

人工智能、人智协作模式的创造力比较研究

——基于广告文案创作的双盲实验

孙铭欣 刘 荣 李育辉

摘要:近年来,人工智能展现出的创造力正在与人类接近,并发展出思维链、指定思维链、使用参考案例、人工优选等人机协作模式。进行人类创造力与人工智能、人智协作展现出的创造力的比较研究具有重要的理论和现实意义。采用实验研究法,邀请 117 位专家对人类独立创作、AI 独立创作、人智协同创作的广告文案进行双盲实验,并与 2023 年的人工智能创造力研究进行对比分析,研究表明:2024 年生成式人工智能的“等同工作年限”从 2023 年的 2.47 年提升至 3.08 年,且上限达 4.9 年,人工智能创造力的平均值基准和离散程度显著提升;多种人智协作方式创作的文案存在明显的定势效应,创造力平均值未进一步提升,这提示人智协作需要更为组织化的创新模式,以避免陷入平庸化的困境。

关键词:生成式人工智能;人智协作;广告文案;创造力;对比研究

中图分类号:F713.8 **文献标志码:**A **文章编号:**2096-5443(2025)04-0055-13

基金项目:国家自然科学基金面上项目(72371242)

一、引言

近年来,人工智能(AI)成为实际应用中发展极快的科学技术之一^[1],且被世界各国认为有潜力成为经济增长和社会进步的主要驱动力^[2-3],且人工智能对企业人力资本策略的影响备受关注。既有研究指出,越来越多“暴露”在人工智能下的工作活动和职业将会被取代,以公开性信息为特征的新闻业、广告业首当其冲面临着“技术性失业”的巨大压力。创造力作为人类难以被替代的核心能力,在行业普遍拥抱人工智能技术的当下,形成了两个关键问题:一是人类与人工智能在创造力上孰高孰低^[4-7];二是人类与人工智能的协作(Human-Intelligence Collaboration,简称“人智协作”)下的创造力提升路径如何变化^[7-8]。随着人工智能能力的不断进化与提高,其对社会产生巨大影响并在众多领域发挥作用^[9],根据中国信息通信研究院最新发布的《人工智能发展报告(2024年)》,全球人工智能产业规模预计将达到 6233 亿美元,预期同比增长 21.5%^[10]。可以预见,人工智能的应用场景会越来越普遍,人类与人工智能的协作将成为重要的生产方式,同时重构我们的工作与社会生活。

在这样的背景下,持续性地评估人工智能与人类创造力的差异显得尤为关键。这种评估不仅有助于我们把握技术适应性中的动态变化,还能指导我们如何顺应这些变化,以优化人工智能在创造力领域的辅助作用,并识别潜在风险。从组织管理的视角来看,这种评估所产生的数据洞察,能够为企业的人力资源配置提供科学的决策参考,特别是在创意人才梯队建设与智能工具配置方面具有显著的指导意义。

在探讨人工智能与人类创造力差异的过程中,人工智能创造力的离散程度成为一个不容忽视的关键变量。传统研究往往集中于创造力的集中程度,也就是关于基线的研究,但随着人工智能技术的飞速发展,离散程度的重要性日益凸显。正如 Dell'Acqua 等人的研究所示,人工智能在提升创造力

基线方面对能力较弱者的帮助更为显著,而离散程度的增加则是推动创造性达到新高度的关键^[11]。因此,在广告领域,创造力的离散程度不仅影响个体和团队的创新能力,也是衡量人工智能技术对广告创作影响的重要维度。

本研究以李育辉等学者2023年的研究为前期研究^[6],其以“广告文案”作为创造力的承载场景,用双盲实验法对2023年的人工智能与人类的创造力差异进行比较,量化二者的创造力与结构差异。同时,研究考虑到工作年限作为工资水平及人力资本的重要影响因素,构建“等同工作年限”这一概念,以此衡量人工智能作为人力资本的替代性要素,得到2023年生成式人工智能创造力正在接近人类的重要结论。

本次研究延续并拓展了前期研究的设计思路,并与之对比分析。2024年以来人工智能大模型在持续地发展,尤其是发展出思维链、指定思维链、参考案例、人工优选等多种人智协作方式,重塑各类创作的边界与格局,也引发新的对创造力水平的疑问:处在技术快速迭代过程中的人类、人工智能、不同人智协作方式下的创造力水平是怎样的?人工智能与人类创造力水平的贴近程度如何?人工智能的创造力是否存在思维定式,是否存在灵感闪现?这些问题构成研究的主要内容与探索方向。

二、文献综述与问题假设

创造力的概念是复杂且有争议的。传统的创造力(creativity)被定义为原创性(original)和有效性(effective)^[12]。一些研究者认为人工智能创造力与人类的真实创造力有着本质的不同,从原创性过程角度认为人类的创造力源于认知过程、情感、经验和意图的复杂相互作用,它通常由内在动机、自我表达的欲望或解决问题的需要所驱动。与之相反,人工智能的创造力是由算法和优化函数所驱动,虽然它所生成的内容往往被称作“创造性”的产出,但它缺乏人类创造性所特有的主观经验和有意识的倾向性^[13]。因此,有研究指出人工智能只具有表面创造力,这种缺乏深度、真实性或个人意义的创造力符合伪创造力(pseudo-creativity)的特征,并非人类真正的创造力^[14]。

与围绕原创性的讨论不同,许多实证研究从有效性出发,对比人类与人工智能在多个场景的创造力。有学者认为,生成式人工智能依托海量数据的深度训练与学习,已具备基本的理解力和综合创造力^[15-16],例如一群来自阿肯色大学的科学家将人类与AI在创造力方面进行了直接比较,认为当前的AI语言模型在发散思维任务上的创造性潜力超过了人类^[17]。也有学者在研究中指出,尽管AI在某些创造力测试中表现出色,但与人类相比,AI的创造性输出在质量上并无明显优势^[18-19],特别是在与当前AI聊天机器人的对比中,人类的表现更为出色^[20],但实践中ChatGPT等智能内容生产工具可替代人类完成某些工作^[21-22],使人类从繁重的重复性工作中解放出来更为直接地进行创新活动。那么,在生成式人工智能快速进化的当下,人工智能所体现出来的创造力与人类创造力之间的差异是什么?人智协作模式未来应如何创新?这一系列问题成为接下来要探究的核心。

