

AI 游戏创作工具用户研究

大学生 AI 游戏创作工具用户研究

附录材料

定性访谈、桌面研究与问卷数据源文件

2026 年 7 月

目录

[附录A 定性访谈源文件 3](#)

[A.1 样本一 | H5-01 3](#)

[A.2 样本二 | H5-02 21](#)

[A.3 样本三 | H5-03 47](#)

[A.4 样本四 | H5-07 65](#)

[A.5 样本五 | H5-08 80](#)

[附录B 桌面研究源文件 103](#)

[研究摘要 103](#)

[一、研究对象与桌面研究任务 103](#)

[二、核心生产端案例：Codex 104](#)

[三、核心平台端案例：TapTap 制造 105](#)

[四、机制对照与补充对象 106](#)

[五、产品化判断框架 107](#)

[六、桌面研究假设如何进入实证研究 107](#)

[七、结论与边界 108](#)

[附录 A：原桌面研究产品扫描与验证问题 108](#)

[附录C 问卷数据 146](#)

附录A 定性访谈源文件

A.1 样本一 | H5-01

样本基本信息

样本编号	H5-01
样本定位	以 Codex 为主力完成 48h Game Jam Demo 的技术型创作样本，涉及 UE、核心玩法代码、多轮调试与人工 Review。
访谈执行人	姓名待补
摘要来源	访谈摘要：H5-01(2).docx
逐字稿来源	访谈逐字稿1.docx
关联答卷	8604f26f-9a81-45ab-baf3-8063962e5976

本文件保留该样本的访谈摘要、逐字稿和源材料附件；最后附上关联问卷中的研究相关作答与既有象限结果。

访谈摘要

原始材料：访谈摘要：H5-01(2).docx

访谈摘要

一、访谈基本信息

受访者编号：H5-01

访谈时间：2026/07/08 14:32:00 - 2026/07/08 15:31:00

导出时间：2026-07-18 0:36

二、样本判定

本访谈对象可判定为 Codex/AI Agent 游戏开发核心样本。受访者为电气专业大三学生，有高中起积累的编程基础，在 Game Jam中以程序员身份参与 48 小时短期 Demo 开发，主要使用 Codex 完成大部分代码工作，同时以 Claude Code 作为轻量级辅助工具。其使用场景涵盖 UE（Unreal Engine）中的钩锁系统物理约束、射线检测、版本控制审计、项目文档拆解等任务，并涉及 AI 生成美术资产（Logo、按钮等）。

受访者提供了较为完整的项目描述和工具使用链路，包括需求拆分、多轮对话生成代码、Bug 修复过程和玩家试玩反馈。材料足以还原一个完整的 Game Jam 项目从需求到交付的过程，因此本样本适合作为核心样本，用于观察“有编程基础、以 Game Jam 快速出 Demo 为目标”的 Codex + Claude Code 联合使用模式。

样本标签：

判定理由：有真实 Game Jam 项目经历，Codex 产出进入短期 Demo，提供了较完整的使用链路和玩家反馈

具体项目：【2026 CIGA Game Jam 作品：《Anchor Drive》-哔哩哔哩】<https://b23.tv/aHwogVi>

项目阶段：短期 Demo / 原型阶段（48 小时开发）

过程材料：逐字稿、B站演示视频链接、Game Jam 作品链接

从 AI 游戏消费者访谈提纲看，也可作为消费端补充样本。受访者以程序员身份参与 Game Jam，其作品《Anchor Drive》在线下站点有玩家试玩环节，受访者能围绕玩家反馈、AI 标签接受度、美术/音效的 AI 生成边界作出具体判断。但受访者并非以纯玩家身份体验一站式 AI 游戏平台作品，因此归为消费端补充样本。

消费端样本标签：

判定理由：有作品试玩反馈和 AI 标签态度判断，但非纯消费端视角，缺少对一站式 AI 游戏平台的完整体验链路

体验作品：2026 CIGA Game Jam 作品《Anchor Drive》（线下试玩）

对应平台：CIGA Game Jam / B站

游玩时长：未明确；为线下站点交流试玩环节

是否知情 AI 参与：是

三、开发者端样本分析

1. 基本背景与开发基础

受访者为电气专业大三学生，从高中阶段起就有较多编程接触，具备扎实的编程基础。其参与的项目为 2026 CIGA Game Jam，一个 48 小时的限时开发活动，受访者以程序员身份参与，负责项目中的代码实现工作。

受访者使用的主要游戏引擎为 Unreal Engine（UE），项目代码量约两三千行。他采用“策划先出策划案 → 自己拆分需求 → 分多个对话喂给 Codex”的工作流，体现了对 AI 工具的成熟使用策略。

该案例支持提纲中的基础假设：用户能否用好 AI 编程工具，很大程度取决于已有的编程基础和对引擎的熟悉程度。受访者明确表示会 review AI 生成的代码或要求 AI 给出逻辑链，说明他不是盲目信任 AI 输出，而是以自身技术能力作为判断和纠错条件。

2. 具体项目还原

受访者的核心项目为 Game Jam 作品。该项目在 48 小时内开发完成，主要代码量约两三千行，使用 Unreal Engine 开发。项目核心玩法涉及钩锁系统的物理约束机制，这是受访者认为在没有 AI 辅助时最容易卡住的技术难点。

项目开发流程为：等待 CGJ 主题公布 → 策划撰写策划案 → 受访者将需求拆分 → 分多个不同对话喂给 Codex 生成代码。美术资产方面，图片类资产（Logo、按钮等）由 AI 生成，材质类资产由美术手动制作。作品完成后在线下站点进行展示，设有玩家交流试玩环节。

玩家反馈主要集中在关卡设计难度过高、难以通关，而对于游戏逻辑和美术表现均为好评。这说明 AI 辅助开发的 Demo 在核心玩法和视觉呈现上达到了可接受的质量水平。

3. 具体任务使用过程

AI 工具使用呈现清晰的双工具策略：

Codex 作为主力代码生成工具：受访者将拆分后的需求描述清楚，使用 Plan 模式先确认方案，然后让 Codex 生成代码。第一次生成通常能完成大部分效果，个别细节 Bug 通过 2-3 轮对话即可修复。例如本次 Demo 中，第一次生成后发现锚点目标的射线检测通道有小 Bug，指正后一轮修改即完成。

Claude Code 作为轻量级辅助工具：主要用于版本控制审计与提交、文件系统操作等边缘任务，不介入项目核心代码。受访者有一次印象较深的经历是 Clone 了一个 UE 插件后用 Claude Code 拆解插件结构，生成完整技术文档以供复刻参考，该任务属于“项目理解”类别。

这一使用链路验证了提纲中的多轮协作假设：真实使用不是一次生成即完成，而是“需求描述 → Plan 确认 → 生成 → 测试 → 指正 → 再生成”的迭代过程。同时也展示了不同 AI 工具在同一项目中的分工协作模式。

4. 效率提升与能力补足

Codex 在 Game Jam 场景中的核心价值是显著缩短从想法到可运行代码的时间。受访者表示，在没有 AI 的情况下，像钩锁系统物理约束这类功能需要查阅大量 UE 文档，耗时较长；而有了 Codex 后，这类“原本不敢做的功能”变得可行。

但受访者并不认为 AI 可以让无基础者独立完成游戏开发。他强调自己会对 AI 生成的代码进行 Review，或要求 AI 给出逻辑链，这说明 AI 的输出仍需有能力的开发者把关。该样本支持“AI 让强者更强”的判断：编程基础是使用 AI 工具产出可靠代码的前提条件。

5. 人工控制、判断与维护成本

受访者对 Codex 生成代码的人工控制策略是"Review + 逻辑链验证"。他不会盲目接受 AI 输出，而是会审视代码逻辑是否正确。在三轮左右的迭代中，主要由受访者发现 Bug 并指正，AI 随后修复。

关于 Claude Code 的代码生成，受访者明确表示"没有拿 Claude Code 做过太多代码工作"，因此人工修改的讨论主要集中在 Codex 上。

后期维护成本方面，受访者认为"有一点"成本，主要体现在 Token 消耗上。对于 48 小时短期 Demo 来说，这一成本可接受；但若迁移到长期项目，Token 成本的累积可能成为考量因素。

该样本呈现的是一种"中审计、快速迭代"的使用模式：不像 H5-02 那样"低审计、结果够用即可"，H5-01 会对 AI 输出进行逻辑审查，但对逐行代码的深度理解可能仍有限，适合短期 Demo 但长期维护存在不确定性。

6. 产出落地与项目影响

Codex 的产出在《Anchor Drive》中直接进入项目，覆盖了钩锁系统、射线检测等核心玩法代码。Claude Code 的产出则用于项目理解（插件拆解文档）和工程辅助（版本控制、文件系统），不进入核心代码链路。

AI 对项目的核心价值在于"快速推进"：将策划案中的需求迅速转化为可运行的代码，使 48 小时内完成两三千行代码量的 Demo 成为可能。同时，AI 也让受访者敢于尝试原本认为技术难度较高的功能（如钩锁物理约束）。

从产出落地角度看，受访者支持一个明确判断：Codex 在短期 Game Jam 项目中不仅能落地，而且可以承担核心玩法代码的生成；但这一落地高度依赖用户自身的编程基础和 Review 能力。

7. 留存、迁移与付费

受访者通过 订阅使用 Codex，为包月制，中间无其他付费环节。使用频率为"只要做项目就用"，无项目时则使用更轻量的工具。

付费决策逻辑清晰：只要没有比 GPT 更强且更便宜的模型出现，就会持续续费。支撑付费的核心因素是"模型和 Agent 的能力"，而非一键生成、导出、发布等附加功能。这说明对有一定技术能力的用户而言，工具的核心竞争力在于模型本身的代码理解和生成质量。

迁移意愿方面，如果未来出现更强且更便宜的模型，受访者表示会考虑更换平台。因此 Codex 的留存基础是"当前最优性价比"，而非不可替代性。

8. AI 参与制作的市场判断

受访者对 AI 参与游戏制作持理性区分态度：

在代码层面，他认为"在保证质量的前提下，程序用 AI 生成和自己写没有任何区别"，玩家无法也不应区分代码来源。

在美术和音效层面，他指出"如果音效和美术是 AI 生成，确实有玩家会介意并且抵制"，并会主动告知玩家作品中 AI 生成的部分。这一区分反映了当前 AI 参与游戏制作的现实边界：代码的后台性质使其不易引发争议，而前台内容（美术、音效）的 AI 生成则更容易触及玩家敏感点。

该样本验证了提纲中的假设：AI 标签不一定损害作品评价，关键在于哪些环节使用了 AI，以及最终质量是否达标。

9. 理想中的 Codex/AI 开发工具

受访者最希望 AI 工具在未来增强的两个方向是：（1）理解能力；（2）支持多人协作。

他认为如果平台未来更强，人类创作者的核心价值仍在于"策划案和架构"，这是 AI 暂时做不好的部分。未来开发者的角色将从手写代码转向"需求拆解、架构规定"等高层工作，偶尔进行 Debug。

这一判断与提纲中的理想工具假设一致：未来的 AI Agent 不应替代人类创作者的策划和架构能力，而应在此基础上进一步提升理解深度和协作能力，使人类能专注于更高层次的设计工作。

四、开发者端访谈记录字段整理

- 受访者编号：H5-01
- 样本类型：Codex/AI Agent 游戏开发者
- 样本标签：TOOL-C
- 是否在校大学生：是
- 专业与年级：电气专业；大三
- 编程基础：有，从高中阶段起积累较多
- 使用引擎或开发工具：Unreal Engine、Codex、Claude Code、GPT Plus
- 具体项目或作品：2026 CIGA Game Jam 作品《Anchor Drive》
- 项目类型：Game Jam（48 小时限时开发）
- 当前阶段：短期 Demo / 原型
- 工具使用时间：按月订阅，做项目时持续使用
- 使用任务类型：代码生成、功能实现、Bug 修复、项目理解/文档拆解、版本控制辅助
- 是否涉及至少两类实质任务：是
- 产出状态：Codex 产出进入核心玩法代码（钩锁系统、射线检测等）；Claude Code 产出用于工程辅助和项目理解
- 人工修改程度：中，Codex 生成后 Review 代码逻辑，约 2-3 轮迭代修复 Bug
- 是否提高效率：是，显著缩短从需求到可运行代码的时间，使原本不敢做的功能变得可行

- 是否补足能力：是，补足了 UE 中复杂物理约束等高级功能的实现能力
- 是否增加维护成本：Token 成本有一定增加，但短期 Demo 中可接受
- 是否持续使用：是，做项目即使用 Codex
- 付费判断：GPT Plus 包月订阅；只要没有更强且更便宜的模型就会续费
- 是否有过程材料：逐字稿、B站演示视频链接、Game Jam 作品链接
- 主要有效场景：Game Jam 原型、UE 核心玩法脚本、物理约束系统、射线检测、插件拆解理解、版本控制辅助
- 主要失效场景：模型"降智"导致理解和工程能力下降、Token 成本累积
- 可用于验证或反驳的假设：验证"Codex 在短期游戏原型中能显著提效"“用户基础影响 AI 产出落地质量”“真实使用为多轮迭代而非一次生成”；补充说明有编程基础的用户会主动 Review AI 输出，形成“中审计、快速迭代”模式

五、消费者补充样本分析

1. 消费端样本定位

受访者符合消费端补充样本条件。受访者以开发者身份参与 Game Jam，但其作品《Anchor Drive》在线下站点有玩家试玩环节，他能提供玩家对 AI 辅助开发作品的反馈，以及自身对 AI 标签、美术/音效 AI 生成边界的判断。但受访者并非以纯玩家身份体验一站式 AI 游戏平台作品，因此归为消费端补充样本而非核心样本。

2. 玩家反馈与作品评价

《Anchor Drive》在线下 CGJ 站点有交流试玩环节，玩家反馈主要集中在关卡设计太难、无法通关，而对游戏逻辑和美术表现均为好评。这说明 AI 辅助开发的 Demo 在核心玩法和视觉呈现上达到了玩家可接受的水平，主要问题出在关卡难度平衡上，而非 AI 生成内容的质量缺陷。

3. AI 标签与接受边界

受访者会主动告知玩家作品中使用 AI 的部分。他对 AI 标签的态度呈现清晰的层次区分：代码层面的 AI 生成在质量达标的前提下与手写无区别，玩家无需介意；但美术和音效的 AI 生成确实可能引发玩家介意甚至抵制。

这一判断揭示了当前 AI 游戏消费端的接受边界：后台技术（代码）的 AI 化不易引发争议，前台内容（美术、音效）的 AI 化则需要更谨慎地处理，创作者在质量和风格统一性上需要承担更多责任。

七、综合研究结论

受访者 展示了一个典型的 Game Jam 场景下 Codex + Claude Code 联合使用模式：用户有扎实的编程基础和 UE 项目经验，以程序员身份在 48 小时内完成《Anchor Drive》Demo 开发。Codex 作为主力承担核心玩法代码生成（钩锁系统、射线检测等），Claude Code 作为轻量辅助处理工程任务和项目理解。两者的分工使开发效率最大化。

从开发者端角度分析，首先，Codex 在短期 Demo 中的核心价值是“快速推进”——将策划需求迅速转化为可运行代码，并让开发者敢于尝试原本技术难度较高的功能。其次，编程基础是必要条件，受访者的 Review 习惯形成了“中审计、快速迭代”的质量控制模式。最后，付费决策以模型和 Agent 能力为核心考量，附加功能（导出、发布等）对技术型用户吸引力有限。

从消费端角度看，揭示了 AI 游戏接受度的层次性：代码层 AI 化在质量达标前提下不影响评价，但美术和音效层的 AI 化需要更谨慎处理。这一区分对 AI 游戏的市场定位和消费者沟通具有参考价值。

该样本同时支持开发端和消费端两组假设：生产端验证 Codex 能让短期原型快速落地且多轮迭代有效，消费端验证 AI 标签的影响取决于使用层次和最终质量，而非“是否使用 AI”本身。

原始附件

本样本原始材料未包含嵌入图片或附件。

访谈逐字稿

原始材料：访谈逐字稿1.docx

访谈逐字稿：

基本信息

- 导出时间：2026-07-18 0：36
- 访谈时间：2026/07/08 14:32:00 - 2026/07/08 15:31：00

逐字稿

访谈者：

同学你好，那我们现在开始访谈吧！你可以通过打字或者语音回复我，我会将访谈内容收集记录！（本次访谈的内容仅用于课题研究与学习，不会用于其它用途）

受访者：

ok

访谈者：

同学你目前是在校学生吗？目前大几？是什么专业啊？

受访者：

大三，电气专业

claude code我一般是拿来解决小的问题，大部分是用的codex

访谈者：

我看到你填写的问卷了，是用Codex做的demo吗？可以仔细介绍一下你的项目吗？

受访者：

最近不是cgj嘛，我就是拿codex完成了大部分的代码工作

<https://gmhub.com/jams/43/game/9934>

等主题出来，我就等策划先把策划案大概写好，我这边把需求拆分一遍，然后分多个不同的对话喂给codex

访谈者：

这个链接是登录的界面，我看不到耶

受访者：

【2026 CIGA Game Jam 作品：《Anchor Drive》-哔哩哔哩】 <https://b23.tv/aHwogVi>

因为是48小时开发嘛，主要的代码量可能就两三千行左右

访谈者：

哦哦，我看你是电气专业，你有编程基础吗？

你主要是用什么游戏引擎开发的呢？

受访者：

有的，从高中阶段接触得就比较多

主要是用ue

访谈者：

你在你这个demo是负责哪一块工作的呢？

受访者：

程序

访谈者：

如果在没有ai的情况下，你做游戏开发你最容易卡在哪个环节？

受访者：

可能是某些比较困难的功能

比如这次钩锁系统的物理约束什么的

可能需要去查阅一下虚幻的文档什么的，会花很多时间

访谈者：

你有没有因为它而尝试原本不敢做的功能呢？

受访者：

有，而且很多，但是基本上代码我都要review一下，或者让ai给我一个逻辑链

访谈者：

你说Claude code，一般是用来解决小问题，在当前项目中主要是解决了什么问题呢？

Claude code是解决项目中的核心问题，还是边缘辅助问题？

受访者：

因为claudecode比较轻量嘛，打开命令行就能用，我就一般快速打开拿来让它帮我审计一下版本控制的状态，帮我提交什么的，或者某些文件系统方面的问题，基本上不会去解决项目核心的问题

访谈者：

请选择最近一次印象比较深的Claude code使用经历，当时你是想解决什么问题？

这个问题属于哪类任务？

-新功能生成

-旧功能修改

- Bug 调试
- 代码重构
- 性能优化
- 项目理解
- 测试或检查
- 其他

受访者：

项目理解吧，有次clone了一个我很感兴趣的ue插件，文件有点多，我就用claudecode帮我拆解这个插件，让它生成一个完整的技术文档供我复刻的时候参考

访谈者：

你如何判断它是否可以进入项目？你能否完全理解它生成的代码吗？在这个过程中有没有进行过人工的修改？

如果你有进行过人工的修改，修改集中体现在哪里？

1.语法错误

2.逻辑错误

3.变量与接口不匹配

4.与项目结构不兼容

5.性能问题

6.可读性或维护性

受访者：

我主力使用的是codex，没有拿claudecode做过太多代码工作

访谈者：

okk

Codex那你第一次生成的结果怎么样？可以运行吗？你一共生成了多少轮才达到自己想要的结果？

回到你做的demo

受访者：

把需求描述清楚后，用plan模式确认一遍，基本上第一次生成就能完成大部分效果，只有个别细节部分会有bug，可能用三轮左右的对话就能完成功能，比如这次demo，第一次生成完后发现对锚点目标的射线检测通道部分有一点小bug，我给它指正后经过一轮修改就完成了效果

访谈者：

这个平台因为我没有使用过，我想了解一下，在这个过程中是否有付费环节？

受访者：

codex是包月制，中间没有付费

访谈者：

它是一个月多钱是吧

受访者：

嗯

就买个gpt plus就能用了

访谈者：

你使用这个平台的频率是什么样的？你会持续愿意付费吗？还是说选择另外的平台？

受访者：

基本上只要做项目就用吧，如果暂时没有比gpt更强且更便宜的模型，应该会一直续下去

访谈者：

是不经常用是吗？没项目就不用了是吧？还是说每个月会用1-2次

在生产的过程中哪些环节是你觉得最好的？

受访者：

对，主要是做项目用，平时如果有需要会用一些更轻量的

访谈者：

哪些能力最能支撑你付费？

- 快速生成可玩原型
- 多轮修改稳定
- 玩法可控
- 美术和音效质量更高
- 可导出项目
- 可发布到平台
- 可多人协作
- 可持续扩展

哪些问题会降低你的付费意愿？

- 生成质量不稳定
- 修改困难
- 作品方向不可控
- 内容同质化
- 无法导出
- 无法长期维护
- 发布价值不明确

受访者：

我只看重模型和agent的能力，其他的什么一键生成，导出，发布之类的我都不关心

访谈者：

你这个demo，有给到玩家试玩吗？玩家反馈是什么样子的？

受访者：

有的，我参加的cgj线下站点，会有交流试玩环节，玩家反馈主要集中在关卡设计太难，没办法通关什么的，对于游戏逻辑，美术什么的都是好评

访谈者：

我看到你demo中也有一些美术资产和音频资产，这些也是ai生成的吗？

受访者：

那些图片类的资产是ai生成的，包括logo，按钮什么的，然后材质之类的是美术自己搓的

访谈者：

那这个作品你主要还是用ai生成，或者辅助完成的，你会主动告诉玩家这是ai生成的吗？

受访者：

会的

访谈者：

你觉得玩家会介意这类作品是由ai生成的吗？你是否担心ai标签让作品廉价？

受访者：

我是程序，我认为在保证质量的前提下，程序用ai生成和自己写没有任何区别。你这个问题应该问音效和美术，如果音效和美术是ai生成，我认为确实有玩家会介意并且抵制

访谈者：

明白

不考虑现实的限制，你希望当前平台未来可以帮助你解决什么问题？你希望它哪些方面可以变强？

- 1.理解能力
- 2.生成更完整的玩法
- 3.支持复杂的系统
- 4.支持导出项目
- 5.支持长期修改

6.支持多人协作

7.提高美术统一性

8.提高音频和文本的质量

9.支持正式发布

受访者:

1, 6

访谈者:

如果这个平台未来更强，甚至一句话可以生成一个游戏，你觉得创作者还需要负责什么？

你觉得人类创作者在这类平台中最核心的价值是什么？

受访者:

策划案和架构吧

核心价值也是

游戏的策划内容，和程序的架构，这个我觉得是ai暂时做不好的

访谈者:

结合我们之前聊的，我再确认一下，这个平台对你最有帮助的地方是，通过这个平台可以快速实现你当前的想法，并且可以快速推进项目进度，方便快捷的整理资料。那有什么弊端呢？有出现什么问题吗？

受访者:

除了付费价格不低和模型有时候会有降智问题，没有其他问题

访谈者:

降智是指它理解能力不行是吧？

受访者:

嗯，有时候理解能力，工程能力什么的都会下降

访谈者:

后期维护成本高吗？你觉得？

受访者:

有一点，毕竟有token成本

访谈者:

如果它未来更强，你觉得自己在项目中的角色会发生变化吗？

受访者:

会吧，以后的工作可能就变成需求拆解，架构规定之类的高层次的工作，或者偶尔debug一下，手写代码的工作可能会越来越少

访谈者:

Okk

我没有其他要问的，如果后续需要补充的话，我会再联系你的，非常感谢你的参与~

问卷作答情况

以下仅收录研究相关题目，排除联系方式、同意、抽奖、路由、质量控制和技术字段；题干与选项按该答卷对应的历史题库版本呈现。

基础信息与筛选

题号与原题干	作答
S1 你当前的学生身份是?	本科在读
S2 你的专业或主要学习方向更接近哪类?	计算机/软件/AI
S3 过去 6 个月，你是否参与过游戏或互动作品创作?	参与过真实项目/课程项目

题号与原题干	作答
S4 你在游戏/互动作品创作中承担过哪些工作？	项目管理；程序/功能实现；调试测试；玩法/关卡策划
S5 你是否使用过以下 AI 工具/平台参与游戏创作？	Codex / Claude Code 等专业 AI 编程代理；Cursor / GitHub Copilot 等 AI 编程辅助工具
S5A 你从什么时候开始较为频繁地使用 AI 工具？	6 个月—1 年
S5B 你目前使用 AI 工具的频率大约是？	几乎每天都会使用
S6 你使用上述工具时，是否有具体游戏/互动项目或可运行原型作为对象？	有真实项目
S7 你是否实际玩过“一站式 AI 游戏制作平台”生成或实质参与制作的作品？	没有
S8 你能否指出具体平台或作品，并说明自己玩到了什么程度？	能指出平台和作品，且完成一轮完整的核心游玩流程

生产端与创作经历

题号与原题干	作答
P1 请概述你最近一个使用 AI 参与的游戏/互动项目。（项目类型、个人/团队、引擎或技术栈、当前阶段）	cgj 游戏demo。团队，ue5，已在48h内开发完成
P2 你的程序基础更接近哪一项？	能独立完成小游戏/模块
P3 你使用过哪些引擎、框架或开发方式？	Unreal；Web/HTML5
P4 如果不使用 AI，你能否独立完成本项目中的程序或功能实现任务？	能完成但效率低/质量不稳
P5 在该项目中，你具体使用过哪些工具？	Codex / Claude Code 等专业 AI 编程代理；Cursor / GitHub Copilot 等 AI 编程辅助工具
MAIN_TOOL 请选择一个你在游戏/互动项目中使用最深入，或最近主要使用的专业 AI 编程/开发工具。	Codex
CC1 【该工具】你主要让它处理哪些程序任务？	代码生成；修改已有代码；提交说明/协作
CC2 【该工具】它进入项目的哪些阶段？	早期原型；核心功能开发
CC3 【该工具】请评价它对程序工作的帮助。	提高速度：5（非常同意/非常高） 降低调试难度：2 帮助理解项目：4 帮助完成原本不会的功能：5（非常同意/非常高） 减少重复劳动：5（非常同意/非常高） 提升代码质量：3
CC4 【该工具】AI 生成内容后，你通常如何处理？	阅读后小改；运行测试；直接使用
CC5 【该工具】你对生成代码的判断与维护能力如何？	能判断是否正确：5（非常同意/非常高） 能定位问题：5（非常同意/非常高） 能解释代码：4 能后续维护：4 能控制依赖风险：3

题号与原题干	作答
CC6 【该工具】相关产出最终是否进入项目？	大部分进入
CC7 【该工具】整体看，它主要是提高已有能力，还是补足原本不足的能力？	两者都有但任务不同
P6 整体来看，AI 工具让你的游戏程序/原型开发发生了哪些变化？	开发速度更快：5（非常同意/非常高） 能做更复杂任务：5（非常同意/非常高） 更依赖 AI：3 人工检查更多：4 返工更多：3 更愿意尝试新玩法：5（非常同意/非常高） 更容易形成可玩结果：5（非常同意/非常高）

开放题

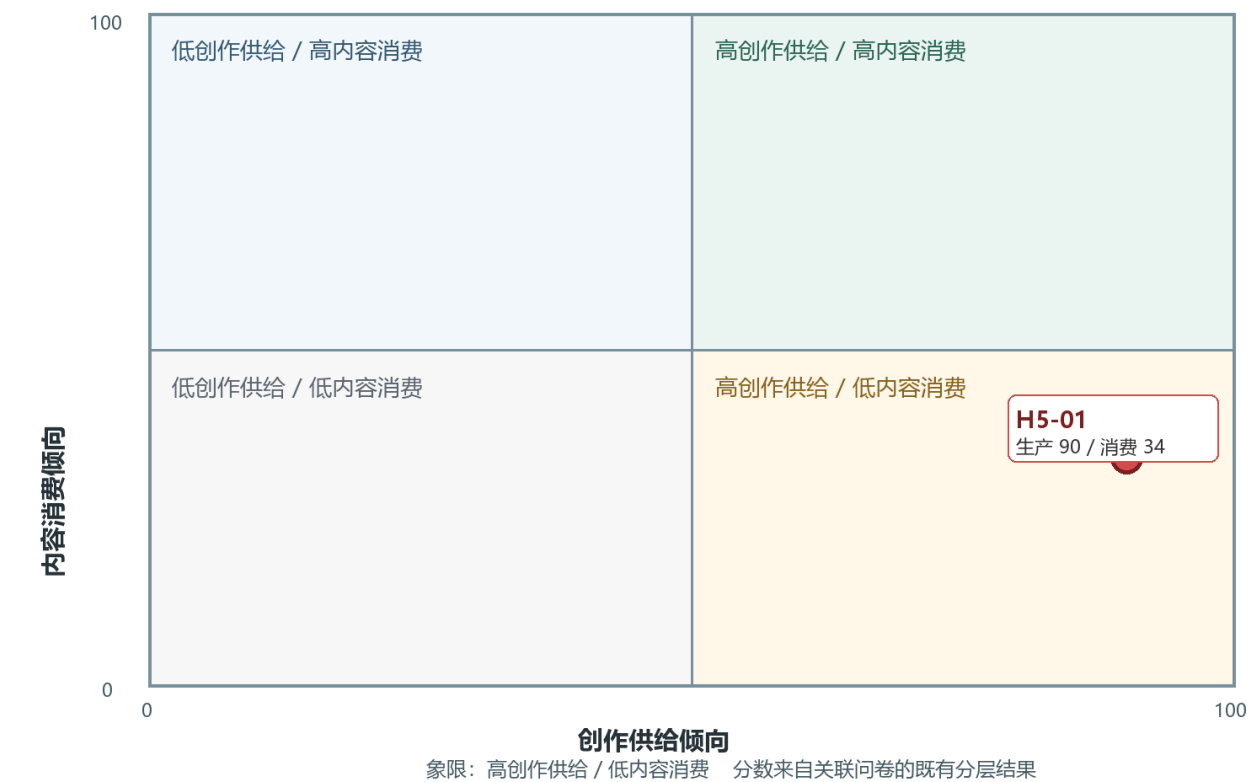
题号与原题干	作答
P7 请还原一个最典型片段：你要做什么程序/功能任务，使用了哪个工具，AI 给出什么，你如何修改，最后结果如何？	实现某类暂时没有具体实现思路的需求，先把需求给ai写出原型，验证产出内容与需求无偏差后会自己或ai辅助阅读源码理解逻辑链路并向ai提出修改意见

研究补充

题号与原题干	作答
G1 你总体上如何看待 AI 进入游戏创作？	提高效率：5（非常同意/非常高） 降低门槛：5（非常同意/非常高） 带来新玩法：3 削弱创作诚意：2 增加同质化：5（非常同意/非常高） 带来版权风险：3
G2 你认为 AI 参与程度是否应该被明确标注？	重要环节应标注
Q4.3 如果你没有实际玩过一站式 AI 平台作品，主要原因是什么？	担心质量差；不感兴趣

生产倾向 × 消费倾向象限图

横轴为创作供给倾向，纵轴为内容消费倾向；50 分为高低分界。



H5-01：生产 90，消费 34，高创作供给 / 低内容消费

关联维度分数

创作	95
效率	88
探索	70
接受	40
付费	39
信任	48

A.2 样本二 | H5-02

样本基本信息

样本编号	H5-02
样本定位	兼具生产端与消费端信息的复合型样本，围绕多个 Game Jam Demo、跨工具协作、AI 美术与 AI 角色扮演产品体验展开。
访谈执行人	姓名待补
摘要来源	访谈摘要：H5-02.md
逐字稿来源	H5-02_transcript.md
关联答卷	0af2e92e-adeb-47df-80ef-9c091ef8f312

