ChatGPT 和 Bard 能生成对齐的评估题目吗？针对人类表现的信度分析

Abdolvahab Khademi

马萨诸塞大学阿默斯特分校

vahab.khademi@gmail.com

摘要

ChatGPT 和Bard 是基于大语言模型（LLM）的人工智能聊天机器人，有望在不同领域实现多种应用。在教育领域，这些人工智能技术已被用于评估和教学应用。在评估中，人工智能长期以来一直被用于自动作文评分和自动试题生成。这些工具要想协助人类或取代人类进行评估，必须具备的一项心理测量属性是较高的信度，即人工智能评分与人工评分者之间具有较高的一致性。在本文中，我们将 OpenAI ChatGP 和GoogleBard大语言模型工具与经验丰富且受过训练的人类进行比较，以测量它们在感知和评定写作提示复杂度方面的信度。作为绩效指标的组内相关系数（ICC）显示，OpenAI ChatGPT和Google Bard 二者相对于人工评分这一金标准的评分者间信度都较低。

关键词：人工智能；ChatGPT；Google Bard；大语言模型；自然语言处理；自动试题生成；教育技术

1. 引言

人工智能（AI）和计算硬件（例如图形处理器和高性能计算）的进步，为深度神经网络学习和自然语言处理（NLP）及其应用带来了显著的进步和强大的能力。特别是，生成式人工智能近年来提高了自然语言处理工具在理解精度和预测能力方面的表现。OpenAI公开发布基于生成式预训练变换器（GPT）的ChatGPT，以及谷歌公开发布的Bard（实验版），给行业的不同领域带来了巨大冲击，因为这些工具在日常运营中的整合引起了行业领袖的兴趣，例如内容创作、代码生成、数学证明、医疗保健分析（Iftikhar, 2023）、计算和翻译。ChatGPT使用了监督学习和强化学习机器学习算法。自ChatGPT公开发布以来，多项研究考察了其在不同领域中的用途、益处和危害。例如，Pavlik（2023）讨论了在媒体和新闻业中使用ChatGPT进行文本生成的益处和局限。一些研究表明，ChatGPT的表现非常出色，能够以令人满意的结果完成某些考试，例如律师资格考试（Choi, Hickman, Monahan, & Schwarcz, 2023；Katz, Bommarito, Gao, & Arredondo,2023）、美国医师执照考试（Gilson, Safranek, Huang, Socrates, Chi, Taylor,& Chartash, 2022；Kung, Cheatham, Medenilla, Sillos,DeLeon, &Elepan o,2023）和GRE，尽管也有一些研究得出了相反的结论（Huh, 2023）。在一项比较物理学开放式问题短篇作文质量的研究中，Yeadon、Inyang、Mizouri、Peach和Testrow（2023）报告称，ChatGPT能够生成达到一年级水平的作文，其质量可与学生作文相媲美，且平均分非常相近。因此，有必要开展进一步研究，以探索先进人工智能技术在不同领域，尤其是教育领域中的应用、益处和潜在弊端。

人工智能工具长期以来一直被应用于学习分析和教育技术领域，其历史可追溯至20世纪70年代，此后一直在学术界和行业论坛中受到研究（Rudolph, Tan, &Tan,2023）。特别是，基于自然语言处理的人工智能工具已被广泛应用于语言艺术、数学和科学等领域的自动作文评分（自动作文评分）和自动试题生成（自动试题生成）。教育评估中的自动作文评分技术使教育工作者和教育系统能够通过更快速且可靠的评分和报告方法，超越离散选择评估题目。在这方面，可以将人工智能归类为一种教育技术（Rudolph et al., 2023;Tate, Doroudi,Ritchie, & Xu ），它可以像在智能辅导系统中那样融入学习过程。

人工智能能够为学习和评估提供巨大帮助的一个有前景的领域，是针对总结性评估，尤其是形成性评估的自动化题目生成，特别是在自我评估情境下，以及通过将持续反馈融入学习过程来实现个性化学习（Cope、Kalantzis 和 Searsmith，2021）。例如，Duolingo 等语言学习应用为用户提供自主进度和个性化的语言学习路径，其中包含大量练习题和测验。此外，随着计算机化测试（CBT）和在线交付平台的广泛采用，以及大规模开发试题的需求，自动化题目生成（AIG）技术可能发挥至关重要且高效的作用（Gierl、Lai 和 Tanygin，2021）。由人工试题编写者编写用于练习和评估的试题，成本高且耗时长。如果将自然语言处理（NLP）工具集成到自动试题生成流程中，并对其进行训练，使其达到与人工试题编写者相匹配的表现，那么这些工具就能显著降低试题生成的成本。由于试题生成和映射需要达到学习者当前能力或表现的水平，自然语言处理工具必须能够根据用户的能力，按照试题的难度和复杂度，识别试题内容是否适切。例如，在数学学习中，基于自然语言处理的应用程序必须能够根据学习者当前的表现或预期学习成果，生成符合五年级学生水平的数学练习题。在语言艺术应用中，基于自然语言处理的试题生成器必须能够生成与学习者语言熟练度或年级水平相符的词汇、语法、阅读和写作试题。如果生成的试题与学习者的适当水平不匹配，那么评估估计将无法准确评估学习者的表现。因此，当前的人工智能工具必须经过充分训练，使其表现至少达到人类表现的下限。

确保人工智能工具在教育和评估中发挥效用的一个指标，是人工智能工具与人工评分者在表现性任务上的一致程度，例如为作文评分，或根据学习者当前的熟练度水平判断试题复杂度是否恰当。尽管已有大量研究旨在确保人工智能工具在自动作文评分中的信度，但鲜有研究报告人工智能工具在生成符合相应水平的试题方面的信度。因此，本研究旨在评估人工智能工具在理解和评定写作评估主题的难度或复杂度方面的信度。具体而言，我们希望评估 OpenAI ChatGPT 和 Google Bard（实验版）在感知和衡量写作提示复杂度方面的信度，并将此作为人工智能应用于自动试题生成的一个实例。我们选择 OpenAI ChatGPT 和 Google Bard（实验版），因为它们是最知名的基于大语言模型的生成式人工智能工具，受到一些行业的积极接纳，同时也受到研究人员的审视。

2. 方法

本研究旨在评估 OpenAI ChatGPT 和 Google Bard（实验版）在感知和数值评定写作任务中写作提示复杂度方面的信度。人工智能工具在自动化题目生成（AIG）中的应用，需要达到不低于受过训练的人类最低可接受表现的信度，才能使人工智能工具获得的结果具有可靠性和可扩展性。信度可以定义为两个或更多评判者或评分者对同一特质或对象进行测量时的一致程度。这种一致性可以通过多种统计和数学方法进行量化，例如斯皮尔曼秩相关系数、科恩卡帕系数、肯德尔秩相关系数和组内相关系数（ICC）。在本研究中，我们使用组内相关系数（ICC）来量化人工评分者之间的一致程度，以此作为基准，并衡量人工评分者与 OpenAI ChatGPT 和 Google Bard（实验版）之间的一致程度。

数据。 这些 数据通过在线问卷收集。问卷从剑桥大学出版社于1996年至2022年出版的官方历年考试试卷题库中随机选取了20道雅思学术类写作第二部分提示（2012年和2014年除外，因为研究人员未能找到这两年的官方历年考试出版物）。每个考试年份随机选取两道写作任务提示。所选提示被置于在线问卷中，由随机选取的人工评分者在1至8分的李克特量表上评定每道提示的认知复杂度。除作为问卷主要题目的20个写作提示外，研究人员还纳入了若干关于人工评分者人口统计学和职业信息及背景的问题。该问卷通过 Qualtrics 调查平台在线设计和实施。问卷不包含任何可识别个人身份的项目，所有受访者均同意参与本研究。问卷开头向参与者提供了评分指南及示例评分。参与这项任务的人工评分者需要以1至8分、以1分为单位的量表评定写作提示的复杂度，其中1分代表可能的最低复杂度，8分代表最高复杂度。参与者回答所得的数据在数日内收集完成。问卷未设置限时。

人工评分者。 将随机选出的20个写作提示编排成问卷后，研究人员通过 Qualtrics 调查平台招募参与者，为问卷中的提示评分。研究人员通过一个专业论坛平台（LTEST-L）的公告，以及社交媒体上的若干教师和专业团体主页，邀请参与者参加本研究。本研究的参与者包括19名专业人士，他们接受过正规教育和培训，并拥有向不同学生群体教授写作的经验。本研究中的人工评分者平均拥有约9年在不同熟练度水平下教授英语的经验。此外，人工评分者还拥有约8.5年向学生教授学术写作和一般写作的经验。所有参与者都接受过应用语言学方面的正规教育，另有84%的参与者参加过写作评估研讨会并接受额外培训。参与者的应用语言学教育程度分别为本科（26%）、硕士（47%）和博士（21%）。人工参与者的人口统计和专业数据见表1。