(一)以广告文案为场景的创造力量化评估

创造力是人类最为重要和独特的能力成分,它不仅与个人和组织绩效密切相关,更是驱动新质生产力的核心因素^[14]。在广告领域,广告文案往往被看作是广告营销人员的核心创造力之一。在生成式人工智能技术的推动下,广告文案的创意潜能被极大地激发,使得创造力不再是少数人的专利。秦雪冰和郭博通过对智能广告文案进行测量,发现智能广告文案在消费者感知参与度和行动参与度上都优于传统广告文案^[13]。而在2023年的前期研究中,研究者通过采用双盲实验,招募专家和消费者来评价人工智能和人类在文案创作能力上的差异。专家和消费者被随机分配评价人类和AI在同等需求下创作的文案。结果发现,普通消费者已无法区分两者,熟悉文案生产过程的专家才有评价其创造力的能力,且具备较高的信度与效度^[6],所以后续创造力的测试需要坚持邀请相关领域的专家作为测试者。

本研究采用“等同工作年限”与“文案专业能力”作为评价的指标测量广告文案创造力,主要基于以下三方面考量:首先,这两个指标构建了“宏观—微观”的双层评估体系——在管理学和人力资源评估的语境下,“等同工作年限”指通过对比人工智能与人类在特定领域(如文案创作)的专业能力表现,推算出人工智能在创作质量、效率、创新性等方面与人类专业人员经过一定年限工作所积累的经验 and 能力相当的时间长度。这一概念旨在从整体效能角度量化人工智能与人类经验值的等效程度;而“文案专业能力”量表则通过清晰完整、容易理解、创造性、消费者洞察、商业洞察 5 个维度对创造力进行精细化解构。其次,二者具有显著的互补优势,“等同工作年限”通过时间维度提供直观的横向比较基准,“文案专业能力”则揭示创造力构成要素的质量差异,这种组合既能反映“量”的等效又能捕捉“质”的特征。最后,采用与李育辉等前期研究一致的测量工具可确保数据的纵向可比性,揭示人工智能文案创造力演进的新趋势。因此,这两个评价指标在本次研究中继续被使用,并与 2023 年的研究形成对比分析。

在文案专业能力量表中,“创造性”特指文案在形式表达上的新颖性(如措辞、修辞、结构创新),是评价文案专业能力的 5 个具体维度之一;而本研究的“创造力”既包含形式层面的创造性,也涵盖策略层面的创新价值(如消费者洞察的独特性、商业目标的实现度)。创造性这一维度是构成广告文案整体创造力的核心要素之一,但创造力还融合量表中其他维度的协同作用,并通过“等同工作年限”的等效经验值得以整体量化。

(二)人工智能的创造力发展及人智协作模式的创新

处在动态演化中的人工智能技术在不断提升自身的算法能力,多个不同的生成式人工智能模型在描述自身的角色及与人类的关系时,以人类“助手”、协作“伙伴”为主要表述。与此相适应的是研究脉络的转向,从前期人类与人工智能二元对立的比较研究转向近年来强调人机协作下的人智协同共创研究。具体来看,包含了以下主题:

1. 人类与人工智能在广告文案创造力上的比较

从李育辉等学者使用 ChatGPT3.5 开始研究以来,人工智能的能力快速发展,国外和中国的新一代生成式人工智能模型通过大规模数据训练,实现了文本、图像、音频、视频等多种内容的自动生成,为内容创作与创新提供了不断革新的工具与平台,如 ChatGPT 4.0 在复杂问题处理、语言理解能力、创意和扩展思维、对话管理以及准确性和细节方面均展现出更优秀的性能^[13]。在中国市场上,文心一言、讯飞星火等生成式人工智能模型也声称达到或超过 ChatGPT4.0 的能力水平。与 2023 年相比,人工智能创造力的提升水平及其具体维度的能力水平成为重点关切的问题。因此,提出相应研究假设:

H1a:2024 年人工智能模型的广告文案创造力显著超越 2023 年的人工智能模型

与此同时,人工智能与人类的创造力也反复被研究。较为常见的看法是,人类目前仍然占据创造力的最优水平,但在不同的细分维度上却不断被人工智能所追赶和超越^[23]。Lee 等通过实证研究探讨了 ChatGPT 对创造力的影响,发现 ChatGPT 能够通过提供新颖且连贯的创意来增强人类的创造力^[24];Koivisto 等比较了 256 名人类参与者和三个人工智能工具(ChatGPT3.5、ChatGPT4.0 和 Copy. Ai)在创造性任务中的差异^[20],研究显示人工智能工具在语义距离和创造力方面的得分高于人类平均值,但在这些任务中得分最高的人类仍然超过表现最好的人工智能工具;Stevenson 等的研究也有类似的结论,人工智能在创造性思维测试中的表现可能已经达到甚至超过了人类的平均水平,但顶尖的人类表现者取得比人工智能更好的结果^[25]。因此,我们大胆地提出以下假设:

H1b:2024 年人工智能模型的广告文案创造力显著超越人类

2. 不同人智协作方式与广告文案创造力的基准提升研究

不同人智协作方式与创造力的关系获得了较多关注^[26]。随着新技术的不断更迭,在经历早期的

提示词工程(Prompt Engineering)、思维链模式(Chain-Thought, CoT)和人工智能代理(AI Agent)之后,当前的人工智能可以实现与人类用户的自然和智能交互^[27]。人智协作是一种特定形式的人机协作形式,通过结合人类真实创造力与批判性思维的能力,以及人工智能在快速数据处理、模式识别及自动化任务等方面的优势,以实现创造力的提升。本研究中,我们将人智协作分为以下4种类型:

协作1:思维链(AI+CoT),是一种改进的提示技术,用以显著提高大型语言模型执行复杂推理的能力,是一种在足够大的语言模型中自然出现的简单方法。该协作模式指的是让人工智能自行使用思维链思考并回答问题,它的关键在于人工智能分解问题的能力,大量的对比研究发现思维链能够提升解题能力与推理能力^[28]。