本文件保留该样本的访谈摘要、逐字稿和源材料附件；最后附上关联问卷中的研究相关作答与既有象限结果。

访谈摘要

原始材料：访谈摘要：H5-02.md

访谈摘要：H5-02

一、访谈基本信息

- 受访者编号：H5-02
- 访谈时间：2026/07/08 14:55:00 - 2026/07/07 16:20:15

二、样本判定

本访谈对象可判定为 Codex/AI Agent 游戏开发补充样本。受访者为软件工程专业学生，有 Game Jam 游戏 Demo 开发经历，主要在短期项目中使用 Codex 生成和修改代码。其使用场景包括 Unity 中的玩家控制、游戏对象创建、脚本挂载、移动逻辑、变量暴露和 Bug 修改等。

但受访者没有提供完整的单一项目链路、具体代码片段、提交记录、Demo 文件或多轮 Codex 输出记录。受访者对使用过程的描述较偏概括，强调“能用就行”“出 bug 再让它改”。因此，本样本更适合作为补充样本，用于观察“有程序基础、以 Game Jam 快速出 Demo 为目标”的 Codex 使用模式。

- 样本标签：TOOL-C
- 判定理由：有真实 Game Jam 项目经历和具体代码任务类型，Codex 产出会进入短期 Demo，但缺少完整过程材料和可追踪代码链路
- 具体项目：多个 Game Jam Demo，包括平台跳跃、类银行与恶魔城、三消、剧情游戏等
- 项目阶段：短期 Demo / 原型阶段
- 过程材料：逐字稿和 1 张 AI 美术相关附件

同时，从 AI 游戏消费者访谈提纲看，H5-02 也可作为消费端补充样本。受访者有较丰富的游戏消费经验，实际体验过 AI 角色扮演/酒馆类产品，并能围绕 AI 游戏的玩法形态、文本表现、模型质量、付费方式、AI 标签和创作者主导性作出具体判断。

- 消费端样本标签：CON-S
- 判定理由：有实际 AI 游戏/AI 角色扮演体验和具体消费判断，但缺少明确一站式 AI 游戏平台作品、作品名称和完整体验链路
- 体验作品：酒馆类 AI 角色扮演产品；Steam 上类似酒馆形态的 AI 游戏被作为评价对象提及
- 对应平台：Steam
- 游玩时长：未明确；受访者确认曾以玩家身份体验 AI 角色扮演
- 是否知情 AI 参与：是

三、开发者端样本分析

1.基本背景与开发基础

受访者为软件工程专业学生，有明确编程基础，以程策身份参与Game Jam 类短期 Demo开发。由于缺少合适的长期团队，他的开发经验集中在命题式、时间受限、以快速做出可运行结果为目标的项目中。

受访者对自身能力的定位比较务实：他能理解项目结构和脚本功能，但不一定会深入理解 Codex 生成代码中的每个循环或具体实现细节。他通常只要求自己“大概明白框架，每个代码实现什么功能即可”。这意味着他不是完全依赖 AI 的零基础用户，而是以已有编程基础作为判断和补救条件，将 Codex 当作快速实现工具使用。

该案例支持提纲中的基础假设：用户能否用好 Codex，很大程度取决于编程基础和项目经验。受访者明确表示，有程序基础是做出可用功能的条件；如果不懂程序，至少也要理解游戏引擎里的项目结构和大致流程，否则小问题不断累积后，人工插手可能接近一次项目重构。

2.具体项目还原

访谈中没有围绕一个完整项目展开细节复盘，而是描述了受访者过往参与的多个 Game Jam Demo。项目类型包括平台跳跃、类银行与恶魔城、三消和剧情游戏。这些项目共同特征是周期短、质量要求相对低、重点是尽快形成可运行内容。

在当前 Codex 使用中，最典型的项目场景是 Unity 游戏原型。受访者会让 Codex 处理玩家控制、创建游戏对象、挂载脚本、实现移动逻辑等基础功能。例如他会要求 Codex 把影响玩家的变量公开声明在脚本上，使速度、跳跃时施加的力等数值能够在 Unity Inspector 中调整。对 Game Jam 来说，这类功能虽然不一定复杂，但直接关系到 Demo 能否跑起来，因此不是完全边缘的辅助任务。

受访者也提到，较复杂的功能机制以前更多依赖手写或网页版 AI 参考。例如早期三消项目中加入机制，对他而言已经属于复杂代码；当时 Codex 还不够智能，所以主要是手搓。这说明当前材料更能证明 Codex 在短期原型和基础功能层面的落地价值，而不是复杂长期项目中的架构能力。

3.具体任务使用过程

H5-02 的 Codex 使用链路较为粗放，核心策略是“先描述功能，再让 Codex 实现”。受访者认为提示词不需要非常复杂，只要把具体功能描述清楚即可。如果自己不知道怎么描述，会先把需求告诉 GPT 网页版，让 GPT 生成适合 Codex 的提示词，再交给 Codex 执行。这形成了一个工作流：通用聊天模型辅助整理需求，Codex 执行代码生成或修改。

Codex 的第一次输出并不总是可用，受访者也并不追求一次成功。他的处理方式是：准不准无所谓，先让 Codex 做；如果没弄好就自己改；如果出现 Bug，就再让 Codex 改。该链路 with 参考提纲中的多轮协作假设一致，但 H5-02 的特点是人工测试和追问过程没有被细化记录，更多体现为速度优先、结果够用即可。

受访者指出，Codex 输出不符合预期非常常见。典型问题是“偷懒”：当一次性描述过多，或者给出一份详细修改文档时，Codex 往往只改用户强调的部分，忽略其他要求。另一个问题是实现方式可能偏离 Unity 的常规组件体系，例如不用 Unity 组件而是自己写脚本，导致后续与其他功能耦合很麻烦。遇到这类问题时，受

访者通常会继续问 AI、找外援，或者在 Game Jam 场景中直接放弃处理。

4.效率提升与能力补足

受访者认为 Codex 在 Game Jam 中的核心价值是提高完成速度。Game Jam 的首要目标是先做出来，代码质量和长期维护不是第一优先级。受访者甚至表示，在 Game Jam 中用了 AI 后，“不管你做出来的好不好，肯定是能做出来的”。

在具体任务上，Codex 对玩家控制、基础脚本、对象逻辑和 Bug 修改最有帮助。受访者作为程策方向的开发者，通常只需要把原型做出来，并处理一些 Bug；这类需求与 Codex 快速生成常规代码的能力高度匹配。因此，Codex 对他既是效率工具，也在一定程度上补足了原型开发中的实现能力。

但受访者并不认为 Codex 可以让完全没有基础的人独立完成游戏开发。他明确提出“AI 只能让强者更强”。在他的理解中，Codex 能加速已有能力，而不是替代用户对项目结构、引擎流程和基础代码逻辑的理解。该样本因此支持“能力补足可能带来依赖风险”的假设：如果用户不理解技术方案，就可能发现不了 Codex 留下的结构性问题。

5.人工控制、判断与维护成本

受访者对 Codex 输出的人工控制主要停留在结果层面，而非逐行审查。他通常只确认脚本大致负责什么功能、游戏对象能否按预期行动、参数能否在 Unity 中调整。对于内部实现细节，他不会强迫自己完全理解。这种判断方式适合短期 Demo，但对长期维护存在风险。

维护风险主要体现在两个方面。第一，Codex 可能遗漏需求或只完成被强调的部分，导致功能表面完成但细节缺失。第二，Codex 可能在用户不了解技术方案时选择不合适的实现方式，例如绕开 Unity 组件自行写脚本，进而增加系统耦合。受访者承认，有些问题在自己不了解技术时甚至不会被及时发现。

因此，H5-02 不是一个“高控制、高审计”的使用样本，而是一个“低审计、快速试错”的样本。对短期 Game Jam 来说，这种模式可以接受，因为目标是快速完成；但如果迁移到长期项目，小问题会不断累积，后续人工维护成本可能显著上升。

6.产出落地与项目影响

Codex 的产出在 H5-02 的 Game Jam 开发中能够进入项目，尤其是玩家控制、基础玩法脚本和 Bug 修改等任务。它对项目的影响不是体现在复杂架构设计，而是体现在把功能快速从想法推进到可运行状态。对受访者来说，Codex 的价值首先是“能做出来”，其次才是“做得多好”。

该样本更能说明 Codex 适合短期 Demo 和原型验证，而不是直接证明其适合长期复杂项目。受访者没有长期团队项目经验，也没有在访谈中展示 Codex 处理大规模工程结构、多人协作或复杂系统重构的过程。相反，他多次强调 Game Jam 的质量要求和维护压力较低，这也是 Codex 能被放开使用的重要前提。

从产出落地角度看，H5-02 支持一个较明确的判断：Codex 产出能否进入项目，不只取决于模型能否生成代码，也取决于项目是否允许“够用即可”的质量标准。短期 Demo 容忍度高，因此更容易落地；长期项目容忍度低，则需要更强的人工控制和结构约束。

7.留存、迁移与付费

受访者尝试过多种 AI 工具。其 Claude 账号已被封禁，目前主要使用 GPT 订阅中的 Codex，同时也尝试过 Grok、Gemini 和 DeepSeek API。受访者对 Grok 的 build 能力评价较低，并表示后悔订阅；Gemini 订阅来自 Pixel 用户的低价方式，但他预计账号稳定性可能不高；DeepSeek API 则更多是在额度不足时作为替代。

他选择 Codex 作为主要代码工具的原因是“用起来简单，教程多”。如果未来不再使用订阅，转向 API 时，他也会考虑别的平台。这说明 Codex 已经进入其常规开发流程，但留存基础主要是便利性、成本、额度和可获得性，而不是不可替代的单一技术优势。

付费方面，受访者当前使用 GPT Plus 订阅，并把 Codex 作为代码开发主力。其付费判断较务实：只要工具能快速完成代码任务、降低 Game Jam 原型开发成本，就有持续使用价值；但一旦价格、额度、封号风险或实际效果不理想，他会迁移到其他 AI 工具。

8.AI 参与制作的市场判断

受访者对 AI 参与游戏制作整体持开放态度。他认为游戏的本质仍然是好玩，核心在于玩法、数值平衡和美术表现力，而不是是否使用 AI。玩家更容易感知 AI 文案和 AI 美术，代码层面的 AI 参与通常不明显。只要开发者仍然主导审美、架构和最终质量，AI 辅助本身不构成问题。

在美术方面，受访者认为当前自己使用 AI 美术效果不好，主要原因是个人绘画基础弱，难以发挥 image2、nanobanana 等工具的能力。他并不排斥 AI 美术，如果 AI 能生成风格统一、符合玩家审美的简易素材，他愿意使用。该判断说明他对 AI 美术的接受条件不是“是否 AI”，而是风格一致性和最终效果。

在文案方面，受访者认为 AI 可以用于角色卡、角色对话、多结局分支和场景描述，再由开发者润色并插入更鲜明的人物表达。他也认可 AI 辅助开发游戏具有付费价值，尤其期待 AI 驱动的动态世界、长期记忆 NPC、玩家自定义 API 接入、行为树战斗逻辑、物理交互和环境解谜结合的游戏形态。

受访者对 AI 参与边界的总结较清晰：如果是 AI 辅助加速，且开发者主导审美和架构，就没有明确边界；但如果是纯 AI 主导，整体效果仍然“不大行”，最多适合做着玩。

9.理想中的 Codex/AI 开发工具

受访者并不期待 Codex 成为全能工具。他认为不同 AI 有不同擅长点，追求单一全能模型并不现实。相比让 Codex 本身同时解决代码、建模、绘图、文案等全部问题，他更期待多个 AI 工具之间形成跨厂商、多智能体协作。

具体来说，他希望 Codex 能与建模 AI、绘图 AI 等外部工具协同，让“Codex 能进入更完整的游戏开发链路”，在开发者主导下协调不同 AI 工具完成代码、素材和其他生产环节。

这一点与提纲中的理想工具假设一致：未来的 Codex/AI Agent 不应只是代码生成器，而应更像开发工作流助手。H5-02 的补充之处在于，他特别强调跨厂商、多智能体协作，而不是单模型全能化。

四、开发者端访谈记录字段整理

- 受访者编号：H5-02
- 样本类型：Codex/AI Agent 游戏开发者
- 样本标签：TOOL-C
- 是否在校大学生：是
- 专业与年级：软件工程专业；年级未提及
- 编程基础：有软件工程背景
- 使用引擎或开发工具：Unity、Codex、GPT/ChatGPT、Grok、Gemini、DeepSeek API、image2、nanobanana
- 具体项目或作品：多个 Game Jam Demo，包括平台跳跃、类银行与恶魔城、三消、剧情游戏
- 项目类型：Game Jam
- 当前阶段：短期 Demo / 原型
- 工具使用时间：GPT plus长期订阅
- 使用任务类型：代码生成、功能实现、代码修改、Bug 修复、提示词辅助生成；部分 AI 文案和 AI 美术尝试
- 是否涉及至少两类实质任务：是
- 产出状态：在 Game Jam Demo 中进入项目，主要进入基础玩法和玩家控制相关代码
- 人工修改程度：低到中，主要通过再次询问 AI、少量手动修改或放弃低优先级问题解决
- 是否提高效率：是，尤其体现在短期 Demo 快速做出可运行功能
- 是否补足能力：是，但强依赖用户已有编程基础和项目结构理解
- 是否增加维护成本：在复杂耦合和长期项目中有明显风险；短期 Game Jam 中通常被容忍
- 是否持续使用：是，代码方面一般使用 Codex
- 付费判断：已使用 GPT Plus 订阅；愿意根据价格、额度、稳定性和实际效果迁移到其他工具
- 是否有过程材料：逐字稿，其余未提供
- 主要有效场景：Game Jam 原型、Unity 基础脚本、玩家控制、可调参数暴露、常规 Bug 修改、AI 辅助文案草稿
- 主要失效场景：AI 美术、复杂功能耦合、一次性需求过多、用户不了解技术方案时的实现偏差、长期项目维护
- 可用于验证或反驳的研究假设：验证“Codex 在短期游戏原型中能显著提效”“用户基础影响 AI 产出落地”“真实使用不是一次生成，而是功能描述、生成、试错和再修改”；同时补充说明，在低质量压力的 Game Jam 中，用户可能并不要求完全理解代码即可让产出进入项目。

五、消费者补充样本分析

1. 消费端样本定位

H5-02 符合消费端补充样本条件，而不是消费端核心样本。其核心原因是：受访者确实以玩家身份体验过 AI 角色扮演/酒馆类产品，并能评价这类产品的模型质量、文本体验、价格、付费方式和未来玩法潜力；但访谈没有锁定一款明确的一站式 AI 游戏制作平台作品，也没有提供作品名称、平台来源、完整游玩时长和逐项体验评价。

因此，H5-02 的消费者价值不在于还原某款 AI 平台作品的具体游玩过程，而在于呈现一类熟悉游戏、熟悉 AI 工具、同时具有开发者视角的玩家如何判断 AI 游戏作品。他的评价标准不是单纯“是否使用 AI”，而是游戏是否好玩、AI 是否服务玩法、创作者是否仍然主导审美和系统设计，以及价格和模型质量是否匹配。

2. 玩家背景与参照系

受访者平时玩独立游戏，也玩手游和网游。单机游戏在投入时可以一天玩 8 小时，忙的时候则几乎不打开；手游和网游更多是碎片时间或有人邀请开黑时游玩。他最近提到购买了《盛世天下》，并表示自己“玩独立游戏，啥主流玩啥”。这说明他不是只从开发者角度讨论 AI 游戏，而是具备较稳定的玩家经验和独立游戏参照系。

在比较不同类型作品时，受访者会区分“业余时间玩什么”和“工作学习时看什么”。如果同类赛道中有独立游戏和 AI 参与的商业游戏，业余时间他更倾向独立游戏，因为独立创作更有趣、更“酷”；但如果出于学习目的，他会选择商业游戏，因为商业项目能提供从立项、宣发、项目管理到版本运营的完整经验。这说明他评价 AI 游戏时会根据消费目的调整参照系：娱乐消费看体验和气质，学习研究看工业流程和可借鉴性。

3. AI 角色扮演与文本体验

受访者实际体验过酒馆类 AI 角色扮演产品。他对这类产品的理解比较具体：可以用角色卡模拟不同角色，让用户扮演主角，通过多轮对话生成对白和场景描述，再由创作者润色，并补充更鲜明的人物表达。他认为在多结局分支剧情中，AI 可以帮助维持不同结局的人设一致性，虽然生成文本的人物特点可能不够鲜明，仍需要人工补写和修饰。

从消费端看，受访者认可 AI 文本互动的潜力，尤其看重开放剧情、角色一致性和持续对话能力。但他也指出当前部分酒馆类产品存在模型质量和价格不匹配的问题：一些 Steam 上的酒馆游戏使用的模型质量不高，却收取较高 token 费用；如果连 DeepSeek 级别模型都算相对较好，那么更弱模型的高价格会显著降低玩家信任和付费意愿。

4. 玩法期待与 AI 游戏潜力

受访者认为 AI 辅助开发的游戏有付费价值，并且预期 AI 游戏会越来越普及。他最期待的不是简单的 AI 生成小游戏，而是 AI 驱动的动态世界：NPC 具有长期记忆和行为逻辑，能记住玩家曾经的行为，并基于这种记忆触发复仇、追踪或其他交互；战斗部分仍可由预设行为树或传统系统承载，AI 主要负责动机、记忆和情境判断。

他进一步设想了一类融合 CRPG、3D ARPG 和开放物理交互的 AI 游戏：玩家可以导入自己的 API，AI 像 MCP 接入游戏引擎一样读取周围可交互元素，NPC 能搭陷阱、解谜、利用环境，并与玩家形成更复杂的互动。这一判断说明，在 H5-02 看来，AI 游戏最有潜力的方向不是替代传统玩法设计，而是创造动态叙事、长期记

忆、个性化交互和开放系统之间的新组合。

5. AI 识别、标签与创作投入感

受访者并不天然反感 AI 参与游戏。他认为玩家最终更在意游戏本身是否好玩，AI 参与与否不是决定性因素。尤其在代码层面，玩家通常察觉不到 AI 参与；感知最明显的是美术和文案。如果这些内容粗糙、模板化或缺乏统一风格，玩家更容易将问题归因于 AI。

受访者强调关键在于“开发者主导 AI 的审美”，而不是让 AI 主导作品。他认为很多小团队使用 AI 是因为缺少强美术或完整分工，AI 至少让原本更差的成品有了被评价甚至被批评的资格。这一表达说明他对 AI 作品的接受边界不是“不能使用 AI”，而是作品背后是否有人进行选择、控制和打磨。如果创作主体消失，只剩下 AI 快速拼接，作品价值会下降；如果 AI 是在人类设计意图下加速生产，则可以接受。

6. 付费、分享与接受边界

受访者愿意承认 AI 游戏有付费价值，但付费前提是玩法、模型能力、内容开放性和价格结构匹配。他对 AI 角色扮演产品的付费设想较具体：纯文字可以采用低价月订阅，加入语音或图片后价格可以提高，也可以按量收费、出售 token 包或算力包。这说明他并不排斥 AI 游戏商业化，但会对模型质量和 token 定价非常敏感。

在接受边界上，受访者的立场较清晰：如果 AI 是辅助加速，且开发者仍主导审美、架构和最终体验，那么 AI 参与没有明确边界；如果是纯 AI 主导，则“啥都不太行”，最多适合做着玩。这与消费者提纲中的假设一致：玩家的接受边界不是简单的“能不能用 AI”，而是 AI 参与程度是否透明、是否影响体验、是否仍然存在创作者控制感。

六、消费者访谈记录字段整理

- 受访者编号：H5-02
- 样本类型：AI 游戏作品消费者补充样本
- 样本标签：CON-S
- 是否为大学生：是
- 平时游戏频率：单机游戏投入时可一天约 8 小时，忙时很少打开；手游和网游多为碎片时间或社交开黑
- 常玩游戏类型：独立游戏、单机游戏、手游、网游；对剧情互动和 AI 角色扮演有体验
- 体验作品名称：提到酒馆类 AI 角色扮演产品和 Steam 上类似酒馆形态的 AI 游戏
- 对应平台：Steam
- 是否实际游玩：是，确认以玩家身份体验过 AI 角色扮演/酒馆类产品
- 游玩时长：未明确
- 是否完整体验：是
- 游玩前是否知情 AI 参与：是
- 游玩中是否识别 AI 参与：是，AI 角色扮演和酒馆类产品本身以 AI 文本互动为核心
- AI 识别依据：文本互动方式、角色卡、模型质量、平台说明、token 计费方式
- 玩法评价：认可 AI 动态叙事、长期记忆 NPC、开放 API 接入和环境交互潜力；对单纯低质量文本互动或弱模型包装的产品评价较低
- 美术评价：认为 AI 美术玩家感知明显，当前个人使用体验不佳；可接受风格统一、审美达标的 AI 美术
- 文本评价：AI 可辅助角色对白、多结局一致性和场景描述，但需要人工润色和强化人物特色
- 完整度评价：未针对具体作品评价完整度；对当前酒馆类产品主要批评集中在模型质量、价格和开放性
- 耐玩性评价：期待长期记忆、动态世界和开放交互提升耐玩性
- 是否愿意继续玩：愿意尝试有玩法、有长期记忆和动态交互的 AI 游戏
- 是否愿意付费：愿意，前提是玩法和模型质量足够好、价格合理、内容体验匹配
- AI 标签是否影响评价：AI 标签本身不是决定性负面因素；若 AI 导致美术、文本或内容明显粗糙，会降低评价
- 主要接受边界：开发者主导审美、架构和质量，AI 用于辅助加速或增强动态内容
- 主要不接受边界：纯 AI 主导、缺乏人工筛选、模型质量低但收费高、内容粗糙或缺乏设计感
- 可用于验证或反驳的假设：验证“玩家不一定反感 AI，真正影响接受度的是质量、完整度、创作者控制感和价格匹配”；补充说明 AI 游戏的消费潜力可能集中在动态叙事、个性化互动、长期记忆和 UGC/开放 API 场景。

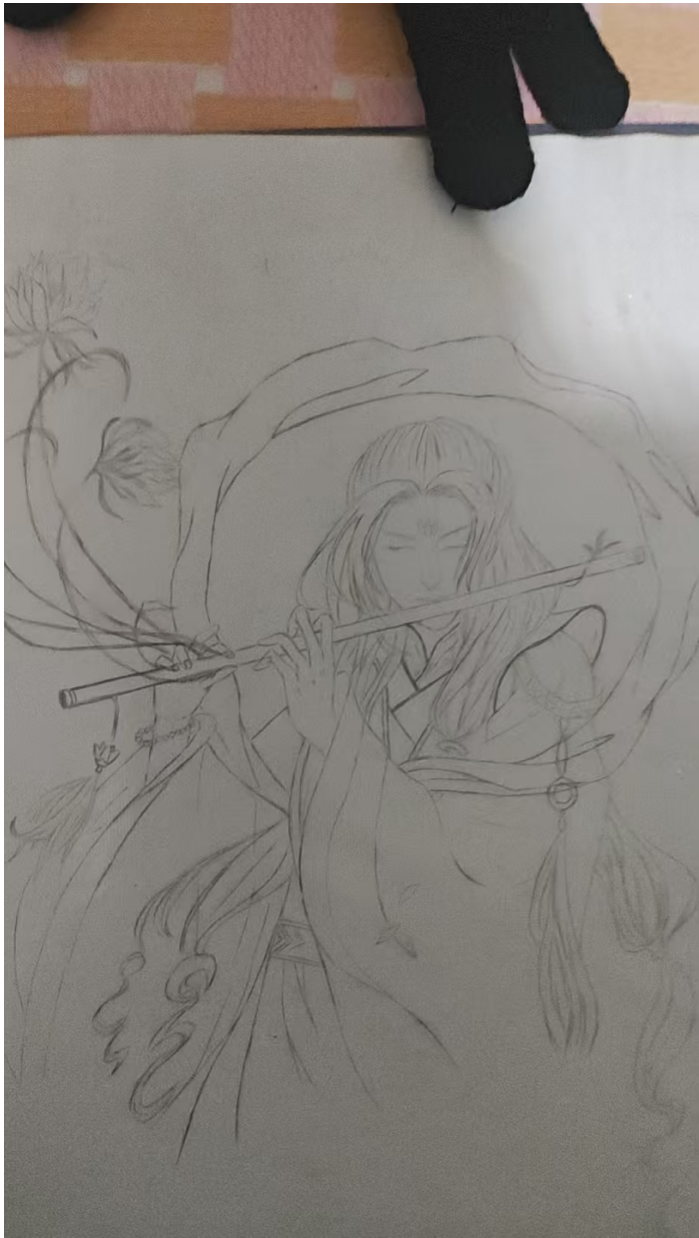
七、综合研究结论

H5-02 展示了一种典型的 Game Jam 场景下的 Codex 使用模式：用户有一定程序基础和 Unity 项目经验，但目标不是构建长期可维护工程，而是在有限时间内快速做出可运行 Demo。Codex 在其中承担基础代码生成、功能实现和 Bug 修改任务，显著降低了原型开发成本。

从开发者端角度分析，首先，Codex 在短期 Demo 中的价值首先是“推进项目跑起来”，而不是生成高质量长期工程代码。其次，用户的编程基础仍然是必要条件，因为只有理解项目结构和基本流程，才能发现或补救 AI 造成的遗漏、耦合和实现偏差。最后，AI 参与游戏制作的接受度取决于开发者是否仍然主导审美、架构和质量；如果只是辅助加速，受访者认为边界很宽，但纯 AI 主导仍难以产生可靠作品。

H5-02 同时呈现出“开发者-玩家双重身份”的样本价值。作为玩家，他并不反感 AI 游戏，也认可 AI 游戏的付费潜力；但其接受前提是玩法成立、模型质量足够、价格合理，并且能感受到人类创作者的选择和控制。该样本因此可以同时支持开发端和消费端两组假设：生产端看 Codex 是否能让短期原型落地，消费端看 AI 作品是否能在质量、完整度、创作者投入感和付费价值上成立。

原始附件



原始附件 1: H5-02_transcript.md (810 × 1440)

访谈逐字稿

原始材料: H5-02_transcript.md

访谈逐字稿: H5\02

基本信息

- 导出时间: 2026\07\08T09:32:18\017Z
- 访谈时间: 2026/07/08 14:55:00 \- 2026/07/07 16:20:15

逐字稿

访谈者:

我会对接下来的聊天记录存档，用于研究分析，会对个人信息做保护，ok的话，那我们开始了

受访者:

可以

访谈者:

你有编程基础吗，或者专业是码类相关的吗

受访者：

我是软件工程专业

访谈者：

具体有参与过哪些游戏开发项目呢，可以说下项目的类型或者主要玩法吗

受访者：

学生，没参与过实际项目，都是game jam的demo

做过平台跳跃，类银行和恶魔城，三消，剧情游戏

访谈者：

参加的都是有命题、短期那种gj吗

受访者：

对的

主要是没有合适的长期团队

访谈者：

这些gj中有涉及ai工具的使用吗

受访者：

那必须的，代码大多数时候都会涉及AI

game jam嘛质量不是那么重要，反而先做出来是第一要务

用了AI，不管你做出来的好不好，肯定是能做出来的

访谈者：

主要使用的是哪些ai呢，有付费情况吗

受访者：

我的Claude账号被封禁了

现在主要是GPT的订阅

codex

grok还没过期，不过grok build不好使，后悔订阅了

访谈者：

codex是plus订阅吗？

受访者：

是的

Gemini的订阅是pixel用户的灰色方式，16块钱卡了一年

不过我估计用不了几个月，账号还会被封禁

另外，deepseek的API也有使用

访谈者：

好的，那几乎都有尝试过啊

可以具体说下codex参与的功能吗，就比如玩家控制部分的功能实现等等

以及都有用这些ai参与过哪些功能的开发呢

受访者：

玩家控制这种功能，创建游戏对象，之后，对象上挂个代码，移动逻辑啥的，直接让codex实现就行

因为我是程策方向的，顶多做一下原型，bug程序改

一般来说，还是codex好使

GPT5\5

deepseek这种一般是没额度之后去用

访谈者：

所以大部分涉及ai的工作还是给codex完成的？

受访者：

是的，我属于能用就行流派，代码方面一般是codex

不过美术不大行

拉跨

访谈者：

好的

可以具体聊下这方面功能的实现过程吗，就比如每轮和codex的对话、输出结果、提示词之类的

受访者：

实话实说，具体一点描述功能即可，主要是，让codex把所有影响玩家的变量，比如速度，跳跃时施加的力等等这种数值，公开声明在脚本上，unity里可以更改准不准无所谓

如果他没弄好就自己改，出bug让他改，就这样

主要是，怎么快怎么省事儿怎么来

有这么复杂吗？如果不知道用啥描述，把功能和gpt网页版说一下，让网页版给提示词

我一般都是这么干的，AI最懂AI

访谈者：

好的，那整个项目的代码实现部分没那么复杂？或者说不涉及一些复杂功能间的耦合之类的？

受访者：

看情况吧，复杂代码的话，我好久没做了，之前三消里面加点机制这种对我而言属于复杂的了

但是当时还没有用AI，那时候还是手搓，用网页版出代码参考，当时codex还不是很智能

访谈者：

那你会要求自己理解codex生成的代码吗

受访者：

一般不会，大概明白框架，每个代码实现什么功能即可，内部的具体循环我不去理解

访谈者：

好的，那有遇到过输出结果不符合要求或者预期的情况吗

受访者：

太常见了

访谈者：

一般会怎么解决呢，会涉及人工修改吗

受访者：

codex经常偷懒，比如你一次给他描述的过多了，或者扔一个详细的要更改的文档给他，然后，他只改你强调的部分，别的很多都忽略了

访谈者：

可以举个具体例子描述一下吗

受访者：

比如不用Unity的组件，自己写了个脚本，之后和其他功能的耦合就很麻烦

有时候你不了解一个技术，没给详细的提示词让他做，之后他就会这样

AI只能让强者更强

访谈者：

那遇到这种问题你会怎么解决呢

受访者：

问AI怎么回事，或者呼叫外援，或者摆烂不去管他了

有时候这种问题，我不了解这个技术，我都没发现，毕竟我不是什么大佬

大多数人都这样吧

访谈者：

那作为有程序基础的开发者，你这方面的能力能帮你更高效地使用ai进行开发吗

受访者：

我觉得，有程序基础，是做出能用的功能的条件

没基础的话，纯AI，至少游戏引擎里，是不大能用的

或者说，你不懂程序，你也要明白项目结构和大致流程

否则一些小问题累积之后，再想去改就很离谱了，人工插手等同于项目重构

访谈者：

刚才你也有提到，你使用过各种ai平台，那具体的使用体验有什么差异吗，是什么让你选择了codex作为主力呢

受访者：

没有，就是Codex用起来简单，教程多

要是以后不用订阅，换API，也会去考虑别的平台

访谈者：

那从整个项目的开发来看，你觉得有哪些部分是可以交给ai的，哪些完全不可以呢

受访者：

我觉得，不管那些部分，AI都可以插手，你给他搭好大致框架，描述一下具体功能，再去改呗

除了美术，真难用，也可能是我用的AI不行

访谈者：

有用过哪些ai美术工具呢

受访者：

image2，我绘画很差

以我的水平，纯AI绘图很难用

![a16a93d02e5c5dbf7d603e8074b0df73\jpg](图片和附件/a16a93d02e5c5dbf7d603e8074b0df73.jpg)

我临摹这样，原创更垃圾

或许专业的会分图层绘画的AI比较好用？我没用过

也可能是我水平垃圾，image2我发挥不出来，nanobanana也差不多

访谈者：

那如果ai美术可以协助生成一个比较简易但很符合很多玩家审美的图，你会使用吗

受访者：

风格统一就用，不挑

访谈者：

因为现在有很多玩家很抵触有ai深度参与的游戏，你会担心自己的游戏因为涉及ai而不受玩家喜欢吗

受访者：

我觉得游戏的本质还是好玩，玩法呀，数据平衡呀，美术的表现力呀，跟AI关系不大。很多人用AI是因为创作的主体不再是开发者了，应该是开发者主导AI的审美，而不是说用AI偷懒，做出一个勉强能用的东西来