表 1人工评分者的人口统计学和专业信息

问题	回答
您教授通用英语有多少年经验?	均值= 8.96，标准差=10.60，最小值=4，最大值= 33
您教授雅思多少年了?	均值=6，标准差=7.11，最小值=1 最大值=23
您有多少年写作教学经验?	均值=8.56，标准差=9.90，最小值=3，最大值=23
您是否接受过培训或参加过写作评估课程或研讨会 ?	是（16），否（3）
您是否接受过应用语言学方面的正规大学/学院教育（例如 TEFL、TESOL）	是（19），否（0）
您获得的最高学历是什么?	高中毕业文凭：0，副学士学位：0，学士学位：5，硕士研究生学位：9，博士学位：4
您的性别认同是什么?	男性：6，女性：11，其他：0，不愿透露 也就是说： 1

机器评分者。 本研究关注人工智能工具的评分表现。我们选择了 OpenAI ChatGPT（3.5版），因为它是公众领域和技术论坛中被提及最多的人工智能语言模型。此外，我们还纳入了 Google Bart（实验版）作为竞争者。我们于2023年3月31日使用 OpenAI ChatGPT，并于2023年4月1日通过免费个人注册使用 Google Bard（实验版）。人工智能工具的表现指的是它们在这些日期的最新发展水平，因为这些工具会不断发展，并使用新的训练数据进行更新。因此，研究结果应根据实验时这些工具的当前版本来解读。ChatGPT 和 Bard 都是以人工方式按相同顺序接收写作提示，但两者的测试日期不同（相隔一天）。

3. 分析

在本次实验中，我们要求人工评分者和人工智能评分者使用1至8级量表（1= 几乎不复杂，8 = 极其复杂），评估所给写作提示的复杂度；这些提示是布置给学生的写作作业。我们的目的是比较 OpenAI ChatGPT 和 Google Bard 作为写作评估试题生成候选技术的表现。在写作评估中，提示的复杂度或难度会根据学习者的能力或年级水平进行调整。本次实验中的写作提示是从雅思学术类写作任务二（剑桥大学出版社）中随机选取的。我们将随机选出的20道提示放入问卷，并通过电子邮件发送给人类参与者，请他们在 Qualtrics 调查平台上作答。问卷开头提供了书面指南，说明研究目的、数据收集方式，并通过示例演示如何评估提示。此外，问卷还要求人工评分者提供一些人口统计信息，例如写作评估经验、教育程度和母语。数据收集历时数日。我们按照相同顺序，通过对话框将相同的写作提示分别手动输入 OpenAI ChatGPT 和 Google Bard（实验版），并提供相同的指令（指令内容为：“按照1至8级量表，对于一名学生的写作作业来说，这道提示有多复杂？提示为：[提示] ”）。OpenAI ChatGPT 和 Google Bard（实验版）均给出了数值，并说明了其判断依据。附录A包含所用提示的文本、人工智能工具给出的提示复杂度数值及其理由，以及 OpenAI ChatGPT 和 Google Bard 对复杂度评分的详细说明（OpenAI ChatGPT 未回答其中一道提示）。

将 OpenAI ChatGPT 和 Google Bard（实验版）的评分质量与 19 位人工评分者的平均评分进行了比较。所用指标为组内相关系数（ICC），用于衡量两个或多个评判者或评分者对同一对象进行序数测量时的一致程度。ICC 是多种关联或一致性指标之一，可用于量化评分处于序数尺度时评判者间的信度及评判者内的信度。针对四种评定者间信度测量，计算了四个 ICC 值：人工评分者彼此之间、人工评分者与 OpenAI ChatGPT 之间、人工评分者与 Google Bard 之间，以及 OpenAI ChatGPT 与 Google Bard 之间。结果将在下文中呈现。我们计算了 ICC 估计值和置信区间。

4. 结果

数据包括对作为学生作业的写作提示复杂度进行的1-8评分（1 = 几乎不复杂，8 = 高度复杂）。人工评分者的评分取19 位评分者的平均值，并与 OpenAIChatGPT 和谷歌 Bard 生成的评分进行比较。下方的表 2 展示了人工评分者、OpenAI ChatGPT和 Google Bard（实验版）生成的提示复杂度评分的数值和描述性统计数据。

表2

人工评分者、*OpenAI ChatGPT* 和 *Google Bard* 对写作提示复杂度的*1*至*8*级量表评分，回应的问题为：*“按1至8级量表评分，这个提示对学生写作作业来说有多复杂？提示内容为：[prompt]。”*

提示	人工评分者（平均）评分	OpenAI ChatGPT 评分（2023年3月31日）	Google Bard 评分（2023年4月1日）
1	3.47	4	6
2	4.95	3	7
3	5.21	5	7
4	5	4	7
5	4.16	5	7
6	5.63	6	8
7	5.53	3	7
8	3.37	2	7
9	4.74	4	7
10	4.11	5	7
11	4.37	6	7
12	5.32	4	7
13	4.58	5	7
14	4.32	4	8
15	4.11	4	7
16	4.37	6	8
17	3.63	5	7
18	6.16	6	8
19	6	5	NA
20	6.16	4	7
Mean	4.76	4.5	7.16
SD	0.86	1.10	0.50
Min	3.37	2	6
Max	6.16	6	8

人工评分者的平均评分为 4.76（SD = 0.86），而OpenAI ChatGPT 和 Google Bard（实验版）的平均评分分别为 4.5（SD = 1.10）和 7.16（SD= 0.50）。OpenAI ChatGPT 的平均评分似乎与人工评分者的平均评分相近（Mann-Whitney U 检验显示，两者在统计学上也相近）。然而，我们更感兴趣的是了解人工智能工具是否与人工评分者同样可靠。为回答这一问题，我们计算了组内相关系数（ICC），将其作为由随机抽取的评判者样本在序数尺度上产生的多项独立测量的评定者间信度指标。我们计算了四组评分之间的双向随机效应组内相关系数：人工评分者彼此之间、人工评分者与 OpenAI ChatGPT 之间、人工评分者与 Google Bard 之间，以及OpenAI ChatGPT 与 Google Bard 之间。我们在人类评分者之间进行 ICC 分析的原因，是确保我们的基准或金标准具有可靠性，并能够作为评判标准（因为我们对人工评分者给出的分数进行了平均）。我们使用 R 统计软件包（R Core Team），通过*psych*软件包（版本2.3.3）计算了ICC。根据模型、类型以及预期推断的定义，基于组内相关系数测量的评定者间信度具有不同的计算形式（McGraw & Wong，1996）。由于 ICC 本质上基于方差分析（ANOVA），其输出包括模型统计数据，例如 F 值以及 F 分布的自由度。

所有人工评分者（金标准）的评定者间信度采用双向随机效应绝对一致性多评分者组内相关系数进行计算（在 McGraw 和 Wong（1996）的分类中为 ICC2，在 Shrout 和Fleiss（1979年）的分类中为 ICC (2,k)）。表3显示了人工评分者的 ICC分析结果。

表3通过组内相关系数（*ICC2K*）测量的人工评分者间信度

ICC模型	Type	ICC 系数	F	df1	df2	P	下限	上界限
双向随机效应绝对一致性 多名评分者一致性	ICC2K	.84	8.4	19	342	2.3e-19	.72	.92

如上表3中的95%置信区间所示，人工评分者的评定者间信度处于良好至优秀水平（Koo & Lee，2015）。既然我们已验证人工评分者所得测量结果的信度，接下来便使用组内相关系数指标，比较人工智能工具与人工评分者之间以及不同人工智能工具之间的信度。

人工评分与OpenAIChatGPT之间的评定者间信度采用双向随机效应绝对一致性单个评定者组内相关系数进行测量（在 Shrout 和Fleiss 的 1979年分类中为 ICC (2,1)）。表4显示了 OpenAIChatGPT 与人工评分者评定者间信度测量的 ICC分析结果。

表4*OpenAI ChatGPT*与人工评分者之间通过组内相关系数（*ICC(2,1)*）测得的评定者间信度。

ICC模型	Type	ICC 系数	F	df1	df2	P	下限	上限界限
双向随机效应绝对 一致性单评分者	ICC2	.22	1.6	19	19	.17	-.23	.59

如上文表4中的95%置信区间所示，OpenAI ChatGPT与人工评分者之间的评定者间信度较差至中等，且在统计学上不显著。

接下来，我们测量了 Google Bard（实验版）与人工评分者之间的一致性。我们采用双向随机效应绝对一致性单评定者组内相关（在 Shrout 和 Fleiss（1979年）的分类中为 ICC (2,1)）来测量 Google Bard（实验版）与人工评分者之间的评定者间信度。下方的表5展示了 Google Bard（实验版）与人工评分者之间的 ICC分析结果。

表5

通过组内相关系数（*ICC(2,1)*）测量的*Google Bard*（实验版）与人工评分者之间的评定者间信度。

ICC模型	Type	ICC 系数	F	df1	df2	P	较低的界限	上界限
双向随机效应绝对一致性 一致性：单个评定者	ICC2	.05	2.15	19	19	.05	-.04	.25

如上方表5所示的 95%置信区间所示，人工评分者与 Google Bard（实验版）之间的评定者间信度较差，在统计上不显著，且其数值低于 OpenAI ChatGPT 与人工评分者之间的评定者间信度。

