

结营答辩 · 2026

大学生 AI 游戏 创作工具用户研究

从程序协作、一站式平台创作到玩家评价

研究报告、项目历程与复盘

汇报人：左涵俊

小组：阈限开拓者



AI已进入游戏生产与发布环节，研究需要转向具体使用过程

课题给出的研究起点

任务一：AI 工具产品化浪潮下的大学生认知与行为研究——从“能用”到“想用”，供需两端的真实诉求究竟在哪里

2024–2026 年，AI 工具从“通用助手”跃迁至“垂直创作平台”：Cursor、Claude Code 重塑编程生产方式，可灵、Seedance 让视频生成走向可用，AI 游戏创作、TapTap 制造、幕间 AI 等）更让零代码用户也能发布完整游戏。

在校大学生是这一浪潮的核心交叉群体——既是 AI 工具高频使用者（消费端），也是最活跃的 UGC 创作参与者（生产端）。然而我们观察到矛盾信号：生流的工具付费率高企，但 AI 游戏平台的创作者留存仍不稳定；消费端，“AI 制作”标签对同龄用户究竟是加分项还是减分项尚无定论，供给趋于无限时消费涉

本课题以在校大学生为样本，从付费意愿、使用深度、留存行为的颗粒度出发，判断 AI 工具产品化的真实天花板与突破口。

（1）供给图谱梳理（桌面研究）

系统分析 AI 游戏创作平台产品矩阵（Yoroll / TapTap 制造 / Astrocade / 幕间 AI 等）聚焦：

- 各产品的核心功能定位差异与创作门槛
- 商业化策略（免费增值/订阅/分成/版权归属）
- 校园渗透方式
- 应用商店评分及社媒口碑，用户反馈等公开数据

（2）生产端诉求挖掘（问卷 + 访谈）

以在校大学生创作者为对象，从“效率提升”“利益驱动”“好奇心驱使”三维拆解涌入动机与行为路径；识别短期尝鲜者（用后即走）与长期留存者（持续创作）差异变量。

（3）消费端诉求验证（问卷 + 桌面研究）

以在校大学生玩家为对象，从内容质量感知（玩法/美术/叙事）、消费心理（“AI 制作”标签的信任度与新鲜感）、体验容忍度（Bug/更新频率）三个维度，对比与传统 UGC 游戏的用户感知差异；探讨供给趋于无限时，消费端的核心诉求转变有哪些？

品类研究报告（PPT 或图文报告），需包含：

问卷数据的关键图表（有效问卷数需大于150份）

访谈洞察摘要（至少5人深度访谈）

访谈洞察的提炼摘要

1、AI 产品深度用户：日常使用 3 款以上 AI 工具（如 ChatGPT、Cursor、Midjourney、Notion AI 等），能从用户体验和付费决策角度输出洞察

需求 2、游戏创作实践者：尝试过使用 AI 工具或平台进行游戏创作/UGC 内容创作，或在校期间参与过 GameJam、校内创意开发等活动

3、用户调研能力：善于设计问卷或访谈提纲，能通过校内社群触达目标受访者（AI 工具活跃用户与流失用户），具备定性 + 定量结合的分析能力

我们观察到的现实背景

AI已经进入实际工作

2026年GDC对2300余名游戏行业从业者的调查显示，36%的受访者已在工作中使用生成式AI，52%认为生成式AI对游戏行业产生了负面影响。Unity AI已进入编辑器操作流程，Steam也要求开发者说明作品中的AI内容。

AI已经进入游戏生产与发布环节，但用户是否接受它，仍与作品质量、责任归属和创作控制有关。

为什么观察大学生

大学生常同时处在学习、课程项目、Game Jam、社团协作、独立创作和游戏体验等情境中。任务周期、技术基础和项目目标并不相同，因而能够呈现AI在不同条件下怎样帮助推进任务、又在哪里遇到问题。

基于原始课题与现实背景，本研究从三个核心概念进入AI游戏创作的 actual 使用过程。

三个核心概念与研究对象、问题和结论边界

01 专业AI工具

进入已有专业工作流程、协助完成某一类具体任务的AI工具。本报告只深入考察其中的程序开发。

本研究如何收束

Codex程序开发路径

以Codex主要使用者为核心样本。Codex能够进入已有项目环境，读取和修改项目文件，并配合命令执行、测试与人工检查，适合观察AI如何嵌入程序开发 workflow。

02 一站式AI游戏创作平台

用户主要通过自然语言提出创意，并在同一平台中继续生成、修改、测试和发布作品的产品。

本研究如何收束

TapTap制造作为案例锚点

同时纳入其他平台创作记录观察共同过程。该路径适合观察创意怎样由自然语言输入进入生成、修改和玩法测试。

03 AI参与作品与玩家感知

AI在开发或内容生成中参与过的游戏作品。消费端关注玩家是否注意到这种参与、依据什么作出判断，以及判断怎样影响评价；这里记录的是玩家的主观感受，不把它当作准确识别结果。

本研究如何收束

相关作品实际体验者

以经过材料复核的实际体验者为核心。观察玩家怎样理解AI参与、评价作品并形成继续游玩、推荐、关注或反馈等后续选择。

在上述概念与范围下，本研究进一步提出核心问题。

生产端——专业 AI 编程工具

01

以 Codex 为核心代表，大学生使用专业 AI 编程工具的核心诉求是什么？这类工具在不同任务中，主要是提升已有能力的效率，还是补足自身尚不具备或不够熟练的能力？在程序 workflows 中，人与 AI 如何协作，哪些环节进展顺利，哪些环节容易受阻？