协作2:指定思维链(AI+指定CoT),指给出人类设计思维链的思考逻辑,然后人工智能按照指定步骤回答问题,回答的质量由AI大模型能力和人类设计CoT的能力决定^[29]。

协作3:参考案例,指通过向人工智能提供优秀案例,帮助其更准确地理解需求和创作方向,并输出作品。

协作4:人工选优,指借助人工智能的生产力,让人工智能回答多次后再由人类参与,参与者根据自己的知识和经验进行优选,确保最终的创作结果符合自己的需求和标准。

以上4种模式存在不同的特点:“思维链”模式提示人工智能模型需要深度思考并具备更大的思维自主性,使其能从更多元的角度探索问题。“指定思维链”能够将人类专业观点与思维方法融入人工智能模型的创新过程中,这对于激发创新思维有着关键作用,但也可能限制了思维的发散性^[30]。“参考案例”模式可使内容创作更符合人类要求,但输出的结果可能会受既有案例框架的束缚^[31]。“人工优选”模式借助人的经验与智慧筛选出更具突破性的方向,但过度依赖参考框架会抑制创新概念的独特生成^[32]。在当前的研究中,以上4种人智协作模式被大量用于解决数学问题、识别与分类等问题的研究,且发现这些模式往往都能增强产出的质量^[33]。那么,将这些模式应用到广告文案领域,可能会进一步增强广告文案的创造力。因此,提出以下假设:

H2a:基于4种人智协作方式的文案创造力高于简单模式下的人工智能广告文案创造力

H2b:基于4种人智协作方式的文案创造力高于人类广告文案创造力

(三)创造力的离散分布研究

当前,人工智能在生产力上具有巨大的优势,形成了数字内容与前沿技术持续交互的模式,这在目前已经初步获得了共识。不少学者认为在创意产业生产过程中,生成式人工智能可大大节省时间和人力成本,提高产出效率,实现内容创作能力的“大爆炸”^[34]。

然而,在以创意为核心的创造性生产中,数量的追求并非核心目标,“在大多数创新环境中,我们宁愿有10个伟大的想法和90个糟糕的想法,也不愿有100个中等质量的想法”^[25]。这推动笔者进一步思考:当下以生产效率和数量优势为特征的人工智能大语言模型,其创造力分布状况究竟如何? Girotra 等学者的研究发现,与顶尖大学的学生相比,ChatGPT 4.0在产品创意生成方面表现出更快的速度和更低的成本,且生成的创意平均质量更高。更重要的是,最佳产品创意的大部分想法是由ChatGPT生成的,且通过提供高质量的产品创意案例可以进一步提高ChatGPT的性能。^[4]

因此,创造力的平均表现固然重要,创造力的差异化也同等重要,是产出多元性的标志,往往也是产业界筛选创意工具的重要标准。在实际的创造性任务中,只有少数优秀的创意能被采纳。在此背景下,对人工智能创作质量最优极值的考察显得十分必要,故提出以下假设:

H3:2024年人工智能和人智协作模式广告文案创造力的差异更高,其创造力能达到或突破人类的最高水平

三、研究方法与研究设计

(一) 实验材料

本研究延续李育辉等前期研究中的任务材料^[6],选择来自2022年“中国大学生广告艺术节学院奖”(以下简称“学院奖”)的品牌命题策略单,并筛选10个品牌的广告文案任务单作为任务列表,覆盖7个行业。针对这10份广告文案任务单,本次实验的材料具体如下表1所示:

1. 继续使用李育辉等采用的人类独立创作的20份文案^[6]。当时,生成式人工智能并未兴起,可以认为人类文案创作水平并没有受到生成式人工智能的影响。

2. 2024年人工智能独立创作文案30份。具体获取方式为:①将专家小组提取的任务单整理为符合人工智能模型输入的要求,且文案长度控制在500字以内;②使用美国和中国当时最新的人工智能平台进行创作,分别为“AI独立创作1”(AI1)(ChatGPT4.0版本,共10份)和“AI独立创作2”(AI2)(文心一言3.5/讯飞星火,共20份)。

3. 人智协作下的文案创作40项。使用最新的ChatGPT4.0完成前述4类人智协作,每种协作方式获取10份文案。

表1 文案创作方式以及文案数量说明

文案创作方式	文案创作说明	文案数量(份)
人类独立创作	直接选用获奖作品文案为样例(学院奖,2022年)	20
AI独立创作1(AI1)	任务委派式生成的文案,来自ChatGPT4.0	10
AI独立创作2(AI2)	任务委派式生成的文案,来自文心一言3.5/讯飞星火	20
协作1:ChatGPT4.0+CoT	在原有的提示词后加上“请一步一步地思考”	10
协作2:ChatGPT4.0+指定CoT	首先询问大模型,“好广告文案的标准是什么”,然后将此标准的回答添加在原有提示词后作为指定CoT	10
协作3:ChatGPT4.0+参考案例	在原有提示词后加上过往2个人类文案创意作为参考	10
协作4:ChatGPT4.0+人工选优	选取人类文案提供给人工智能作为参考,每个品牌生成10份文案,从中选1份最优文案	10

(二) 专家评价者

通过志愿者招募和专家遴选的方法获取被试,并对其从业经历和教育背景进行筛选,确认其对广告文案创造力及指标体系的理解。筛选后,一共117位专家成为实验中的“评价者”,其中75.2%为高等院校的广告营销相关专业老师,其余为广告营销行业从业人员。专家群体年龄均值为42.4岁,标准差为8.7岁;从业年限均值为6.0年,标准差为2.1年。在最高教育程度方面,拥有博士学位者占比达51.6%,其中70.2%的专家有广告学专业学习经历,参与者的性别分布较为均衡,男性占比为54.8%,女性占比为45.2%。

(三) 实验程序

所有测试随机分为4组问卷^①,每组问卷至少包括5个品牌,20个文案测试任务。每个品牌都有多个被测文案,其创作方式随机分配,但人类独立创作的文案为必测选项。测试开始后,专家们需随机选择其中一组问卷并进行评价,完成5个以上文案测评的为有效样本。作为被试的专家和作为执行投放的编程人员均不知文案的创作方式。测试共获得有效评价1631次,其中对人类独立创作文