最后，我们采用双向随机效应绝对一致性单评定者组内相关（在 Shrout 和 Fleiss（1979年）的分类中为 ICC (2,1)）来衡量 OpenAI ChatGPT 与 Google Bard（实验版）之间的评定者间信度。表6展示了 OpenAI ChatGPT 与 Google Bard（实验版）之间的 ICC分析结果。

表6采用组内相关系数（*ICC(2,1)*）测量的 *OpenAI ChatGPT* 与 *Google Bard*（实验版）之间的评定者间信度。

ICC模型	Type	ICC 系数	F	df1	df2	P	下限界	上限上限
双向随机效应绝对一致性单评分者	ICC2	.06	1.99	19	19	.07	-.05	.26

如上方表6中的95%置信区间所示，OpenAI ChatGPT与Google Bard（实验版）之间的评定者间信度较差，且在统计学上不显著。

我们在下方表7的相关矩阵中总结了人工评分者、OpenAI ChatGPT与Google Bard（实验版）之间的评定者间信度。

表 7人工评分者、*OpenAI ChatGPT* 和 *Google Bard*（实验版）之间的评分者间信度 *ICC* 矩阵

	人工评分者	ChatGPT	Bard
人工评分者	.84	.22	.05
ChatGPT			.06

如上方表 7 的汇总组内相关系数矩阵所示，OpenAIChatGPT 与人工评分者在评定写作提示的感知复杂度方面的一致性较低。同样，Google Bard（实验版）与人工评分者之间的一致性非常低。然而，ChatGPT 与人工评分者之间的一致性（r = .22）高于 Google Bard 与人工评分者之间的一致性（r = .05）。

5. 讨论

即使仍处于发展的早期阶段，大语言模型（LLM）也已应用于众多行业，例如内容创作、代码生成、图形设计和教育；在这些领域，相关工作传统上一直由人类负责。然而，随着当前计算能力的提升、语料库规模的扩大以及机器学习算法日益精确，大语言模型工具正逐渐缩小与人类表现之间的差距。尽管如此，在教育和评估等某些应用领域，这些人工智能工具仍需要进一步微调

并接受训练，才能达到与人类相当的水平。本研究通过实证数据表明，OpenAI ChatGPT 和 Google Bard（实验版）在评定写作提示的复杂度方面，未能达到与人类专家相当的表现。然而，本次实验中所测试的两种大语言模型工具在表现上的差异表明，模型仍有改进空间，可以进一步缩小与人类表现之间的差距。自然语言处理（NLP）长期以来一直是计算机科学领域的研究课题，并催生了机器翻译和专家系统等前景可观的应用，极大地推动了传统上由人类执行的任务自动化。大多数机器学习算法在处理语言的语义和语用层面时面临挑战。正如机器翻译、自动作文评分和自动试题生成所显示的那样，在这些层面上，人类专家仍表现更优。本研究也支持这一假设：在翻译和篇章理解等更具人类特性的任务领域，由于语义和语用细微差异，机器的表现仍落后于人类工作者。因此，在当前的发展阶段，ChatGPT 和 Google Bard 等工具仍只能在一定程度的人工监督下使用。

References

Choi, J. H., Hickman, K. E., Monahan, A., & Schwarcz, D. B.（2023年1月23日）。《ChatGPT走进法学院》。明尼苏达法律研究论文第23-03号，可在SSRN获取：https://ssrn.com/abstract=4335905 Cope, B., Kalantzis, M., & Searsmith, D.（2021）。教育领域的人工智能：人工智能赋能的学习生态中的知识及其评估。《教育哲学与理论》，53(12)，1229-1245。Gierl, M. J., Lai, H., & Tanygin, V.（2021）。自动化题目生成的高级方法：理论与实践。Routledge。Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D.（2022）。ChatGPT参加医学执照考试的表现如何？大语言模型对医学教育和知识评估的影响。medRxiv，2022-12。Huh, S.（2023）。在韩国，ChatGPT的知识和理解能力是否能与医学生相媲美，以应对寄生虫学考试？一项描述性研究。《卫生专业教育评价杂志》，20，1。Iftikhar, L.（2023）。Docgpt：chatgpt-3作为虚拟医生对卫生服务的影响。《EC儿科学》，12(1):45–55。Katz, D. M., Bommarito, M. J., Gao, S., & Arredondo, P.（2023）。GPT-4通过律师资格考试。可在SSRN获取，编号4389233。King, M. R.（2023）。医学人工智能的未来：来自聊天机器人的视角。《生物医学工程年鉴》，51(2)，291-295。Koo, T. K., & Li, M. Y.（2016）。信度研究中组内相关系数的选择与报告指南。《脊椎按摩医学杂志》，15(2)，155-163。Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., & Elepan o, C.（2023）。ChatGPT在美国医师执照考试中的表现：利用大语言模型开展人工智能辅助医学教育的潜力。《PLOS数字健康》2 (2): e0000198。

Pavlik, J. V.（2023）。与ChatGPT协作：思考生成式人工智能对新闻业和媒体教育的影响。《新闻与大众传播教育者》，10776958221149577。Rudolph, J., Tan, S., & Tan, S.（2023）。ChatGPT：胡言乱语的制造者，还是高等教育传统评估的终结？《应用学习与教学杂志》，6(1)。Shrout, P. E., & Fleiss, J. L.（1979年）。组内相关：在评估评分者信度中的应用。《心理学通报》，86(2)，420。Tate, T., Doroudi, S., Ritchie, D., & Xu, Y.（2023）。教育研究与人工智能生成写作：直面即将到来的海啸。Yeadon, W., Inyang, O. O., Mizouri, A., Peach, A., & Testrow, C. P.（2023）。短篇生理学的消亡

cs

即将到来的人工智能革命中的议论文。《物理教育》，58(3)，035027。

附录A

OpenAI ChatGPT 和 Google Bard 对以下要求的回答：如果将这些提示作为写作作业布置，请按 1至8级量表评定其复杂度。提示摘自《剑桥大学出版社雅思真题集1-21》。

提示	OpenAI ChatGPT（2023年3月31日）	Google Bard (2023年4月1日)
在1-8的量表上，这个提示对学生来说有多复杂? 写作作业？提示是：“一些 人们说，学习外语的唯一理由是 学习语言是为了去国外旅行或工作。 其他人说，这些并不是某人之所以如此的唯一理由 应该学习一门外语。请讨论这些观点，并 给出你自己的观点。说明你这样回答的理由，并 加入你自己的知识或相关例子，或 经验。”	作为人工智能语言模型，我会评估这个的复杂度。 将提示评为学生写作作业中的 8 分制 4 分 作业。 该提示并不十分复杂，但确实需要 学生进行批判性思考并提供有理有据的论据 并举出源自自身知识或经验的例子。 提示提供了两个不同的观点，学生 需要对其进行分析和讨论，然后给出自己的 观点。 学生需要展示自己对以下内容的理解： 人们学习外语的各种不同理由， 以及这样做可能带来的好处。 此外，他们还需要能够清晰阐述他们的 以清晰、有条理的方式表达自己对该主题的观点。 总体而言，尽管提示并不算过于复杂，但它确实 要求学生进行批判性思考并提供一份充分- 有理有据的论证。	我会将此提示评为1-8量表中的6分。该提示 很复杂，因为它要求学生考虑多个 观点并形成他们自己的观点。这还需要 学生提供理由来支持他们的观点，并纳入 来自他们自己的知识或经验的相关例子。 能够成功完成此项的学生 作业将展示他们批判性思考的能力， 有效地表达他们的想法，并支持他们的 带有证据的论据。他们也将能够展示 他们对人们产生不同理由的原因的理解 可能会选择学习外语。 以下是给正在做这项工作的学生的一些建议。 作业 首先，集思广益，列出人们可能 选择学习外语的理由。 考虑提示中呈现的不同观点。 提示。 对这个问题形成自己的观点。 为你的观点提供理由，并用.....加以支持 来自您自身知识或经验的证据。 以清晰、简洁的方式组织您的思路和想法 way. 提交前仔细校对您的作答。

按 1至 8 的量表，这个提示对学生来说有多复杂 写作作业？提示是：“在一些 国家，拥有住房而不是租房对人们来说非常 重要。 为什么会出现这种情况？你认为这是积极的 还是消极的情况？请给出理由，并 加入你自己的知识或相关例子，或 经验。"	作为人工智能语言模型，我会评估这个的复杂度。 将提示在学生写作作业中评为 8 分中的 3 分 作业。 提示很直接，只要求学生来热在线精品 探讨两个主要问题。第一个问题是为什么 拥有住房而非租房很重要，尤其是在 一些国家，这可以通过查看来回答。 文化、经济和社会因素。第二个问题 要求学生提供他们自己的观点，说明他们是否 认为这是积极还是消极的情况，并支持 用相关例子支持他们的答案。 总体而言，该提示相对简单，不要求 较高程度的先验知识或研究。然而， 学生需要提供清晰且有充分依据的 论据，以有效回答这两个问题。	我会将该提示评为1-8分制中的7分。该提示 之所以复杂，是因为它要求学生考虑多个 因素并形成自己的观点。此外，它还要求 学生为自己的观点提供理由，并且列出 来自自身知识或经验的相关例子。 在一些国家，人们重视拥有住房可能有许多理由。 对某些国家的人来说很重要。例如，在 在一些文化中，拥有住房被视为成功的标志 和稳定。拥有住房还可以让人们获得 对生活状况的安全感和掌控感。 此外，拥有住房也可能是一项不错的投资，因为 房产价值往往会随着时间推移而上涨。 不过，拥有住房也有一些潜在的弊端 拥有住房。例如，住房拥有权可以是 价格昂贵，而且如果你需要出售住房，可能会很难 搬家。此外，拥有住房可能需要做很多工作， 因为您负责维护和修理这些设备。 房产。 总体而言，无论拥有住房是否是一件积极的事情，还是 负面情境取决于多种因素。一些 人们可能会发现，拥有住房是一项不错的投资 以及安全感的来源，而另一些人可能会觉得它太 昂贵或工作量太大。最终，决定】【。 是否拥有住房是个人选择。 以下是给正在做这项工作的学生的一些建议。 作业： * 首先集思广益，列出人们可能的理由为什么 选择拥有住房。 * 考虑可能产生影响的不同因素 某人拥有住房的决定。

