条件句理论出错的地方在于没有理解理解反事实条件句的方法是将其视为普通条件句‘如果’加上所谓的模态算子‘会’的结果。我认为如果你像那样将它们拆开，你实际上可以很好地理解它们的工作原理。”

谈到可能性，你希望什么是真的？我的意思不是政治之类的，而是其他哲学问题中什么是真的？

“嗯，我希望我的观点是真的。”

让我换个说法。你有很多证据，有很多知识。但你也有很多不确定的事情。让我们谈谈那些你不确定但希望是真的事。你希望什么是真的？如你所见，这是一个好问题，因为它不容易回答。

“我猜有各种样的事情，你知道，我希望一些相对简单、优雅、强大的理论是真

的，而不是一团糟到很难说出任何真理。这就是我的普遍态度。我希望语境在自然语言语义中的作用相对有限，因为否则将很难对很多事情获得任何理论控制。所以从某种意义上说，这是一种希望事物是美丽的愿景，尽管由于当然很多事物都丑陋得可怕。”

我很抱歉我的摄像头开了——噢。无论如何，你一直给学生的建议是什么？

“嗯，我不会给所有学生同样的建议，因为不同的学生需要不同的东西。例如，有些学生缺乏信心，而另一些则过度自信，所以我对那些需要更多自信的人和那些需要减少自信的人说的话是不一样的。我认为对几乎所有人来说都成立的一点是，努力工作很重要。这听起来很显而易见，但偶尔有人认为在哲学中，真正重要的是拥有可以用一句话总结的某种天才洞见。而你只需要在那儿坐等它发生。我的意思是，当然，那从来都不是事实。

我经常告诉学生不要担心原创性，只要担心他们所说的话是否真实。如果他们足够关注准确性，他们实际上会被迫在许多领域说出原创的话，因为公认的观点是不正确的。”

你能详细说明一下吗？如果你强迫自己只说准确的话，那么你最终会变得具有原创性，因为公认的观点通常是不正确的？

“是的。如果公认的观点都是不正确的，并且通过仔细观察你会发现它们确实是不正确的，那么如果你要做到准确，你所说的正面观点就不可能是任何那些公认的观点，因为那些观点是错误的。这迫使你想出另一种不同于它们的观点，且该观点不受同样类型的批评。我认为这经常驱动人们——迫使他们想出一些新的东西，因为所有的老方案都无法解释某些困难的例子或其他什么，或者不具有正确的结构特征。”

你是否发现对于你个人而言，你的大部分见解更多来自我列出的以下几项之一？也许你的大部分见解来自洗澡，或散步，或与同事交谈，或阅读，或写作——就像我刚说的那几样。或者也许你想到的是别的。教授，有没有什么东西一直以来比其他东西给你更多的见解？

“我喜欢散步，但我并不觉得那是一个获得新想法的特别好的方式。我最喜欢的散步是在苏格兰的山区。在那里，如果你只关注某个哲学问题，你可能会从悬崖上掉下去。所以这并不是个好主意。事实上，我喜欢它是因为那是一个我基本上无法思考哲学的机会和情境。有几次休息真是太好了。

我的想法经常在讨论中产生，但通常是在讨论中我会说一些话，说完之后，我会想：‘嘿，这主意不错。这里面可能有戏。’它给了我一个跟进的线索。我认为有人说过，你知道，‘在听到我说了什么之前，我怎么知道自己在想什么？’

但是，有些环境适合产生新想法，也有一些环境适合解决某些问题，比如当你有一种感觉好像就在嘴边却极难表达清楚想法的时候。有时洗澡是一个解决问题的好地

方。我曾有过这种情况，我觉得：‘这是一个我几乎可以抓取的想法，但还不完全在那儿。’所以有一次，我泡了一个热水澡并告诉自己：‘我要一直坐在这个浴缸里直到我想通为止。’两个小时后，水变得挺凉了，但我确实想通了。