^①测试文案链接:<https://docs.qq.com/sheet/DUEpCcGdHY3RyZHZJN>;问卷地址:<https://mzsurvey.miaozhen.com/r/eeze>。

案的测评 412 次,对 AI 独立创作的文案测评 354 次,对 4 种人智协作创作文案的测评 865 次。具体过程如下:

1. 分组及展示:问卷随机分组,然后向被试展示不同品牌的广告文案需求信息及与之对应的广告文案。

2. 评价过程:专家针对每个广告文案的清晰完整、容易理解、创造性、消费者洞察、商业洞察等维度进行 5 等级 Likert 量表评价,如表 2 所示。计算文案专业能力的平均分,并以其作为得分;请专家推测文案创作者的工作经验年限,并计算其平均值,并以其作为“等同工作年限”的得分。

表 2 广告文案的评价过程示例

评价过程	信息展示
1. 阅读广告文案需求任务单	甲方:江中制药 产品:江中健胃消食片广告 主题:年轻,吃得消 品牌调性:健康、好吃、药食同源、美食伴侣
2. 阅读随机分配的广告文案	文案:大食化小,小食化了
3. 对文案专业能力进行评分	清晰完整 容易理解 创造性 消费者洞察 商业洞察
4. 推测等同工作年限	1 年、2 年、3 年、4 年、5 年、6 年、7 年、8 年、9 年、10 年及以上

(四) 数据分析

本次实验数据结合李育辉等^[6]的研究结果,使用 SPSS26.0 软件进行方差分析和 t 检验进行显著性检验。检验将文案作者分辨力、文案专业能力评估整体及各维度、专家感知整体及各维度分别作为因变量;将任务以及文案作者(人类/人工智能/人智协作)作为自变量,并纳入两者的交互作用,将文案的品牌、评价专家的年龄、性别、工作年限作为控制变量。

四、研究结果

(一) 人工智能的广告文案创造力有显著提升,但尚未超越人类

对比 2024 年和 2023 年的专家对相同的人类独立创作的文案的测评结果发现:两年的专家测评在“文案专业能力”与“等同工作年限”上的评价极为接近,相差仅为 -0.04 和 0.13,t 检验结果 $p=0.87, p=0.81$ 。可以认为,两次测评的标准没有发生显著性的变化。这为后续的跨年度比较提供了基础,具体见表 3。

表 3 2023 年、2024 年对人类独立创作的广告文案的创造力评估总体比较

测量变量	2023 年测试 (M±SD)	2024 年测试 (M±SD)	差值	显著性(t 检验)
文案专业能力总分	3.55±0.95	3.51±1.00	-0.04	$p=0.87$
等同工作年限(年)	3.83±2.19	3.96±2.28	0.13	$p=0.81$

如图 1 所示,整体而言,2024 年人工智能在文案创造力上的提升显著。2024 人工智能的等同工作年限提升至 3.08 年(n=354),显著性高于 2023 年 AI 的 2.47 年(n=560, $p<0.01$),且文案专业能力得分也显著高于 2023 年的得分($p<0.01$)。

值得注意的是,2024 年 AI 的等同工作年限仍然显著性低于 2023 年人类的 3.83 年($p<0.01$)。这意味着人工智能的创造力进化在平均水平上仍没有达到专业人类的高度。对比以 OpenAI 为代表的国外模型和文心一言、讯飞星火等国内模型,两者在等同工作年限和文案专业能力上极为接近,验证了各个主要模型的进化程度接近,因此,后续的实验和分析不再区分国内和国外 AI 模型。

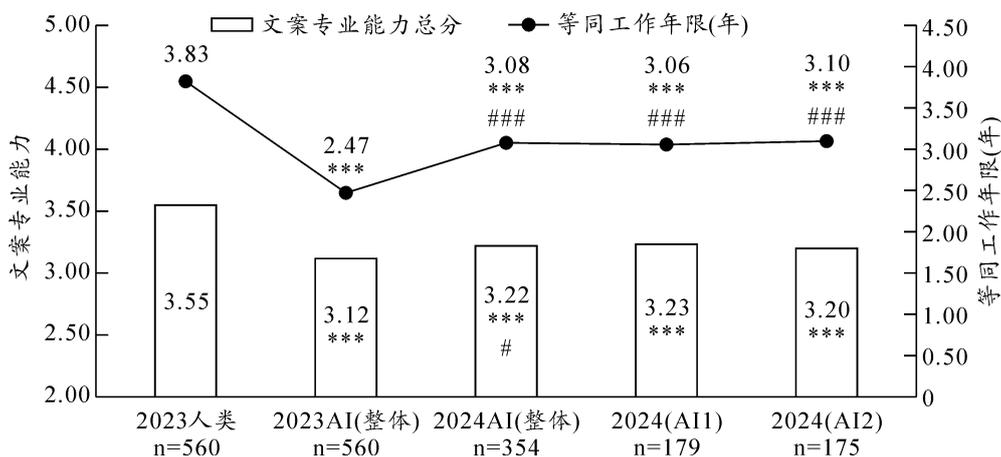


图 1 人工智能 (AI) 广告文案创造力发展追踪对比

备注: vs. 2023 人类: * $p<0.10$, ** $p<0.05$, *** $p<0.01$; vs. 2023AI(整体): # $p<0.10$, ## $p<0.05$, ### $p<0.01$

如图 2 所示,分别对文案专业能力各维度进行比较,发现 2024 年的人工智能在关键弱项(创造性、消费者洞察、商业洞察)上有显著提升,且逐渐接近人类水平,然而其在清晰完整和容易理解维度上得分下降。这一现象源于 2024 年人工智能数据规模与推理能力的增强,使其更倾向于生成创新性、深度化的内容,但可能因信息过载和结构复杂化而牺牲了可读性。这表明人工智能的创造力结构正逐步向人类靠拢,但在平衡深度与简洁性上仍需优化。因此,假设 H1a 得以证明,假设 H1b 不成立。