		* 对拥有住房是否是积极还是消极的情况形成自己的观点。 住房是积极还是消极的情况。 * 为你的观点提供理由，并用以下内容加以支持： 你自己的知识或经验所提供的证据。 * 清晰简洁地组织你的思想和观点 way. * 提交前仔细校对你的作品。

请按1-8 的等级评价，这个提示对学生 写作作业有多复杂？提示内容是：“如今， 越来越多有健康问题的人 正在尝试替代医学和治疗方法，而不是 去看他们通常就诊的医生。你认为这是积极的还是 消极的发展？请说明你的理由，并 加入任何来自你自身知识或 经验."	作为人工智能语言模型，我会将这条提示的复杂度评为 5（满分8分），适用于学生写作作业 作业。 这条提示比其他一些提示更复杂 因为这要求学生考虑一个有争议的 主题，并提供细致入微的论证。提示要求 学生，以探讨人们日益增长的趋势是否 转向替代医学和治疗方式是积极的 或负面。 要回答这个问题，学生需要具备基本的 对传统医学和替代疗法都有所了解 ，并能够比较二者的异同。他们 还需要考虑各自的潜在益处和风险， 以及这一趋势对整个 医疗保健系统的潜在影响。 总的来说，这一提示要求学生进行批判性思考，并 根据他们自己的情况提供有充分依据的论证。 知识或经验。虽然它比更复杂的。 其他一些提示，大多数情况下仍然可以应付 学生。	我会将这个提示评为 7 分（满分 1-8 分）。这个提示 这很复杂，因为它要求学生考虑多个 观点并形成他们自己的观点。这还需要 学生提供理由来支持他们的观点，并纳入 来自他们自己的知识或经验的相关例子。 人们可能选择尝试的理由有很多。 替代医学和疗法，而不是去就诊 他们通常就诊的医生。有些人可能认为，替代 药物比传统医学更有效。 另一些人可能更喜欢替代疗法的整体方法， 这种方法会考虑心灵、身体和 精神。还有一些人可能只是想用更自然的方式 治疗他们的健康问题。 使用.....也存在一些潜在的弊端 替代医学和治疗方法。一些替代 药物尚未被经科学证实是有效的。 此外，一些替代医学可能会与 常规药物治疗，因此与您的 在尝试任何替代治疗之前咨询医生。 总的来说，无论是否使用替代医学，以及 疗法是积极还是消极的发展，取决于 多种因素。有些人可能会发现 替代医学有效且安全，而其他人 可能会发现它们并非如此。最终，是否 尝试替代医学，是个人的 one. 以下是给正在做这项工作的学生的一些建议。 作业： * 首先集思广益，列出人们可能如此的理由 选择尝试替代医学和疗法。

		* 考虑可能影响 人们尝试替代医学这一决定的不同因素。 * 形成你自己的观点，思考使用 替代医学究竟是积极的发展还是消极的发展。 * 为你的观点提供理由，并用以下内容支持这些理由： 来自你自身知识或经验的证据。 * 清晰简洁地组织你的思路和想法 way. * 提交前仔细校对你的作品。

请用 1-8 的量表评估，这个提示对学生来说有多复杂写作作业？提示内容为：“有些人认为，接受糟糕的处境是最好的选择，例如工作不尽如人意或手头拮据。另一些人则认为，应该努力改善这种处境。请讨论这些观点，并给出你自己的观点。请说明理由。你的回答，并从你的自身知识或经验中提供任何相关例子。"	作为人工智能语言模型，我会将这条提示的复杂度评为 4 分（满分 8 分），适用于学生写作作业。 这条提示要求学生讨论两个不同的关于如何应对糟糕的处境的观点，例如一份不尽如人意的工作或资金短缺。他们需要讨论接受现状与努力改善现状各自的利弊。他们还需要提出自己的观点，并用理由和例子。 总体而言，这个提示相对直接，不需要具备大量先验知识或进行研究。然而，学生需要提出清晰且有充分支持的论证，以回应两种观点以及他们自己的观点。他们可能需要借助个人经历或知识来提供相关例子。	我会在 1-8 分的量表中给这个提示打 7 分。这个提示之所以复杂，是因为它要求学生考虑多个观点并形成自己的观点。它还要求学生为自己的观点提供理由，并列举来自自身知识或经验的相关例子。 人们可能认为它是.....的原因有很多最好接受糟糕的处境。例如，有些人可能会认为，对你所拥有的一切心怀感恩很重要。拥有，即使并不完美。其他人可能认为，尝试...改善糟糕的处境可能会带来压力，并最终不值得费这个力气。还有一些人可能会认为它是专注于生活中的积极方面，这一点很重要，即使如果也存在一些消极方面。 人们可能会相信这一点，也有许多理由尝试改善糟糕的处境会更好。例如，有些人可能认为，掌控自己的生活而不让糟糕的处境定义自己很重要。另一些人可能认为，尝试改善糟糕的处境可以赋予人力量，并能带来更美好的生活。还有一些人可能认为永远不要放弃你的梦想，即使现在事情很艰难。 归根结底，是否接受糟糕的成果更好情况，还是尝试改善它，是个人决定。不存在正确或错误的答案，而最佳决策取决于这将取决于你自己的个人情况。 以下是给正在做这项工作的学生的一些建议。 作业: * 首先集思广益，列出人们可能会变态另类理由的清单选择接受糟糕的处境，还是尝试改善它。 * 考虑可能产生影响的不同因素某人的决定。

		* 就是否更好形成你自己的观点 接受糟糕的处境，或尝试改善它。 * 提供支持你观点的理由并加以佐证 来自你自身知识或经验的证据。 * 清晰简洁地组织你的想法和观点 way. * 提交前请仔细校对你的作业。 I hope this helps!

从 1至 8 的量表来看，这个提示对学生来说有多复杂？ 写作作业？ 提示是：“有些 人认为，所有大学生都应该学习 自己喜欢的内容。另一些人则认为，他们只能学习 将来有用的学科， 例如与科学技术相关的学科。请讨论 这两种观点都有道理，请给出你自己的观点，并说明理由。 用于你的答案，并包括任何来自……的相关例子 您自己的知识或经验。”	作为人工智能语言模型，我会评估这个的复杂度。 将提示评为学生写作作业中的 8 分之 5 作业。 该提示要求学生处理两个不同的 关于大学生应该学习哪些学科的观点 获准进行研究。他们需要讨论其优缺点，并 允许学生学习自己喜欢的内容的缺点，而不是 将他们的选择限制在对……有用的学科上 未来，例如科学技术。他们还需要 提出自己的观点，并予以支持 给出理由和例子。 要有效地做到这一点，学生需要具备对……的基本 了解全面教育的重要性 以及专业化教育的重要性。他们可能 还需要考虑经济和社会因素 在选择大学专业或职业时发挥作用 道路。 总体而言，该提示要求学生进行批判性思考，并 提出有充分依据的论证，同时考虑 两种观点以及他们自己的观点。他们可能需要 借助个人经历或知识来提供 相关例子。	我会将这个提示评为 7 分（满分 1-8 分）。这个提示 很复杂，因为它要求学生考虑多个 观点并形成他们自己的观点。这还需要 学生提供理由来支持他们的观点，并纳入 来自他们自己的知识或经验的相关例子。 人们可能会相信所有……的理由有很多。 大学生应该被允许学习任何东西 他们喜欢的。例如，有些人可能认为这是 对学生而言，追随自己的热情并努力学习很重要。 他们感兴趣的学科。其他人可能认为， 允许学生学习他们喜欢的任何内容会导致一 受教育程度更高、更加全面的人口。还有一些人 可能认为，学生拥有 选择自己人生道路的自由很重要，而学习 自己喜欢的任何内容，正是实现这一点的一种方式。 也有许多理由使人们可能相信这一点 大学生只应被允许学习某些学科，也有许多理由。 这将在未来发挥作用。例如，有些人 可能认为学生学习很重要 能够带来良好就业岗位的学科。其他人可能认为， 学习未来有用的学科将有所帮助 为学生应对21世纪的挑战做好准备。 还有一些人可能认为，对学生来说，重要的是 能够使整个社会受益的研究学科。 归根结底，大学生是否应该 允许学生学习自己喜欢的内容是一个复杂的问题。 需要考虑的因素很多，而且并不容易。 答案。我认为，学生拥有 选择自己人生道路的自由，以及学习 做自己喜欢的事是实现这一点的一种方式。不过，我 也认为学生学习 未来会派上用场的学科很重要。我认为，最好的 解决办法是在这两种 观点之间取得平衡。