但问题是，通常人们——当某件事已经思考了一段时间，但需要时间沉淀时，他们需要做的就是放松，等他们不再担心它时，事情就会来到他们身边。当然，数学也是如此。但如果你对此过于自觉，设置了（可以说）能敲定想法的最佳条件，那么你就很容易对‘瞧，这些是最佳条件，如果我现在还没想通，我永远都不会想通了，所以这赌注太大了’产生焦虑。所以你必须想办法让赌注相对较低，以便足够放松，让事情在脑海中理顺。这在解决问题的某些阶段是必要的。但当然，你也必须花很多时间。嗯，如果是逻辑方面的问题，那么你就得一直在纸上写写画画。”

你对自由意志有什么特别的（sordid，此处指不寻常或被视为世俗的）看法吗？特别是自由意志主义（libertarian free will）？

“我是少数几个并不真的在乎自己是否有自由意志的哲学家之一。我倾向于认为自由意志与决定论是相容的。我的意思是，如果我们讨论的是愿景，那也许是我希望的情况。但如果一个神谕告诉我自由意志与决定论不相容，我不会伤心。即使神谕还告诉我决定论是真的。

我的意思是，如果你思考艺术家及其与作品的关系，（你知道）如果一个艺术家说：‘我必须这样画这幅画，因为我没法以其他方式画它’，这似乎很正常且一点也不坏。我是说，假设这是一幅伟大的画。在这种情况下，艺术家可能会说：‘我只是画了它——我在这件事上没有自由意志。我必须那样画。’我们不会对此反应说：‘噢，好吧，如果你没有自由意志，那么这幅画就不能归功于你了。’我们甚至可能印象更深刻，并认为：‘实际上你无法以其他方式画它，说明你对美学约束等的理解非常牢固。’

我不确定普通行为的情况有多大不同。例如，如果你看到有人救了一个遇险的孩子——比如溺水——他没有多想，直接跳进池塘救了孩子。后来他可能会说：‘嗯，我不觉得在这件事上我有任何自由意志。这就是我必须做的事。’这似乎比那个在那儿权衡半天、意识到自己可以不管孩子溺水的家伙要好。你会开始怀疑那个权衡的人脑子是不是有问题。我不认为自由意志——假设它不与决定论相容——是多么值得追求的东西。当然，人们会喋喋不休地谈论责任等问题，但我不是一个过度讲道德（moralizing）的人。所以我并不太受这种考虑的影响。”

教授，感谢你花了两小时——超过两小时——的时间陪伴我。

“嗯，也谢谢你花了两小时听我说话。这是我的荣幸。”

你好，我是库尔特。如果你想从《万物理论》中获得更多内容和最好的听觉体验，请务必查看我的 Substack。主要好处包括每周你都能提前听到全新的节目，还能获得专为会员提供的额外文字内容。地址是 CurtJaimungal.org。你也直接在 Google

上搜索我的名字和 Substack。自那个 Substack 创办以来，它在某种程度上已经成为科学类别的第二名。

对于不熟悉的人来说，Substack 就像一份通讯邮件，格式精美且零垃圾邮件。这是关注本频道不曾出现在 YouTube 或 Patreon 上的独家内容的最佳场所。它是免费的。如果你愿意，也有支持我的方式，并且如果你这样做会获得特别奖励。

有些人问我：“嘿，库尔特，你和这么多理论物理、哲学、意识领域的专家交谈过，你自己是怎么想的？”虽然我在采访中保持中立，但 Substack 是一个了解我目前对这些话题思考的窗口。这是直接支持我的完美方式。CurtJaimungal.org 或在 Google 搜索。噢，我还收到了一些教授和研究人员的邮件和评论，说他们向学生推荐本节目。太棒了。如果你是教授或讲师，且有某期节目让学生受益，请务必分享。

当然，非常感谢我们的广告赞助商《经济学人》。访问 economist.com/TOE 获取年度订阅的大幅折扣。我也订阅了它，你也会喜欢的。本节目是他们目前合作的唯一播客。对我来说这是巨大的荣幸。而你则获得了独家折扣。地址是 economist.com/TOE。

最后，你需要知道这个播客在 iTunes、Spotify 等所有音频平台上都有。你只需要输入 Theories of Everything 找到。我知道我的姓氏很复杂，所以你可以直接搜索节目名。我个人从重看讲座和播客中受益匪浅，我也在评论中读到听众同样也喜欢重听。所以如果你愿意，你可以在 iTunes 或 Spotify 上重新听一遍。感谢你的收听。