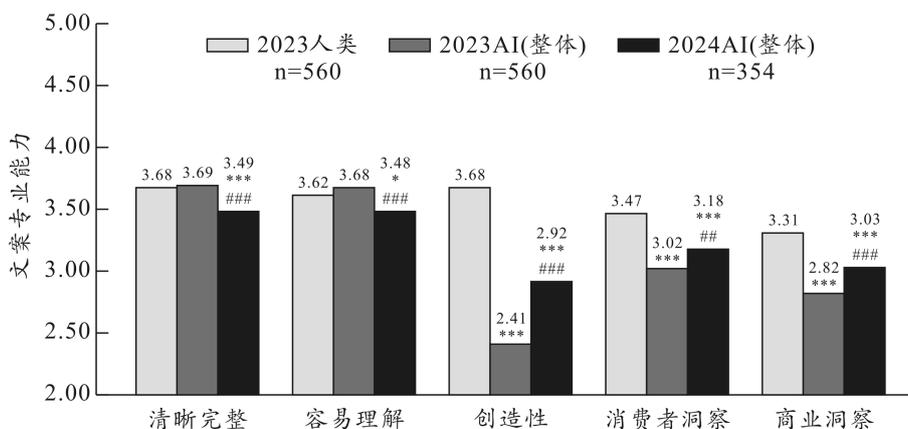


图 2 人工智能文案专业能力分维度对比

备注: vs. 2023 人类: * $p<0.10$, ** $p<0.05$, *** $p<0.01$; vs. 2023AI(整体): # $p<0.10$, ## $p<0.05$, ### $p<0.01$

(二)人智协作方式无法提升人工智能的广告文案创造力

出乎意料的是,如图3所示,4种人智协作方式(n=865)都没有显著提升人工智能的文案创造力。协作1(ChatGPT4.0+CoT)(n=181)作为最常用思维链模式,等同工作年限为2.93年,与2024年人工智能模型的3.08年无显著性差异。但协作2模式(ChatGPT4.0+指定CoT)(n=227)的等同工作年限降低为2.51年,显著性低于2024年简单使用人工智能模型的水平,这或许是因为指定思维链模式约束了人工智能模型的内在创造力空间。协作3(ChatGPT4.0+参考案例)(n=231)的等同工作年限为2.74年,仍然低于2024年简单模式下人工智能模型的水平,这或是因为参考案例的输入约束了大模型的内在创造力空间,或是因为没有提供更好的案例作为参考。协作4(ChatGPT4.0+人工选优)(n=226)的创造力等同工作年限上升到2.82年,但显著性仍然低于2023年人类和略低于2024年简单模式下人工智能模型的水平。因此,假设H2a和H2b都不成立。

以上结果揭示这样一个结论:当前人智协作模式确实提供了更多的参考和思维链,这对于提高计算、推理能力等有一定效果,但无法有效提升创造力,这也从侧面印证了创造力在原创性方面存在内在结构。

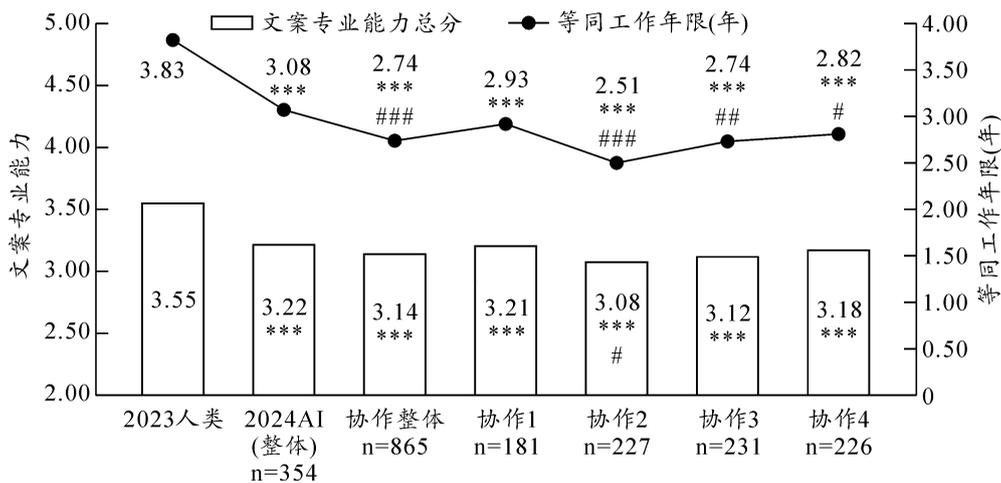


图3 不同协作创作方式的文案创造力对比

备注:vs. 2023人类: *p<0.10, **p<0.05, ***p<0.01; vs. 2023AI(整体): #p<0.10, ##p<0.05, ###p<0.01

(三)基准趋近与离散加剧:人工智能广告文案创造力出现类人化的“灵感闪现”

除了对比平均数外,我们将每个文案的等同工作年限进行分析。发现2023年每个AI生成文案的等同工作年限都低于人类独立创作文案的中位数,且内部差异性不大(标准差=0.25)。而人类独立创作文案的等同工作年限呈现较大的差异性(标准差=0.65)。2024年AI的文案等同工作年限不仅中位数增长,而且内部差异性远大于2023年AI的作品(方差齐性检验显著)与人类创造的内部差异无显著性差异,见表4。

值得关注的变化是,2024人工智能在等同工作年限的分布上,拉近了与人类创作的差距。不仅从2023年相对集中的趋势转变为接近人类创作的分布状态,更为重要的一点是,人工智能的上限值与人类的最优表现已趋于吻合(评分数值:4.9年)。这意味着在人工智能大模型技术的类人化趋势不断加深的情况下,人工智能也出现了文案创作中类人的“灵感闪现”现象。因此,研究结果部分支持假设H3,即人工智能创造力差异性变大,有少量创意超过人类平均水平,甚至接近人类较高创意水平,见图4。

表 4 不同创作方式的文案等同工作年限的分布统计

	人类	2023AI(整体)	2024AI(整体)	人智协作(整体)
中位数	4.00	2.48	3.00	2.77
标准差	0.65***	0.25	0.61***	0.38*
标准差/中位数	16.3%	9.9%	20.3%	13.7%
文案数量	20	20	30	40