加上很多的小团队，他们没有很详细的分工，比如没有一个很强的纯美术，就只能让AI来主导美术，也算是迫不得已吧，之前他们的成品更差，但现在有AI让他们的成品有被骂的资格了，这也算是AI的一个进步

其实我觉得AI代码这方面玩家察觉不太出来，感知明显的只有文案和美术

而且吧，大厂的游戏开发也用AI啊，不用AI的团队真的很少了，我觉得

访谈者：

那你个人有尝试过ai文案吗

受访者：

就是酒馆吗？角色卡，然后你用角色卡来模拟你出场的每个角色，之后用户模拟主角，然后一顿对话，再把AI写的对白和场景描述什么的润色一下，就可以拿去当文案了，之后再把每个角色鲜明的对话插进去，突出人物角色特点的对白插进去就好了

比如多结局分支的话，一个多结局分支的游戏，保持不同结局人设的一致性。这方面可能AI做的也不比人差，只不过可能没有那么鲜明的对话特点，可以自己去插吗这方面

这样就行了，主要是这样方便，快

反正做着玩

访谈者：

从玩家视角来看，你觉得ai参与的游戏有付费价值吗，或者会为哪种ai游戏付费呢

受访者：

很多种吧，我觉得AI做的游戏会越来越普及，付费价值肯定有的

我最想玩AI驱动的游戏，世界是动态的，最好支持玩家自己导入API

访谈者：

像酒馆这种的？

受访者：

又不一样，要有玩法，然后行为逻辑和长期记忆是AI，战斗逻辑啥的，可以说写好的行为树啥的，（这俩冲突吗？我了解不深）

比如AI判断，玩家曾经抢劫过我，然后要去复仇，到玩家周围，调用战斗部分的行为树

我觉得，大厂有机会实现

访谈者：

相当于用ai做npc的角色扮演这种吗

受访者：

差不多，更强的博德之门和3DARPG的结合体

当然，如果还有很强的物理效果，环境交互就更好了，玩家的API像是MCP接入游戏引擎一样，进入游戏中，之后get到周围那些元素可以交互

会自己搭陷阱，做一些和玩家一样的解密，这样博德之门这种CRPG\+塞尔达的物理交互

不过我估计没人会去做，还需要很多年

访谈者：

你有玩过别的ai游戏吗

受访者：

不太玩，我玩独立游戏，啥主流玩啥，前几天刚买了盛世天下

访谈者：

那你每天或每周的游戏时长是多久呢

受访者：

这要拆开看了，手游网游和单机，单机，玩进去的时候，一天8小时也有，忙的时候，根本不去打开单机游戏

手游网游，碎片时间和有人找我开黑

访谈者：

假设有两个类型一样、赛道都差不多的游戏，一个是独游小团队的，另一个是ai参与的商业游戏，你会选择哪款呢

这里的ai参与不仅仅指代码参与，文案啥的都有

受访者：

如果是业余时间的话，我选独立游戏

如果是工作的话，我选商业游戏

访谈者：

可以说下原因吗

受访者：

商业游戏，能更多理解，游戏从立项到宣发，到规范的项目管理，到版本运营等方面的知识以后做别的游戏，也能用上

学习为主的话，商业游戏更合适

当然，如果是业余时间创作，谁去做商业游戏啊，一点也不好玩，不酷

访谈者：

话说你有以玩家身份玩过酒馆吗，就纯roleplay来玩这样

受访者：

涩涩的时候是这样的，某些欧洲模型不限制涩涩，就是文案翻译不太好，欧洲不限制涩涩的模型中文本地化不好

我甚至觉得，AI可以把剧情向黄油市场干掉一大半

访谈者：

那相比较线下那种放的比较开的跑团，你觉得和酒馆有什么区别呢，有喜好偏向吗

受访者：

我反而不是很喜欢线下跑团，主要是规则麻烦，语言交流太碎了，信息的精炼程度，代入感反而不如CRPG

个人喜好问题

我甚至觉得，AI可以把剧情向黄油市场干掉一大半

访谈者：

可以说下原因吗

受访者：

欧洲有的小模型不限制涩涩，如果Steam有人做高度开放的剧情向类似酒馆的黄油，支持创意工坊，一键导入人设场景等等，就用欧洲那些不限制涩涩的开源模型，订阅收费啥的，十几块钱一个月纯文字，三四十同步支持语音，一百左右还支持图片

我觉得可以搞

或者按量收费，卖算力宝，token包啥的

访谈者：

那你觉得目前像酒馆这种形式的ai游戏有什么明显的缺点吗

受访者：

还好吧，最大的缺点是收费高，token卖的太贵了

而且，还不支持涩涩，用一堆垃圾模型，一看token价格，快赶上GLM了

主要是，Steam的酒馆游戏，很少用优质模型的，用DeepSeek都算好的了

这种小模型，凭啥那么贵啊，把玩家当猪宰，价格看起DeepSeek都是赚钱了

访谈者：

那么分别从玩家和开发者的视角来看，你觉得ai参与游戏开发的内容有边界吗

受访者：

我觉得，如果是AI辅助加速，开发者主导审美架构等等，那就没有边界

纯AI的话，啥都不太行，搞着玩还凑合

访谈者：

你对未来像codex这种ai工具有什么期待吗，比如希望能有更强的物理建模功能等等

受访者：

不同的AI有不同的擅长点，没必要追求全能，全能的话不太现实，主要是多个AI的配合吧。比如现在的Codex能不能和别的AI配合呢？像和一些建模的AI啊，或者跟一些绘图的AI，而不是调用它那个智障image2

跨厂商的多智能体协作

可能大厂已经有这东西了

访谈者：

好的，那么问题都问完了，我们访谈到这里就结束了

问卷作答情况

以下仅收录研究相关题目，排除联系方式、同意、抽奖、路由、质量控制和技术字段；题干与选项按该答卷对应的历史题库版本呈现。

基础信息与筛选

题号与原题干	作答
S1 你当前的学生身份是？	本科在读
S2 你的专业或主要学习方向更接近哪类？	计算机/软件/AI
S3 过去 6 个月，你是否参与过游戏或互动作品创作？	做过个人 Demo/原型
S4 你在游戏/互动作品创作中承担过哪些工作？	玩法/关卡策划；美术/视觉；项目管理

题号与原题干	作答
S5 你是否使用过以下 AI 工具/平台参与游戏创作？	Codex / Claude Code 等专业 AI 编程代理；普通 Claude/ChatGPT 等对话工具
S5A 你从什么时候开始较为频繁地使用 AI 工具？	1 年以上
S5B 你目前使用 AI 工具的频率大约是？	每周使用 3—5 次
S6 你使用上述工具时，是否有具体游戏/互动项目或可运行原型作为对象？	有可运行/可展示原型
S7 你是否实际玩过“一站式 AI 游戏制作平台”生成或实质参与制作的作品？	玩过其他一站式 AI 平台作品 其他填写：unity+MCP+Codex 加上各种工具做美术
S8 你能否指出具体平台或作品，并说明自己玩到了什么程度？	能指出平台和作品，且完成一轮完整的核心游玩流程

生产端与创作经历

题号与原题干	作答
P1 请概述你最近一个使用 AI 参与的游戏/互动项目。（项目类型、个人/团队、引擎或技术栈、当前阶段）	gamejam作品，生而为锚，unity，做完demo扔了不管了
P2 你的程序基础更接近哪一项？	能写小功能
P3 你使用过哪些引擎、框架或开发方式？	Unity
P4 如果不使用 AI，你能否独立完成本项目中的程序或功能实现任务？	只能完成很小一部分
P5 在该项目中，你具体使用过哪些工具？	Codex / Claude Code 等专业 AI 编程代理；普通 Claude/ChatGPT
MAIN_TOOL 请选择一个你在游戏/互动项目中使用最深入，或最近主要使用的专业 AI 编程/开发工具。	Codex
CC1 【该工具】你主要让它处理哪些程序任务？	需求拆解；代码生成；项目理解
CC2 【该工具】它进入项目的哪些阶段？	早期原型；关卡/玩法迭代
CC3 【该工具】请评价它对程序工作的帮助。	提高速度：5（非常同意/非常高） 降低调试难度：3 帮助理解项目：3 帮助完成原本不会的功能：3 减少重复劳动：5（非常同意/非常高） 提升代码质量：1（非常不同意/非常低）
CC4 【该工具】AI 生成内容后，你通常如何处理？	直接使用；阅读后小改；请队友检查；运行测试
CC5 【该工具】你对生成代码的判断与维护能力如何？	能判断是否正确：2 能定位问题：2 能解释代码：4 能后续维护：1（非常不同意/非常低） 能控制依赖风险：1（非常不同意/非常低）
CC6 【该工具】相关产出最终是否进入项目？	只少量进入

题号与原题干	作答
CC7 【该工具】整体看，它主要是提高已有能力，还是补足原本不足的能力？	主要提效
P6 整体来看，AI 工具让你的游戏程序/原型开发发生了哪些变化？	开发速度更快：5（非常同意/非常高） 能做更复杂任务：2 更依赖 AI：3 人工检查更多：3 返工更多：2 更愿意尝试新玩法：3 更容易形成可玩结果：4

开放题

题号与原题干	作答
P7 请还原一个最典型片段：你要做什么程序/功能任务，使用了哪个工具，AI 给出什么，你如何修改，最后结果如何？	我是懂一点程序的策划，一般用unity6+MCP连接Codex，实现可互动的原型，但是质量很差，比如，如果流程中的页面素材不提前生产好，他会同时生成背景图片和文字，UI的文字清晰度很低

消费端体验与评价

题号与原题干	作答
C1 你体验过的平台或作品是什么？	我玩AI酒馆，角色卡什么的，算是AI游戏吗？
C2 该作品来自哪类平台？	不确定但知道是 AI 平台生成
C3 你实际玩到了什么程度？	完成一轮完整的核心游玩流程
C4 你过去的游戏体验更接近哪类？	商业游戏；独立游戏
C5 请评价该作品的基础游玩体验。	玩法清晰：4 操作顺畅：3 目标反馈：3 规则平衡：3 稳定性：5（非常同意/非常高） 内容量：5（非常同意/非常高） 耐玩性：5（非常同意/非常高）
C6 请评价该作品的作品内容品质评价。	美术质量：2 视觉统一性：2 音效/音乐：1（非常不同意/非常低） 文本自然度：3 内容重复感：2 整体完成度：4
C7 和你平时玩的普通小游戏、独立游戏相比，这款 AI 制作的作品有没有给你带来不一样的游玩感受？如果有，请简单描述具体体现在哪里。	怎么说呢？大多数是垃圾，除非是酒馆这种高度自定义的
C8 体验前，你是否知道该作品有 AI 参与制作？	明确知道
C9 如果没有被告知，你能否从作品中感知 AI 参与？依据是什么？	美术如果是纯AI肯定能感知出来，文案程序就够呛了
C10 知道或感知 AI 参与后，对你的评价有什么影响？	略减分

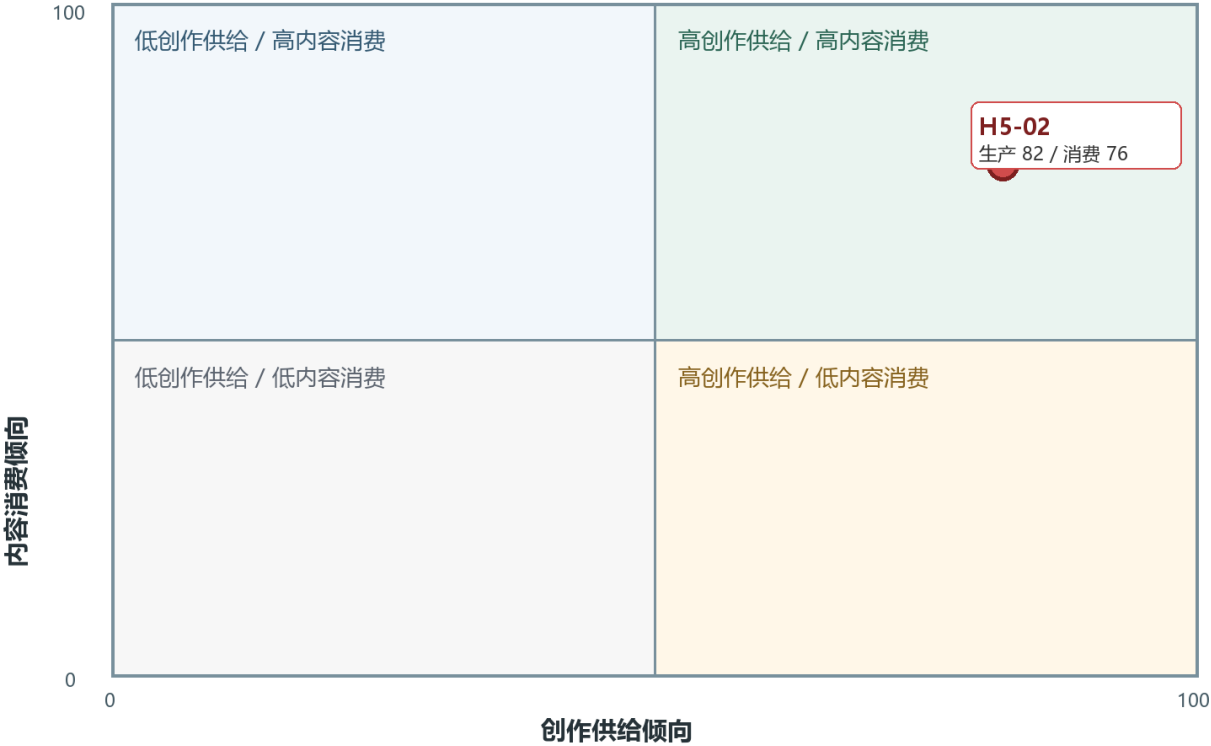
题号与原题干	作答
C11 以下因素会如何影响你对这类作品的信任与接受？	AI 标注透明：5（非常同意/非常高） 版权清晰：5（非常同意/非常高） 创作者的用心程度：5（非常同意/非常高） 作品稳定性：5（非常同意/非常高） 玩法原创性：4 价格/付费方式：5（非常同意/非常高）
C12 你是否愿意继续游玩、分享或付费支持这类作品？	继续游玩：2 推荐给朋友：3 关注作者/平台：2 小额付费：4 提交游玩反馈与建议：5（非常同意/非常高）
C13 与普通小游戏、独立游戏或 UGC 作品相比，你觉得它最大的优势和问题分别是什么？	如果一个AI游戏，还有AI驱动的话，支持玩家自己导入自己的API，那就非常棒了

研究补充

题号与原题干	作答
G1 你总体上如何看待 AI 进入游戏创作？	提高效率：5（非常同意/非常高） 降低门槛：5（非常同意/非常高） 带来新玩法：2 削弱创作诚意：4 增加同质化：4 带来版权风险：5（非常同意/非常高）
G2 你认为 AI 参与程度是否应该被明确标注？	必须明确标注

生产倾向 × 消费倾向象限图

横轴为创作供给倾向，纵轴为内容消费倾向；50 分为高低分界。



H5-02：生产 82，消费 76，高创作供给 / 高内容消费

关联维度分数

创作	95
效率	85
探索	80
接受	75
付费	46
信任	48

A.3 样本三 | H5-03

样本基本信息

样本编号	H5-03
样本定位	以 Godot/GDScript 独立项目为背景的创作端样本，重点涉及旧项目逻辑迁移、多轮功能迭代、手感调试与 AI 代码审查。
访谈执行人	姓名待补
摘要来源	访谈摘要：H5-03.md
逐字稿来源	访谈逐字稿：H5-03.md
关联答卷	dc9ec0cc-0c06-46d9-8354-b10c1dbad338

本文件保留该样本的访谈摘要、逐字稿和源材料附件；最后附上关联问卷中的研究相关作答与既有象限结果。

访谈摘要

原始材料：访谈摘要：H5-03.md

访谈摘要：H5-03

一、访谈基本信息

- 受访者编号：H5-03
- 访谈时间：2026/07/07 14:55-16:33

二、样本判定

按原访谈大纲的严格口径，本访谈对象可直接判定为 Codex 核心样本。其在 Game Jam 项目中全程使用 Codex 生成和迭代代码，并在长期 Godot 沙盒项目中使用 AI 进行项目结构理解、资源系统梳理和配置规则说明，能够提供较具体的提示词、输出结果和人工测试反馈。

- 样本标签：TOOL-C
- 判定理由：有真实游戏项目、具体任务链路和过程材料
- 具体项目：Game Jam 平台乱斗项目；长期个人沙盒项目
- 项目阶段：Game Jam 项目已形成核心玩法原型；沙盒项目为长期开发中项目
- 过程材料：访谈中提供了多轮提示词、AI 输出摘要、测试反馈与后续纠偏记录

三、受访者画像

受访者并非码类相关专业，但具备一定 Godot/GDScript 基础。其编程能力不是“完全依赖 AI 生成”的状态，而是能够提出项目约束、引用既有代码、判断功能手感、识别不符合预期的改动，并通过多轮提示词推动 AI 修正。受访者对 AI 生成代码不要求逐行完全理解，但要求整体架构、设计模式和功能结果符合自己的目标；对于图形学 shader 等复杂领域，则接受较低理解程度。

受访者的 AI 使用方式偏“项目协作者/多轮执行代理”，不是简单问答。他会让 Codex 读取其他项目代码作为参考，迁移已有系统逻辑，再在新项目中扩展功能。例如在 Game Jam 项目中，他要求 Codex 参考沙盒项目里的“脐带物理”和人物物理，实现平台乱斗玩法中的抛锚、绳索连接、虚拟 Z 轴、影子表现和手感调参。

四、项目与任务链路

Game Jam 项目是一个平台乱斗类型游戏，核心循环是玩家在平台内抛出锚，将敌人砸出平台后再捡回锚。该项目中代码“全部都是 Codex 写的”。Codex 参与的不是边缘辅助功能，而是核心玩法：玩家、锚、抛物运动、连接绳、碰撞、力量属性、血量属性、可配置参数和手感调整。

典型使用链路如下：

- 第一轮，受访者要求 Codex 不改动指定目录外的项目，只读取沙盒项目中的脐带物理和基础人物物理，在 Game Jam 项目中实现玩家和锚的初始原型。Codex 输出了可抛锚的简单原型，但错误地添加了空气墙和背景，且未正确处理素材导入和锚头绑绳。
- 第二轮，受访者指出素材、锚的节点类型、绳子连接位置和再次抛锚逻辑问题。Codex 补全了交互逻辑和素材接口，并删除了不需要的空气墙和背景。
- 第三轮，受访者要求增加玩家属性组件，使力量影响可捡锚范围、抛锚高度和绳索限制，并要求参数便于策划配置。Codex 基本实现了抛锚/捡锚的物理碰撞玩法，但出现飞行过程中碰撞箱未禁用、锚被玩家控制方向出错等问题。
- 第四轮，受访者用人工测试发现的问题继续追问，要求修正玩家输入对锚飞行落点的影响逻辑，并禁用飞行时碰撞。Codex 修复了这些问题。
- 后续微调中，Codex 曾误解“力量影响水平位移”的程度，使力量越大水平速度越大、手感偏离预期。受访者再次反馈后，Codex 将 strength 对虚拟 Z 轴高度和水平速度影响拆开，限制水平速度上限，提高空中 XY 阻力，并降低玩家输入对落点的影响。

长期沙盒项目中，AI 的主要价值不只在写代码，而在帮助理解已有架构。受访者让 AI 只读 Godot 4.6 项目中的 OrganData、ItemData、Rule、RuleProcessor、OrganTrigger、OrganSystem 等脚本和器官资源，输出字段结构、枚举顺序、资源引用方式、目录数量、缺失字段和损坏引用风险。随后又要求 AI 生成“器官效果配置说明书”，把器官系统概括为器官本体、静态规则、事件触发器、器官电路、攻击修改器等层次，辅助受访者理解和设计器官流派。

五、效率、能力补足与人工控制

受访者认为 Codex 最大的帮助是效率提升：减少复制粘贴，让他可以并行推进多个功能。他明确表示不太担心依赖 AI 让自身能力变弱，因为没有 AI 时，个人开发时间会被拉长好几倍。

但受访者也认为深度使用 Codex 有门槛，需要基础语法和框架理解。原因是如果 AI 改不动报错，用户自己也不会纠错，开发就会堵住。这一判断支持大纲中的假设：AI Agent 的价值不只取决于生成能力，也取决于用户是否有能力判断、测试和约束输出。

在人工控制上，受访者主要通过以下方式降低风险：

- 限定 AI 可修改的目录，允许读取其他项目作为参考，但不允许跨项目随意改动。
- 对长期项目更谨慎，文件夹和资源管理主要手动处理，不太让 AI 自己新增文件夹。
- 对短期项目更开放，因为资源少、结构简单，可以允许 AI 自由建文件夹。
- 不要求自己逐行理解所有 AI 生成代码，但会检查整体架构、设计模式、功能表现和手感是否符合预期。
- 通过人工试玩和手感判断发现偏差，再把问题转化为下一轮提示词。

六、留存、付费与迁移判断

受访者表示未来会继续使用 Codex。目前付费为 ChatGPT 20 美元额度，继续付费的核心理由是 AI 的上下文理解能力强。其他能力在他看来相对次要。如果 Gemini 后续“不降智”，他也可能重新选择 Gemini，说明其工具留存并不绑定品牌，而是绑定上下文理解、稳定性和性价比。

受访者从 Gemini 转向 GPT/Codex 的原因包括：早期 Gemini 免费，且不像 GPT 那样“过度设计”；后来 Gemini 3.0 体验下降，同时发现 ChatGPT 会员也能使用 Codex，因此迁移到 GPT/Codex。

七、AI 参与制作的市场判断

受访者对代码层面的 AI 参与较为接受，对美术层面的 AI 参与更谨慎。他认为玩家抵触的更多是 AI 美术，只要美术不让 AI 参与，问题不大。作为玩家，他是否愿意为 AI 参与的游戏付费主要看质量；如果画面有明显 AI 生图痕迹，他不会付费。

他不愿让 AI 深度参与美术的理由包括：AI 美术生成太靠运气，追求整体效果和整体精度，导致画面节奏平均、过于精致，明度、对比度、灰度层次缺少人为控制，细节刻画也容易用力过猛。不过对于实验性小项目，例如中世纪复古风格等，他认为可以尝试 AI 美术。

八、访谈记录字段整理

- 受访者编号：H5-03
- 样本类型：AI Agent 游戏开发者补充样本；非严格 Claude Code 样本
- 样本标签：X（严格 Claude Code 口径）
- 是否在校大学生：访谈记录中未明确确认
- 专业与年级：专业非码类相关；年级未提及
- 编程基础：中
- 使用引擎或开发工具：Godot、GDScript、Codex、ChatGPT、Gemini
- 具体项目或作品：Game Jam 平台乱斗项目；长期沙盒项目“活体重组”
- 项目类型：Game Jam；个人长期项目
- 当前阶段：Game Jam 项目形成核心玩法原型；沙盒项目长期开发中
- Claude Code 使用时间：未确认；记录中为 Codex/ChatGPT/Gemini 使用
- 使用任务类型：代码生成、修改、调试、项目理解、资源/配置规则梳理
- 是否涉及至少两类实质任务：是
- 产出状态：Game Jam 项目中进入项目；沙盒项目中至少进入项目理解和配置设计流程
- 人工修改程度：低到中，主要通过测试反馈和提示词纠偏控制输出

- 是否提高效率：是
- 是否补足能力：是，但依赖基础语法和框架理解
- 是否增加维护成本：在长期项目中存在结构失控风险，因此受访者限制 AI 改动文件夹和项目结构
- 是否持续使用：是
- 付费判断：已为 ChatGPT 20 美元额度付费；核心付费理由是上下文理解强
- 是否有过程材料：有，包含提示词、输出摘要和人工测试反馈；未见截图、提交记录或 Demo 文件
- 主要有效场景：短期原型、核心玩法代码生成、多轮手感调试、读取旧项目代码迁移逻辑、理解复杂资源系统
- 主要失效场景：AI 美术、长期项目文件结构管理、对细微手感需求的误解、缺乏用户纠错能力时的报错处理
- 可用于验证或反驳的研究假设：验证“AI Agent 价值取决于是否进入真实项目链路”“用户基础影响产出落地”“真实使用是多轮协作而非一次生成”；部分修正“必须完全理解生成代码才能落地”的假设，受访者采用的是架构和结果层面的可控理解。

九、研究结论

该受访者展示了一种典型的“有一定技术基础的个人游戏开发者使用 AI Agent”的模式：用户负责提出玩法目标、选择参考项目、限定修改范围、测试手感和控制项目结构；AI 负责快速生成、迁移和修改大量代码。AI 的价值不只是写代码，而是把原本耗时的工程实现、配置梳理和项目理解压缩成多轮可迭代过程。

同时，该案例也提示研究中需要区分短期项目和长期项目。短期 Game Jam 中，受访者可以接受 AI 大范围生成代码，甚至整个项目代码都由 Codex 完成；长期沙盒项目中，受访者明显收紧权限，不让 AI 随意管理文件夹和项目结构，只允许它在指定范围内读取、理解或批量填资源。这说明 AI Agent 的落地并非单纯取决于生成质量，还取决于项目周期、结构复杂度、用户控制策略和维护风险。

原始附件

本样本原始材料未包含嵌入图片或附件。

访谈逐字稿

原始材料：访谈逐字稿：H5-03.md

访谈逐字稿：H5-03

基本信息

- 导出时间：2026-07-07T09:22:20.068Z
- 访谈时间：2026/07/07 14:55:00 - 2026/07/07 16:33:50

逐字稿

访谈者：

我会对接下来的聊天记录存档，用于研究分析，会对个人信息做保护，ok的话，那我们开始了

受访者：

ok

访谈者：

首先，你有编程基础吗，或者专业是码类相关的吗

受访者：

有的以前学过godot的gds语言，专业和码类没啥关系

访谈者：

具体有参与过哪些游戏开发项目呢，可以说下项目的类型或者主要玩法吗

受访者：

就拿最近一次的gj以及自己在做的沙盒项目来说，gj的做的游戏是平台乱斗类型，主要玩法是玩家在平台内抛锚把敌人砸出去之后再捡锚的循环。沙盒项目的主要玩法是玩家和怪物都是一套胸腔模板，玩家要不断开胸取怪物器官生存

访谈者：

项目都有ai参与吗

受访者：

gj和沙盒项目代码都有ai参与

访谈者：

那我们就主要聊这两项目好了

主要使用的是什么ai呢

受访者：

gj全程codex，沙盒项目前期gemini，后期转chatgpt和codex

访谈者：

可以说下这些ai具体参与的部分吗，比如codex来实现玩家控制部分的功能等等

受访者：

gj的codex具体实现的是玩家抛锚和锚砸下来是物理效果和手感，都是通过让codex参考其他项目已有的代码，比如锚和玩家之间的连接物理是让codex参考我的沙盒项目里是脐带物理，然后我再提出增加虚拟z轴的方案，让codex在参考基础上增加z轴和相应的视觉表现比如影子会随z轴高低变化颜色

沙盒项目ai参与很多功能，我全都说还是挑几个说？

访谈者：

可以具体描述一下这里使用codex的开发过程吗，就比如每轮的输入提示词和codex的输出结果

以及是怎么处理这些输出结果的呢

挑几个核心的吧

受访者：

- 第一轮提示词输入：请不要动除了26gj文件夹（D盘中）以外的项目改动（可以读取），找到活体重组的脐带物理和基础的人物物理，我现在想在26Gj里做个玩家，做个锚，玩家类似可以随便变化自己的史莱姆（用程序性动画？）我画个纯色像素球应该可以把，然后锚的物理属性要求细致，玩家可以变换自身（程序性动画），像脐带和动力储罐的结合体效果一样把锚抛出去（这是2d俯视角的，带一点点视觉重力吧），然后不会像活体重组里面一样收回来，至于26gj项目里的把脐带变为史莱姆自身（程序性动画），不要动除了26Gj文件夹以外的项目改动（但可以读取）
- 第一轮输出：codex输出的是一个能够抛出去锚的简单原型，但ai没有做到锚的竖头绑绳子和素材导入问题，并且自己添加了空气墙和背景

第二轮输入：可以但，素材不对，按照我的唯一——张A1，按16x16为一个单位，第二个单位开始的2x2单位是吗锚，第四个单位的1x1单位是史莱姆，而且锚也应该用characterbody2d吧，用刚体以后会出问题吧。把背景删去，然后请把射的位置绑在这个的竖头上，参考脐带的重量之间的动量关系，这个重量应该在物理上重心在素材下方，然后把脐带加回来，类似绳子连着这玩家和这个的竖头，然后再次左键的逻辑不是让锚在一个位置生成，而是把玩家爬到锚的一个范围才能再次抛出

- 第二轮输出：codex根据我的要求把交互逻辑和素材接口补全了，并且删去空气墙和背景
- 第三轮输入：删去墙体，然后给玩家加入属性组件，参考活体重组里的属性组件和写法(不过不需要器官等中间件)，这个游戏里玩家只有力量和血量的属性，力量影响可抛锚的范围（力量越大可捡取锚的范围越大）同时这些东西要策划配置容易，还有绳子的最大限制策划也要容易配置，对了，同时力量越大，锚能够抛的越高，第二点，锚被抛起的过程和锚静止的不一样，锚被抛起时玩家达到绳子最大限度时不会被锚控制完全不能动，而是玩家能减速到原速度百分之八十，同时可以控制锚的掉落方向（一定的影响，抛锚的掉落点和抛锚的力量，方向，玩家在锚飞行时的方向都有一定影响，类似抛体运动应该把），修改一下，但现在的手感和目前功能不要删掉了

第三轮输出：codex基本实现我的抛锚捡锚的物理碰撞玩法，但会出现一些小问题，比如飞行过程中碰撞箱没有禁止，飞行过程种锚被玩家控制方向出错。

- 第四轮输入：还有个问题，比如玩家玩家移动输入会轻微影响锚的飞行/落点方向这个拆开来是比如玩家在（上北下南）玩家在北，锚在南，玩家往北会影响，但如果此时玩家也往南同时没有超过锚的Y，那么怎么会影响呢，还有锚在飞行时的碰撞怎么没禁用

第四轮输出：直接修复上述问题

gj做的比较快，输出结果都是直接在项目内修改后我再进行测试并输入下一轮的提示词

这个是gj主要玩法实现的输入输出

访谈者：

gj项目中输出结果的人工修改部分主要修改了哪些内容呢

或者说，在人工测试部分，有测试出codex输出内容不符合提出要求的结果吗

受访者：

有的，比如后面一轮（第四轮）的微调中AI把原本抛东西之后掉落的手感改了，变成力量越大水平速度更大了

我这一轮反馈给他的提示词是：当我的力量（强度为1），水平位移也能仍很远，你理解错我意思了，力量确实能影响水平位移，但影响水平位移（说实话，xy轴其实都算水平）的效果有限（会衰减），然后抛锚的也是只是玩家的那个输入会影响有限，力量影响抛得高不高（不知道要不要做个虚拟z轴）

codex输出：你说得对，我之前让“力量加高度”变相扩大了水平距离，因为飞行时间变长但水平初速还太大。现在拆开了strength 主要影响虚拟 Z 轴高度水平速度有上限：默认最多 380，不会力量越大就无限远。空中 XY 阻力从 80 提到 190，水平位移会衰减得更明显。玩家输入对锚飞行影响从 135 降到 42，只做微弱落点修正，不再像推着锚飞。