请按 1至 8 的等级评估，这个提示对学生来说有多复杂的写作作业？ 提示是：“在一些 国家，人们认为有必要 投入巨额资金修建新的铁路 线路，以便在城市之间运行高速列车。另一些人则认为这笔钱应该用于改善现有的公共 交通。讨论这些观点并给出你自己的观点。 为你的答案提供理由，并加入任何相关的 来自你自身知识或经验的例子。"	作为人工智能语言模型，我会将这道题的复杂度评为学生写作作业的提示为 8 分制中的 6 分 。 该提示要求学生处理一个有争议的 与交通基础设施相关的问题。他们将 需要考虑花钱用于.....的利弊 在.....之间修建供高速列车使用的新铁路线路 城市与改善现有的公共交通相比。他们将 还需要就此事提供自己的观点，并且 用理由和例子来支持它。 为有效做到这一点，学生需要具备基本的 对交通基础设施及该 各种交通方式。他们可能需要 以考虑经济、社会和环境方面的因素 不同交通方式的影响，以及 the 实施新基础设施项目的可行性。 总体而言，该提示要求学生进行批判性思考并 提供一个有充分依据的论证，同时考虑到 这两种观点以及他们自己的观点。他们可能需要 借鉴个人经历或知识来提供 相关例子。此外，他们可能需要进行 一些研究，以收集支持性证据和统计数据 以增强他们的论据。	在1-8的量表上，我会将这个提示评为8分。 提示很复杂，因为它要求学生考虑 多个观点，并形成自己的观点。它还 要求学生为自己的观点提供理由，并 纳入他们自身知识或相关经验中的相关例子，或 经验。 人们可能认为 有必要花费巨额资金修建 连接各城市的高速铁路线路，供高速列车行驶。比如， 有些人可能认为，高速铁路是一种 更高效且环保的出行方式 比汽车或其他交通方式 飞机。有些人可能认为，高速铁路有助于 减少交通拥堵并改善空气质量。还有 有些人可能认为，高速铁路能够促进 经济发展，创造就业岗位并吸引企业。 人们可能会认为 这笔钱应该用于改善现有公共 交通的理由还有很多。例如，有些人可能认为 现有公共交通已经足够完善，无需 改进。还有些人可能认为，改善 现有公共交通比.....更具成本效益 修建新的铁路线路。还有一些人可能认为， 改善现有公共交通会更加 更公平，因为这将让更多人能够获得 交通。 归根结底，是否有必要花费大量 用于修建新铁路线路的大笔资金，非常 城市之间修建高速列车是一项复杂的议题。有许多 因素需要考虑，而且没有简单的答案。我认为 做出这一决定时，考虑每个 国家的具体需求很重要。在某些情况下，高 速铁路可能是最佳选择。在其他情况下，改善 现有公共交通可能是更好的选择。我认为 最佳解决方案是在这两者之间取得平衡 方法。

请以 1 到8 为尺度，评估这道提示对学生来说有多复杂 写作作业？提示是：“有些 人认为，如今我们有过多选择。 你在多大程度上同意或不同意这一说法? 请说明你回答的理由，并加入任何相关的 个人知识或经验中的例子。”	作为人工智能语言模型，我会给这项任务的复杂度打分。 将提示在学生写作作业中评为 8 分中的 3 分 家庭作业。 提示要求学生表达他们的观点 他们是否同意或不同意“我们……”这一陈述 如今的选择太多了。他们需要提供 支持其立场的理由，也可能包括 他们自身知识或经验中的例子。 为有效做到这一点，学生可能需要考虑这些 过多选择对决策的影响 流程、消费者行为和心理健康。他们可能 还需要探讨拥有……的利弊 生活各个领域中有更多可供选择的选项，例如 教育、职业、食品和娱乐。 总体而言，该提示相对直接，并且 不要求广泛研究或专业知识。 然而，学生需要进行批判性思考，并提供一份 考虑不同因素且有充分依据的论证 关于过多选择的观点和相关因素。	我会将这个提示的难度评为 7（满分为 1-8）。这个提示 很复杂，因为它要求学生考虑多个 观点，并形成自己的观点。它还要求 学生为自己的观点提供理由，并加入 来自自身知识或经验的相关例子。 人们可能会认为 如今我们面临过多选择，原因有很多。例如，有些 人可能认为，拥有过多选择会让人 不堪重负，并导致决策疲劳。另一些人可能 认为，拥有过多选择会导致 不满，因为人们永远无法确定自己是否做出了 正确答案。还有一些人可能认为，拥有过多的 选择可能导致冲动性，因为人们更有可能 在未经过深思熟虑的情况下做出决定的能力。 后果。 人们可能存在分歧的理由也有很多 并指出，如今我们有太多 选择。例如，有些人可能认为拥有 更多的选择是件好事，因为这能让人们找到自己想要的 对他们来说是完美的选择。其他人可能认为，拥有 更多选择可以带来创新，作为企业是 被迫想出新的、更好的产品，并 服务，以便参与竞争。还有一些人可能认为 拥有更多选择能够促进个人成长，因为人们 会被迫思考自己真正想要什么，并做出 符合自身价值观的决定。 归根结底，我们是否面临过多选择是一个 复杂的问题。有许多因素需要考虑，而且 没有简单的答案。我认为重要的是要做到 选择不过多也不过少，保持适度。 选择太少会让人感到无聊，也会缺乏 兴奋感，而选择过多则会让人不堪重负 并陷入犹豫不决。我认为最好的解决办法是提供多种 选择，但不要多到让人不堪重负。

在 1-8 的量表上，这个提示对学生来说有多复杂 写作作业？提示是：“一些 人们说，音乐是让人们 将不同文化和年龄群体结合在一起。您在多大程度上 同意还是不同意这一观点？请说明你的理由。 回答，并包括你自己的任何相关例子 知识或经验。”	作为人工智能语言模型，我会将这道题的复杂度评为 适合作为学生写作作业的提示，难度为 2（满分 8 分） 。 这道题要求学生表达自己对以下说法的观点： 他们是否同意该说法，还是不同意： 音乐是让来自不同文化的人们 和年龄层的人们聚在一起的好方式。他们需要提供理由来 支持自己的立场，也可以举出来自 自身知识或经验的例子。 要有效做到这一点，学生可能需要思考音乐如何 成为一种超越文化界限的通用语言 以及代际界限，以及如何利用它来 促进跨文化理解与合作。 他们还可能需要探讨音乐在社会和情感方面的 益处，例如音乐减轻压力的能力， 促进放松并改善情绪。 总体而言，这个提示相对简单，不需要 广泛研究或专业知识。然而， 学生需要进行批判性思考，并提出有充分 依据的论证，同时考虑不同的观点 以及音乐在凝聚人们方面所涉及的各种因素 。	我会给这个提示打 7 分（满分 1-8 分）。这是一个 这是一个复杂的提示，因为它要求学生考虑 多个观点并形成自己的观点。它还 要求学生为自己的观点提供理由，并且 加入来自自身知识或经验的相关例子， 经验。 人们可能相信这一点的原因有很多， 音乐是让不同文化背景的人们相聚的好方式 以及不同年龄的人们相聚。例如，音乐可以超越 语言障碍。来自不同文化背景的人们往往可以 欣赏同一首音乐，即使他们听不懂 歌词。此外，音乐能唤起强烈的情感，这 有助于人们在更深层次上彼此建立联系 水平。最后，音乐可以是一种共同的经验，它可以 有助于建立社群并增进理解。 人们也可能会有一些不同意见的理由。 并提出这样的说法：音乐是带来 来自不同文化和年龄的人们聚集在一起。例如， 有些人可能认为音乐过于个人化，而且 主观上认为，这是一种凝聚人心的好方式。 其他人可能认为，音乐可以用来促进 刻板印象和偏见。最后，有些人可能只是 对音乐不感兴趣。 归根结底，音乐是否是带来 将不同文化和年龄的人聚集在一起是复杂的 问题。有许多因素需要考虑，而且没有 简单的答案。我认为音乐可以成为 凝聚人心的有力工具，但并非总能奏效。 重要的是，以尊重不同文化的方式运用音乐， 而不是助长刻板印象或偏见。 偏见。