备注:与 2023AI 对比方差齐性分析,* p<0.1,*** p<0.01;方差齐性检验,F(19,29)=1.135,p>0.05

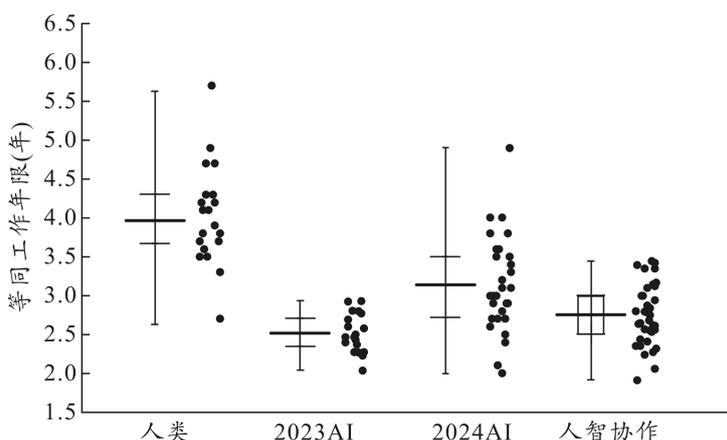


图 4 不同创作方式的文案等同工作年限区间与散点图

虽然在等同工作年限得分上,人工智能尚未超越人类,但从变化的趋势来看,随着大语言模型的演化迭代逐步加深,人工智能的上限值突破将更为常见,并进一步提升其整体的数值。但 4 种协作方式在某种程度上约束了人工智能的想象力,差异性反而降低了。

通过对 2024 年等同工作年限评分排名前 30 位(占文案总量的 37.5%)的文本分析发现,讯飞星火生成的一则文案得分 4.9 与人类创作文案在评分上并列第二位。进一步的数据分析显示,人类创作者的表现呈现出更为广泛的分布特征:其评分区间从 2.7 年至 5.7 年不等,具体可分为高值区(4.7 年—5.7 年)、中值区(3.8 年—4.3 年)和低值区(2.7 年—3.7 年)3 个明显区间。相较而言,人工智能协作产出的文案评分分布则相对集中,其评分主要集中于 2.7 年至 3.2 年区间。

这一分布特征可能揭示了人工智能在广告文案生产过程中存在的“定势效应”(Einstellung Effect)。该效应指使用者在人工智能辅助创作时,倾向于依赖既有的、熟悉的解决方案,从而限制了创意表达的多样性和突破性。这种认知偏见的持续作用,可能导致人工智能辅助创作陷入中等化、平均化的困境。研究表明,虽然人工智能在某些个案中可以达到与人类相当的创造力水平,但在整体创意表现的广度与深度上仍存在明显局限。

五、结论、讨论与展望

本次研究的实验设计与分析,呈现了当前生成式人工智能与不同类型人智协作模式在创造力方面的差异化进程。通过对比前后两年的实验结果可以发现,在针对广告文案的创造力比较中,人类整体的创造力得分仍然占优势,但伴随着技术的不断演化升级,生成式人工智能在以下特定维度展现出创造力的再提升:

其一,在等同工作年限上,对比2023、2024年的专家评价数据发现,2024年生成式人工智能的等同工作年限从2023年的2.47年显著提升至3.08年。这意味着人工智能模型在不断创新与优化,并从技术层面深化了其对数据的理解和运用,从而实现了等同工作年限上的进一步提升。对该项指标的连续性测量,为我们理解生成式人工智能作为人力资本的替代性要素提供了直观的理解视角。

其二,对文案专业能力的各项指标进行比较,可以看到2024年生成式人工智能在关键弱项(创造性、消费者洞察、商业洞察)上有显著提升,且逐渐接近人类的水平。可以预见,生成式人工智能的创造力将与人类越来越相似,并在拓展创意边界、提供灵感启发等方面提供更多的价值。

其三,在等同工作年限的分布上,2024年生成式人工智能的分布形态呈现出更显著的离散化特征,突破了2023年集中性较强的同质化状态,与人类创作的分布形态趋于接近。具体表现为:一方面在等同工作年限峰值上持续突破,已有偶发的类人化“灵感闪现”的创作表现;另一方面在清晰完整等基础性指标上波动增大,这种分布演变使我们更清晰地认识到,生成式人工智能创意系统已开始模拟人类创作中“突破与妥协”的动态平衡机制,标志着其从技术性输出向真正的创造性产出的关键转变。

随着人工智能成为人类智力工作的合作伙伴,人机关系的不确定性与复杂程度可能远超人们的预测与想象^[35-36]。这一发展趋势促使我们不仅需要适应与人工智能的常态化协作,理解人机互惠式增强关系^[37],更需要深入反思:在技术赋能的背景下,人类创造力的本质价值将如何彰显?创作内容中的人类真实性应当居于何种地位?当生成式人工智能成为创意生产的前沿工具时,我们与创造力的关系将面临根本性重构。这或许要求我们回归哲学层面的基础性思考——重新回到“人之为人”与“机器之为机器”的思考原点?^[38]爱因斯坦曾指出:“创造力是智力的乐趣所在。”^[39]创造性工作不仅产出成果,更承载着人类独特的意义建构与情感体验。从这个角度来看,找到人类与人工智能创造力的平衡点也是未来创造性工作需要面临的重要挑战。这一平衡点的探寻,不仅关乎生产效率,更涉及人类在智能时代的存在意义与发展方向。