访谈者：

所以这个物理手感功能的后续维护阶段（比如手感调整，在其他功能做完后的微调、资源合并等等）都是有codex参与的吗

受访者：

这个gj的项目全部都是codex写的代码

访谈者：

那么可以聊下沙盒项目里ai参与度比较高的地方吗，以及你刚刚提到前期gemini，后期转chatgpt和codex，为什么做出了这个转变

受访者：

后面那个问题主要是前期gemini免费而且不太像gpt会过度设计，当后面3.0降智转gpt，发现gpt充的会员codex也能用

沙盒项目ai参与度较高的地方是对于之前器官计算系统搭好的理解，我等下发下输入输出结果

沙盒项目我就说一个后期帮助我理解架构的例子

- 第一轮输入：res://resource/物品/，要求命名中文且内部的器官tags以及item id 和item name 均为中文，至于物品tags则沿用（如果是英文已有tag就继续沿），图片不用填，只要先填一个流派的器官。现在先理解稳定项目先只读项目，不要修改任何文件。这是一个Godot 4.6项目。请先扫描并理解以下系统：1. OrganData.gd2. ItemData.gd3. Rule.gd4. RuleProcessor.gd

5\。OrganTrigger.gd6. TriggerValueSource.gd7. OrganTriggerRelay.gd8. OrganPropertyTransfer.gd9. AttackModifier.gd10. OrganSystem.gd11. 当前所有器官.tres资源

请输出：- 每种Resource的字段结构- 各枚举当前数字顺序- .tres中引用Rule、Trigger和子资源的方式- 当前器官资源目录与数量

- 哪些资源缺少description、manual_descriptions、rules、triggers、trigger_relays、property_transfers或attack_modifiers
- 哪些资源存在损坏引用或字段不兼容风险 绝对不要修改文件。
- 输出结果按照输出要求列举了我的项目信息
- 第二轮输入：那我在设计器官流派的时候，请输出器官效果配置的参数规则让我明白，不要修改当前文件和脚本，输出官效果配置的参数规则说明书
- 第二轮输出：（太长了截取一部分）器官效果配置说明书器官效果目前可以理解成 5 层：器官本体、静态规则、事件触发器、器官电路、攻击修改器。核心资源是 [器官资源.gd (line 1)]
(D:/godot%E9%A1%B9%E7%9B%AE/%E6%B4%BB%E4%BD%93%E9%87%8D%E7%BB%84/resource/%E7%89%A9%E5%93%81/%E5%99%A8%E5%也就是 OrganData。1. 器官本体 OrganData每个器官 .tres 主资源至少关心这些字段：
 - item_id # 物品唯一 ID，建议中文
 - item_name # 显示名称，建议中文
 - item_tags # 物品标签，例如 organ、nutrition，沿用现有英文
 - tags # 器官构筑标签，建议中文，例如 猎手、植物、易碎、根系
 - attributes # 基础属性字典
 - unique_properties # 动态变量、叠层、储能等
 - rules # 静态规则
 - triggers # 事件触发器
 - manual_descriptions # 手写效果描述
 - trigger_relays # 触发转发，当前资源基本未用
 - property_transfers # 属性搬运，当前资源基本未用
 - attack_modifiers # 攻击修改器，当前资源基本未用
 - description # 物品描述
 - item_tags 是物品系统用的，比如 organ、nutrition。
 - tags 是器官流派系统用的，比如 猎手、骨质、易碎。以后设计流派时，主要靠 tags 做联动。
 - attributes 常见键：

health

speed

damage

defence

nutrition_rate

urgent_ratio

urgent_threshold

可以加新键，但要确认读取方支持。规则系统本身可以读任意字典键。

访谈者：

感觉你对ai的输出结果有一定理解上的要求，那么在ai参与的开发过程中，你会要求自己搞懂ai生成部分的东西吗，如果会的话，一般会理解到什么程度呢

受访者：

我不要求自己搞懂，只要在大体架构和运用的设计模式符合我给的并且帮助我完成功能就可以，有一些则完全不要求搞懂比如图形学shader之类的东西

访谈者：

这个沙盒项目是长期项目吗

受访者：

是

访谈者：

那你有尝试过让ai进行长期的项目管理相关的内容吗（比如资源管理、文件内容管理、自动debug等等）

或者长期和短期项目的ai使用方式有区别吗，比如长期项目会更关注可靠度之类的

受访者：

文件管理都是我手动添加文件夹管理，文件这块长期项目我不太会让ai自己加文件夹，但在同一个文件夹批量填资源

长期项目文件夹不太会给ai动

短期项目资源少ai随便建文件夹

访谈者：

那么从项目整体来看，codex对你最大的帮助主要是在哪一块呢

你有绝对不会让ai参与的部分吗

受访者：

更有效率了，不用我再复制粘贴，我可以多线做很多个功能

美术不会让ai参与

但如果是实验性的小项目比如中世纪复古风格之类的美术也可以尝试用ai

访谈者：

可以说下美术这块的理由吗

受访者：

美术ai生成太靠运气了

ai在美术这方面一般是追求整体效果和整体精度

导致画面的节奏看上去都是平均的都太精致了

而现实的画作明度对比度灰度在画面的每个部分都不太会是一样的

还有细节ai有点刻画用力过猛

访谈者：

那么回到刚才codex的问题，你觉得使用codex进行游戏开发需要门槛吗，就比如要懂程序、懂游戏项目的框架等等

这里指codex深度参与游戏开发的情况

受访者：

基础的语法和框架还是要会

不然报错了ai不会改个人本身也纠错不了就堵了

访谈者：

那你会担心自己太过依赖于codex，让自己的开发能力变弱吗

受访者：

不太担心，没ai个人开发估计时间要被拉长好几倍

访谈者：

那从项目的可落地性来看，你觉得codex更适合长期项目还是短期项目，团队项目还是个人项目呢

受访者：

我只能说codex适合的程度和这个团队或者个人的背后的经验或者工作流有关，适合长期还是短期项目看codex使用者的目标

访谈者：

你觉得如果要让codex深入参与到可落地的项目中，最理想的环境是怎样的（就比如一个工业化很强的以商业化为目标的团队等等）

受访者：

最理想的环境是把之前的工业流程内比较琐碎且不是很重要的环节先让codex替代试水应该吧

访谈者：

那从个人开发者的角度看，你会在未来的项目里继续使用codex吗

受访者：

会的

访谈者：

你目前付费的额度是多少呢

受访者：

chantgpt的20美元额度

访谈者：

那么gpt的功能中，最能支持你继续付费下去的功能有哪些，比如写代码方便，上下文理解能力强等等

受访者：

ai上文理解强其他都无所谓，如果gemini新的不降智了我可能会重新选择gemini

访谈者：

因为现在有很多玩家对ai深度参与游戏开发非常抵触，你会担心自己的游戏因为ai参与不受玩家喜欢吗

受访者：

玩家抵触的更多是美术方面的吧，美术这边不让ai动应该问题不大

访谈者：

那么从玩家角度看，你会愿意为有ai参与的游戏付费吗

受访者：

分情况吧，如果游戏画面有很明显我之前提到的AI生图的问题我不会付费，主要看质量

访谈者：

不考虑现实因素，你最希望未来codex能在游戏开发中解决什么问题呢

受访者：

一些图形学，物理模拟优化攻克吧

这方面还是有点差

访谈者：

好的，那我们整体的访谈就到这里，非常感谢参与

问卷作答情况

以下仅收录研究相关题目，排除联系方式、同意、抽奖、路由、质量控制和技术字段；题干与选项按该答卷对应的历史题库版本呈现。

基础信息与筛选

题号与原题干	作答
S1 你当前的学生身份是？	本科在读
S2 你的专业或主要学习方向更接近哪类？	文科社科
S3 过去 6 个月，你是否参与过游戏或互动作品创作？	做过个人 Demo/原型
S4 你在游戏/互动作品创作中承担过哪些工作？	程序/功能实现；调试测试；玩法/关卡策划；美术/视觉；文本/叙事；项目管理
S5 你是否使用过以下 AI 工具/平台参与游戏创作？	Codex / Claude Code 等专业 AI 编程代理；普通 Claude/ChatGPT 等对话工具
S5A 你从什么时候开始较为频繁地使用 AI 工具？	6 个月—1 年
S5B 你目前使用 AI 工具的频率大约是？	几乎每天都会使用
S6 你使用上述工具时，是否有具体游戏/互动项目或可运行原型作为对象？	有真实项目
S7 你是否实际玩过 “一站式 AI 游戏制作平台” 生成或实质参与制作的作品？	没有
S8 你能否指出具体平台或作品，并说明自己玩到了什么程度？	能指出平台和作品，且完成一轮完整的核心游玩流程

生产端与创作经历

题号与原题干	作答
P1 请概述你最近一个使用 AI 参与的游戏/互动项目。（项目类型、个人/团队、引擎或技术栈、当前阶段）	过去一年都在做一个围绕核心玩法的沙盒个人项目，使用的是godot引擎，过去的半年以上用的都是普通对话式AI，当时的工作流程是由我来设计项目架构，再通过对话让AI生成代码，底层的架构设计实现和策划编辑工具是ai搭完的，后面三个月用codex在已经搭完的架构上实现更多小功能，同时让它批量把策划的文本文件转换为引擎内部资源文件进行铺量
P2 你的程序基础更接近哪一项？	能独立完成小游戏/模块
P3 你使用过哪些引擎、框架或开发方式？	Godot
P4 如果不使用 AI，你能否独立完成本项目中的程序或功能实现任务？	能完成但效率低/质量不稳
P5 在该项目中，你具体使用过哪些工具？	Codex / Claude Code 等专业 AI 编程代理；普通 Claude/ChatGPT
MAIN_TOOL 请选择一个你在游戏/互动项目中使用最深入，或最近主要使用的专业 AI 编程/开发工具。	Codex
CC1 【该工具】你主要让它处理哪些程序任务？	代码生成；修改已有代码；调试报错；项目理解；接口/工具链接入；需求拆解
CC2 【该工具】它进入项目的哪些阶段？	性能/结构优化；核心功能开发
CC3 【该工具】请评价它对程序工作的帮助。	提高速度：5（非常同意/非常高） 降低调试难度：2 帮助理解项目：5（非常同意/非常高） 帮助完成原本不会的功能：5（非常同意/非常高） 减少重复劳动：5（非常同意/非常高） 提升代码质量：3
CC4 【该工具】AI 生成内容后，你通常如何处理？	阅读后小改；运行测试；只采纳思路
CC5 【该工具】你对生成代码的判断与维护能力如何？	能判断是否正确：4 能定位问题：4 能解释代码：3 能后续维护：5（非常同意/非常高） 能控制依赖风险：5（非常同意/非常高）
CC6 【该工具】相关产出最终是否进入项目？	修改后部分进入
CC7 【该工具】整体看，它主要是提高已有能力，还是补足原本不足的能力？	两者都有但任务不同
P6 整体来看，AI 工具让你的游戏程序/原型开发发生了哪些变化？	开发速度更快：5（非常同意/非常高） 能做更复杂任务：5（非常同意/非常高） 更依赖 AI：3 人工检查更多：3 返工更多：2 更愿意尝试新玩法：5（非常同意/非常高） 更容易形成可玩结果：5（非常同意/非常高）

开放题

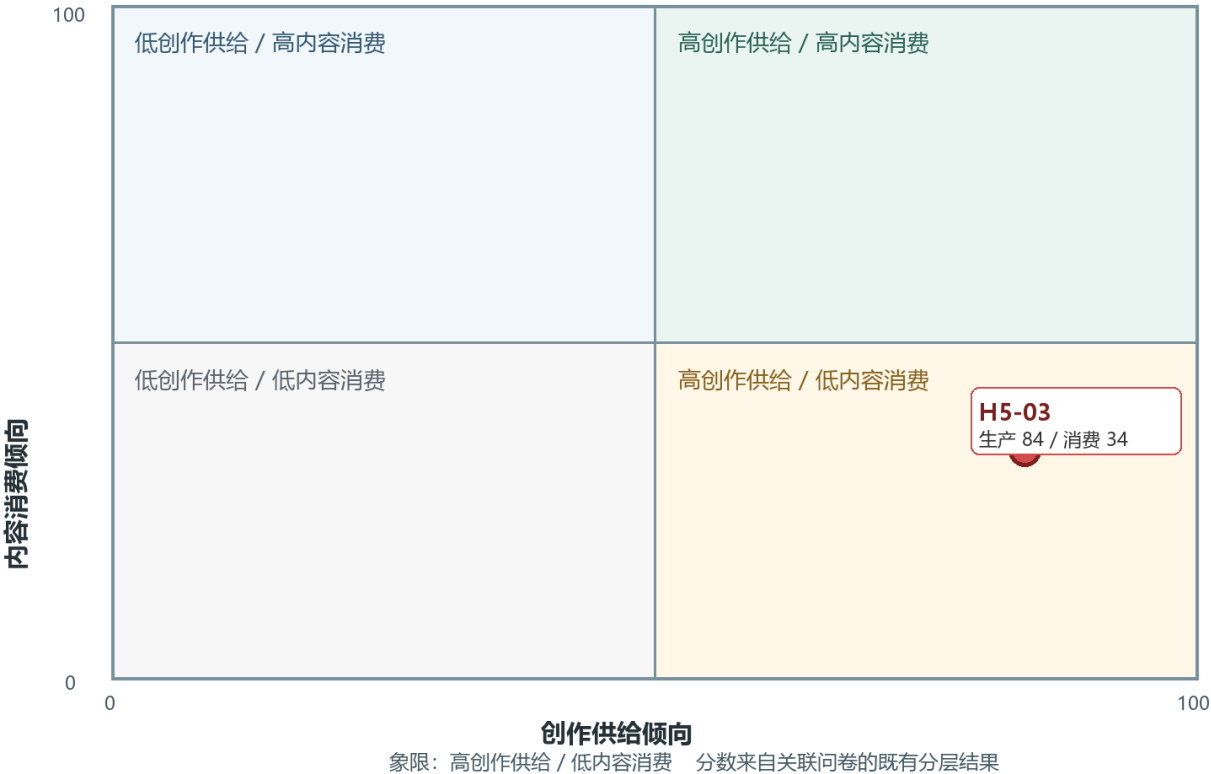
题号与原题干	作答
P7 请还原一个最典型片段：你要做什么程序/功能任务，使用了哪个工具，AI 给出什么，你如何修改，最后结果如何？	就比如我要给原本的地图编辑器加上一个新的关卡编辑器，我使用codex，先让他只理解当前的架构并限制它过度设计，然后给出不同的相应方案，选择更改到我满意方案，最后让它直接修改，最后我自己测试能够在当前项目里运行并无bug即可

研究补充

题号与原题干	作答
G1 你总体上如何看待 AI 进入游戏创作？	提高效率：5（非常同意/非常高） 降低门槛：5（非常同意/非常高） 带来新玩法：2 削弱创作诚意：2 增加同质化：2 带来版权风险：2
G2 你认为 AI 参与程度是否应该被明确标注？	可由作者选择
Q4.3 如果你没有实际玩过一站式 AI 平台作品，主要原因是什么？	担心质量差

生产倾向 × 消费倾向象限图

横轴为创作供给倾向，纵轴为内容消费倾向；50 分为高低分界。



H5-03：生产 84，消费 34，高创作供给 / 低内容消费

关联维度分数

创作	95
效率	86

创作	95
探索	70
接受	40
付费	39
信任	60

A.4 样本四 | H5-07

样本基本信息

样本编号	H5-07
样本定位	偏消费端的补充样本，重点讨论一站式 AI 游戏作品体验、玩家对画面与玩法完整度的判断，以及对 AI 美术和创作者投入的接受边界。
访谈执行人	姓名待补
摘要来源	访谈摘要：H5-07.docx
逐字稿来源	H5-7逐字稿.docx
关联答卷	31c6ebdd-9929-4527-8c10-6317fa233d59

本文件保留该样本的访谈摘要、逐字稿和源材料附件；最后附上关联问卷中的研究相关作答与既有象限结果。

访谈摘要

原始材料：访谈摘要：H5-07.docx

访谈摘要：H5-07

一、访谈基本信息

受访者编号：H5-07（原始文件名使用 H5-7）

访谈时间：逐字稿未记录

材料来源：H5-7逐字稿.docx；包含访谈前置字段、8 组主问题和追问内容

样本类型：AI 游戏作品消费者

受访者背景：在校大学生；每日玩游戏；同时自述为使用 AI 的游戏开发者

二、样本判定

按《AI游戏创作与消费研究_样本筛选与执行口径表》的消费端标准，H5-07 可判定为 TapTap 制造消费端直接对应访谈样本。受访者不是只听说或观看宣传材料，而是在 TapTap 制造专场的线下活动中实际体验过作品；前置字段记录其游玩时长少于 2 小时、完成了体验，并能对玩法、操作、美术、完整度、稳定性、耐玩性以及继续游玩、分享和付费意愿作出回答。该材料满足“具体平台 + 实际游玩 + 完成最低体验节点 + 能评价作品”的直接对应条件。

与 H5-01、H5-02、H5-08 主要作为消费端补充材料不同，H5-07 的有效信息来自一站式平台作品的直接消费体验，因此不应仅因受访者同时具有开发者身份而改判为普通玩家补充样本。开发者身份需要作为解释变量保留：它可能使受访者更能理解 AI 的工具属性，也可能降低其对“是否使用 AI”本身的敏感度。这里先确认消费端证据的对应关系，不把该访谈当作生产端项目案例。

来源管理标签：TOOL-C | CP-T（CP-T 为 TapTap 制造消费端直接对应；管理表当前 is_core_sample=false）

判定理由：有明确 TapTap 制造活动来源、实际游玩和完整体验记录，能提供具体作品评价与消费意愿判断

体验对象：TapTap 制造专场中的多人拍照聚会游戏；受访者提到活动题目/主题“54321”，具体作品名未确认

体验平台：TapTap 制造专场线下活动

游玩深度：少于 2 小时；前置字段记录为完整体验

过程材料：逐字稿；未提供作品链接、截图或独立游玩记录

管理表与开放题对应

两份 CSV 通过同一个 record_id 对应到 H5-07 的前置样本：31c6ebdd-9929-4527-8c10-6317fa233d59。管理表中该记录的基础信息为“student、数字媒体/游戏/动画、primary_tool=Codex、played_platforms=TapTap 制造”，case_tags 为“TOOL-C|CP-T”。在所有 CP-T 候选中，它是唯一的学生记录，因此与逐字稿中的大学生、AI 游戏开发者和 TapTap 制造体验能够相互印证。

开放题记录也提供了同一受访者的前置参照：C1 只写“Tap tap”；C7 认为“AI 制作游戏对于我来说，目前还是不够成熟的”；C9 提到作品“没有完全照顾到玩家的内心的感觉”；C13 则概括其优势为“内容更加丰富，画质更加好”。这些回答与访谈中的具体体验并不矛盾：前置开放题更接近对 AI 制作游戏总体成熟度的判断，逐字稿则针对线下专场的一款具体作品给出玩法和美术上的正面评价。P1、P7 还记录了 AI 短视频/漫剧尝试，但不足以把 H5-07 扩写成游戏生产端案例。

对应样本 UUID：31c6ebdd-9929-4527-8c10-6317fa233d59

管理表状态：quality_status=valid；compatible_path=professional_tool；compatible_persona=efficiency_builder；is_core_sample=false

开放题证据：C1、C7、C9、C13 直接涉及 TapTap 制造消费体验；P1、P7 涉及 AI 短视频/漫剧尝试

样本的主要限制也需要保留。访谈没有给出具体作品名称、活动日期、实际分钟数、音效和文本评价，也没有足够材料支持长期留存或真实付费行为判断。管理表当前将 is_core_sample 记为 false，因此本摘要使用“直接对应访谈样本”而不直接覆盖为“核心样本”；是否在访谈复核后升级管理状态，应由研究口径统一决定。“愿意付费”属于受访者在假设作品完成度足够时的意愿，不等同于已经支付过费用。

三、消费端样本分析

1. 玩家背景与参照系

受访者平时每天玩游戏，类型跨度较大。线上游戏方面提到《三角洲行动》、APEX 等 FPS，也会玩动作游戏、RPG、文字冒险和独立游戏，举例包括《杀手47》《艾尔登法环》《黑暗之魂》《巫师 3》以及《Beholder》等。这样的游戏经验使其评价不只停留在“好不好玩”的总体印象，而会区分不同类型的价值：玩热门 FPS 时更看重解压、社交和热度；面对单机或独立游戏时，则更关注表达、机制、整体艺术风格、叙事以及这些元素如何落实到 UI 和具体体验中。

因此，H5-07 的参照系不是单一的商业大作，也不是只对 AI 作品抱有新鲜感的轻度玩家。其评价中同时存在竞技/社交游戏的即时反馈标准，以及动作和独立游戏对风格、机制和完整性的要求。这使得他对一款 48 小时作品的正面评价更有区分度：他认可作品的创意和可玩性，但没有把短期开发的细节缺口直接等同于 AI 造成的质量问题。

2. 实际体验对象与核心玩法

受访者回忆自己在“聚光灯”的 TapTap 制造线下专场中接触过多款作品，其中印象较深的是一款拍照聚会游戏。其核心操作非常明确：多人使用简单手柄，在场景中争抢拍照镜头。受访者将它评价为“非常有趣，有创意”，说明作品至少完成了一个可理解、可参与的核心循环，而不是只停留在展示概念或观看演示的层面。

需要注意的是，原话中的“54321”被表述为活动题目或主题，逐字稿没有确认它就是作品名称。因此本摘要只把它作为活动语境记录，不把“54321”扩写为确定的游戏标题。关于关卡数量、局内时长、玩家人数上限和是否有完整结算循环，材料没有进一步证据，不做补充推断。

3. 玩法与操作评价

玩法评价明显偏正面，受访者认可“抢拍照镜头”这一简单但有社交互动的机制，认为它有趣且有创意。操作层面的意见则更具体：整体“简单易懂”，但键位仍有优化空间，可能没有充分遵循玩家常见的操作习惯。这个反馈将问题定位在输入设计，而不是泛泛地说作品“不够好”。对于线下多人试玩来说，上手成本低是优点；但如果键位与常规习惯存在偏差，玩家仍可能在第一次尝试时把注意力耗在理解操作上。

这组评价说明，短流程 AI 作品的消费门槛首先由核心循环和操作反馈决定。受访者并没有因为作品规模小而否定其创意，也没有把“48 小时完成”当成可以忽略操作问题的理由；他接受的是核心玩法成立、细节仍可继续打磨的状态。

4. 美术、完成度与 48 小时语境

美术评价带有明确限定：“不看细节的话，Q 版像素加上有限场景的配合，是比较和谐美观的。”这不是无条件的高评价，而是对整体风格统一性的认可，同时保留了细节层面的不足。受访者在前置字段中将完整度记为“完整，细节可进一步优化”，稳定性评价为“好”，耐玩性评价为“好”。其中“细节可进一步优化”与其对美术的限定相互印证：作品已经形成可体验的整体，但距离更长时间游玩或更高完成度仍有提升空间。

受访者能够把这些缺口放回 48 小时开发的现实条件中理解。他没有要求短期作品达到商业产品的素材量和细节精度，但也没有因此放弃对风格、操作和完成度的判断。这种评价对于消费端研究有价值：玩家可以容忍小型作品的规模和细节限制，但容忍并不意味着不再判断，核心体验是否成立、风格是否协调仍然是必要条件。

音效与文本：逐字稿前置字段均标记为不涉及，不能据此推断为正面或负面

5. AI 知情、识别与评价影响

受访者在游玩前就知道作品与 AI 参与有关，游玩中也识别到 AI 参与，识别依据被记录为“活动类型”。这意味着他的 AI 认知不是在看到明显的生成瑕疵后才形成，而是由 TapTap 制造专场这一平台语境直接提供。尽管如此，前置字段仍记录“AI 标签是否影响评价：否”，访谈中也明确表示自己不会因为作品是 AI 参与制作就改变期待。

受访者把 AI 放在工具层面理解，并明确说自己也是使用 AI 的游戏开发者。在他的判断里，玩家最终应关注作品质量；创作者如何使用 AI、如何把结果呈现出来，则是创作团队需要承担的责任。这并不等于“只要好玩，所有 AI 参与都可接受”，而是把评价重点从标签本身转向玩法、审美、完成度和创作者控制。由于受访者具有开发者背景，这一判断需要与纯消费端样本分开比较，不能直接代表所有玩家。

6. 创作者投入感与接受边界

受访者给出的接受边界不是“禁止使用 AI”，而是要求创作者对结果负责。他接受“创作者有诚意地使用 AI 工具创作”，不接受没有诚意、制作粗糙的作品。对于直接调用通用大模型生成一个“美少女”并原样放进游戏的做法，他认为玩家很容易认出这是通用生成结果，并产生反感。问题不只在于素材来自 AI，还在于素材没有进入游戏整体的美术基调和情绪系统。

他进一步提出了较具体的创作控制要求：先确定游戏的美术基调，例如像素、油画或现代艺术；再确定整体情绪是欢快、压抑还是悲伤，并据此决定调色板和冷暖关系；最后让生成素材服从这些约束，再通过色调和风格后处理统一呈现。由此可见，受访者判断“诚意”的依据是有没有明确的创作意图、筛选和后处理，而不是创作者是否完全拒绝 AI。

访谈原话：“通常这些都是直接通用大模型生成的绘画，玩家一眼就能认出，非常没有诚意，容易引起反感。”这句话具体指向的是可感知的通用化和突兀感，不能被简化为对所有 AI 美术的普遍否定。

7. 继续游玩、分享与付费意愿

受访者愿意继续体验这类 AI 平台生成作品，也愿意分享；对付费的回答则是条件式的：作品完成度高时愿意考虑付费。他不预设广告收入或买断收入哪一种模式更正确，而是尊重制作组选择，最终仍看具体质量。这说明 AI 标签没有阻断继续游玩和传播意愿，但付费需要更高的质量证明。

这里的“愿意付费”是访谈中的陈述意愿，没有对应的实际支付记录、价格区间或付费实验结果。因此它可以用于理解消费门槛，不能用来推算价格敏感度、付费转化率或具体商业模式。对 H5-07 而言，完成度、玩法成立和创作者投入感是付费前置条件，AI 本身不是付费理由。

8. 研究解释与证据边界

H5-07 对消费端研究的直接贡献，是把“AI 标签是否影响评价”落到了一个具体的一站式平台作品体验上：受访者知道 AI 参与，也能识别来源，但仍然依据玩法、操作、整体美术和完成度判断作品。这与只听说 AI 游戏、只看宣传视频或评价自己制作的 Mod 有明显区别。

同时，材料更适合支持以下有限判断：第一，AI 标签可以保持中性，前提是作品的核心体验和质量没有明显问题；第二，玩家对 AI 作品的反感往往由通用化、粗糙和缺少创作控制触发；第三，操作键位和核心循环等传统游戏质量因素，可能比抽象的 AI 争议更直接地影响试玩感受。材料不支持音效、文本、长期留存、真实付费和跨作品偏好等更宽结论。

四、消费端访谈记录字段整理

受访者编号：H5-07

样本类型：AI 游戏作品消费者

样本标签：TOOL-C | CP-T（来自样本管理表；CP-T 为 TapTap 制造消费端直接对应）

是否为大学生：是

平时游戏频率：每日

常玩游戏类型：线上 FPS/PVP、动作游戏、RPG、文字冒险和独立游戏

体验作品：TapTap 制造专场中的多人拍照聚会游戏；具体作品名未确认

对应平台：TapTap 制造专场线下活动

是否实际游玩：是

游玩时长：少于 2 小时

是否完整体验：是；前置字段记录如此，但未提供更细的关卡或局数记录

游玩前是否知情 AI 参与：是

游玩中是否识别 AI 参与：是

AI 识别依据：活动类型 / TapTap 制造专场语境

玩法评价：非常有趣，有创意；拍照抢夺机制有明确互动性

操作评价：简单易懂，但键位仍有优化空间

美术评价：Q 版像素与有限场景整体协调；细节仍可优化

音效评价：不涉及

文本评价：不涉及

完整度评价：完整，细节可进一步优化

稳定性评价：好

耐玩性评价：好；缺少更具体的长期游玩说明

是否愿意继续玩：是

是否愿意分享：是

是否愿意付费：作品完成度高则愿意考虑

AI 标签是否影响评价：否

主要接受边界：创作者有明确审美、情绪和风格控制，并有诚意地使用 AI 工具

主要不接受边界：通用模型素材未经筛选直接使用、制作粗糙、风格和情绪突兀

可用于验证或反驳的假设：原始字段标注 H5、H2；具体假设映射需以正式研究假设编码表为准

五、可用于验证或修正的研究判断

AI 标签与质量评价：H5-07 知情且能识别 AI 参与，但评价未因此下降，支持“标签本身不是决定性负面因素”的判断

平台作品的最低成立条件：核心玩法要能被理解和参与，操作不能形成明显阻碍，整体美术需要保持基本统一

创作者控制：玩家接受的是创作者主导下的 AI 辅助，不是通用素材的无筛选拼接

消费转化：继续玩和分享门槛低于付费；付费需要更高完成度，不能仅靠 AI 标签或新鲜感推动

样本使用边界：可作为 TapTap 制造消费端直接对应案例；不单独承担音效、文本、长期留存和真实付费结论

六、综合研究结论

H5-07 是一名具有较丰富游戏经验的 AI 游戏消费端受访者，实际体验了 TapTap 制造专场中的作品，并对一款多人拍照聚会游戏形成了相对完整的判断。其评价结构很清楚：玩法有趣、创意成立，操作容易理解但键位仍可优化；Q 版像素与有限场景整体协调，作品完成度已经足以游玩，但细节还有提升空间。玩家并没有因为作品规模小或使用 AI 就直接否定它。

更重要的是，受访者把 AI 参与放在“创作者是否真正控制作品”的框架中判断。AI 作为工具本身不会降低评价，但通用模型素材直接拼接、缺少美术基调和情绪设计，会被感知为没有诚意。由此形成的消费端链路可以表述为：知道或识别 AI 参与 -> 观察作品是否有创作控制和质量完成度 -> 评价核心玩法与整体风格 -> 决定继续玩、分享或在完成度足够时付费。

该结论仍应限定在单个 TapTap 制造直接对应样本内。受访者同时是 AI 游戏开发者，可能比普通玩家更能理解 AI 的工具属性；作品名称、音效、文本、实际付费和长期留存信息也未被逐字稿记录，管理表的 is_core_sample 也尚未更新。因此，H5-07 适合与 H5-01、H5-02、H5-08 的消费端补充材料以及后续纯玩家样本并列比较，用来检验“质量、创作者投入感和使用层面”是否比 AI 标签本身更能解释接受度。

原始附件

本样本原始材料未包含嵌入图片或附件。

访谈逐字稿

原始材料：H5-7逐字稿.docx

受访者编号：H5-7

样本类型：AI 游戏作品消费者

样本标签：CON-TM / CON-OTHER / CON-S / X

是否为大学生：是

平时游戏频率：每日

常玩游戏类型：线上pvp，动作游戏，RPG，以及诸多独立游戏

对应平台：TapTap 制造是否实际游玩：

游玩时长：<2h

是否完整体验：是

游玩前是否知情 AI 参与：是

游玩中是否识别 AI 参与：是

AI 识别依据：活动类型

玩法评价：非常有趣，有创意

操作评价：简单易懂，键位有待优化（可能未使用常规操作习惯类型）

美术评价：不看细节的话比较和谐美观

音效评价：（不涉及）

文本评价：（不涉及）

完整度评价：完整，细节可进一步优化

稳定性评价：好

耐玩性评价：好

是否愿意继续玩：是

是否愿意分享：是

是否愿意付费：完成度高则愿意

AI 标签是否影响评价：否

主要接受边界：创作者有诚意地使用ai工具创作

主要不接受边界：没有诚意，制作粗糙

可用于验证或反驳的假设：H5,H2

1:你平时玩游戏的频率大概是多少？ 主要玩什么类型的游戏？

受访者:大概是每天都玩吧，主要玩比较火的线上游戏，比如三角洲行动或者APEX，还有一些单机的游戏，这部分比较多，我更倾向于玩一些我认为比较有意思的游戏，比方说杀手47，艾尔登法环，黑魂，巫师三这类的动作游戏，还有类似beholder，请出示证件，别观察猴子之类得文字冒险，然后还有一些独立游戏，比如太吾绘卷，for the king，魔王将至这样的。