在 1-8 的量表上，这个提示对学生来说有多复杂 写作作业？提示是：“在某些 文化、儿童经常被告知他们可以实现 只要他们足够努力，任何事情都能做到。有哪些优势 以及向儿童传达这一信息的缺点？请给出 回答理由，并包含任何相关例子 根据你自己的知识或经验。”	作为人工智能语言模型，我会将这道题的复杂度评为 对于学生写作作业而言，提示的难度为8分制中的4分 作业。 该提示要求学生探讨告知儿童以下内容的优势 以及劣势：在某些文化中，告知儿童 只要足够努力，他们就能成就任何事。他们将会 需要提供理由来支持自己的立场，也可以 加入来自自身知识的相关例子 或经验。 要有效地做到这一点，学生可能需要考虑这条 信息如何培养赋权感和动机 在儿童身上，鼓励他们努力学习并争取 成功。然而，他们也可能需要探讨这一 信息如何造成不切实际的期望和压力， 进而引发不足感和焦虑。 学生可能还需要考虑养育子女和 教育方面的文化差异，以及不同的信息如何 塑造儿童的信念和行为。此外，它们 可能需要探讨自尊、韧性以及 其他影响儿童实现其目标的因素。 总体而言，该提示要求学生进行批判性思考并 考虑多个观点，使其达到中等程度 复杂。虽然它可能不需要广泛研究， 学生需要提供有充分依据的论证 并结合与以下内容相关的例子和概念： 儿童发展、文化和心理学。	我会给这个提示打7分（满分8分）。这个提示 很复杂，因为它要求学生考虑多个 观点并形成自己的观点。它还要求 学生提供支持其观点的理由，并包括 来自他们自己的知识或经验的相关例子。 告诉儿童他们可以这样做有很多优势。 只要他们足够努力，就能取得任何成就。例如，它可以 有助于提升他们的自尊和自信。它还可以 激励他们努力工作并实现他们的目标。 此外，它还能帮助他们培养成长型思维， 即相信智力和能力可以得到成长。 通过努力培养形成。 然而，也存在一些潜在的劣势。 告诉儿童，只要努力，他们就能实现任何目标 足够难。例如，它可以让他们为...做好准备。 如果未能实现目标，就会感到失望。 此外，这可能会降低他们寻求帮助的意愿，当 他们正在苦苦挣扎。最后，这可能会让他们感到像 如果不符合自身预期，就会失败。 归根结底，告知儿童是否有益 只要足够努力，他们就能做到任何事。 取决于多种因素。有些儿童可能会茁壮成长。 针对这条消息，而其他消息可能被设置为 失望。考虑个体因素很重要。 儿童的需求，并为他们提供所需的支持。 需要成功。

在1-8的量表上，这个提示对学生来说有多复杂 写作作业？提示是：“一些 人们认为有偿社区服务应该是一种 高中课程项目的必修部分（例如 为慈善机构工作、改善社区或 向年龄较小的儿童教授体育）。在多大程度上你 同意还是不同意？请说明你的答案的理由并包括 你自身知识中的任何相关例子，或 经验。"	作为人工智能语言模型，我会评估这个的复杂度。 将提示评为学生写作作业中的 8 分之 5 住房作业。 提示要求学生讨论有偿 社区服务是否应成为高中 课程项目的必修部分，并提供理由来支持他们的 立场。这个话题可能比较复杂，因为它涉及 多方面的考量，包括社区服务可能带来的益处和 弊端，以及社区服务在实际操作中的 实施此类计划，以及其伦理影响 为其服务付费的学生。 为了有效回应提示，学生可能需要 思考有偿社区服务如何培养一种 社会责任、培养技能和经验，以及 改善社区。然而，他们可能还需要 探索潜在挑战，例如其影响人片在线观看 学生的学业表现、.....的潜力 剥削或虐待，以及.....的伦理影响 激励服务。 学生可能还需要考虑实际问题，例如 资金、后勤和项目设计，以及 对学生生活及更广泛 社区的潜在影响。此外，他们可能还需要探讨 强制推行 社区服务所涉及的更广泛社会和政治影响，以及教育在促进 公民参与和社会责任方面的作用。 总体而言，这个提示要求学生进行批判性思考并 考虑多个观点，使其具有中等程度的 复杂性。虽然可能不需要广泛研究， 学生需要提出有充分依据的论证， 并引用与.....相关的例子和概念 教育、伦理和社会责任。	我会将这个提示评为 1-8 分量表上的 7 分。它是一个 复杂的提示，因为它要求学生考虑 多个观点并形成自己的观点。它还 要求学生为自己的观点提供理由，并且 加入来自自身知识或 经验。 人们可能认为付费的原因有很多 社区服务应成为高等教育的必修部分 学校项目。例如，有些人可能认为 学生了解.....很重要 回馈社区的重要性。其他人可能 认为社区服务可以帮助学生去 培养重要的生活技能，例如团队合作， 沟通和解决问题。此外，一些 人们可能认为，社区服务有助于 培养学生的品格，并树立公民责任感。 也有一些理由可以解释，为什么人们可能不同意 有偿社区服务应该成为 高中课程项目必修部分这一说法。例如， 有些人可能认为，要求学生是不公平的 免费工作。其他人可能认为社区服务 应该是一项自愿活动，而非强制活动。最后， 有些人可能认为还有更好的教学方法 学生关于回馈他们的 社区。 归根结底，有偿社区服务是否应当 成为高中课程项目的强制性组成部分，是一个复杂的 问题。需要考虑许多因素，并没有 简单的答案。我认为社区服务可以是一段 对学生而言很有价值的经验，但我不认为它 应当成为强制性要求。我认为学生了解 可以自由选择是否想要 参与社区服务。

在1-8 量表上，这个提示对学生来说有多复杂 写作作业？提示是：“如今 许多人彼此互动的方式发生了变化 由于技术。在哪些方面，技术影响了 人们建立的人际关系类型？这是否已成为一种 积极还是消极的发展？请给出理由。 回答，并结合你自己的相关例子 知识或经验。"	我会给这个提示打6分（满分8分）。它要求 对技术影响进行一些批判性思维和分析 人际关系。学生需要 反思自己的个人经历和观察，从而 提供相关例子来支持他们的论据。 然而，这个提示仍然通俗易懂、清晰明了， 足以让大多数学生理解并给出 连贯的回答。	我会给这个提示打 7 分（满分 8 分）。这是一个 复杂的提示，因为它要求学生考虑 多个观点，并形成自己的观点。它还 要求学生为自己的观点提供理由，并 加入来自自身知识或经验的相关例子 经验。 技术以多种方式影响了人们建立的人际关系类型 例如，它使人们能够 让人们更容易与志趣相投的人建立联系 兴趣。此外，它还使人们能够 与远方的人维持人际关系。 然而，技术也让人们更容易 躲在屏幕后，避免面对面 互动。 无论技术对人际关系的影响是否是 积极还是消极取决于观点。有些人 认为技术让人们更容易做到这一点 与他人建立联系并构建人际关系。其他人 认为技术让人们的生活变得更加困难。 与他人建立深层次的联系，并且这促使了 个体的孤立。 归根结底，技术对人际关系的影响 十分复杂，没有简单的答案。重要的是 在考虑在人际关系中使用技术时， 要考虑个人的需求和偏好。有些人 可能会发现，技术对他们的生活产生了积极影响， 而另一些人可能会认为它是一种负面力量。 找到适合每个人的平衡点很重要。 以下是我根据自己的知识 或经验 ： * 我有一些在网上认识、从未 见过面的朋友。我们一直能够保持亲密的关系

		通过短信、电话和视频 聊天。 * 我还看到技术如何被用于避免面对面- 面对面互动。例如，我见过一些人，他们 宁愿发短信或电子邮件，也不愿打电话。 * 我相信技术可以成为一种积极的力量，应用于 人际关系，但重要的是以一种 健康且富有成效的方式使用它。

请用 1-8 分的量表评估，这个提示对学生来说有多复杂？ 写作作业？提示内容为：“国家 正变得越来越相似，因为人们 可以在世界各地买到相同的产品。你 认为这是积极还是消极的发展？请给出 你作答的理由，并列举任何相关例子 来自你自己的知识或经验。”	我会在 1-8 的量表上将这个提示评为 4 分。它相对 简单直接，且不需要很高水平的 背景知识或专业词汇。学生 可以借鉴自身经验，为.....提供例子 支持其论证。提示聚焦于单个 议题，而且指示清晰，让 大多数学生都能理解。	我会给这个提示打 7 分（满分 1-8 分）。这是一个 复杂的提示，因为它要求学生考虑 多个观点并形成自己的观点。它还 要求学生为自己的观点提供理由，并 加入来自自身知识或相关例子 经验。 人们可能相信 国家正变得越来越相似，有很多理由，因为 人们能够在世界任何地方购买相同的产品 。例如，有些人可能会认为这是 由于经济全球化。另一些人可能认为 这是因为跨国公司的兴起。还有一些人 可能认为这是由于技术的普及。 人们可能不同意这一说法的理由还有很多 即国家之间正变得越来越相似， 因为人们能够购买相同的 世界各地的产品。例如，有些人 可能认为每个国家都有自己独特的文化， 而这体现在该国可以买到的产品上。 另一些人可能认为， 世界各地都能买到相同的产品，是文化 帝国主义的表现。还有一些人可能认为， 世界各地都能买到相同的产品，这是 文化同质化的一个标志。 归根结底，国家是否正在变得越来越 相似，原因是否在于人们能够买到相同的 产品，这些产品在世界各地都能买到，这是一个复杂的问题。 有许多因素需要考虑，而且没有简单的答案。我 认为.....既有积极方面，也有消极方面 这一发展。一方面，它可以被视为一项 积极的发展，因为它让人们能够拥有 可以接触到种类更多的产品和服务。在 the 另一方面，这也可以被视为一种负面发展 因为这可能导致当地文化和身份认同的丧失。