未来的研究可从如下方面深化展开:首先,实验中采用的生成式人工智能模型(文心一言和讯飞星火)虽然代表了2024年的主流技术水平,但随着DeepSeek和元宝等为代表的人工智能技术快速迭代,其创造力特征也会持续变化发展,因此,后续研究需要保持连续性,以持续获得最新的数据,确保对生成式人工智能创造力的全面理解和准确把握。其次,本研究选取的大广赛优秀广告文案样本虽然具有代表性,但可能无法全面反映人类广告文案创作的整体水平,且广告文案仅作为一个单一任务进行分析,这可能无法充分揭示创造力的多维度特性,因为创造力在不同任务和情境中可能表现出不同的特征和要求,之后的研究需要进一步拓宽研究对象。再次,被试群体主要由广告行业的专家组成,虽然他们的专业性较强,但其对广告文案的评价标准可能与普通消费者的实际感知存在一定的偏差,这可能影响到研究结果的普适性和适用性。最后,创造力作为一个多维度的复杂构念,其提升路径尚未在现有的人机协作模式中得到充分实现。主流的4种协作方式均未能有效突破创造力评分的基准线,这一结果凸显出现有模式在创造力离散度提升方面的局限性。通过对2024年等同工作年限排名前30的文案进行溯源分析发现两个显著特征:一是基准化聚集效应;二是高离散度缺失,仅有讯飞星火的个案(4.9年)突破了这一基准范围。这种分布格局印证生成式人工智能在创意生产中存在的“定势效应”,其产出往往收敛于基准水平,难以实现高离散度的创造性突破。这也提示后续研究应发展更为系统化的协作模式,进一步拓展人智协作在创造力上的边界,以动态适应生成式人工智能日益融入工作场景的背景下对传统工作模式的冲击与重塑,为创意生产的范式转型提供一定参考。

(致谢:研究获得明略科技秒针营销科学院谭北平、王云、杨菲、刘子平对AI工具的实验支持;感

谢中国高等教育学会广告教育专业委员会 2024 年学术年会与会专家的支持;宁波大学人文与传媒学院庞菊爱副教授,北京工商大学语言与传播学院硕士生胡家欢、肖晴、刘欣雨、曹胜利对本研究亦有贡献。)

参考文献:

- [1] N. Anantrasirichai, D. Bull. Artificial Intelligence in the Creative Industries: A Review. *Artificial Intelligence Review*, 2022, 55: 589-656. [2024-10-07] <https://doi.org/10.1007/s10462-021-10039-7>.
- [2] D. W. Hall, J. Pesenti. Growing the Artificial Intelligence Industry in the UK. 2017-10-15. [2024-10-07] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf.
- [3] Executive Office of the President National Science and Technology Council Committee on Technology of the United States. Preparing for the Future of Artificial Intelligence. 2016-10-12. [2024-10-07] https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.
- [4] K. Girotra, L. Meincke, C. Terwiesch, et al. Ideas are Dimes a Dozen: Large Language Models for Idea Generation in Innovation. *SSRN Electronic Journal*, 2023. 2023-08-02. [2024-10-08] <https://doi.org/10.2139/ssrn.4526071>.
- [5] G. Amato, M. Behrmann, F. Bimbot, et al. AI in the Media and Creative Industries. 2019: 1905. 04175. [2024-10-08] <https://arxiv.org/abs/1905.04175v1>.
- [6] 李育辉, 庞菊爱, 谭北平. 人工智能与人类的创造力比较研究: 基于专家和消费者的双重视角. *商业经济与管理*, 2023, 10: 23-35.
- [7] 张进澳, 卢新元, 孙冰悦等. 人智协同知识创新机制的建模与仿真——基于模拟实验的系统动力学分析. *情报理论与实践*, 2024, 12: 143-153+96.
- [8] 吴小龙, 肖静华, 吴记. 人与 AI 协同的新型组织学习: 基于场景视角的多案例研究. *中国工业经济*, 2022, 2: 175-192.
- [9] 谭北平, 金立印, 王云等. 生成式营销研究报告. 哲学社会科学预印本平台. 2024-09-10. [2024-10-07] <https://zsyb.cn/abs/202409.00654>.
- [10] 中国信息通信研究院. 人工智能发展报告(2024年). 2024-11-29. [2025-05-28] http://www.caict.ac.cn/kxyj/qwfb/bps/202412/t20241210_647283.htm.
- [11] F. Dell'Acqua, E. McFowland, E. R. Mollick, et al. Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. *Harvard Business School Technology & Operations Mgt*, 2023, Unit Working Paper No: 24-013. [2024-10-07] <https://www.hbs.edu/faculty/Pages/item.aspx?num=64700>.
- [12] M. Runco, G. Jaeger. The Standard Definition of Creativity. *Creativity Research Journal*, 2012, 24(1): 92-96.
- [13] 秦雪冰, 郭博. 智能广告文案的消费者参与度研究——基于汽车之家 App 的实证检验. *新闻与传播研究*, 2022, 6: 56-72+127.
- [14] M. A. Runco. AI can Only Produce Artificial Creativity. *Journal of Creativity*, 2023, 33(3): 100063. [2024-10-07] <https://doi.org/10.1016/j.yjoc.2023.100063>.
- [15] 朱毅鑫, 高涛, 范丽凤等. “暗”, 不止于“深”——迈向认知智能与类人常识的范式转换. *Engineering*, 2020, 3: 230-306.
- [16] L. Chen, L. Sun, J. Han. A Comparison Study of Human and Machine-generated Creativity. *Journal of Computing and Information Science in Engineering*, 2023, 23(5): 051012. [2024-10-07] <https://doi.org/10.1115/1.4062232>.
- [17] F. K. Hubert, N. K. Awa, L. D. Zabelina. The Current State of Artificial Intelligence Generative Language Models is More Creative than Humans on Divergent Thinking Tasks. *Scientific Reports*, 2024, 14(1): 3440. [2024-10-07] <https://doi.org/10.1038/s41598-024-53303-w>.
- [18] E. E. Guzik, C. Byrge, C. Gilde. The Originality of Machines: AI Takes the Torrance Test. *Journal of Creativity*, 2023, 33