2:尝试这些游戏时，你比较在意哪种因素？ 玩法，题材，或者热度之类

受访者:当然，在玩比较火的fps游戏的时候比较注重解压和社交的属性，也能归类于热度的部分，但是刨除这个之外个人觉得可能是表达还有与之结合的游戏机制吧，打比方说会很看中游戏的整体的艺术的风格，叙事，这些东西会作用于游戏的机制还有ui，这会给人带来很棒的体验

3:你是否会因为作品是 AI 参与制作而改变期待？

受访者:不会，因为实际上ai参与游戏制作已经是正在进行时了，我也是使用ai的游戏开发者，从立场上来说，我更希望玩家可以关注游戏最终完成的质量

4:你在平台上是否玩过ai参与制作的游戏？如果有，他是什么玩法？

受访者:玩过，我记得当时聚光灯的taptap制造专场的线下场，有很多很惊艳的游戏，题目是"五四三二一"我印象比较深刻的是有一个拍照的聚会游戏，用简单的手柄操作，几个人去抢拍照镜头的玩法，非常有趣。

5:有没有某个机制让你觉得不合理？ 如果和普通小游戏或独立游戏相比，它的玩法差距在哪里？

受访者:除了键位来说，都没什么太大问题，比对的话只有完成度和机制变化了，但是毕竟是48小时的作品，这个属于正常的范畴

6：你如何评价这款作品的美术？ 比如角色、场景、UI 是否协调？

受访者:整体的风格不看细节的话，q版像素加上有限场景的配合，是比较和谐美观的

7:你是否愿意继续玩这类 AI 平台生成作品？ 如果足够完善，你是否愿意为这类作品付费？

受访者:如前文所提，我对是否ai平台生成或者掺杂ai制作的这个底层关注并不大，我比较看中的方面是几个，一是制作人的审美，整体美术风格如何，ui交互如何 二是玩法类型，是否是我感兴趣的类型，流程是否完整可玩 这两个点做好了，就会游玩这个游戏。对于付费我会考虑完成度高的游戏，对于制作组在广告收入和买断收入之间的抉择我表示尊重，当前需要看具体的游戏质量。

8:刚才你提到的这款游戏，它在玩法、操作、完整度、美术、处理得当，很有可玩性 我的理解是：玩家是否接受一站式 AI 平台生成或实质参与制作的作品，并不只取决于 “是否使用 AI”，而取决于作品本身是否好玩，整体美术符不符合审美，创意是否得当。

受访者:我认为是这样的，当然，不排除目前部分玩家对ai生成式内容抱有怀疑态度，我个人更多承认的是ai的工具属性，如何去使用ai，是主创团队的问题，但是如何呈现作品，这个是主创团队需要思考的问题。

:也就是有创作者投入感，并不是快速生成的内容？

受访者:嗯，对的，但是详细来说.....我将其称之为审美吧，不论是对美术的，还是整体玩法，还有音乐音效的。这是整一个游戏工程的体现。我举个例子，比如你想要一个角色，直接和ai说，我要一个美少女，然后ai给你使用通用模型给你跑了一个角色出来，然后把他做进游戏里，这个是通常的做法，但是这个有什么问题呢。

通常这些都是直接通用大模型生成的绘画，玩家一眼就能认出，非常没有诚意，容易引起反感

没有考虑整体基调，这个角色为什么是这样的，会显得突兀 我认为比较妥当的做法是在做这些事情之前，要先定几个东西 美术基调，你的游戏是什么美术基调？像素风？油画？现代艺术？解构艺术？这是决定整体的美术风格。然后是情绪，你的游戏整体的情绪是欢快，还是压抑，还是悲伤，这决定你的游戏的调色板，冷暖颜色的碰撞会影响你游戏的情绪

然后带着这些东西，过去去做符合你游戏基调的素材，加上色调和风格的后处理。或许会让玩家感觉到主创团队的用心之处。

问卷作答情况

以下仅收录研究相关题目，排除联系方式、同意、抽奖、路由、质量控制和技术字段；题干与选项按该答卷对应的历史题库版本呈现。

基础信息与筛选

题号与原题干	作答
S1 你当前的学生身份是？	本科在读
S2 你的专业或主要学习方向更接近哪类？	数字媒体/游戏/动画
S3 过去 6 个月，你是否参与过游戏或互动作品创作？	做过个人 Demo/原型
S4 你在游戏/互动作品创作中承担过哪些工作？	程序/功能实现；美术/视觉
S5 你是否使用过以下 AI 工具/平台参与游戏创作？	普通 Claude/ChatGPT 等对话工具；Codex / Claude Code 等专业 AI 编程代理
S5A 你从什么时候开始较为频繁地使用 AI 工具？	1 年以上

题号与原题干	作答
S5B 你目前使用 AI 工具的频率大约是？	每周使用 3—5 次
S6 你使用上述工具时，是否有具体游戏/互动项目或可运行原型作为对象？	有可运行/可展示原型
S7 你是否实际玩过“一站式 AI 游戏制作平台”生成或实质参与制作的作品？	只看过视频/宣传
S8 你能否指出具体平台或作品，并说明自己玩到了什么程度？	能指出但只浅尝

生产端与创作经历

题号与原题干	作答
P1 请概述你最近一个使用 AI 参与的游戏/互动项目。（项目类型、个人/团队、引擎或技术栈、当前阶段）	最近在使用ai个人产生短视频，漫剧之类的
P2 你的程序基础更接近哪一项？	学过基础课程
P3 你使用过哪些引擎、框架或开发方式？	Unity
P4 如果不使用 AI，你能否独立完成本项目中的程序或功能实现任务？	能完成但效率低/质量不稳
P5 在该项目中，你具体使用过哪些工具？	普通 Claude/ChatGPT
MAIN_TOOL 请选择一个你在游戏/互动项目中使用最深入，或最近主要使用的专业 AI 编程/开发工具。	Codex
CC1 【该工具】你主要让它处理哪些程序任务？	需求拆解；代码生成
CC2 【该工具】它进入项目的哪些阶段？	关卡/玩法迭代
CC3 【该工具】请评价它对程序工作的帮助。	提高速度：4 降低调试难度：4 帮助理解项目：4 帮助完成原本不会的功能：4 减少重复劳动：4 提升代码质量：4
CC4 【该工具】AI 生成内容后，你通常如何处理？	阅读后小改
CC5 【该工具】你对生成代码的判断与维护能力如何？	能判断是否正确：4 能定位问题：4 能解释代码：4 能后续维护：4 能控制依赖风险：4
CC6 【该工具】相关产出最终是否进入项目？	只少量进入
CC7 【该工具】整体看，它主要是提高已有能力，还是补足原本不足的能力？	两者都有但任务不同

题号与原题干	作答
P6 整体来看，AI 工具让你的游戏程序/原型开发发生了哪些变化？	开发速度更快：3 能做更复杂任务：3 更依赖 AI：3 人工检查更多：3 返工更多：3 更愿意尝试新玩法：3 更容易形成可玩结果：3

开放题

题号与原题干	作答
P7 请还原一个最典型片段：你要做什么程序/功能任务，使用了哪个工具，AI 给出什么，你如何修改，最后结果如何？	我使用ai帮我生成一段抗战时期的视频，结果他说血腥暴力不能生成

消费端体验与评价

题号与原题干	作答
C1 你体验过的平台或作品是什么？	Tap tap
C2 该作品来自哪类平台？	TapTap 制造
C3 你实际玩到了什么程度？	完成一关/一局/一回合
C4 你过去的游戏体验更接近哪类？	小游戏/网页游戏；UGC/创意工坊
C5 请评价该作品的基础游玩体验。	玩法清晰：3 操作顺畅：3 目标反馈：3 规则平衡：3 稳定性：3 内容量：3 耐玩性：3
C6 请评价该作品的作品内容品质评价。	美术质量：3 视觉统一性：3 音效/音乐：3 文本自然度：3 内容重复感：3 整体完成度：3
C7 和你平时玩的普通小游戏、独立游戏相比，这款 AI 制作的作品有没有给你带来不一样的游玩感受？如果有，请简单描述具体体现在哪里。	AI制作游戏对于我来说，目前还是不够成熟的
C8 体验前，你是否知道该作品有 AI 参与制作？	大概知道
C9 如果没有被告知，你能否从作品中感知 AI 参与？依据是什么？	他没有完全照顾到玩家的内心的感觉
C10 知道或感知 AI 参与后，对你的评价有什么影响？	略加分
C11 以下因素会如何影响你对这类作品的信任与接受？	AI 标注透明：3 版权清晰：3 创作者的用心程度：3 作品稳定性：3 玩法原创性：3 价格/付费方式：3

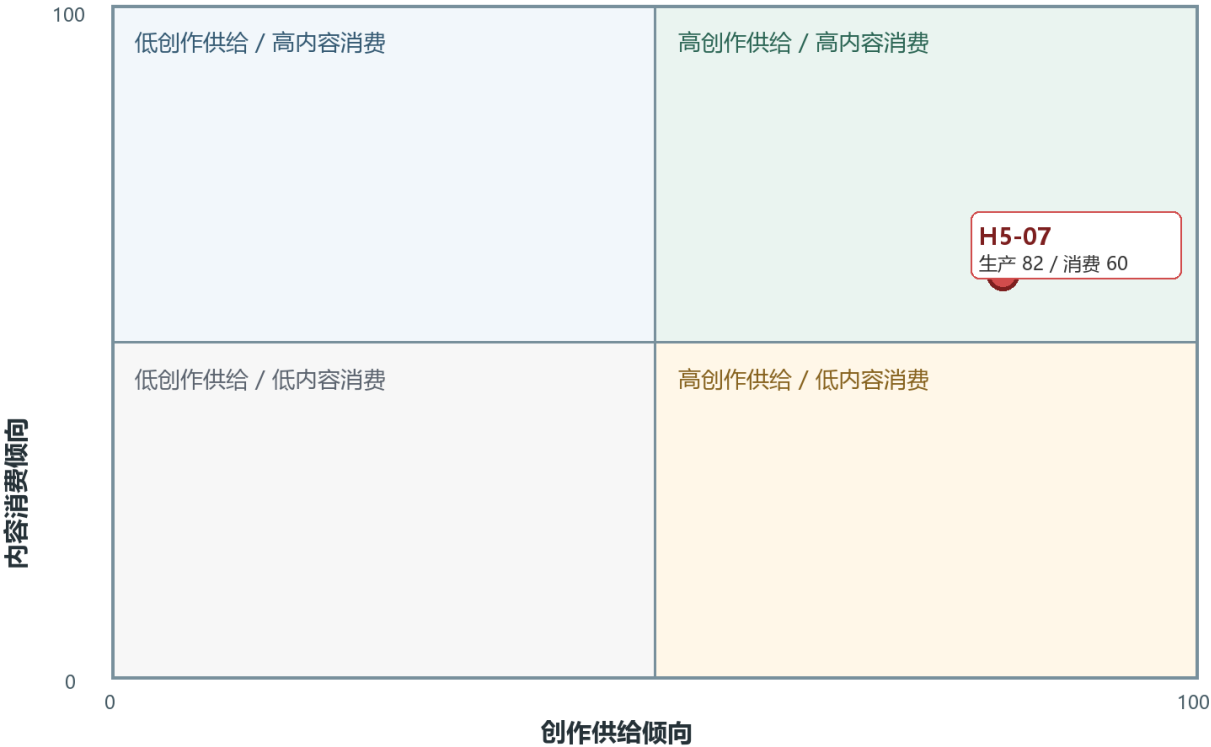
题号与原题干	作答
C12 你是否愿意继续游玩、分享或付费支持这类作品？	继续游玩：3 推荐给朋友：3 关注作者/平台：3 小额付费：3 提交游玩反馈与建议：3
C13 与普通小游戏、独立游戏或 UGC 作品相比，你觉得它最大的优势和问题分别是什么？	内容更加丰富，画质更加好

研究补充

题号与原题干	作答
G1 你总体上如何看待 AI 进入游戏创作？	提高效率：3 降低门槛：3 带来新玩法：3 削弱创作诚意：3 增加同质化：3 带来版权风险：3
G2 你认为 AI 参与程度是否应该被明确标注？	可由作者选择
Q4.3 如果你没有实际玩过一站式 AI 平台作品，主要原因是什么？	找不到作品

生产倾向 × 消费倾向象限图

横轴为创作供给倾向，纵轴为内容消费倾向；50 分为高低分界。



H5-07：生产 82，消费 60，高创作供给 / 高内容消费

关联维度分数

创作	95
效率	79
探索	73
接受	66
付费	39
信任	59

A.5 样本五 | H5-08

样本基本信息

样本编号	H5-08
样本定位	以 AI 辅助 Mod 开发为背景的补充样本，涉及反编译、存档功能扩展、AI 编程工具使用和独立调试。
访谈执行人	姓名待补
摘要来源	访谈摘要：H5-08(1).docx
逐字稿来源	访谈逐字稿2.docx
关联答卷	e63661ca-24f5-464e-8488-2634b0ddf043

本文件保留该样本的访谈摘要、逐字稿和源材料附件；最后附上关联问卷中的研究相关作答与既有象限结果。

访谈摘要

原始材料：访谈摘要：H5-08(1).docx

访谈摘要

一、访谈基本信息

受访者编号：H5-08

访谈时间：2026/07/07 14:06:00 - 2026/07/07 15:19:00

导出时间：2026-07-18 0:36

二、样本判定

本访谈对象可判定为 WorkBuddy/Trea/VSCode AI 游戏 Mod 开发补充样本。受访者为风景园林专业大四学生（五年学制），有一定的 Python 编程基础和 PBR 建模美术基础，以独立开发者身份为 Steam 上的小型独立游戏制作 Mod（存档功能扩展）。其使用场景包括反编译游戏代码、梳理逻辑、让 AI 生成功能代码、自行调试和测试维护。

受访者提供了项目文件夹截图和 Mod 运行能力，但缺少完整的多轮 AI 对话记录,更多项目截屏。该样本更适合作为补充样本，用于观察"非计算机专业、编程基础较弱、以功能实现和练手为目标"的 AI 辅助 Mod 开发模式。

样本标签：TOOL-C

判定理由：有真实 Mod 项目经历和具体 AI 工具使用链路，产出可运行，但缺少完整过程材料（AI 对话记录未保留）和可追踪代码链路

具体项目：Steam小游戏Mod

项目阶段：功能完整、可运行，但未发布

过程材料：逐字稿、项目文件夹截图 1 张、Mod 项目文件（.csproj 及 .cs 源文件）

三、开发者端样本分析

1. 基本背景与开发基础

受访者为风景园林专业大四学生（五年学制），非计算机科班出身。其编程基础有限：有一些 Python 基础，但自述"代码能力弱"，在 Mod 开发中基本全部授权 AI 生成代码。美术方面有一定基础，曾报班学习过 PBR 建模流程，但在当前项目中未涉及美术资产的原创，仅聚焦代码功能实现。

该案例呈现了一种典型的“跨专业独立开发者”画像：有审美/设计专业背景（风景园林），具备初步的编程能力但不足以独立开发，借助 AI 工具填补代码实现能力缺口。受访者对自身定位清晰：“第一是练手，第二是确实想要实现功能”，目标不是做出商业级作品，而是验证想法和实现具体功能。

该样本支持提纲中的基础假设：用户能否用好 AI 开发工具，取决于编程基础和对项目结构的理解程度。受访者虽然代码能力弱，但有 Python 基础，能够理解反编译后的代码逻辑，并通过提示词工程与 AI 协作。如果完全没有编程基础，即使是 AI 工具也难以产出可运行结果。

2. 具体项目还原

受访者的项目是为 Steam 上的一款小型独立游戏制作存档功能 Mod。该游戏体量小，未开放创意工坊，Nexus Mods（N网）也没有相关 Mod，受访者因此决定自己尝试制作。

项目具体信息：

项目名称：NightSaveMod（存档功能 Mod）

目标游戏：Steam 上的独立游戏（体量较小）

功能目标：为游戏添加及时存档功能

开发工具链：反编译工具 → Trea/VSCode 自带 AI（前期梳理逻辑）→ WorkBuddy（主力生成代码）

项目文件：NightSaveMod.csproj（C# 项目）、Plugin.cs 等 C# 源文件，包含 bin、Data、Manager、Obj、Patches、Serialization、Utils 等标准 Mod 工程目录结构

项目开发流程为：反编译游戏 → 用 Trea/VSCode AI 梳理代码逻辑 → 提需求让 AI 生成代码 → 使用 WorkBuddy 进行授权调试和依赖管理 → 测试维护。

3. 具体任务使用过程

的 AI 工具使用链路可分为两个阶段：

第一阶段：使用 Trea 和 VSCode 自带的 AI。受访者将游戏反编译后，把代码丢给 AI，让 AI 梳理逻辑，然后提需求让 AI 跑。这一阶段的问题是“不像 WorkBuddy 可以授权自行调试”，调试效率低。

第二阶段（主力）：切换到 WorkBuddy。WorkBuddy 可以授权后自行调试，做好必要的调用和依赖，显著减少了“开发→报错→复制日志→粘贴日志→解决→开发”的循环。受访者描述其 workflow 为：重新把源文件给 WorkBuddy → 理解 → 需求 → 制作。

受访者的提示词策略较为简单，主要通过文本描述功能需求。当被问及人工控制环节时，受访者明确表示“只有提示词工程”，没有其他人工操作环节。

关于首次生成结果，偏差不大，主要只是一个功能的实现，有 Bug 需要后续测试维护。处理方式是重新把源文件给 WorkBuddy，走“理解→需求→制作”的循环，共 8 轮后达到可运行状态。

4. 效率提升与能力补足

WorkBuddy 对受访者的核心价值是“不用自己手动调试”和“自察解决能力强”。受访者形容其最佳环节是“减少了开发→报错→复制日志→粘贴日志→解决→开发的循环”，这对应 Mod 开发中最耗时的调试阶段。

效率数据：WorkBuddy 使用时间共 4 小时，8 轮生成，最终得到可运行的 Mod 项目。对一名代码能力较弱的跨专业学生而言，这一产出效率具有实质意义——如果没有 AI 工具，受访者很可能无法独立完成此项目。

但受访者也指出了能力补足的代价：“因为自己代码能力弱，基本都是全部授权 AI，后期维护成本高”。这验证了提纲中的“能力补足可能带来依赖风险”假设——AI 让非专业开发者能产出可运行代码，但代码的长期维护和理解成本被转嫁到未来。

5. 人工控制、判断与维护成本

受访者对 AI 输出的人工控制程度较低。当被明确问及“人工控制的内容只有通过文本调试了吗？还有什么是你人工操作的环节？”时，受访者回答“没有，只有提示词工程”。这意味着受访者不逐行审查 AI 生成的代码，也不手动修改代码逻辑，完全依赖 AI 的 self-debugging 能力。

维护风险在此样本中表现明显：

调试困难：Trea 和 VSCode AI 不像 WorkBuddy 可以授权自行调试，前期调试体验差

后期维护成本高：因为全部授权 AI 生成代码，受访者自述“后期维护成本高”

审查困难：代码能力弱导致难以独立审查 AI 产出，发现问题后仍需依赖 AI 修复

该样本是一个典型的“低审计、高依赖”使用模式。对短期功能实现来说可以接受，但如果项目需要扩展功能或长期维护，缺乏代码理解将构成显著风险。

6. 产出落地与项目影响

WorkBuddy 的产出在项目成功落地：Mod 可以运行，实现了存档功能。产出状态为“能跑通的项目”，包含完整的 C# 项目文件（.csproj）和插件源文件（.cs），目录结构规范（包含 bin、Data、Manager、Obj、Patches、Serialization、Utils 等标准模块）。

但项目未发布，原因有二：一是社群小，目标游戏的玩家社群规模有限；二是“开发者明确表示不会做这个功能”，受访者尊重开发者的选择，仅在线下朋友间传递使用。

该样本说明：AI 工具能让跨专业独立开发者完成功能级 Mod 开发并产出可运行结果，但产出的传播和影响力受限於社群规模、开发者意愿和发布渠道等多重因素。AI 工具解决了“能不能做”的问题，但“能不能被更多人用”涉及更复杂的生态条件。

7. 留存、迁移与付费

受访者使用 WorkBuddy 的频率为"每月约 1 个项目"。使用过程中没有付费，原因是"腾讯新出的，额度给得宽裕，而且本来工程量就小"。

付费意愿方面，受访者态度务实：如果付费，会先对比有无平替，没有的话还是会付费。付费支撑项为"快速生成可玩原型"和"多轮修改稳定"；降低付费意愿项为"生成质量不稳定""修改困难""无法导出""无法长期维护"。

该样本反映出一个典型的个人独立开发者付费画像：使用频率不高（月均 1 个项目），对价格敏感，会先评估平替方案，但工具确实能解决问题时愿意付费。留存基础是"能帮我实现功能"的实用价值，而非品牌忠诚度。

8. AI 参与制作的市场判断

受访者对 AI 参与游戏/Mod 制作持明确的分层态度：

在艺术性维度，受访者认为"游戏的美术、音乐等艺术层面上，使用 AI 削减了艺术性；但在代码上使用 AI，并没有多大区别"。这一区分非常清晰：代码是功能性层面，不影响作品的艺术表达；而美术和音乐直接构成艺术性本身，AI 参与会削弱原创艺术价值。

在商业价值维度，受访者认为"使用 AI 会增加开发效率，自然就减少社会生产时长，即价值，完全可以说是廉价了"，但同时表示"商业廉价并无什么不好，但最好不要艺术廉价"。这一立场兼顾了产业效率逻辑和艺术原创保护。

关于 Mod 的 AI 标签，受访者表示"AI 参与度极高，会主动告诉"玩家，并认为"Mod 的话，玩家容忍度蛮高的"。这说明在 Mod 生态中，玩家对 AI 参与的接受度高于对游戏本体的接受度。

在美术参与上，受访者明确表示"不太希望 AI 参与美术、音乐这两个方向"，当前项目也"不会考虑 AI 生成代码结合自己原创美术资产"，因为"功能只是一个拓展，没有涉及美术、音乐、玩法的艺术层面"。

9. 理想中的 AI 开发工具

受访者希望平台未来在以下方面变强：

生成更完整的玩法

支持复杂的系统

支持长期修改

当被问及"如果平台未来更强，甚至一句话可以生成一个游戏，创作者还需要负责什么"时，受访者回答：

玩法创意

用户需求提炼

受访者进一步阐述："对于创作者的价值，平台为其表达想法提供的实现途径，也为探索 AI 边界、探索游戏边界提供的场景。对于平台的价值，创作者的活动、需求提供了迭代方向。同时为人机交互、AI 伦理等议题提供实际依据也是平台自身对商业实现进行过渡。"

这一回答体现了受访者对"人类创作者核心价值"的清晰认知：AI 负责实现，人类负责创意和需求定义。同时，受访者也从平台生态角度理解了创作者反馈对 AI 工具迭代的意义。

四、开发者端访谈记录字段整理

受访者编号：H5-08

样本类型：AI 辅助 Mod 开发者

样本标签：TOOL-C

是否在校大学生：是

专业与年级：风景园林专业；大四（五年学制）

编程基础：有少量 Python 基础，自述代码能力弱

美术基础：有 PBR 建模流程学习经历

使用工具：WorkBuddy（主力）、Trea、VSCode 自带 AI

具体项目或作品：Steam 小游戏 Mod

项目类型：游戏 Mod（功能扩展）

当前阶段：功能完整、可运行，未发布

工具使用时间：WorkBuddy 使用约 4 小时，8 轮生成

使用任务类型：代码生成、功能实现、调试、依赖管理

是否涉及至少两类实质任务：是（代码生成 + 调试）

产出状态：可运行的 C# Mod 项目，含完整工程结构

人工修改程度：极低，仅通过提示词工程控制，不逐行审查代码

是否提高效率：是，避免了手动调试的繁琐循环

是否补足能力：是，让代码能力弱的跨专业学生能独立完成 Mod 开发

是否增加维护成本：是，自述后期维护成本高，全部授权 AI 导致审查困难

是否持续使用：是，频率约每月 1 个项目

付费判断：当前免费（额度宽裕）；如有平替会对比，无平替仍会付费

是否有过程材料：逐字稿、项目文件夹截图、Mod 源文件

主要有效场景：小型 Mod 功能开发、快速原型验证、独立开发者练手

主要失效场景：复杂系统调试、长期维护、美术/音乐方向（受访者不希望 AI 参与）

可用于验证或反驳的研究假设：验证"AI 工具能让非科班开发者完成功能级产出"能力补足伴随维护风险"提示词工程可替代部分代码审查",补充说明 Mod 场景中玩家对 AI 标签容忍度较高

五、消费端补充样本分析

1. 消费端样本定位

受访者符合消费端补充样本条件，而非消费端核心样本。核心原因是：受访者是 Steam 游戏玩家和 Mod 使用者，对 AI 参与游戏/Mod 制作有明确态度和判断；但其消费体验主要集中在传统独立游戏和自研 Mod 上，缺少对一站式 AI 游戏平台作品的完整体验链路。

因此，受访者的消费者价值不在于评价某款 AI 平台作品，而在于呈现"跨专业独立开发者兼玩家"如何判断 AI 参与游戏制作的边界，以及 Mod 生态中玩家对 AI 标签的接受度。

2. 玩家背景与参照系

受访者是 Steam 平台独立游戏玩家，对 Mod 生态有实际参与经验（制作并使用 Mod）。虽然专业是风景园林，但其游戏消费经验和 Mod 开发实践使其对"玩家视角"和"开发者视角"都有切身理解。

在判断 AI 参与时，受访者能清晰区分不同层面：代码层面的 AI 参与不影响艺术性，玩家不易察觉；美术和音乐层面的 AI 参与则直接削减艺术性，玩家容易感知。这种分层判断说明其评价标准不是简单的"是否使用 AI"，而是 AI 参与的是哪个层面、是否影响作品的核心体验。

3. AI 标签与接受边界

受访者对 AI 标签的态度明确且细致：

Mod 场景：玩家容忍度蛮高；会主动告知玩家 AI 参与度极高

游戏本体美术/音乐：AI 参与削减艺术性，不希望 AI 参与这两个方向

代码层面：AI 参与与艺术性无多大区别

关于"AI 作品让玩家付费最需要证明什么"，受访者回答：

证明人在艺术创作中处于主导地位

证明游戏作品具有游玩价值

这一判断与提纲中的假设一致：玩家接受 AI 参与的前提是能感受到人类创作者的主导性和作品本身的游玩价值，而非 AI 参与的多寡。

4. 付费、分享与接受边界

受访者认为 AI 辅助开发的游戏有付费价值，但前提是"具有游玩价值"和"人在艺术创作中处于主导地位"。

在 Mod 发布问题上，受访者选择不发布，理由是社群小且"开发者明确表示不会做这个功能"，尊重开发者选择。这反映了 Mod 开发者的发布决策不仅取决于技术可行性，还受开发者意愿、社群规模和平台政策共同影响。

接受边界总结：

可接受：AI 辅助代码实现功能，开发者主导创意和需求

可接受：Mod 场景中的高 AI 参与度（玩家容忍度高）

不可接受：AI 主导美术和音乐创作（会削减艺术性）

不可接受：纯 AI 主导、无人工筛选和控制的作品

六、消费者访谈记录字段整理

受访者编号：H5-08

样本类型：AI 游戏/Mod 消费者补充样本

样本标签：CON-S

是否为大学生：是

平时游戏频率：Steam 独立游戏玩家，具体时长未明确

常玩游戏类型：Steam 独立游戏、Mod 使用

体验作品名称：Steam 上的独立游戏（未具名）及其 Mod

对应平台：Steam

是否实际游玩：是

游玩时长：未明确

是否完整体验：是（Mod 可运行）

游玩前是否知情 AI 参与：是（自身为开发者）

游玩中是否识别 AI 参与：是

AI 识别依据：自身开发经验、代码层面判断

玩法评价：Mod 功能实现成功，存档功能可运行

美术评价：不希望 AI 参与美术，认为会削减艺术性

文本评价：未涉及

完整度评价：Mod 功能完整可运行，但未发布

耐玩性评价：未明确

是否愿意继续玩：是（线下朋友间传递使用）

是否愿意付费：认为 AI 作品可付费，前提是具有游玩价值和人为主导

AI 标签是否影响评价：Mod 场景容忍度高；美术/音乐场景敏感

主要接受边界：代码 AI 辅助可接受；美术/音乐 AI 主导不可接受

主要不接受边界：AI 主导艺术创作、纯 AI 拼接无人工控制

可用于验证或反驳的假设：验证"玩家对 AI 标签的接受度因场景而异：Mod > 游戏本体；代码 < 美术/音乐""付费前提是游玩价值 + 人为主导"

七、综合研究结论

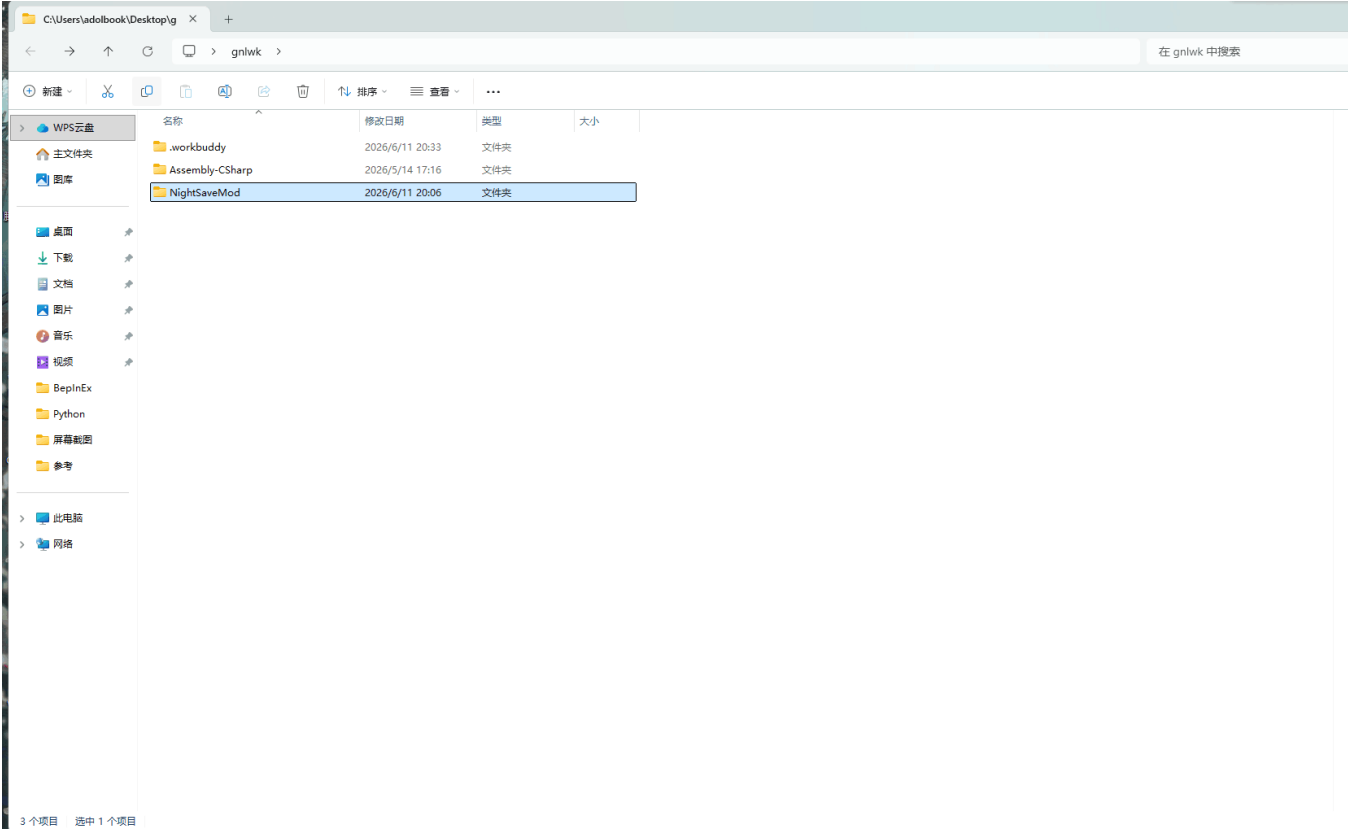
受访者 展示了一种典型的跨专业独立开发者使用 AI 工具完成 Mod 开发的模式：风景园林专业大四学生，有少量 Python 基础和 PBR 建模经历，代码能力弱，通过 WorkBuddy 的授权调试和自察能力，在 4 小时、8 轮生成内完成了 Steam 独立游戏存档 Mod 的可运行项目。