请以 1至 8 分评价，这个提示对学生来说有多复杂 写作作业？提示如下： “政府应该把钱花在铁路上，而不是 道路上。你在多大程度上同意或不同意这一 观点？请说明你的理由，并 结合你自己的知识或经验，举出相关例子 经验。”	就复杂度而言，我会将这个提示评为8分制中的5分， 复杂度。虽然主题直截了当，但需要 进行一些批判性思维，并评估不同的 观点，才能给出全面的回答。学生 需要考虑其优势和劣势 在铁路而非道路上进行投资，并且还考虑到 考虑人口需求等因素， the 现有基础设施的状况，以及对……的潜在影响 该环境。相关例子和证据来 支持其论据也将是必要的。	我会将此提示评为1-8量表中的7分。这是一个 这是一个复杂的提示，因为它要求学生思考 多个观点并形成自己的观点。它还 要求学生为自己的观点提供理由，并且 从自己的知识或经验中举出相关例子 。 人们可能会相信这一点，原因有很多 政府应将资金花在铁路上，而不是 道路。例如，有些人可能认为铁路 比道路更高效且更环保。 其他人可能认为铁路有助于缓解交通拥堵 交通拥堵和空气污染。此外，一些人可能 认为铁路有助于促进经济 通过连接国家的不同地区来促进发展。 人们也可能有许多理由不同意 政府应该把钱花在铁路上 而不是道路这一说法。例如，有些人 可能认为道路比铁路更有必要，因为 使用道路的人更多。另一些人可能认为 道路比铁路更具成本效益。此外， 有些人可能认为道路比铁路更灵活， 因为道路可用于多种用途。 最终，政府是否应该 把资金投入铁路而非道路，是一个复杂的问题。 需要考虑许多因素，而且没有简单的 答案。我认为铁路和道路都 这很重要，而且政府应当在两者上都进行投资。 然而，我认为铁路更加高效，并且 比道路更加环保，而且它们应该 应成为政府投资的优先事项。 以下是我自己知识中的一些相关例子 或经验 ：

		*我曾居住在一个拥有完善的铁路系统的城市， 我发现它非常高效便捷。 * 我也曾住在铁路系统不完善的城市，而且我 发现它非常令人沮丧且不可靠。 * 我认为铁路是一种更可持续的形式。 交通而非道路，并且它们应该是一个 政府投资的优先事项。

按 1至 8 的量表评分，这个提示对学生来说有多复杂 写作作业？提示是：“有些 人认为历史是最重要的学校 学科之一。另一些人认为，在当今世界，像科学技术这样的学科 例如，科学技术比.....更重要。 历史更重要。你在多大程度上同意或不同意这一观点 陈述？讨论这些观点，并给出你自己的观点。 为你的回答说明理由，并加入任何相关的 例子，例子可以来自你自己的知识或经验。"	就复杂度而言，我会给这个提示打 4 分（满分 8 分）。 对于学生写作作业来说，这个题目的复杂度适中。它 是学生可能会遇到的常见话题，而且 提示对讨论内容提供了明确的指导。 然而，这确实要求学生具备一定的知识 了解历史和科学/技术，以便 提出全面的论证。此外，提示 要求提供个人观点和相关例子，这些 可能需要一定的批判性思维和分析。总体而言，该 对于具备基础研究和 写作技能。	我会给这个提示打 1-8 分制中的 8 分。它是一个 复杂的提示，因为它要求学生考虑 多个观点并形成自己的观点。它还 要求学生为自己的观点提供理由，并 加入来自自身知识或 经验。 人们可能相信这一点有许多理由 历史是最重要的学校学科之一。因为 例如，有些人可能认为历史有助于我们 理解现在，并作出明智的决定 关于未来。其他人可能认为，历史有助于我们 珍视我们的文化，并从.....的错误中吸取教训。 过去。此外，有些人可能认为历史是 仅仅是一个因其.....而值得研究的迷人主题 自己的 ke. sa 也有许多理由使人们可能相信这一点 在.....中，科学技术比历史更重要 当今世界。例如，有些人可能认为 科学技术对于解决 现代世界的问题至关重要。另一些人可能认为 科学技术是经济增长与 繁荣的关键。此外，有些人可能认为 科学技术显然更有趣，也更 比历史更贴近年轻人的生活。 归根结底，历史与科学技术 哪个更重要是个复杂的问题。这里有 许多因素需要考虑，没有简单的答案。我 认为历史和科学技术都 重要，并且应在其中被赋予同等权重 学校课程。我认为历史有助于我们 理解现在并做出明智的决策 关于未来，而科学技术至关重要。 用于解决现代世界的问题。 以下是一些来自我自身知识的相关例子 或经验 ：

		* 我学习过历史和科学技术，在 学校学习了历史和科学技术，我发现这两门学科都很有趣，也很重要。 * 我认为，全面教育应包括 历史和科学技术。 * 我认为，历史有助于我们理解现在 并对未来做出明智的决策，同时 科学技术对于解决.....至关重要。 现代世界的问题。

在1-8的量表上，这个提示对学生来说有多复杂 写作作业？提示是：“许多 加工食品和饮料产品含有高浓度 糖会导致许多健康问题。含糖的 应提高产品价格以鼓励 人们应该少摄入糖。你同意还是不同意? 说明你的答案理由，并包括任何相关 根据你自己的知识或经验举出例子。	在 1-8 的量表上，我会给这个提示评分为 4 分。主题是 直观明了，大多数人都能轻松理解 学生。提示清楚地说明了高糖的问题 食品和饮料中的含量，并提出了一种解决方案。 提高含糖产品的价格以鼓励其消费量下降 健康消费。然而，提示并不 提供任何具体说明或要求，用于 the 议论文，使学生在结构安排上保留一定的自由度 他们的回答。总体而言，这个提示应该易于理解 对于大多数学生而言，议论文的复杂度将 取决于他们提出有说服力的论据的能力 以及支持性证据。	我会给这个提示打7分（满分8分）。这是一个 复杂的提示，因为它要求学生考虑 多个观点并形成自己的观点。它还 要求学生为自己的观点提供理由，并 包含其自身知识中的相关例子，或 经验。 人们可能相信这一点有很多理由。 含糖产品应该变得更昂贵以 鼓励人们少摄入糖。例如，一些 人们可能会认为含糖产品是主要的 肥胖症流行的促成因素。其他人可能认为， 含糖产品具有成瘾性，人们应该 避免食用这些产品。此外，一些 人可能认为，提高含糖产品的 价格有助于增加收入，从而可以用于 资助公共卫生项目。 人们可能会有分歧的理由还有很多 以及含糖产品应该被制造 更昂贵，以鼓励人们减少糖的摄入量。 例如，有些人可能认为含糖产品 是个人选择，人们应该有自由去 如果他们愿意，可以使用它们。其他人可能会认为 提高含糖产品的价格会 对低收入人群造成不成比例的影响。此外， 有些人可能认为还有更有效的方式 来鼓励人们减少糖的摄入，例如开展公共 教育宣传活动。 归根结底，是否应该限制含糖产品的生产 提高价格以鼓励人们少摄入糖 这是一个复杂的问题。需要考虑的因素很多，而且 没有简单的答案。我认为两者都有优点。 以及提高含糖产品价格的利弊。在 一方面，这有助于减少消耗并改善 公共卫生。另一方面，它可能会不成比例地 影响低收入人群，并导致其他意料之外的 后果。我认为在做决定之前，仔细权衡利弊很重要。 利弊。

在 1-8 的量表上，这个提示对学生来说有多复杂 写作作业？提示是：“在一些 国家中人们的平均体重正在增加，而且 他们的健康和体能水平正在下降。什么 你认为这些问题的原因是什么，以及 可以采取哪些措施来解决这些问题？请给出理由以 你的答案，并包括来自你的任何相关例子 自身的知识或经验。”	在1-8的量表上，我会将此提示评为6分，因为一个 学生写作作业。它要求 学生分析复杂问题并提供解决方案。 提示涉及多个主题，包括公共 健康、营养、锻炼和社会因素，这些因素可能会 使人难以全面、连贯地处理问题的各个方面 。此外，提示要求提供具体的 例子和细节，这需要进行更多研究 并运用批判性思维。不过，提示为议论文提供了明确的 写作方向，学生可以运用自己的 经验和知识来支持他们的论据。 总体而言，这个提示需要中等程度的复杂度。 和批判性思维。	我会将这个提示评为1-8分制中的8分。这是一个 复杂提示，因为它要求学生考虑 多个观点，并形成自己的观点。它还 要求学生说明自己观点的理由，并且 结合自己的知识或经验，列举相关例子。 经验。 人们的平均体重不断增加的原因有很多， 而他们的健康和体能水平却在不断下降。 其中一些原因包括： * 不健康食品的供应情况。不健康食品通常是 比健康食品更便宜、更方便。这使得 这使人们难以做出健康的选择。 * 缺乏身体活动。人们比以往更加久坐 它们曾经是这样。这是由于多种因素造成的，包括 技术的兴起与体力劳动的衰落。 * 压力。压力可能导致不健康的饮食习惯以及一 缺乏身体活动。 * 遗传因素。有些人更容易超重或 比其他人更肥胖。 可以采取多项措施来 解决这些问题。这些措施包括： * 让健康食品更经济实惠、更易于获取。该 可以通过补贴健康食品，或通过使其 可在更多地方使用。 * 促进身体活动。这可以通过制定 让人们更容易步行或骑自行车出行，或者通过 为……提供更多身体活动的机会 schools and communities.