- (3):100065. [2024-10-07] <https://doi.org/10.1016/j.yjoc.2023.100065>.
- [19] J. Haase, P. Hanel. Artificial Muses; Generative Artificial Intelligence Chatbots Have Risen to Human-Level Creativity. *Journal of Creativity*, 2023, 33(3): 100066. [2024-10-08] <https://doi.org/10.48550/arXiv.2303.12003>.
- [20] M. Koivisto, S. Grassini. Author Correction: Best Humans still Outperform Artificial Intelligence in a Creative Divergent Thinking Task. *Scientific Reports*, 2024, 14(1): 4239-4239.
- [21] 陆小华. 智能内容生成的实质、影响力逻辑与运行范式——ChatGPT等智能内容生成现象透视与解析. *新闻大学*, 2023, 4: 17-24+119-120.
- [22] 李建会, 夏永红. 人工智能会获得创造力吗? *国外社会科学*, 2020, 5: 52-60.
- [23] M. Farina, W. Pedrycz, A. Lavazza. Towards a Mixed Human-Machine Creativity. *Journal of Cultural Cognitive Science*, 2024, 8: 151-165. [2024-10-07] <https://doi.org/10.1007/s41809-024-00146-6>.
- [24] B. C. Lee, J. J. Chung. An Empirical Investigation of the Impact of ChatGPT on Creativity. *Nature Human Behaviour*, 2024, 8: 1906-1914. [2024-10-07] <https://doi.org/10.1038/s41562-024-01953-1>.
- [25] C. Stevenson, I. Smal, M. Baas, et al. Putting GPT-3's Creativity to the (Alternative Uses) Test. 2022-06-10. [2024-10-08] <https://doi.org/10.48550/arXiv.2206.08932>.
- [26] D. Wang, E. Churchill, P. Maeset. From Human-Human Collaboration to Human-AI Collaboration; Designing AI Systems That can Work Together with People. *Proceedings of CHI'20: CHI Conference on Human Factors in Computing Systems*, 2020. 2020-04-25. [2024-10-08] <https://dl.acm.org/doi/10.1145/3334480.3381069>.
- [27] 郭全中, 苏刘润薇. 作为新基础设施的 AGI: 以 GPT-4.0 等新一代生成式人工智能为例. *新闻爱好者*, 2024, 7: 16-21.
- [28] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, et al. Generative Adversarial Nets. *Proceedings of the 28th International Conference on Neural Information Processing Systems*, 2014, 2: 2672-2680. [2024-10-07] <https://dl.acm.org/doi/10.5555/2969033.2969125>.
- [29] R. Rombach, A. Blattmann, D. Lorenz. High-resolution Image Synthesis with Latent Diffusion Models. 2022-04-13. [2024-10-07] <https://doi.org/10.48550/arXiv.2112.10752>.
- [30] J. McGuire, D. D. Cremer, D. V. T. Cruys. Establishing the Importance of Co-creation and Self-efficacy in Creative Collaboration with Artificial Intelligence. *Scientific Reports*, 2024, 14: 18525. 2024-08-09. [2025-10-08] <https://doi.org/10.1038/s41598-024-69423-2>.
- [31] A. A. Cabrera, A. Perer, J. I. Hong. Improving Human-AI Collaboration with Descriptions of AI Behavior. 2023-01-06. [2024-10-08] <https://doi.org/10.48550/arXiv.2301.06937>.
- [32] R. A. Finke, T. B. Ward, S. M. Smith. *Creative Cognition: Theory, Research, and Applications*. Cambridge: MIT Press, 1992: 241.
- [33] Y. H. Liu, T. L. Han, S. Y. Ma. Summary of ChatGPT-related Research and Perspective Towards the Future of Large Language Models. *Meta-Radiology*, 2023, 1(2): 100017. [2024-10-08] <https://www.sciencedirect.com/science/article/pii/S2950162823000176?via%3Dihub>.
- [34] 蔡子凡, 蔚海燕. 人工智能生成内容(AIGC)的演进历程及其图书馆智慧服务应用场景. *图书馆杂志*, 2023, 4: 34-43+135-136.
- [35] 朱恬骅. 人工智能会创作吗? ——“类文学文本”生成的创造力疑难与集体性方案. *南京社会科学*, 2024, 7: 113-121.
- [36] S. Aaron, H. Jess, L. Lindsay, et al. Vero: An Accessible Method for Studying Human-AI Teamwork. *Computers in Human Behavior*, 2023, 141: 107606. [2024-10-07] <https://doi.org/10.1016/j.chb.2022.107606>.
- [37] 姚曦, 姚俊. 第三种劳动: 智能品牌传播中的人机关系. *湖北大学学报(哲学社会科学版)*, 2024, 1: 154-161+176.
- [38] 洪杰文. 混合智能时代重返新闻传播的“人机之界”. *新闻与传播评论*, 2024, 6: 1.
- [39] G. S. Viereck. What Life Means to Einstein; An Interview by George Sylvester Viereck. *The Saturday Evening Post*, 1929-10-26. [2025-06-12].

A Comparative Study on Creativity between Artificial Intelligence and Human-AI Collaborative Models: A Double-blind Experimental Study Based on Advertising Copywriting

Sun Mingxin (Beijing Technology and Business University)

Liu Rong (Tianjin Foreign Studies University)

Li Yuhui (Renmin University of China)

Abstract: In recent years, the creativity demonstrated by artificial intelligence is approaching that of humans, developing human-machine collaboration modes such as Chain-Thought, Designated ChainThought, Good Examples, and Artificial Selection. It is of great theoretical and practical significance to conduct comparative studies on human creativity and the creativity shown by artificial intelligence and human-intelligence collaboration. The study employed an experimental research method, inviting 117 experts to conduct a double-blind experiment on advertising copy generated independently by humans, AI, and human-AI collaboration. It was compared with the 2023 artificial intelligence creativity research. The findings indicate that in 2024, the “equivalent working years” of generative AI increased from 2.47 years in 2023 to 3.08 years, with a maximum of 4.9 years. The average benchmark and dispersion of AI creativity have significantly improved. There is an obvious set effect in the copywriting created by various ways of human intelligence collaboration, and the average creativity is not further improved. This suggests that human-intelligence collaboration needs a more organized innovation mode to avoid falling into the predicament of mediocrity.

Key words: generative artificial intelligence; human-AI collaboration; advertising copy-writing; creativity; contrastive study

■ 收稿日期: 2024-11-07

■ 作者单位: 孙铭欣, 北京工商大学语言与传播学院; 北京 100048

刘 荣, 天津外国语大学国际传媒学院; 天津 300204

李育辉(通讯作者), 中国人民大学劳动人事学院; 北京 100872

■ 责任编辑: 肖劲草