从开发者端角度分析，首先，AI 工具（WorkBuddy）在小型功能级 Mod 开发中的核心价值是替代手动调试，显著减少“开发→报错→复制日志→粘贴日志→解决→开发”的循环。其次，用户的编程基础虽弱但非零，能够理解反编译代码逻辑并通过提示词工程与 AI 协作，这是产出能落地的必要条件。最后，全部授权 AI 的模式虽然降低了入门门槛，但带来了后期维护成本高的风险，验证了其“能力补足伴随依赖风险”的假设。

从消费端角度分析，受访者 呈现了清晰的分层判断：代码层面 AI 参与不影响艺术性，玩家不易察觉；美术和音乐层面 AI 参与削减艺术性，玩家敏感。Mod 场景中玩家对 AI 标签容忍度较高，但游戏本体的艺术层面需要人类主导。付费前提是“游玩价值 + 人为主导”。

受助者 呈现了“开发者-玩家双重身份”的样本价值：作为 Mod 开发者，他证明了 AI 工具能让非科班开发者独立完成功能级产出；作为玩家，他对 AI 参与的分层态度反映了实际消费者对 AI 游戏/Mod 的接受边界。该样本因此可以同时支持开发端和消费端两组假设：生产端看 AI 工具能否让跨专业开发者完成功能级产出，消费端看 AI 参与层面（代码 vs. 美术/音乐）和场景（Mod vs. 游戏本体）如何塑造玩家接受度。

原始附件



原始附件 1：访谈逐字稿2.docx（1657 × 1018）

名称	修改日期	类型	大小
bin	2026/6/11 23:21	文件夹	
Data	2026/6/11 19:55	文件夹	
Manager	2026/6/11 19:57	文件夹	
obj	2026/6/11 23:21	文件夹	
Patches	2026/6/11 20:01	文件夹	
Serializers	2026/6/11 19:57	文件夹	
Utils	2026/6/11 19:55	文件夹	
NightSaveMod.csproj	2026/6/11 20:07	C# Project 源文件	4 KB
Plugin.cs	2026/6/11 22:49	C# 源文件	3 KB

原始附件 2：访谈逐字稿2.docx（735 × 339）

访谈逐字稿

原始材料：访谈逐字稿2.docx

访谈逐字稿：

基本信息

- 导出时间：2026-07-18 0：36
- 访谈时间：2026/07/07 14:06:00 - 2026/07/08 15:14：00

逐字稿

访谈者：

同学你好，那我们现在开始访谈吧！你可以通过打字或者语音回复我，我会将访谈内容收集记录！（本次访谈的内容仅用于课题研究与学习，不会用于其它用途）

受访者：

ok

访谈者：

同学你目前是在校学生吗？目前大几？是什么专业啊？

受访者：

在校，大四（五年学制），风景园林

访谈者：

我看到你填写的问卷了，可以仔细介绍一下你的项目吗？

受访者：

是mod那个吗

访谈者：

是的

受访者：

就是我自己steam游玩上的一个独游，体量也比较小

然后玩的时候觉得功能没做好，就是没有及时存档功能

然后我就想着自己试试，做一个mod解决一下（因为体量小，没有开放创意工坊，N网也没有）

然后工具上，先是用trea和vscode自带的ai，

我自己把游戏反编译后，丢个他们，梳理逻辑，然后让提需求，让ai跑

访谈者：

你有这个过程的截屏吗，我想了解一下

受访者：

现在可以截屏

访谈者：

可以的

受访者：

你要那一部分呢

反编译的部分，还是ai对话部分

访谈者：

你都可以截屏一下，你那个mod也可以给我看看呢

受访者：

这是后面用workbuddy做的

这是项目文件夹

然后对话部分太吃那个了就不截屏了

还有什么需要的嘛

访谈者：

我看你是风景园林专业的同学，请问你有做代码，编程，或者做美术资产的基础吗？

受访者：

有一点美术资产的基础，之前报过班学过pbr建模流程

有一点python基础

访谈者：

你觉得你在做这个项目的时候，你遇到最大的困难是什么？

受访者：

就是调试吧，trea和vscode不像workbuddy，w可以授权然后让他自行调试，做好必要的调用和依赖

访谈者：

在做这个项目的时候，你最主要的动机是什么？是为了快速验证自己的想法，还是想做一个很完整的作品，或者说是用于练习/课程学习？

受访者：

当然还有一点就是审查，因为自己代码能力弱，基本都是全部授权ai，后期维护成本高

第一是，练手。第二是，确实想要实现功能

访谈者：

那你第一次生成的结果怎么样？可以运行吗？你一共生成了多少轮才达到自己想要的结果？

受访者：

用workbuddy生成的可以，算得上一个能跑通的项目。总共花了八轮，一共4小时（只计算了使用workbuddy的时间）

访谈者：

在第一次生成的时候，与你预期的结果有多少偏差，主要在哪些地方不符合你预期

你当时是选择继续修改，还是重新输入生成？

受访者：

偏差没有多少，主要只是一个功能的实现。非要说的话就是有bug，需要后续测试维护

是重新把源文件给workbuddy

然后理解→需求→制作

访谈者：

这个平台我没有使用过，我想了解一下，在这个过程中是否有付费环节？

受访者：

没有，因为是腾讯新出的，额度给得宽裕

而且本来工程量就小

访谈者：

你使用这个平台的频率是什么样的？假如这个平台需要付费，你会愿意付费吗？还是说选择另外的平台？

在生产的过程中哪些环节是你觉得最好的？

受访者：

频率不定，一般每月1个项目。如果付费，还是先对比有无平替，没有的话还是会付费

不用自己手动调试，出现代码问题，他的自察解决能力强

就是减少了开发→报错→复制日志→粘贴日志→解决→开发的循环

访谈者：

哪些能力最能支撑你付费？

-快速生成可玩原型

-多轮修改稳定

-玩法可控

-美术和音效质量更高

-可导出项目

-可发布到平台

-可多人协作

-可持续扩展

-哪些问题会降低你的付费意愿？

-生成质量不稳定

-修改困难

-作品方向不可控

-内容同质化

-无法导出

-无法长期维护

-发布价值不明确

受访者：

支撑项：1 2 8

降低项：1 2 5 6

访谈者：

然后我还想问一下你这个mod，是模组吗？目前可以运行或者试玩吗？

受访者：

可以运行

访谈者：

你发布了吗？

受访者：

没有，一是社群小

二是开发者明确表示不会做这个功能，我尊重开发者的选择

但是我会和线下朋友传递使用

访谈者：

了解，那这个作品你主要还是用ai生成，或者辅助完成的，你会主动告诉玩家这是ai生成的吗？

受访者：

AI参与度极高，会主动告诉

访谈者：

你觉得玩家会介意这类作品是由ai生成的吗？

受访者：

这类作品是指mod，还是游戏本体

访谈者：

mod

受访者：

mod的话，玩家容忍度蛮高的

访谈者：

你是否担心ai标签让作品廉价？（从游戏本身和mod来说吧）

受访者：

廉价是指游戏的艺术性还是商业价值

访谈者：

可以从这两个维度都说一下吧

就是假如这个你说的这个游戏，作者允许添加你的模组，作为另一个板块的游玩方式，你觉得ai生成的模组会让它游戏本身会掉价吗？

受访者：

艺术性，我认为游戏的美术、音乐等艺术层面上，使用AI削减了艺术性，但是在代码上使用AI，并没有多大区别。

在商业价值上，使用AI会增加开发效率，自然就减少社会生产时长，即价值，完全可以说是廉价了

最后说一句，我觉得商业廉价并无什么不好，但最好不要艺术廉价

访谈者：

明白，那你有考虑过就是ai生成一部分代码，结合自己原创的美术资产吗？？

受访者：

以我的角度，不会，因为我认为我的功能只是一个拓展，没有涉及美术、音乐、玩法的艺术层面，只是利用AI写代码实现功能而已

没有

但我认为开发者可能会认为会，因为开发者明确表示不希望sl，以减少游玩体验，游玩乐趣

访谈者：

你认为ai作品可以让玩家付费最需要证明什么？

受访者：

证明人在艺术创作中处于主导地位，

以及证明游戏作品具有游玩价值

访谈者：

不考虑现实的限制，你希望当前平台未来可以帮助你解决什么问题？你希望它哪些方面可以变强？

1.理解能力

2.生成更完整的玩法

3.支持复杂的系统

4.支持导出项目

5.支持长期修改

6.支持多人协作

7.提高美术统一性

8.提高音频和文本的质量

9.支持正式发布

受访者：

2 3 5

访谈者：

如果这个平台未来更强，甚至一句话可以生成一个游戏，你觉得创作者还需要负责什么？

你觉得人类创作者在这类平台中最核心的价值是什么？

受访者：

玩法创意

用户需求提炼

访谈者：

结合我们之前聊的，我再确认一下，这个平台对你最有帮助的地方是，因为自己代码能力相对来说比较弱，通过这个平台可以快速实现你当前的想法。主要的问题是还是调试和后期维护这一块相对来说比较高是吧？

受访者：

对于创作者的价值，平台为其表达想法提供的实现途径。也为探索AI边界、探索游戏边界提供的场景

对于平台的价值，创作者的活动、需求提供了迭代方向。同时为人机交互、AI伦理等议题提供实际依据也是平台自身对商业实现进行过渡

暂时只想到这么多

是的

当然美术、音乐也有

因为不太希望AI参与这两个方向

访谈者：

你人工控制的内容只有通过文本调试了吗？还有什么你人工操作的环节？

受访者：

没有，只有提示词工程

访谈者：

还有使用其他工具吗？

你还有用到其他ai工具吗？

受访者：

游戏制作只使用过上面三个

访谈者：

我没有其他要问的，如果后续需要补充的话，我会再联系你的，非常感谢你的参与～

问卷作答情况

以下仅收录研究相关题目，排除联系方式、同意、抽奖、路由、质量控制和技术字段；题干与选项按该答卷对应的历史题库版本呈现。

基础信息与筛选

题号与原题干	作答
S1 你当前的学生身份是？	本科在读
S2 你的专业或主要学习方向更接近哪类？	理工科非计算机
S3 过去 6 个月，你是否参与过游戏或互动作品创作？	只尝试过工具
S4 你在游戏/互动作品创作中承担过哪些工作？	没有承担过
S5 你是否使用过以下 AI 工具/平台参与游戏创作？	Codex / Claude Code 等专业 AI 编程代理
S5A 你从什么时候开始较为频繁地使用 AI 工具？	最近 1 个月内
S5B 你目前使用 AI 工具的频率大约是？	每周使用 3—5 次
S6 你使用上述工具时，是否有具体游戏/互动项目或可运行原型作为对象？	有可运行/可展示原型
S7 你是否实际玩过“一站式 AI 游戏制作平台”生成或实质参与制作的作品？	没有
S8 你能否指出具体平台或作品，并说明自己玩到了什么程度？	仅打开页面未玩

生产端与创作经历

题号与原题干	作答
P1 请概述你最近一个使用 AI 参与的游戏/互动项目。（项目类型、个人/团队、引擎或技术栈、当前阶段）	游戏mod 已完成，个人，测试更新阶段，
P2 你的程序基础更接近哪一项？	学过基础课程
P3 你使用过哪些引擎、框架或开发方式？	Python/Pygame
P4 如果不使用 AI，你能否独立完成本项目中的程序或功能实现任务？	基本不能
P5 在该项目中，你具体使用过哪些工具？	Codex / Claude Code 等专业 AI 编程代理
MAIN_TOOL 请选择一个你在游戏/互动项目中使用最深入，或最近主要使用的专业 AI 编程/开发工具。	其他专业 AI 编程/开发工具____ 其他填写：trae workbuddy
CC1 【该工具】你主要让它处理哪些程序任务？	代码生成；修改已有代码；调试报错；项目理解
CC2 【该工具】它进入项目的哪些阶段？	早期原型；调试测试
CC3 【该工具】请评价它对程序工作的帮助。	提高速度：5（非常同意/非常高） 降低调试难度：3 帮助理解项目：5（非常同意/非常高） 帮助完成原本不会的功能：5（非常同意/非常高） 减少重复劳动：5（非常同意/非常高） 提升代码质量：2

题号与原题干	作答
CC4 【该工具】AI 生成内容后，你通常如何处理？	直接使用
CC5 【该工具】你对生成代码的判断与维护能力如何？	能判断是否正确：1（非常不同意/非常低） 能定位问题：1（非常不同意/非常低） 能解释代码：2 能后续维护：1（非常不同意/非常低） 能控制依赖风险：1（非常不同意/非常低）
CC6 【该工具】相关产出最终是否进入项目？	大部分进入
CC7 【该工具】整体看，它主要是提高已有能力，还是补足原本不足的能力？	两者都有但任务不同
P6 整体来看，AI 工具让你的游戏程序/原型开发发生了哪些变化？	开发速度更快：5（非常同意/非常高） 能做更复杂任务：3 更依赖 AI：5（非常同意/非常高） 人工检查更多：2 返工更多：1（非常不同意/非常低） 更愿意尝试新玩法：5（非常同意/非常高） 更容易形成可玩结果：5（非常同意/非常高）

开放题

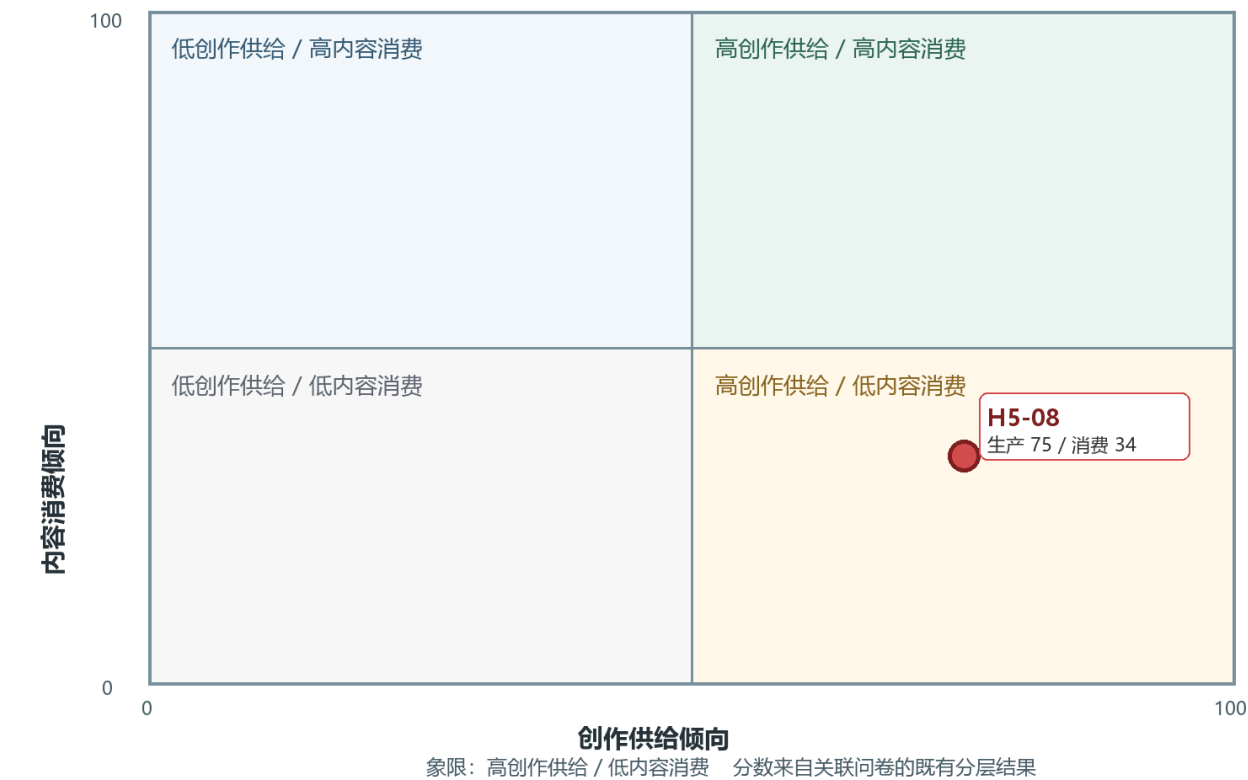
题号与原题干	作答
P7 请还原一个最典型片段：你要做什么程序/功能任务，使用了哪个工具，AI 给出什么，你如何修改，最后结果如何？	提供反编译后文件，利用ai理解游戏功能，使用trae理解并实现mod制作，但效果不佳，且自己无法查出问题，最后使用workbuddy成功完成

研究补充

题号与原题干	作答
G1 你总体上如何看待 AI 进入游戏创作？	提高效率：4 降低门槛：5（非常同意/非常高） 带来新玩法：3 削弱创作诚意：2 增加同质化：4 带来版权风险：4
G2 你认为 AI 参与程度是否应该被明确标注？	重要环节应标注
Q4.3 如果你没有实际玩过一站式 AI 平台作品，主要原因是什么？	没听说过

生产倾向 × 消费倾向象限图

横轴为创作供给倾向，纵轴为内容消费倾向；50 分为高低分界。



H5-08：生产 75，消费 34，高创作供给 / 低内容消费

关联维度分数

创作	95
效率	84
探索	64
接受	40
付费	39
信任	31

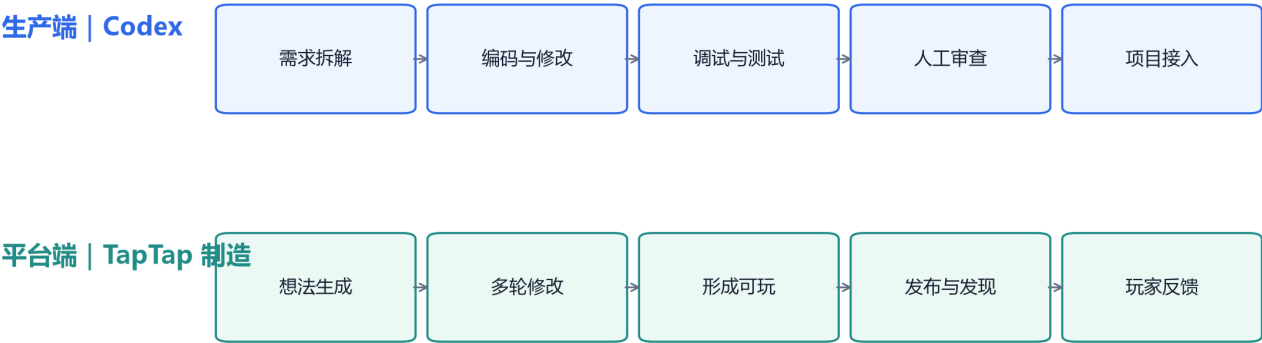
附录B 桌面研究源文件

研究摘要

研究对象声明 本研究并非泛化讨论所有 AI 产品。生产端以 Codex 为核心案例，考察 AI 编程工具如何嵌入大学生游戏开发中的需求拆解、编码、调试、人工审查与项目接入；平台端以 TapTap 制造为核心案例，考察一站式平台如何支持作品生成、修改和可玩产出，并观察玩家对相关作品的体验与反馈。Claude Code、WorkBuddy、其他平台及造化工坊定向样本作为机制对照或补充证据，不与核心对象等量解释。

桌面研究负责建立产品背景、能力边界和待验证假设。原始材料中的广泛产品扫描保留在附录；正文以 Codex 为生产端核心案例、TapTap 制造为平台端核心案例，其他产品只承担机制对照。

两条核心案例链路



核心判断并非“AI 是否会生成”，而是生成内容能否被人控制、进入项目或作品，并在后续使用和反馈中继续成立。

图 1 Codex 与 TapTap 制造两条核心案例链路

一、研究对象与桌面研究任务

1. 核心案例、类别测量与补充证据

表 1 产品对象在研究中的层级

层级	对象	观察重点	进入最终结论的方式
核心生产端案例	Codex	需求拆解、编码、调试、人工审查、项目接入	结合问卷主力组和样本1-3解释真实工作流
核心平台端案例	TapTap 制造	作品生成、多轮修改、可玩产出、玩家体验与反馈	结合平台题与样本4解释创作和消费两端
机制对照	Claude Code、WorkBuddy	项目理解、授权调试、人工控制差异	只比较机制，不推导 Codex 品牌能力
补充平台	造化工坊及其他平台	MVP 门槛、精确修改、稳定性、传播与付费顾虑	小样本方向证据，不做平台排名
背景候选池	通用助手、内容生成工具、其他开发工具	价格、生态、使用深度和产品化天花板	提出假设与解释背景，不等量进入结论

2. 桌面研究与实证材料的分工

公开资料能够说明产品定位、潜在门槛和争议，却不能证明大学生的实际使用或玩家接受度。桌面研究位于证据链上游，问卷确认为行为结构和分母，访谈解释过程与条件。

研究证据如何进入总体结论

研究方案规定对象和边界；桌面研究提出假设；问卷提供分布与实际分母；访谈解释过程与条件。

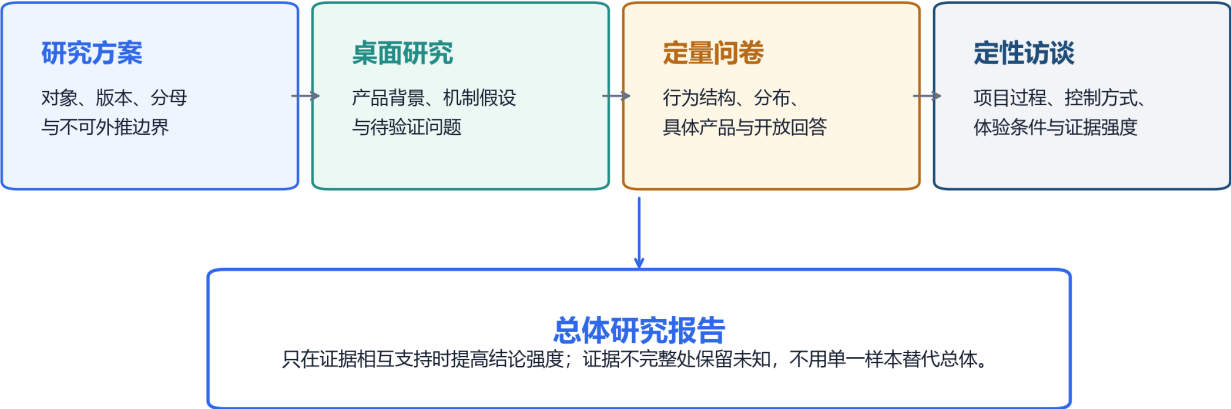


图 2 桌面研究在总体证据链中的位置

二、核心生产端案例：Codex

1. 为什么把 Codex 作为核心案例

Codex 的研究价值不在单次代码生成，而在围绕真实工程读取上下文、理解需求、修改文件、调试并配合验证。本研究据此观察需求拆解、编码修改、错误定位、测试和项目接入。

官方资料口径：Codex 可通过终端、代码编辑器和其他 Codex 工作界面参与软件开发任务。本文只使用这一公开定位建立分析维度，不把官方能力描述直接写成大学生用户的实际使用事实。用户行为仍以问卷与访谈为准。

2. 五个需要被实证检验的环节

- 1. 需求拆解：用户是否把策划或功能目标转化为可执行任务，是否由 AI 协助澄清边界。
- 2. 编码与修改：AI 是生成孤立片段，还是进入已有工程、修改现有代码并理解项目结构。
- 3. 调试与测试：用户如何发现错误、提供日志、运行测试并判断结果，而非只接受生成文本。
- 4. 人工审查：用户是否阅读代码、检查架构和依赖，技术基础较弱时又通过什么方式保持控制。
- 5. 项目接入：产出是停留在聊天窗口，还是经过修改与验证进入可运行项目。

3. 产品化机会与风险

表 2 Codex 工作流的产品化判断

维度	潜在价值	主要风险	实证需要确认
效率	减少重复编码和日志往返，缩短原型周期	速度掩盖遗漏与耦合	任务类型、轮次、项目接入程度
能力补足	帮助跨专业或基础较弱用户推进功能	理解不足导致维护依赖	人工验证方式与失败节点
上下文	在已有工程中理解文件、结构和需求	错误上下文扩大修改范围	目录约束、权限和 Review
持续使用	项目资产和工作流形成回访理由	额度、价格、稳定性与替代品影响迁移	继续使用条件和付费依据

桌面研究判断 Codex 的产品化基础更可能来自工作流嵌入，而不是一次性生成的新鲜感；但嵌入越深，人工审查、测试、权限和维护责任越不能省略。该判断需要由项目接入和控制方式数据共同验证。
--

三、核心平台端案例：TapTap 制造

1. 一站式平台的研究对象不是生成按钮

一站式平台把玩法、逻辑、素材、预览和发布放在同一链路。TapTap 制造的核心问题是：创作者能否从首次生成推进到修改和可玩产出，玩家又如何发现、体验并评价相关作品。

平台能生成 Demo 并不等于能够承接稳定作品。真正影响深度使用的往往是生成之后：是否能够精确修改，错误是否稳定复现，玩法是否有继续迭代空间，作品能否发布、被发现并获得有效反馈。桌面研究因此把“生成—修改—测试—发布—反馈”作为完整链路，而不是只比较首轮生成速度。

2. 创作者端与玩家端必须分别测量

表 3 TapTap 制造两端观察维度

端别	核心问题	可观察证据	不能直接推断
创作者端	能否从想法推进到可运行、可试玩或可展示作品	任务、修改轮次、测试、作品状态、停止原因	仅凭平台接触推断深度创作
玩家端	是否实际玩过、玩到什么程度、如何评价	具体平台/作品、体验深度、AI 知情、品质评价、支持意愿	由抽象 AI 态度推断具体作品体验
连接层	创作结果能否被发现、反馈是否回流	发布、推荐、评论、迭代线索	没有同一作品链路时认定闭环成立

3. 平台产品化的关键断点

- 精确修改：自然语言可以快速启动，但用户需要控制局部玩法、数值、UI、素材和逻辑。
- 稳定性：重复生成、Bug 和性能问题会把低门槛优势转化为返工成本。
- 玩法深度：首轮可玩不等于耐玩，持续使用需要进一步设计和迭代空间。
- 发布与发现：作品只有进入可见渠道并获得反馈，才可能形成持续创作动力。
- 信任：玩家仍通过玩法、完成度、透明度、评论和创作者投入判断是否继续体验。

四、机制对照与补充对象

表 4 补充对象的限定用途

对象	用于观察	不用于
Claude Code	项目理解、文件系统和版本控制等周边协作方式	与 Codex 等量解释或替代 Codex 核心样本
WorkBuddy	授权调试减少日志复制与人工往返的机制	把样本5的表现归入 Codex
造化工坊相关样本	MVP 门槛、精确修改、稳定性、视觉一致性和消费顾虑	宣称已完成定向招募或比较平台优劣
其他平台与工具	价格、生态、内容生成和使用深度背景	共同代表核心研究对象

执行边界 造化工坊在研究设计中承担定向补充角色；截至 2026 年 7 月 20 日，当前可核验的 5 份造化工坊产品关联记录均来自直接访问，网站中尚无无带造化工坊 campaign 标记的记录。因此，最终分析将其称为“造化工坊相关增补样本”，不写成已经完成的定向招募批次，也不据此推断总体平台使用率。

五、产品化判断框架

产品化判断回到两条具体链路：Codex 是否持续产生可接入成果；TapTap 制造是否帮助创作者跨过生成后的控制成本，并让作品进入可发现、可评价的反馈环境。价格和免费额度影响尝试，任务成果与作品沉淀才支撑持续使用。

表 5 产品化判断的可观察指标

判断主题	Codex 路径	TapTap 制造路径	研究边界
初次价值	需求、代码和调试任务能否推进	能否快速形成可见、可运行结果	不能由使用过直接推断成功
使用深度	是否进入已有项目、人工修改与测试	是否多轮修改、测试并形成可玩作品	按实际作答分母解释
持续使用	项目资产、稳定性、额度和替代品	精确修改、发布、发现和反馈	横截面数据不能证明长期留存
付费条件	能力优势、效率收益和项目需求	高级编辑、稳定产出、传播或收益	意愿不等于真实支付

六、桌面研究假设如何进入实证研究

表 6 从公开资料命题到问卷与访谈证据

桌面研究命题	问卷观察	访谈观察	最终报告处理
专业工具进入真实开发流程	任务、阶段、项目接入、人工处理	样本1-3的拆解、编码、测试与 Review	用分布确认结构，用个案解释机制
一站式平台降低原型门槛	作品状态、修改轮次、停止原因	样本4只提供消费体验，不补写创作事实	方向性支持，保留平台小样本限制
生成后控制影响深度使用	精确修改、稳定性、玩法深度、Bug	项目纠偏和授权调试方式	区分专业工具与平台的控制成本
玩家会依据“AI 痕迹”形成判断	S8、C8、C9、C10-C12	样本4的美术基调、风格控制与人工筛选判断	区分事先知情、主观感知和评价变化，不写成识别准确率
平台反馈形成留存闭环	发布、发现、评论与支持意愿	现有样本无同一作品闭环	保留为未充分验证命题

AI 参与识别假设的最终处理

桌面研究同时保留两种可能：玩家会借助同质化、拼贴感和文本模板等线索判断 AI 参与；低成本作品也可能被误判，经过人工打磨的 AI 辅助内容则可能不被识别。最终问卷显示视觉、美术、UI 和模型是最集中的主观线索，但没有解决判断是否正确的问题。

C8 显示多数消费端回答者已有不同程度的事先知情，C9 又没有盲测作品、制作真值或随机披露条件。因此，数据只能说明受访者调用了哪些线索，不能形成识别准确率结论；详细分布见定量报告。

七、结论与边界

桌面研究形成两条核心判断：Codex 需要以需求、调试、人工审查、测试和项目接入的完整 workflow 来观察；TapTap 制造需要以生成、修改、测试、发布和反馈的完整作品链路来观察。

Claude Code、WorkBuddy、造化工坊和其他产品只承担机制对照或补充证据。报告不做品牌排名，也不把公开资料、问卷自报和访谈个案视为同等证据；关于 AI 识别，当前只支持主观线索判断。

附录 A：原桌面研究产品扫描与验证问题

附录说明
以下内容保留原桌面研究中的广泛产品候选、产品化分析、消费端公开现象与待验证问题。它们作为上游背景扫描保留，不与 Codex、TapTap 制造两个核心案例等量进入最终结论。为统一阅读，仅调整标题层级、字体、表格宽度和分页，不据此新增产品事实。

- 一、研究范围与产品候选清单
- 二、通用 AI 对话产品研究卡片
- 三、编程 / 效率 AI 工具研究卡片
- 四、内容生成 AI 产品研究卡片
- 五、AI 游戏开发 / UGC 平台研究卡片
- 六、横向对比表
- 七、产品天花板与突破口分析
 - 7.1 共性天花板信号
 - 7.2 突破口信号
 - 7.3 付费意愿深度分析
 - 7.4 使用深度分析
 - 7.5 留存行为分析
 - 7.6 综合结论
- 八、消费端诉求验证
 - 8.1 研究目的与公开现象
 - 8.2 内容质量感知
 - 8.3 消费心理与AI标签效应
 - 8.4 体验容忍度与供给变化
- 九、待验证假说与验证问题清单