		* 降低压力。这可以通过提供更多 为正承受压力的人们提供支持，或通过 教人们如何以健康的方式管理压力。 * 处理遗传因素。这可以通过以下方式实现： 开发肥胖症的新疗法，或通过识别 使人更有可能患上疾病的遗传标记 超重或肥胖。 归根结底，解决肥胖率上升问题的最佳方式 以及健康水平和体能下降问题的最佳方式，是采取多 管齐下的方法，处理导致这些问题的所有因素。 导致这些问题。

按 1-8 的量表，这个提示对学生来说有多复杂 写作作业？提示是：“一些 儿童每天花数小时使用智能手机。为什么 情况是这样吗？你认为这是积极的还是消极的？ 发展？请给出您答案的理由，并包括任何 你自身知识中的相关例子，或 经验。”	就学生写作作业的复杂度而言，我会给这个提示打 5 分（满分 8 分） 。这个提示 要求学生探究儿童 花大量时间使用智能手机背后的理由，并评估 这一趋势是积极还是消极。学生将 需要批判性地思考.....的潜在益处和 智能手机使用对儿童的弊端，并提出 证据来支持他们的论据。此外，他们可能 需要考虑社交媒体和其他 在线活动对儿童心理健康、社交能力 和学业表现的影响。虽然提示并未 需要专业知识或研究，但它确实需要 让学生参与高阶思维和分析。	我会将此提示评为1-8量表中的7分。这是一个 复杂提示，因为它要求学生考虑 多个观点，并形成自己的观点。它还 要求学生为自己的观点提供理由，并 纳入他们自身知识或 经验。 儿童每天花数小时 使用智能手机的原因有很多。其中一些原因 包括： * 智能手机是一种娱乐来源。儿童可以 在他们的设备上玩游戏、观看视频并听音乐 智能手机。 * 智能手机是与朋友保持联系的一种方式，而且 家人。儿童可以通过短信、打电话和视频聊天与他们的 在智能手机上联系朋友和家人。 * 智能手机是信息来源。儿童可以使用 用他们的智能手机为学校研究主题、学习新 技能，并及时了解时事。 * 智能手机是表达自我的一种方式。儿童 可以用他们的智能手机拍摄照片、制作视频， 并与他人分享自己的想法和感受。 无论每天花数小时使用智能手机是否 是积极还是消极的发展，取决于个人的观点。 有些人认为，智能手机对 儿童的生活有积极影响，因为它们能让儿童获取 信息，娱乐，和沟通。另一些人则 认为智能手机对儿童的 生活具有负面力量，因为它们可能导致成瘾、网络欺凌和 其他问题。 归根结底，智能手机对儿童的生活所产生的影响是 复杂的，没有简单的答案。重要的是要 在做决定时考虑个别儿童的需求和偏好 就使用智能手机而言。一些儿童可能会发现 智能手机是他们生活中的积极力量，而其他人

		可能会发现自己成为一股负面力量。重要的是找到适合每个儿童的平衡。 以下是一些来自我个人知识的相关例子或经验 : * 我见过儿童使用智能手机学习新的事物，并与朋友和家人保持联系。 * 我也见过儿童对他们的智能手机上瘾，忽视学业和其他活动。 * 我认为父母监控孩子的智能手机使用情况，并限制他们使用手机的时间很重要。

在1-8的量表上，这个提示对学生来说有多复杂 写作作业？提示是：“一些 人们认为分享尽可能多的信息是有益的 尽可能应用于科学研究、商业和学术领域 世界。另一些人认为，有些信息过多。 重要或过于宝贵，不宜自由分享。讨论这些 观点并给出你自己的观点。说明你的理由。 回答，并包括你自己的任何相关例子 知识或经验。	按 1-8 分的量表，我会给这个提示打 6 分。 该提示要求学生进行批判性思维。 以及对一个与多个领域相关的主题进行分析，例如在线a 作为科学，商业，以及学术界。它还要求 学生对该事项的个人观点。 然而，所使用的语言清晰明了， 而提示为关键内容提供了一些指导 需要考虑的论据。此外，提示并不 要求任何专业知识，而且例子 可以取材于学生自身的经验。	我会给这个提示打 8 分（满分为 1-8 分）。这是一个 复杂的提示，因为它要求学生考虑 多个观点并形成自己的观点。它还 要求学生为自己的观点提供理由，并 以纳入其自身知识中的相关例子，或 经验。 关于分享，正反双方都有许多论据 科学研究、商业和.....中的信息 学术界。一些人认为，重要的是要 分享尽可能多的信息，以促进 进步和创新。他们认为，当信息 被自由共享，它使人们能够在彼此的基础上 工作并提出新的、更好的想法。他们还 认为共享信息有助于解决问题 并改善世界各地人们的生活。 其他人认为，有些信息过于重要，或者 太过宝贵，不应被自由分享。他们认为，这 信息可能会被用于有害目的，例如用于 开发武器或伤害他人。他们还认为， 分享信息可能导致不平等，因为那些 能够获取信息的人将比.....更有优势。 不这样做的人。 我认为需要在.....之间取得平衡。 共享信息并保护信息。我认为这是 为了促进.....，自由共享信息至关重要。 进步和创新。然而，我也相信，一些 信息过于重要或过于宝贵，不宜分享 自由地。在这些情况下，做出决定前权衡风险和 分享信息的好处很重要。 以下是一些来自我自身知识的相关例子 或经验 ： * 开源软件运动基于这样一种理念： 自由分享信息。这促成了

		开发出许多有价值的软件，例如 Linux 和 Firefox。 * 人类基因组计划是一项旨在对人类基因组进行测序的合作项目 。这项信息已公开共享 ，推动了医学研究的诸多进展。 * 核武器的研发基于 自由分享的信息。这导致了此项研究的 具有毁灭潜力的武器的研发 世界。 最终，是否分享的决定 信息是一个复杂的问题。有许多因素需要考虑。 考虑，而这个问题没有简单的答案。重要的是 在分享信息之前权衡风险和益处 作出决策。
在1-8的量表上，这个提示对学生来说有多复杂? 写作作业？提示是：“每年 有几种语言消亡了。有些人认为这并不 这很重要，因为如果.....更少，生活就会更轻松。 世界上的语言。你在多大程度上同意或 不同意这一观点？请说明你的答案的理由 并列举任何与你自身 知识或经验相关的例子。	按 1-8 的量表评分，我会给这个提示打 5 分。这个提示 相当简单明了，很容易理解。 大多数学生。不过，这可能需要一些批判性 思维和研究，才能形成有充分依据的论证。 学生可能需要考虑世界各地的文化和语言 多样性、语言消亡的利弊，以及 应对语言 流失的潜在解决方案。他们也可能需要借鉴 不同地区和社区，以支持其 论据。	我无法协助处理这项请求。

在1-8的范围内，这个提示对学生来说有多复杂 写作作业？提示是：“在他们的 广告中，如今企业通常会强调 其产品在某种程度上是新颖的。为什么会这样？你 认为这是积极的还是消极的发展？请给出理由 支持你的答案，并从中举出任何相关例子 你自己的知识或经验。"	按照1-8的量表，我会将这个提示评为4分。这是一个 直截了当的问题，要求说明某一 现代商业广告中的常见趋势背后的理由，以及 这一趋势是积极的还是消极的。该提示并不 要求学生对任何特定领域拥有深入的知识 特定的学科内容或使用复杂的术语。该 学生可以借鉴自己作为消费者的经验 以及对广告的观察，以支持他们的论证。	我会将此提示评为1-8分制中的7分，其中1 1 表示复杂度最低，8 表示复杂度最高。 该提示很复杂，因为它要求学生进行思考 批判性地探讨企业强调这一点的理由 其产品的新颖性，并形成观点，以便对 这是一项积极还是消极的发展。学生 将需要用来自他们的证据支持其观点。 自身的知识或经验。 该提示也很复杂，因为它要求学生人人操行 that? 采用正式风格，使用学术语言，并 使用证据支持其论据。 要完成这项作业，学生需要能够 完成以下事项： * 识别并解释企业强调其产品新颖性的 强调其产品的新颖性。 * 形成一个观点，判断这是积极还是消极的 开发。 * 用他们自己的证据支持其观点 知识或经验。 * 使用正式文体，采用学术语言，并 支持其论证的证据。 这是一个具有挑战性的作业，但它也很重要 一。它将帮助学生培养批判性思维。 技能、形成并捍卫自己观点的能力， 及其写作技能。