十、结论

附录：信源汇总

一、研究范围与产品候选清单

1.1 研究范围界定

本报告聚焦于"AI 工具"与"AI 游戏开发平台"两大品类。其中：

- AI 工具：指面向终端用户或开发者的 AI 驱动型软件产品，涵盖通用对话、编程辅助、内容生成三大子类。
- AI 游戏开发平台：指利用 AI 技术降低游戏创作门槛的平台，支持自然语言/零代码生成可运行游戏。

1.2 产品候选清单

二、通用 AI 对话产品研究卡片

2.1 ChatGPT (OpenAI)

2.2 Claude (Anthropic)

2.3 Kimi (月之暗面 / Moonshot AI)

2.4 豆包 (字节跳动)

三、编程 / 效率 AI 工具研究卡片

3.1 Cursor (Anysphere)

3.2 GitHub Copilot (GitHub / Microsoft)

四、内容生成 AI 产品研究卡片

4.1 Midjourney

4.2 Runway

4.3 可灵 (Kling AI)

4.4 即梦 (Jimeng / Dreamina)

4.5 Seedance 2.0 (字节跳动)

五、AI 游戏开发 / UGC 平台研究卡片

5.1 Yoroll

5.2 TapTap 制造

六、横向对比表

6.1 功能定位与创作门槛对比

6.2 商业模式对比

6.3 校园渗透对比

七、产品天花板与突破口分析

7.1 共性天花板信号

天花板信号	具体表现与影响
免费额度不足	ChatGPT 免费版仅 10 条/5hrs；Midjourney 无免费试用；Runway 仅 125 一次性积分 → 学生和轻度用户难以深度体验
付费价格偏高	Claude Pro \$20/mo、ChatGPT Plus \$20/mo、Midjourney Standard \$30/mo → 对中国学生用户价格敏感
AI 生成质量不稳定	即梦快速动作生成“异形”肢体；Midjourney 早期 V6 手部问题；可灵物理模拟偶发失真
同质化竞争严重	各 AI 视频工具（可灵/即梦/Runway/Pika）功能高度重叠，差异化不足
版权/商用限制	免费版多数带水印且禁止商用；Stealth 模式需 Pro 以上（Midjourney）；版权归属不明确
语言/本地化门槛	Midjourney 中文提示词支持有限；Runway/Cursor 对中文用户支持较弱
AI 标签化引起的反感	部分用户对 AI 生成内容有“廉价感”偏见；社交媒体上存在 AI 负面舆论

7.2 突破口（机会）信号

突破口方向	具体策略
针对学习/实习场景优化	Kimi (无限免费+超长上下文)和豆包(基础免费)已证明“免费+优质”在中国学生群体中的巨大吸引力
明确降低创作门槛	Yoroll 和 TapTap 制造实现“自然语言→可玩游戏”，门槛降至零代码→可打开全新用户群体
校园渗透差异化	GitHub Student Pack 是 Copilot 和 Cursor 在学生市场成功的关键 → 国产工具可借鉴
生态整合优势	豆包×抖音、即梦×剪映×抖音的生态闭环是核心竞争力 → 工具×分发×变现一体化
分层定价精细化	豆包的 68-200-500 元分层、Cursor 的 \$20-60-200 分层 → 需要更细的梯度匹配不同用户
学生专属优惠价值	GitHub Education 模式证明了“获客→养成→转化”路径 → 先让学生免费/低价用，毕业后自然付费

7.3 付费意愿深度分析

用户是否愿意为 AI 工具付费，取决于产品价值与使用成本之间的关系。产品价值来自任务效率、产出质量、门槛降低、生态承接和使用回报；使用成本包括价格、访问、注册、支付、学习、试错、版权和平台迁移等因素。

对学生用户而言，价格是重要门槛，但价格是否构成阻碍，取决于产品是否解决了具体任务。能够帮助用户完成作业、资料整理、PPT、代码项目、视频作品、游戏 Demo、作品集或比赛成果的工具，更容易形成付费理由。若免费或低价替代品已经能够完成同类任务，高价产品的付费空间会被压缩。

从八个因素判断付费意愿：

这些因素可以解释不同产品为什么采取不同商业化策略，也能帮助判断哪些产品具备从免费使用走向付费转化的条件。

现有样本中，免费策略分化明显。

Kimi、豆包的策略重点是通过免费、易访问、中文友好、移动端可用覆盖高频轻任务，包括资料整理、课程作业、PPT 大纲、社团文案、短内容生成等。它们的模型能力弱于 GPT / Claude，但在中国学生的日常使用场景中，低门槛和稳定够用本身就具有较强产品化价值。

ChatGPT、Claude 的优势集中在复杂推理、代码、长文写作、多步骤任务拆解等高复杂度场景。其付费理由更依赖高价值任务，例如科研写作、复杂报告、代码项目、专业创作等。对于普通轻任务，访问、支付、价格和国产免费替代品会削弱其付费转化。

Midjourney、Runway 的限额策略体现了内容生成产品的成本结构。图像、视频生成需要用户多次试错，但平台必须控制每次生成的算力成本。Runway Free 仅 125 一次性积分，Midjourney 无免费试用，表明这类产品很难像文本对话产品一样长期提供宽松免费额度。

天花板：

免费额度不足会压缩学生用户的深度体验空间。尤其是图像、视频、游戏生成类产品，用户往往需要多次试错才能得到满意结果，过早触达付费墙会降低继续探索的意愿。对厂商而言，完全放开免费额度又会带来算力成本和滥用风险。

突破口：

更可行的方向是让免费额度服务具体任务，例如学生专属额度、课程 / 比赛专项额度、低清预览免费、高清导出付费、创作挑战换额度、失败生成部分返还等。这样既能控制成本，也能让用户完成一次相对完整的体验。

值得进一步验证的是：

学生用户在免费额度耗尽后，更可能先转向免费替代品，而不是直接升级高价订阅；但当产品能服务课程作业、作品集、比赛或商业接单时，付费接受度会提高。

豆包是典型案例。文档记录豆包为国产月活第一 AI 原生应用，月活 3.45 亿。其模型能力在复杂推理、代码能力和长链路规划上弱于 GPT / Claude，但用户体量更大。

这一现象表明，产品化竞争并不只由模型能力决定。对多数学生日常场景而言，产品是否易访问、是否免费、是否中文友好、是否能在熟悉平台中使用，直接影响用户选择。

学生的日常需求中，许多任务并不需要最强模型。例如课程总结、PPT 大纲、资料整理、简单翻译、课堂笔记、社团文案、短视频脚本等，更强调快速、方便、稳定、低成本。豆包背靠字节生态，具备平台熟悉度和内容场景优势；Kimi 的长文档能力也适合资料整理和课程材料处理。

天花板：

高能力模型在中国学生市场中的付费转化，会受到访问、支付、价格和免费替代品影响。能力优势主要在复杂任务中体现，在日常轻任务中不一定能转化为用户体量和付费意愿。

突破口：

高能力产品需要绑定更高价值场景，例如科研写作、复杂报告、代码项目、专业创作、作品集制作等。国产产品则需要从“轻任务够用”继续提升到“复杂任务可靠”，否则容易停留在免费入口和日常轻工具层。

值得进一步验证的是：

学生用户在低复杂度任务中更倾向选择免费、易访问、本地化工具；在复杂写作、编程、科研和项目任务中，更可能转向 GPT、Claude、Cursor 等高能力工具。

付费意愿还受替代品环境影响。用户会比较某个产品是否比免费替代品更值得付费。

当 Kimi、豆包等免费工具已经可以完成大量日常任务时，ChatGPT Plus、Claude Pro 的付费理由会集中到更高复杂度场景。类似地，如果即梦、可灵、豆包生图等工具已经能满足课程展示、社媒图文或短视频作业，普通学生长期订阅 Midjourney 或 Runway 的动力会减弱。

高价工具的付费空间主要来自免费替代品难以覆盖的任务，例如复杂分析、稳定高质量输出、代码项目、专业视觉创作、商用授权、长期项目支持等。这些需求通常与课程成果、作品集、实习开发、比赛作品、商业接单等高价值场景相关。

天花板：

高价订阅在轻任务场景中容易被免费替代品稀释。若用户的需求只是作业辅助、资料整理、简单文案或基础图文生成，免费工具已经具备较强替代性。

突破口：

高价工具需要证明其在高价值场景中的不可替代性，例如复杂课程项目、科研写作、专业设计、实习开发、商业视频、作品集制作等场景，更容易支撑付费意愿。

值得进一步验证的是：

学生用户是否愿意为某个 AI 工具付费，取决于该工具是否在具体任务中显著优于免费替代品，并能带来效率、质量、成果展示或收益上的明确回报。

付费意愿的核心是需求满足程度。价格、免费额度和学生优惠是关键条件，但真正决定付费的是产品能否在具体场景中提供清晰价值。

Kimi、豆包通过免费、易访问、中文友好、移动端和本地生态覆盖高频轻任务；ChatGPT、Claude 的能力优势集中在复杂任务，但在中国学生日常场景中受到访问、支付和免费替代品影响；Midjourney、Runway 受限于图像 / 视频生成成本，免费额度难以长期放开；Cursor、GitHub Copilot 的学生优惠逻辑则来自开发者教育获客和未来职业转化。

付费意愿需要回到三个：

产品是否满足了用户的具体任务需求；

产品是否明显优于免费或低价替代品；

付费后是否能带来效率、质量、成果展示或收益上的明确提升。

7.4 使用深度分析

使用深度关注用户是否会把产品纳入学习、创作、开发或发布流程。影响使用深度的因素包括任务复杂度、链路完整度、学习成本、重复使用频率、二次加工成本、生态联动和资产沉淀。

轻任务更依赖免费、易访问和本地化，例如资料整理、PPT 大纲、社团文案和课堂笔记。复合任务和项目型任务更依赖稳定能力、工作流嵌入和完整链路，例如代码项目、视频作品、游戏 Demo、作品集和比赛成果。

产品能力强弱只是其中一个变量。真正影响使用深度的是产品是否能持续服务一个真实任务流程，并减少用户从生成到完成之间的加工成本。

不同任务对应不同使用深度。

Kimi、豆包更适合高频轻任务，ChatGPT、Claude 更适合复杂写作和深度分析，Cursor、GitHub Copilot 更适合项目型开发任务，Midjourney、Runway、可灵、即梦等更适合视觉创作和内容生成任务。

天花板：

如果产品只服务轻任务，用户容易被其他免费工具替代；如果产品能进入复合任务或项目型任务，使用深度和付费可能性都会提高。

突破口：

产品需要从“快速生成”走向“任务完成”，例如支持多轮修改、格式输出、项目管理、编辑导出和作品沉淀。

值得进一步验证的是：

能用于课程项目、作品集、比赛和实习任务的 AI 工具，比只服务日常轻任务的工具更容易形成深度使用。

使用深度取决于产品是否支持从生成到完成，而不只支持生成。

完整链路通常包括：

生成 → 修改 → 编辑 → 导出 → 发布 → 反馈 → 再迭代 不同产品覆盖的链路不同。

内容生成工具最常见的问题是二次加工成本。图像工具生成的是素材，视频工具生成的是镜头或片段，AI 游戏制作平台生成的是 Demo 或原型。它们要进入深度使用，需要补齐编辑、测试、发布、反馈等环节。

天花板：

单点生成难以形成深度使用。用户得到素材或 Demo 后，如果还需要大量人工处理，产品容易停留在尝鲜阶段。

突破口：

产品需要提供模板、局部修改、版本管理、导出格式、发布渠道和反馈机制，降低从生成到完成的加工成本。

值得进一步验证的是：

支持生成、编辑、导出、发布的一体化产品，比单点生成工具更容易形成深度使用。

使用深度还受到学习成本影响。

Midjourney 需要用户理解提示词、风格控制和参数逻辑；Runway 需要用户理解视频生成、镜头控制和素材管理；Cursor 需要编程基础；AI 游戏制作平台即使降低代码门槛，也仍需要基本玩法设计、关卡组织和测试意识。

Kimi、豆包学习成本较低，因此更容易进入普通学生的高频日常任务。Cursor、Midjourney、Runway 等工具学习成本较高，但在专业用户或明确项目场景中，学习成本也可能转化为更强的用户粘性。

天花板：

学习成本越高，普通学生从尝鲜进入深度使用的门槛越高。

突破口：

降低学习成本的方式包括模板化流程、示例工程、新手引导、中文教程、课程场景预设、自动排版、自动剪辑、自动发布等。

值得进一步验证的是：

对普通学生而言，低学习成本会显著影响 AI 工具使用深度；对专业用户而言，工具一旦嵌入项目流程，高学习成本对留存的负面影响会下降。

产品接入用户熟悉的生态，会提高使用深度。

豆包依托字节生态，接近抖音、剪映、短内容生产和移动端使用。即梦、可灵、Seedance 等国内视频生成工具，如果能与剪辑、字幕、配音、发布链路结合，更容易进入短视频创作流程。Cursor、GitHub Copilot 依托 IDE 和 GitHub 生态，更容易进入开发流程。TapTap 制造如果能与 TapTap 社区形成连接，则比单纯生成工具更有机会形成游戏创作和分发闭环。

天花板：

孤立工具在生成之后缺少承接场景，用户需要额外寻找编辑、导出、发布和传播渠道，使用深度容易受限。

突破口：

生态联动可以降低用户从生成到完成的转换成本，并沉淀作品、项目、账号和社区关系。

值得进一步验证的是：

接入剪映、抖音、GitHub、TapTap 等熟悉生态的 AI 产品，比孤立工具更容易形成深度使用。

使用深度取决于产品是否进入真实任务流程。轻任务主要依赖免费、易访问和本地化；复合任务和项目型任务更依赖稳定能力、工作流嵌入、完整链路和资产沉淀。

产品能力强弱只是其中一个变量。真正决定使用深度的，是产品能否帮助用户从“生成内容”走向“完成任务”。对内容生成工具和 AI 游戏制作平台而言，突破口在于补齐编辑、导出、发布、反馈和再迭代链路。

7.5 留存行为分析

留存行为来自持续任务、资产沉淀和外部反馈。不同产品的留存来源不同：通用 AI 助手依赖高频任务，编程 AI 工具依赖项目流程和代码资产，内容生成工具依赖作品沉淀和发布需求，AI 游戏制作平台依赖创作者端与消费者端的反馈闭环。

通用 AI 助手的留存依赖高频任务，例如资料整理、写作、问答、翻译、PPT、作业辅助等。编程工具的留存依赖项目，用户一旦将其纳入代码开发流程，迁移成本会提高。内容生成工具的留存依赖作品需求，如果用户持续做图、做视频、做社媒内容，就会反复使用。AI 游戏制作平台的留存依赖平台闭环，创作者需要玩家反馈，玩家需要内容筛选和消费理由。

天花板：

如果产品只带来一次性尝鲜，没有任务复用、作品沉淀或反馈机制，留存会明显受限。

突破口：

产品需要制造持续回访理由，例如任务历史、项目沉淀、创作者主页、作品库、评论反馈、推荐机制、比赛活动和收益机制。

留存会反向影响付费意愿。用户越频繁使用、越依赖产品、越能沉淀成果，越可能接受付费。

不同品类的逻辑不同：

通用 AI 助手通过高频任务提高使用频率；

编程工具通过项目依赖形成长期订阅理由；

内容生成工具通过持续创作带来额度和高级功能需求；

AI 游戏制作平台通过创作者反馈、作品沉淀和玩家消费形成持续投入。

天花板：

只靠单次生成或短期好奇心，很难支撑长期付费。用户尝鲜后没有继续使用理由，付费和续费都会受限。

突破口：

产品需要把一次性生成变成连续任务，例如持续项目、作品更新、社群互动、创作挑战、课程任务、比赛活动和收益回报。

值得进一步验证的是：

用户越能在产品中沉淀项目、作品或社交反馈，越可能形成持续使用和付费意愿。

留存还取决于用户离开产品的成本。

Kimi、豆包这类通用 AI 助手迁移成本相对较低，因为通用问答和资料整理替代品较多。Cursor、GitHub Copilot 的迁移成本更高，因为它们嵌入代码项目和开发习惯。Midjourney 用户如果形成固定视觉风格和作品积累，迁移成本也会提高。AI 游戏制作平台如果沉淀了作品、创作者主页、玩家评论和发布数据，迁移成本会明显增强。

天花板：

没有项目、作品、账号、社区关系沉淀的工具，用户更容易迁移到替代品。

突破口：

提高留存的关键是沉淀用户资产，例如历史任务、项目文件、作品集、创作者主页、评论数据、粉丝关系和发布记录。

值得进一步验证的是：

用户在某一 AI 平台中沉淀的作品、项目或社交反馈越多，迁移到其他工具的意愿越低。

AI 游戏制作平台不能只看创作者是否能生成作品，还要看玩家是否愿意试玩、评价和传播。

创作者端需要解决：

是否能低门槛生成游戏；

是否能编辑玩法；

是否能测试和修 Bug；

是否能发布作品；

是否有玩家反馈；

是否有收益或曝光；

是否能沉淀作品和创作者身份。

消费者端需要解决：

玩家是否愿意尝试 AI 制作游戏；

玩家是否认为 AI 游戏有新鲜感；

玩家是否担心质量低；

玩家是否认为内容同质化；

玩家是否接受 Bug、粗糙美术和简单玩法；

玩家是否因为“AI 制作”标签降低信任；

平台是否有评分、评论、精选和推荐机制降低筛选成本。

AI 游戏平台的核心闭环是：

创作者生成作品 → 平台支持编辑和发布 → 玩家试玩和评价 → 创作者获得反馈 → 作品继续迭代 → 平台形成内容循环 天花板：

如果平台只有生成能力，没有玩家入口、反馈机制、质量筛选和创作者沉淀，创作者容易停留在 Demo 尝鲜阶段。

突破口：

平台需要从 AI 生成工具走向 UGC 游戏平台。评分、评论、精选、推荐、创作者主页、比赛活动和收益机制，是提升留存的关键。

值得进一步验证的是：

AI 游戏制作平台的留存不只取决于生成能力，更取决于玩家反馈、平台分发、作品质量和创作者持续迭代机制。

留存行为的核心是用户为什么持续回来。通用 AI 助手依赖高频任务，编程工具依赖项目流程，内容生成工具依赖作品沉淀和发布需求，AI 游戏制作平台依赖创作者端与消费者端闭环。

对 AI 游戏制作平台而言，生成能力只是起点。玩家试玩、评分评论、精选推荐、创作者主页、比赛活动和收益机制，决定平台能否从 Demo 工具发展为持续创作平台。

7.6 综合结论

现有数据表明，AI 工具产品化的天花板并不只来自技术能力。用户是否选择、是否付费、是否深度使用、是否长期留存，还受到价格、免费额度、可访问性、本地化、具体场景、生态入口、二次加工成本、作品沉淀和消费端反馈等因素影响。

在付费意愿上，产品需要证明自己在具体任务中明显优于替代品，并且付费后能带来效率、质量、成果展示或收益上的明确提升。对学生用户而言，免费和低价可以促成尝试，但持续付费通常需要更强的任务回报，例如课程项目、作品集、比赛、实习任务或内容发布。

在使用深度上，产品需要进入真实任务流程。单点生成只能带来初始体验，完整链路才能带来深度使用。对内容生成工具和 AI 游戏制作平台而言，编辑、导出、发布、反馈和再迭代是从尝鲜走向深度使用的关键。

在留存行为上，产品需要提供持续回访理由。通用 AI 助手依赖高频任务，编程工具依赖项目和代码资产，内容生成工具依赖作品沉淀和发布需求，AI 游戏制作平台依赖创作者端与消费者端的闭环。对于 AI 游戏制作平台，真正的突破口在于建立从创作、发布、试玩、评价到再创作的循环。

本章基于第七章对 **付费意愿、使用深度、留存行为** 的分析，提炼后续问卷与访谈需要验证的问题。验证重点不是泛泛询问“是否使用 AI”，而是围绕用户为什么选择某类工具、为什么愿意或不愿意付费、是否进入深度使用、是否会持续留存展开。

后续调研需要验证三类问题：

1. **付费意愿是否成立**

用户是否愿意为 AI 工具付费，付费理由来自价格可接受、任务效率、产出质量、使用便利、生态承接，还是作品展示、比赛、收益等回报。

2. **使用深度是否成立**

用户是否只是浅层尝鲜，还是已经将 AI 工具纳入学习、写作、编程、视频制作、游戏创作等真实任务流程。

3. **留存行为是否成立**

用户是否会持续回到同一产品，留存原因来自高频任务、项目依赖、作品沉淀、社区反馈、平台生态，还是创作者收益和玩家消费闭环。

现有资料显示，不同产品付费门槛差异明显：Claude Pro 为 \$20/月，Cursor Pro 为 \$20/月，Runway Pro 为 \$35/月，Midjourney 无免费试用，Runway 免费层仅 125 一次性积分。对学生用户而言，这些价格和额度可能直接影响首次付费。

学生用户能接受的 AI 工具月付价格区间是多少；

用户更接受订阅制、点数包，还是按次付费；

免费额度用完后，用户更倾向于付费、换工具，还是减少使用；

小额付费是否比长期订阅更适合学生轻度创作；

学生优惠是否会显著提高试用和付费意愿。

学生用户并非完全不愿意为 AI 工具付费，但价格需要与明确任务回报匹配。若产品只服务轻任务，高价订阅会被免费替代品稀释；若产品服务课程项目、作品集、比赛或实习任务，付费接受度会提高。

Kimi、豆包等产品通过免费、易访问和中文友好覆盖大量轻任务。豆包在现有资料中被记录为国产月活第一 AI 原生应用，表明低门槛产品在用户体量上具有明显优势。

用户在日常作业、PPT、资料整理、文案、翻译等场景中，是否优先选择免费国产工具；

用户是否认为 Kimi、豆包等工具已经足够完成多数学习任务；

用户在哪些任务中会觉得免费工具 “不够用” ；

用户是否会为 ChatGPT、Claude、Cursor、Midjourney、Runway 等高能力或专业工具付费；

免费替代品存在时，高价工具需要提供什么优势才能促成付费。

免费或低价替代品会压缩高价工具在轻任务场景中的付费空间。高价工具的付费理由主要来自复杂写作、代码项目、专业视觉、视频作品、作品集或商业用途等高价任务。

付费意愿需要落到具体需求，而不是抽象地询问 “是否愿意为 AI 付费” 。

需求包括：

需求类型	需要验证的问题
作业与课程	用户是否愿意为提高作业、PPT、报告质量付费
资料整理	用户是否愿意为长文档、论文、课堂材料处理付费
编程项目	用户是否愿意为代码生成、Debug、课程项目、实习开发付费
内容生成	用户是否愿意为图片、视频、短视频、社媒内容付费
作品集 / 比赛	用户是否愿意为可展示、可参赛、可发布成果付费
游戏创作	用户是否愿意为低门槛生成游戏 Demo 或完整小游戏付费
社交娱乐	用户是否会因表情包、趣味视频、同学互动等好奇心场景小额付费

付费意愿与任务价值强相关。越接近课程成绩、作品展示、实习能力、比赛成果或商业收益，用户越可能接受付费；单纯好奇心和娱乐尝鲜更容易带来短期试用，但难以支撑长期付费。

内容生成工具和 AI 游戏制作平台涉及作品发布、商用、参赛和收益问题。版权规则不清、免费版水印、商用限制、创作者收益缺失，都可能影响用户是否将其用于正式创作。

学生是否关心 AI 生成内容的版权归属；

用户是否会因为版权不清而不愿公开发布作品；

去水印、高清导出、商用授权是否会提升付费意愿；

创作者是否在意平台是否提供分成、曝光、推荐或作品沉淀；

对学生创作者而言，作品集、比赛、同伴反馈是否能替代现金收益。

对创作者而言，付费意愿不仅来自工具能力，也来自作品能否被正式使用。版权清晰、可发布、可展示、可参赛、可获得反馈或收益，会提高用户付费和持续创作的可能性。

不同产品可能对应不同使用深度。通用 AI 助手可能用于日常问答，编程工具可能嵌入项目开发，视频和图像工具可能停留在素材生成，AI 游戏制作平台可能停留在 Demo 生成。

用户使用 AI 工具的频率；

用户每次使用是否只完成单点任务，还是完成完整项目；

用户是否会围绕同一个任务进行多轮修改；

用户是否将 AI 工具用于课程、项目、比赛、作品集等正式任务；

用户是否会长期使用同一工具，而不是每次随机选择不同工具。

AI 工具的使用深度与任务复杂度相关。轻任务更容易形成高频但浅层使用，项目型任务更容易形成低频但深度使用。

使用深度的关键是产品是否进入真实流程，而不是用户是否“试过”。

用户是否将 AI 工具纳入写作、资料整理、代码开发、视频制作、游戏制作流程；

产品是否帮助用户完成从生成到修改、导出、发布的完整链路；

用户是否还需要大量二次加工；

用户是否因为二次加工成本高而放弃继续使用；

用户是否更偏好一站式平台，而不是多个工具来回切换。

支持生成、编辑、导出、发布的一体化产品，比单点生成工具更容易形成深度使用。二次加工成本越高，用户从尝鲜走向深度使用的转化越弱。

Midjourney、Runway、Cursor、AI 游戏制作平台都存在不同程度学习门槛。普通学生是否愿意投入学习成本，会影响其使用深度。

用户是否因为不会提示词、剪辑、代码、参数设置而放弃使用；

用户是否愿意学习复杂工具；

模板、新手引导、案例工程是否能降低学习门槛；

专业学生和非专业学生的使用深度是否不同；

工具一旦嵌入项目流程后，用户是否愿意继续投入学习成本。

对普通学生而言，低学习成本会显著影响 AI 工具使用深度；对专业用户而言，若工具能嵌入项目流程，高学习成本的负面影响会下降。

豆包、即梦、剪映、抖音、GitHub、TapTap 等平台生态，会影响用户是否愿意持续使用某个工具。

用户是否更愿意在熟悉平台中完成 AI 创作；

用户是否偏好“生成—编辑—发布”在同一生态内完成；

抖音、剪映、B站、小红书等发布场景是否影响 AI 视频工具使用；

GitHub 和 IDE 生态是否提高编程 AI 工具使用深度；

TapTap 社区是否能提高 AI 游戏制作平台的创作和发布动力。

接入熟悉生态的平台，比孤立工具更容易形成深度使用。生态承接可以降低从生成到完成的转换成本，并沉淀作品、项目和社交反馈。

留存行为来自持续任务、资产沉淀和外部反馈。不同品类的留存逻辑不同。

用户持续使用某个 AI 工具的原因；

用户是否因为作业、项目、资料整理等高频任务反复使用；

用户是否因为代码项目、作品集、视频作品、游戏作品而持续使用；

用户是否因评论、点赞、评分、推荐、收益等外部反馈持续创作；

用户是否会把同一平台作为长期创作工具。

通用 AI 助手的留存来自高频任务，编程工具的留存来自项目依赖，内容生成工具的留存来自作品沉淀，AI 游戏制作平台的留存来自创作者端与消费者端的反馈闭环。

用户是否容易更换工具，取决于产品是否沉淀了项目、作品、账号、风格、社区关系和反馈数据。

用户是否会频繁更换 AI 工具；

用户更换工具的原因是什么；

用户是否因为作品、历史记录、项目文件、创作者主页而留在某个平台；

用户是否因为熟悉某个平台而不愿迁移；

用户是否因为替代品更多而降低对单一工具的依赖。

产品中沉淀的用户资产越多，迁移成本越高，留存越强。没有项目、作品、账号或社区关系沉淀的工具，更容易被替代品取代。

留存会反过来影响付费。用户使用越频繁、越依赖某个工具、越能从中获得成果，越可能接受付费。

高频使用用户是否更愿意付费；

用户是否因为项目依赖而持续订阅；

用户是否因为作品沉淀而不愿离开平台；

用户是否因为付费后体验提升而增加使用频率；

用户是否会因免费额度不足而停止使用或转向替代品。

留存和付费之间存在相互强化关系。用户越能从产品中持续获得任务成果、项目进展或外部反馈，越可能接受付费；付费后若能明显提升体验，也可能增强留存。

AI 游戏制作平台涉及创作者端和消费者端，需要单独验证。

用户是否愿意尝试 AI 游戏制作平台；

用户尝试的动机是好奇、课程、比赛、作品集，还是想发布作品；

用户是否认为 AI 降低了游戏制作门槛；

用户生成 Demo 后，是否愿意继续编辑和完善；

用户是否需要玩法设计、关卡设计、UI、数值、测试等进一步支持；

用户是否因为作品没人看而停止创作；

用户是否需要玩家反馈、推荐、比赛、收益或创作者主页；

用户是否愿意为高级编辑、更多生成额度、发布能力或曝光机会付费。

AI 游戏制作平台的创作者留存不取决于生成 Demo 本身，而取决于后续编辑能力、发布渠道、玩家反馈、作品沉淀和创作者成长机制。

玩家是否愿意试玩 AI 制作游戏；

“AI 制作” 标签会提高尝试兴趣，还是降低质量预期；

玩家是否更关注 AI 标签，还是更关注玩法是否好玩；

玩家是否能接受 AI 游戏中的 Bug、粗糙美术、简单玩法；

玩家是否认为 AI 游戏容易同质化；

玩家是否愿意为 AI 制作游戏付费或投入时间；

评分、评论、精选、推荐是否能提高玩家信任；

玩家反馈是否能促进创作者持续迭代。

AI 游戏制作平台的消费端接受度，取决于作品质量、玩法趣味和平台筛选机制。“AI 制作” 标签可能带来新鲜感，也可能带来低质预期；最终影响留存的，是玩家是否愿意试玩、评价和传播。

AI 游戏制作平台需要验证创作者端与消费者端是否能形成循环。

创作者是否因为玩家反馈继续创作；

玩家是否因为平台推荐机制发现更多作品；

平台是否能让优质作品被看见；

创作者是否愿意围绕玩家反馈迭代作品；

玩家评论、评分、收藏是否能增强创作者留存；

作品发布后是否形成再创作动力。

AI 游戏制作平台的核心突破口是形成“创作—发布—试玩—反馈—再创作”的循环。缺少消费者端反馈，平台容易停留在低门槛 Demo 工具阶段。

优先级	假设	对应维度	验证价值
高	学生用户在日常轻任务中更倾向选择免费、易访问、本地化工具	付费意愿 / 使用深度	解释豆包、Kimi 等产品的用户体量优势
高	免费额度不足时，学生更可能转向免费替代品，而不是直接付费	付费意愿	判断免费策略是否阻碍付费转化
高	高价 AI 工具的付费意愿集中在课程项目、作品集、比赛、实习等高价值任务	付费意愿	判断价格是否与任务价值匹配
高	支持生成、编辑、导出、发布的一体化产品更容易形成深度使用	使用深度	判断内容生成工具和 AI 游戏平台的产品化潜力
高	AI 游戏制作平台的留存取决于玩家反馈和平台分发，而不仅是生成能力	留存行为	判断 Yoroll、TapTap 制造等平台的关键问题
中	学习成本会影响普通学生进入深度使用，但对专业用户影响较小	使用深度	判断不同用户分层
中	版权清晰度会影响学生是否将 AI 生成内容用于比赛、作品集或公开发布	付费意愿	判断版权规则的影响
中	生态承接能提高使用深度，例如剪映、抖音、GitHub、TapTap 等平台联动	使用深度 / 留存	判断平台生态价值

优先级	假设	对应维度	验证价值
中	作品集、比赛、同伴反馈等非现金回报也能增强学生创作者留存	留存行为	判断学生创作者动力
中	AI 制作标签对消费端存在双重影响：新鲜感与低质预期并存	消费端	判断 AI 游戏消费者接受度

第八章的核心任务是把第七章中的产品化判断转化为可验证问题。后续问卷与访谈需要围绕付费意愿、使用深度和留存行为展开，不应只询问“是否使用 AI”，而应进一步追问用户在什么场景使用、为什么选择某个工具、是否愿意付费、是否会持续使用、是否受到替代品影响，以及是否认可 AI 制作内容。

对 AI 游戏制作平台而言，验证重点需要同时覆盖创作者端和消费者端。创作者端关注低门槛创作、编辑、发布、反馈和收益；消费者端关注是否愿意试玩、如何评价 AI 制作标签、是否接受作品质量，以及平台筛选和推荐机制是否能建立信任。只有这两端形成闭环，AI 游戏制作平台才有可能从低门槛生成工具发展为持续创作平台。

八、消费端诉求验证

8.1 研究目的与公开现象

关注 AI 生成游戏与传统 UGC 游戏在消费端的感知差异。研究对象为在校大学生玩家，桌面研究阶段基于 Steam、TapTap、itch.io、B站、小红书、知乎等公开平台资料，整理玩家对 AI 制作游戏的评价、争议和消费态度线索。

核心问题是：

当 AI 降低游戏生产门槛，使游戏供给快速增加时，玩家如何判断一款 AI 游戏是否值得试玩、付费、评价和持续关注？

从三个维度展开：

此外，补充一个关键问题：

玩家是否能够识别 AI 游戏，以及“AI 味”“没灵魂”“同质化”等标签是如何形成的。

TapTap 制造是目前国内 AI 游戏创作平台中创作者反馈较集中的样本。

早期反馈显示，AI 游戏创作工具能显著降低启动门槛。有创作者评价：

“通过口令制作和打磨游戏，帮助0基础家人们实现万丈高楼平地起……”

实测者记录：

“只用了10分钟，游戏主体和操作系统就搞定了，几乎不用debug……”

这类反馈表明，AI 工具在快速生成 Demo、搭建基础操作系统、减少代码检索和调试时间方面有实际价值。

但 4 月底之后，TapTap 制造引入人工审核和 Token 配额审批制，引发创作者不满。有创作者批评：

“人人能做游戏的愿景变成了：人人得先写申请，再等审批，获准后才能做游戏。”

该案例暴露了 AI 游戏制作平台的结构矛盾：免费模式吸引创作者，但算力成本由平台承担；平台引入审核与配额控制成本，又会降低创作者自由度和创作热情。

同时，公开反馈也指出 AI 游戏制作平台存在隐性门槛：

“你得先知道怎么做游戏，才能告诉AI怎么做游戏。”

这表明 AI 降低了启动门槛，但没有完全消除游戏设计、玩法调试、内容维护和长期运营的门槛。

Game Oracle 对 2025 年 1—10 月 Steam 平台 9879 款游戏的研究显示：

声明使用 AI 的游戏，首月评论数量平均比未使用 AI 的游戏低 52.6%；

AI 游戏首月评论中位数为 4 条，非 AI 游戏为 7 条；

近 20% 的 AI 辅助游戏完全零评论，非 AI 游戏约 15.2%；

在获得至少 100 条评论的游戏中，AI 游戏好评率中位数为 84.6%，低于非 AI 游戏的 88.3%。

这些数据表明，AI 标签在 Steam 消费端存在一定负面影响。它可能降低玩家点击、购买、评论或信任意愿。

Steam 新品节的数据也显示类似问题。2026 年 6 月 Steam 新品节共上线 4227 个游戏 Demo，其中 1694 个搭载 AI 制作内容，占比约 40%。玩家对 AI 内容大量涌入产生反感，有玩家表示：

“这游戏看着好棒！哦，用了AI啊。”

还有主播称：

“本届Steam新品节的AI内容数量离谱得过分。”

玩家提出的建议包括：封面图增加 AI 标签、将 AI 声明放在描述上方、增加筛选功能，甚至像过滤成人内容一样屏蔽 AI 内容。

这表明，AI 标签在供给少时可能带来新鲜感；当 AI 内容大量出现时，玩家更需要筛选机制和质量保障。

itch.io 的讨论显示，独立游戏社区对 AI 使用更加敏感。

在《My Fluffy Neighbor》的评论中，用户认可开发者使用 AI 美术的前提是：开发者仍然投入心思，确保 AI 产出不是混乱和不一致的垃圾素材。

这表明部分玩家并不绝对排斥 AI，而是看重：

是否透明披露；

是否经过人工筛选；

是否提升了作品表现；

是否没有破坏作品整体质量。

但在《WelCUM To The City》的评论区，有用户批评开发者隐瞒 AI 使用：

“看起来太糟糕了，完全拉低了游戏质量.....描述写着‘未使用生成式AI’，但那是个谎言。”

itch.io 的案例表明，在独立游戏社区中，隐瞒 AI 使用比使用 AI 本身更容易引发信任危机。AI 标签不仅是技术信息，也和创作者诚信、作品质量、社区价值观有关。

B站、小红书、知乎等平台上的讨论显示，AI 游戏的消费端吸引力主要来自三个方向：

好奇心和技术围观

玩家想知道 AI 能把游戏做到什么程度。个性化体验

小红书 AI 文游案例显示，玩家可以定制剧情、人设和即时文本反馈，获得情绪价值。社交传播

AI 生成内容适合做梗图、趣味视频、同学互动和评论区讨论。例如，小红书 AI 文游被包装成 9.9 元 Prompt，用户复制到 DeepSeek 或豆包即可开始游戏。玩家认可其个性化剧情和大量即时反馈，但也很快发现叙事质量、结构稳定性和内容重复的问题。

知乎讨论则更关注 AI 游戏制作平台的真实能力。有用户指出：

“如果你想做一个大型的游戏，融合的系统过多，AI是做不出来的。”

这表明消费端与创作者端都逐渐意识到：AI 游戏的价值不只是“能不能生成”，而是能否形成稳定、可持续、可维护的游戏体验。

Astrocade 上线 6 个月积累 2000 万用户，月新增游戏超过 2 万款，总游玩次数超过 5 亿次。App Brain 显示其评分为 4.53/5。

但实测反馈也指出：

“这些游戏质量普遍比较一般，大部分甚至达不到微信小游戏的水平。”

这表明 AI 平台可以迅速扩大内容供给，但供给扩张并不自动等于消费端满意。

幕间 AI 则展示了另一类风险。其融资后被创作者质疑“抄袭、白嫖创作者、将用户创作内容作为资产变现”，引发大量玩家和创作者撤离。这表明 AI 内容平台的信任基础非常脆弱。平台如果不能处理好创作者权益、内容归属和社区信任，很容易从技术创新转化为舆论风险。

8.2 内容质量感知

TapTap 制造的早期反馈表明，AI 平台能快速生成游戏主体和操作系统，降低从想法到 Demo 的门槛。但公开讨论也指出，真正做游戏仍然需要用户理解玩法设计、系统结构和维护逻辑。

玩家消费游戏时，核心仍是：

是否有明确目标；

是否有规则反馈；

是否有可重复游玩的价值；

是否有挑战、成长或策略；

是否只是一次性技术展示。

AI 游戏的玩法天花板在于：如果平台只能生成基础 Demo，缺少玩法迭代、关卡设计、平衡调整和长期目标，玩家的新鲜感会很快消耗。

AI 生成游戏的玩法评价不取决于“是否能生成”，而取决于生成后是否能形成可玩的规则系统。

AI 美术可以快速生成角色、场景和视觉素材，降低创作者门槛。但玩家对美术的评价并不只看单张图是否精美，而是看整体是否协调。

玩家可能通过以下特征判断“AI 味”：

风格不统一；

立绘、UI、场景不协调；

素材拼贴感；

角色细节异常；

宣传图精美但实际游戏粗糙；

画面缺少整体设计。

itch.io 案例显示，玩家可以接受经过人工筛选和整体控制的 AI 美术；但未经打磨、影响质量、甚至隐瞒 AI 使用，会显著降低信任。

AI 美术能提高初次吸引力，但消费端更看重整体风格统一和人工打磨感。

AI 文游和 AI 叙事的优势在于个性化和即时反馈。小红书 AI 文游案例显示，玩家可以定制剧情、角色人设，并获得持续文本回应。

但玩家对 AI 叙事的批评集中在：

模板化；

狗血化；

情绪浅；

缺少作者表达；

长期体验重复；

“AI 味”明显。

叙事体验的关键不是文本数量，而是角色一致性、情绪推进、选择后果和记忆点。

AI 叙事适合提供即时互动和情绪反馈，但长期体验需要结构设计、角色塑造和作者表达。

8.3 消费心理与AI标签效应

AI 制作标签一方面能带来好奇心，另一方面会降低质量预期。

吸引试玩的原因	降低期待的原因
好奇 AI 能做到什么程度	认为 AI 内容粗糙
想围观技术演示	担心同质化
想体验个性化内容	觉得缺少人工打磨
想找社交话题	担心版权问题
想测试 AI 游戏能否替代传统开发	认为创作者偷懒
	认为 AI 内容低成本却仍然收费
	AI 标签能带来初始注意力，但会提高玩家对质量和透明度的审查
	消费端研究不能只问“看到 AI 标签你怎么看”，还要问玩家是否能在没有标签时识别 AI 使用

玩家可能根据以下线索判断游戏是否“像 AI 做的”：

美术风格不统一；

文本模板化；

叙事缺少情绪；

素材拼贴明显；

UI 和场景不协调；

玩法浅；

内容多但缺少设计重点；

宣传图与实际体验落差大；

更新快但质量不稳定。

这里存在误判风险。玩家可能把低成本独立游戏误认为 AI 制作，也可能无法识别经过人工打磨的 AI 辅助内容。

AI 标签的影响既来自平台披露，也来自玩家的主观识别。玩家是否能准确识别 AI 内容，本身就是后续问卷和访谈需要验证的问题。

玩家对 AI 的抽象态度和具体购买行为并不完全一致。

Quantic Foundry 调研显示，1799 名玩家中，63% 对游戏使用 AI 持“非常负面”态度。

GameDiscoverCo 对 3800 名 Steam 用户的调查则显示，68.6% 对 AI 游戏持接受或中立态度，仅 8.1% 明确拒绝。这表明玩家在价值判断上可能反感 AI，但在具体游戏消费中仍会给作品机会。真正影响消费行为的仍是玩法、价格、质量、口碑和平台信任。

问卷需要区分“对 AI 的态度”和“是否愿意试玩 / 购买具体 AI 游戏”。两者不能混为一谈。

8.4 体验容忍度与供给变化

传统 UGC 游戏的粗糙度可能被玩家理解为个人创作者能力有限，只要作品有创意或社区氛围，玩家会给予一定宽容。

AI 游戏的容忍逻辑更复杂。AI 被宣传为高效率工具，玩家可能反而期待更快修复和更新。若 AI 游戏长期保持低完成度，玩家容易将问题归因为“AI 低质生成”或“创作者偷懒”。

AI 游戏的新鲜感可以带来短期宽容，但无法长期抵消 Bug、粗糙体验和低完成度。

TapTap 制造和 Astrocade 的案例表明，AI 工具可以显著提高内容生成速度。Astrocade 月新增游戏超过 2 万款，表明 AI 平台具有高产能供给能力。

但供给速度提高后，玩家会自然提高更新期待。如果 AI 游戏长期不修复、不迭代、不回应反馈，玩家会质疑其 AI 制作优势。

AI 提高了内容生成速度，也提高了玩家对更新和维护的期待。快速生成只是起点，持续迭代才影响留存。

Steam 新品节中 AI Demo 占比达到 40%，Astrocade 月新增游戏超过 2 万款，表明 AI 游戏供给扩张已经出现。供给增加后，玩家的问题从 “有没有游戏可玩” 转向 “哪些值得玩”。

玩家会更依赖：

评分；

评论；

精选；

推荐；

标签；

排行榜；

创作者信誉；

主播或朋友推荐。

AI 生成游戏如果大量涌入平台，玩家需要判断哪些内容经过筛选、哪些内容仍在更新、哪些创作者可信。平台需要建立质量信任机制，包括：

AI 使用披露；

评分评论；

精选推荐；

创作者主页；

更新记录；

玩家反馈；

内容审核；

玩法标签。

B站 Toy 和 AI 公开赛案例显示，AI 内容如果进入社区共创，消费体验会发生变化。UP 主上传作品后，弹幕和评论区成为需求池，观众在追更中提出建议，创作者根据反馈迭代产品。

这表明 AI 生成内容的消费价值不只在作品本身，也在于玩家是否能参与反馈和迭代。

AI 游戏平台要形成持续消费，需要从 “内容供给平台” 转向 “创作—反馈—迭代” 的社区系统。

九、待验证假说与验证问题清单

以下假说与验证问题综合自产品桌面研究、产品化天花板分析与消费端研究，供后续问卷与访谈验证。

9.1 基于产品桌面研究的假说

以下假说供后续问卷/访谈验证：

来源模块：模块 3/6

依据：ChatGPT 10条/5hrs、Copilot Free 2000补全/月、Midjourney 无免费试用

建议问题：当免费额度用尽后，你最可能采取的行为是什么？(A) 付费升级 (B) 换用其他免费工具 (C) 减少使用频率

优先级：高

来源模块：模块 2/6

依据：Copilot 学生免费、Cursor 1年Pro免费；而 Midjourney/Runway 无学生优惠

建议问题：你更愿意为以下哪类 AI 工具付费？(A) AI 编程辅助 (B) AI 图像/视频生成 (C) AI 写作/翻译

优先级：高

来源模块：模块 6.3

依据：Kimi 免费无限+国产直连；豆包月活 3.45 亿为国内第一；即梦与抖音生态打通

建议问题：你日常使用频率最高的 AI 助手是？(A) ChatGPT (B) Claude (C) Kimi (D) 豆包 (E) 其他

优先级：高

来源模块：模块 6.2

依据：豆包标准版 68元/月 = ~\$9.5；对标 ChatGPT Plus \$20 + 中国购买力平价

建议问题：你每月最多愿意为 AI 工具支付多少？(A) 0元 (B) 1-30元 (C) 30-68元 (D) 68-200元 (E) 200元以上

优先级：中

来源模块：模块 5

依据：TapTap 制造”对话即创作，完成即发布”；Yoroll ”成本降至传统 1/100”

建议问题：如果有工具让你用自然语言就能创作一个游戏，你会尝试吗？(A) 一定会 (B) 可能会 (C) 不会

优先级：中

来源模块：模块 2/4

依据：Kimi 凭”免费+超长上下文”在中国走红；豆包免费版覆盖 3.45 亿月活

建议问题：选择 AI 工具时，你最重要的三个因素是什么？(排序) 免费/中文支持/功能全面/生成质量/速度

优先级：中

9.2 基于产品化分析的验证问题

本章基于第七章对 **付费意愿、使用深度、留存行为** 的分析，提炼后续问卷与访谈需要验证的问题。验证重点不是泛泛询问 “是否使用 AI” ，而是围绕用户为什么选择某类工具、为什么愿意或不愿意付费、是否进入深度使用、是否会持续留存展开。

后续调研需要验证三类问题：

1. **付费意愿是否成立**

用户是否愿意为 AI 工具付费，付费理由来自价格可接受、任务效率、产出质量、使用便利、生态承接，还是作品展示、比赛、收益等回报。

2. **使用深度是否成立**

用户是否只是浅层尝鲜，还是已经将 AI 工具纳入学习、写作、编程、视频制作、游戏创作等真实任务流程。

3. **留存行为是否成立**

用户是否会持续回到同一产品，留存原因来自高频任务、项目依赖、作品沉淀、社区反馈、平台生态，还是创作者收益和玩家消费闭环。

现有资料显示，不同产品付费门槛差异明显：Claude Pro 为 \$20/月，Cursor Pro 为 \$20/月，Runway Pro 为 \$35/月，Midjourney 无免费试用，Runway 免费层仅 125 一次性积分。对学生用户而言，这些价格和额度可能直接影响首次付费。

学生用户能接受的 AI 工具月付价格区间是多少；

用户更接受订阅制、点数包，还是按次付费；

免费额度用完后，用户更倾向于付费、换工具，还是减少使用；

小额付费是否比长期订阅更适合学生轻度创作；

学生优惠是否会显著提高试用和付费意愿。

学生用户并非完全不愿意为 AI 工具付费，但价格需要与明确任务回报匹配。若产品只服务轻任务，高价订阅会被免费替代品稀释；若产品能服务课程项目、作品集、比赛或实习任务，付费接受度会提高。

Kimi、豆包等产品通过免费、易访问和中文友好覆盖大量轻任务。豆包在现有资料中被记录为国产月活第一 AI 原生应用，表明低门槛产品在用户体量上具有明显优势。

用户在日常作业、PPT、资料整理、文案、翻译等场景中，是否优先选择免费国产工具；

用户是否认为 Kimi、豆包等工具已经足够完成多数学习任务；

用户在哪些任务中会觉得免费工具 “不够用” ；

用户是否会为 ChatGPT、Claude、Cursor、Midjourney、Runway 等高能力或专业工具付费；

免费替代品存在时，高价工具需要提供什么优势才能促成付费。

免费或低价替代品会压缩高价工具在轻任务场景中的付费空间。高价工具的付费理由主要来自复杂写作、代码项目、专业视觉、视频作品、作品集或商业用途等高价任务。

付费意愿需要落到具体需求，而不是抽象地询问 “是否愿意为 AI 付费” 。

需求包括：

需求类型	需要验证的问题
作业与课程	用户是否愿意为提高作业、PPT、报告质量付费
资料整理	用户是否愿意为长文档、论文、课堂材料处理付费
编程项目	用户是否愿意为代码生成、Debug、课程项目、实习开发付费
内容生成	用户是否愿意为图片、视频、短视频、社媒内容付费
作品集 / 比赛	用户是否愿意为可展示、可参赛、可发布成果付费
游戏创作	用户是否愿意为低门槛生成游戏 Demo 或完整小游戏付费

需求类型	需要验证的问题
社交娱乐	用户是否会因表情包、趣味视频、同学互动等好奇心场景小额付费

付费意愿与任务价值强相关。越接近课程成绩、作品展示、实习能力、比赛成果或商业收益，用户越可能接受付费；单纯好奇心和娱乐尝鲜更容易带来短期试用，但难以支撑长期付费。

内容生成工具和 AI 游戏制作平台涉及作品发布、商用、参赛和收益问题。版权规则不清、免费版水印、商用限制、创作者收益缺失，都可能影响用户是否将其用于正式创作。

学生是否关心 AI 生成内容的版权归属；

用户是否会因为版权不清而不愿公开发布作品；

去水印、高清导出、商用授权是否会提升付费意愿；

创作者是否在意平台是否提供分成、曝光、推荐或作品沉淀；

对学生创作者而言，作品集、比赛、同伴反馈是否能替代现金收益。

对创作者而言，付费意愿不仅来自工具能力，也来自作品能否被正式使用。版权清晰、可发布、可展示、可参赛、可获得反馈或收益，会提高用户付费和持续创作的可能性。

不同产品可能对应不同使用深度。通用 AI 助手可能用于日常问答，编程工具可能嵌入项目开发，视频和图像工具可能停留在素材生成，AI 游戏制作平台可能停留在 Demo 生成。

用户使用 AI 工具的频率；

用户每次使用是否只完成单点任务，还是完成完整项目；

用户是否会围绕同一个任务进行多轮修改；

用户是否将 AI 工具用于课程、项目、比赛、作品集等正式任务；

用户是否会长期使用同一工具，而不是每次随机选择不同工具。

AI 工具的使用深度与任务复杂度相关。轻任务更容易形成高频但浅层使用，项目型任务更容易形成低频但深度使用。

使用深度的关键是产品是否进入真实流程，而不是用户是否“试过”。

用户是否将 AI 工具纳入写作、资料整理、代码开发、视频制作、游戏制作流程；

产品是否帮助用户完成从生成到修改、导出、发布的完整链路；

用户是否还需要大量二次加工；

用户是否因为二次加工成本高而放弃继续使用；

用户是否更偏好一站式平台，而不是多个工具来回切换。

支持生成、编辑、导出、发布的一体化产品，比单点生成工具更容易形成深度使用。二次加工成本越高，用户从尝鲜走向深度使用的转化越弱。

Midjourney、Runway、Cursor、AI 游戏制作平台都存在不同程度学习门槛。普通学生是否愿意投入学习成本，会影响其使用深度。

用户是否因为不会提示词、剪辑、代码、参数设置而放弃使用；

用户是否愿意学习复杂工具；

模板、新手引导、案例工程是否能降低学习门槛；

专业学生和非专业学生的使用深度是否不同；

工具一旦嵌入项目流程后，用户是否愿意继续投入学习成本。

对普通学生而言，低学习成本会显著影响 AI 工具使用深度；对专业用户而言，若工具能嵌入项目流程，高学习成本的负面影响会下降。

豆包、即梦、剪映、抖音、GitHub、TapTap 等平台生态，会影响用户是否愿意持续使用某个工具。

用户是否更愿意在熟悉平台中完成 AI 创作；

用户是否偏好“生成—编辑—发布”在同一生态内完成；

抖音、剪映、B站、小红书等发布场景是否影响 AI 视频工具使用；

GitHub 和 IDE 生态是否提高编程 AI 工具使用深度；

TapTap 社区是否能提高 AI 游戏制作平台的创作和发布动力。

接入熟悉生态的平台，比孤立工具更容易形成深度使用。生态承接可以降低从生成到完成的转换成本，并沉淀作品、项目和社交反馈。

留存行为来自持续任务、资产沉淀和外部反馈。不同品类的留存逻辑不同。

用户持续使用某个 AI 工具的原因；

用户是否因为作业、项目、资料整理等高频任务反复使用；

用户是否因为代码项目、作品集、视频作品、游戏作品而持续使用；

用户是否因评论、点赞、评分、推荐、收益等外部反馈持续创作；

用户是否会把同一平台作为长期创作工具。

通用 AI 助手的留存来自高频任务，编程工具的留存来自项目依赖，内容生成工具的留存来自作品沉淀，AI 游戏制作平台的留存来自创作者端与消费者端的反馈闭环。

用户是否容易更换工具，取决于产品是否沉淀了项目、作品、账号、风格、社区关系和反馈数据。

用户是否会频繁更换 AI 工具；

用户更换工具的原因是什么；

用户是否因为作品、历史记录、项目文件、创作者主页而留在某个平台；

用户是否因为熟悉某个平台而不愿迁移；

用户是否因为替代品更多而降低对单一工具的依赖。

产品中沉淀的用户资产越多，迁移成本越高，留存越强。没有项目、作品、账号或社区关系沉淀的工具，更容易被替代品取代。

留存会反过来影响付费。用户使用越频繁、越依赖某个工具、越能从中获得成果，越可能接受付费。

高频使用用户是否更愿意付费；

用户是否因为项目依赖而持续订阅；

用户是否因为作品沉淀而不愿离开平台；

用户是否因为付费后体验提升而增加使用频率；

用户是否会因免费额度不足而停止使用或转向替代品。

留存和付费之间存在相互强化关系。用户越能从产品中持续获得任务成果、项目进展或外部反馈，越可能接受付费；付费后若能明显提升体验，也可能增强留存。

AI 游戏制作平台涉及创作者端和消费者端，需要单独验证。

用户是否愿意尝试 AI 游戏制作平台；

用户尝试的动机是好奇、课程、比赛、作品集，还是想发布作品；

用户是否认为 AI 降低了游戏制作门槛；

用户生成 Demo 后，是否愿意继续编辑和完善；

用户是否需要玩法设计、关卡设计、UI、数值、测试等进一步支持；

用户是否因为作品没人看而停止创作；

用户是否需要玩家反馈、推荐、比赛、收益或创作者主页；

用户是否愿意为高级编辑、更多生成额度、发布能力或曝光机会付费。

AI 游戏制作平台的创作者留存不取决于生成 Demo 本身，而取决于后续编辑能力、发布渠道、玩家反馈、作品沉淀和创作者成长机制。

玩家是否愿意试玩 AI 制作游戏；

“AI 制作” 标签会提高尝试兴趣，还是降低质量预期；

玩家是否更关注 AI 标签，还是更关注玩法是否好玩；

玩家是否能接受 AI 游戏中的 Bug、粗糙美术、简单玩法；

玩家是否认为 AI 游戏容易同质化；

玩家是否愿意为 AI 制作游戏付费或投入时间；

评分、评论、精选、推荐是否能提高玩家信任；

玩家反馈是否能促进创作者持续迭代。

AI 游戏制作平台的消费端接受度，取决于作品质量、玩法趣味和平台筛选机制。“AI 制作” 标签可能带来新鲜感，也可能带来低质预期；最终影响留存的，是玩家是否愿意试玩、评价和传播。

AI 游戏制作平台需要验证创作者端与消费者端是否能形成循环。

创作者是否因为玩家反馈继续创作；

玩家是否因为平台推荐机制发现更多作品；

平台是否能让优质作品被看见；

创作者是否愿意围绕玩家反馈迭代作品；

玩家评论、评分、收藏是否能增强创作者留存；

作品发布后是否形成再创作动力。

AI 游戏制作平台的核心突破口是形成“创作—发布—试玩—反馈—再创作”的循环。缺少消费者端反馈，平台容易停留在低门槛 Demo 工具阶段。

优先级	假设	对应维度	验证价值
高	学生用户在日常轻任务中更倾向选择免费、易访问、本地化工具	付费意愿 / 使用深度	解释豆包、Kimi 等产品的用户体量优势
高	免费额度不足时，学生更可能转向免费替代品，而不是直接付费	付费意愿	判断免费策略是否阻碍付费转化
高	高价 AI 工具的付费意愿集中在课程项目、作品集、比赛、实习等高价值任务	付费意愿	判断价格是否与任务价值匹配
高	支持生成、编辑、导出、发布的一体化产品更容易形成深度使用	使用深度	判断内容生成工具和 AI 游戏平台的产品化潜力
高	AI 游戏制作平台的留存取决于玩家反馈和平台分发，而不仅是生成能力	留存行为	判断 Yoroll、TapTap 制造等平台的关键问题
中	学习成本会影响普通学生进入深度使用，但对专业用户影响较小	使用深度	判断不同用户分层
中	版权清晰度会影响学生是否将 AI 生成内容用于比赛、作品集或公开发布	付费意愿	判断版权规则的影响
中	生态承接能提高使用深度，例如剪映、抖音、GitHub、TapTap 等平台联动	使用深度 / 留存	判断平台生态价值
中	作品集、比赛、同伴反馈等非现金回报也能增强学生创作者留存	留存行为	判断学生创作者动力
中	AI 制作标签对消费端存在双重影响：新鲜感与低质预期并存	消费端	判断 AI 游戏消费者接受度

第八章的核心任务是把第七章中的产品化判断转化为可验证问题。后续问卷与访谈需要围绕付费意愿、使用深度和留存行为展开，不应只询问“是否使用 AI”，而应进一步追问用户在什么场景使用、为什么选择某个工具、是否愿意付费、是否会持续使用、是否受到替代品影响，以及是否认可 AI 制作内容。

对 AI 游戏制作平台而言，验证重点需要同时覆盖创作者端和消费者端。创作者端关注低门槛创作、编辑、发布、反馈和收益；消费者端关注是否愿意试玩、如何评价 AI 制作标签、是否接受作品质量，以及平台筛选和推荐机制是否能建立信任。只有这两端形成闭环，AI 游戏制作平台才有可能从低门槛生成工具发展为持续创作平台。

9.3 消费端验证问题

玩家看到“AI 制作”标签时，是更想试玩、更不想试玩，还是无所谓；

更想试玩的原因是好奇、猎奇、低成本尝试、社交谈资，还是创作兴趣；

降低期待的原因是低质预期、缺少作者性、同质化、版权疑虑，还是商业反感；

如果游戏免费、由平台精选、朋友推荐或明确披露 AI 使用范围，态度是否改变。

玩家最关注玩法、美术、叙事、稳定性，还是更新频率；

玩家如何判断一款游戏“像 AI 做的”；

玩家能否准确识别 AI 使用；

玩家是否会把低成本独立游戏误判为 AI 游戏；

玩家是否认为 AI 游戏更容易同质化；

玩家是否认为 AI 叙事缺少情绪和作者表达。

AI 游戏出现 Bug 时，玩家是否比传统 UGC 更宽容；

免费或低价是否提高容忍度；

新鲜感是否提高短期容忍度；

AI 游戏是否被期待更快更新；

玩家是否愿意继续关注未完成的 AI 游戏 Demo；

玩家认为 AI 游戏的 Bug 是技术限制，还是创作者不负责。

玩家面对大量 AI 游戏时如何选择；

玩家是否希望平台提供 AI 内容筛选功能；

玩家更信任平台精选、用户评分、朋友推荐，还是主播试玩；

玩家是否需要 AI 使用范围的明确标注；

高分和真实评论是否能降低 AI 标签带来的不信任。

创作者是否因为玩家反馈继续创作；

玩家是否愿意给 AI 游戏反馈；

玩家是否愿意参与未完成 AI 游戏的测试；

创作者是否愿意围绕玩家反馈迭代；

平台推荐和玩家评论是否能增强创作者留存；

作品发布后是否能形成再创作动力。

十、结论

10.1 消费端研究结论

消费端桌面研究显示，AI 生成游戏的核心问题已经从“能不能生成”转向“玩家是否信任、愿意试玩、愿意反馈并持续消费”。AI 制作标签能制造新鲜感，但也会触发低质预期。玩家对 AI 游戏的长期评价仍然落在玩法、美术、叙事、完成度、稳定性和社区反馈上。

当供给趋于无限时，玩家的核心诉求会从内容数量转向筛选效率、质量信任和社区反馈。AI 游戏制作平台要形成消费端闭环，需要让玩家能找到好内容、相信好内容、评价好内容，并让这些反馈回流到创作者端。

10.2 产品化分析综合结论

现有数据表明，AI 工具产品化的天花板并不只来自技术能力。用户是否选择、是否付费、是否深度使用、是否长期留存，还受到价格、免费额度、可访问性、本地化、具体场景、生态入口、二次加工成本、作品沉淀和消费端反馈等因素影响。

在付费意愿上，产品需要证明自己在具体任务中明显优于替代品，并且付费后能带来效率、质量、成果展示或收益上的明确提升。对学生用户而言，免费和低价可以促成尝试，但持续付费通常需要更强的任务回报，例如课程项目、作品集、比赛、实习任务或内容发布。

在使用深度上，产品需要进入真实任务流程。单点生成只能带来初始体验，完整链路才能带来深度使用。对内容生成工具和 AI 游戏制作平台而言，编辑、导出、发布、反馈和再迭代是从尝鲜走向深度使用的关键。

在留存行为上，产品需要提供持续回访理由。通用 AI 助手依赖高频任务，编程工具依赖项目和代码资产，内容生成工具依赖作品沉淀和发布需求，AI 游戏制作平台依赖创作者端与消费者端的闭环。对于 AI 游戏制作平台，真正的突破口在于建立从创作、发布、试玩、评价到再创作的循环。

附录：信源汇总

— 报告完 —

本报告数据采集于 2026 年 6 月，所有价格与功能信息以各产品官方最新公布为准。

附录C 问卷数据

表 C-1 问卷原始数据文件登记

原始文件	samples-all-mixed-filtered-answers-20260726-0110.csv
文件格式	CSV（UTF-8 with BOM）
数据结构	长表，一条记录对应一道题目或一个矩阵行作答
数据记录	6,736 条（另含 1 行字段名）
字段数量	18 个
样本编号	117 个
SHA-256	D70727971DA6DE902AF32F2D664E6F504630D936A77DE018AE5AE3B4A6C4FEB9

表 C-2 问卷原始数据字段

序号	原始字段名
1	来源
2	来源类型
3	数据库ID
4	样本编号
5	提交时间
6	来源版本
7	题库版本
8	映射版本

序号	原始字段名
9	身份范围
10	题目ID
11	题目角色
12	题型
13	题干
14	矩阵行ID
15	矩阵行名称
16	答案序号
17	选项ID或原值
18	选项文本