生产端——一站式 AI 游戏创作平台

02

以 TapTap 制造为核心代表，大学生使用一站式 AI 游戏创作平台的核心诉求是什么？从想法生成到可玩或发布，用户如何修改、测试并推进作品，哪些环节容易受阻？

消费端——作品体验

03

大学生玩家体验 AI 参与创作的游戏时，对游戏中的 AI 使用有怎样的感受？他们是否能够感知 AI 参与、是否在意，并依据什么评价作品？哪些条件会影响他们继续游玩、关注、推荐、提供反馈或付费的意愿？

H5问卷入口负责分流，定量、开放回答和访谈承担不同证据任务

光核高校共创课题小组

AI 游戏创作画像测试

看看 AI 以哪种路径进入你的游戏世界

无论你用专业 AI 工具开发、用一站式平台制作，还是实际玩过平台作品，问卷都会按经历自动跳题，并生成一张行为画像。

通常约 10 分钟

创作 + 游玩双经历约 15-20 分钟

研究结果匿名呈现

联系方式仅在自愿参与后续环节时填写

抽奖报名已截止

问卷仍可填写并生成画像

开始测试

[活动规则与隐私说明](#)

画像用于趣味反馈和研究回看，不构成心理测评或能力评价。抽奖与后续访谈互不绑定。

双端桥接路径

双端桥接者

连接创作工具与玩家体验

材料如何进入分析



- 定量题 描述当前材料中的现象
- 开放回答/访谈 解释部分任务和评价过程
- 桌面研究 提供背景与初始假设

15 Codex主要使用者 6 平台创作记录
6 相关作品实际体验者 5 访谈对象

分别依据不同的使用或体验经历形成

本章用问卷定位现象，用三个项目案例解释程序结果如何进入项目

15

名 Codex 主要使用者

15名以Codex为主要工具并完整回答程序协作问题，构成本节的核心样本。
回答任务进入、作用评价、生成后处理和项目接入

开发者 A / B / C

48小时 Unreal 限时开发
多个 Unity 短期演示项目
Godot 限时作品与长期沙盒项目

三份访谈用于还原任务拆分、运行检查、玩法调整和目录控制。

6 名已毕业用户

完成相同问题，只用于补充观察大学生材料中出现的协作过程。

不与15名大学生核心样本合并计算，作为文末补充

材料链条：问卷现象 → 开放回答与访谈过程 → Codex对象结论

速度认可、能力补足与质量分化同时出现，诉求不是单一“替代能力”

Codex进入什么任务

代码生成

13 / 15

需求拆解

12 / 15

已有代码修改

11 / 15

调试错误

10 / 15

14/15会阅读后小幅修改，10/15会运行测试；
11/15表示大部分产出进入项目。

作用评价怎样分化

14人

提高速度、减少重复劳动均给4—5分

11人

认为提高已有能力与补足不足对应不同任务，能力补足给4—5分

5 / 5 / 5

代码质量评价在低、中、高三档各5人

提高开发速度和减少重复劳动是评价最集中的两项作用
而当前样本用户对代码质量的评价存在分化，不过即使5名代码质量低评价者也全部高度评价速度，其中4人的产出仍有大部分进入项目

“先让它只理解当前的架构并限制它过度设计……最后我自己测试能够在当前项目里运行并无Bug即可。”

开放文本|大学生用户 | 样本 dc9ec0cc

“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。”

开放文本|大学生用户 | 样本 8604f26f

三个访谈案例中，人工控制分别落在需求、原型和资源边界

三名开发者访谈还原任务如何进入项目、问题如何暴露并被修正

项目情境	如何进入项目	问题暴露与人工控制
开发者 A 48小时 Unreal 限时开发 完成速度优先	策划先形成基本方案；开发者将程序工作拆成多个任务，分别放入不同对话，先确认实现方案，再让Codex生成钩锁系统的大部分功能。 原话：“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。”	首次运行发现射线检测通道错误，开发者指出问题并要求继续修改。 人工控制 需求、方案、逻辑与运行验收
开发者 B Unity 短期 Game Jam 项目 可运行演示优先	提出创建游戏对象、挂载脚本、实现移动逻辑和公开Unity属性面板变量等要求；复杂需求先经通用对话工具整理。Codex生成玩家控制、对象逻辑和基础脚本。 原话：“Game Jam嘛，质量不是那么重要，反而先做出来是第一要务。”	多项要求同时输入会发生遗漏；部分实现偏离Unity常用组件，增加连接步骤。随后继续询问AI、自己修改或寻求其他开发者帮助。 人工控制 框架、参数、原型运行与调整
开发者 C Godot 短期与长期项目 迁移并验证抛锚玩法	读取长期沙盒项目的人物物理和“脐带物理”，迁移为Godot平台乱斗的抛锚玩法；短期项目允许创建文件，长期项目限定目录和资源范围。 原话：“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。”	首轮原型在素材、锚节点、绳索连接和输入方向上出现偏差；继续调整力量、绳索长度、位移和空中控制，直至玩法手感符合预期。 人工控制 目录、资源范围与玩法手感

Codex的核心作用是更快形成可继续处理的程序结果，而不是省略后续判断

诉求与作用

当前核心样本主要希望更快形成可继续检查和修改的程序结果，并减少重复劳动。提效与能力补足随任务并存：熟悉任务压缩实现时间，不熟悉或暂时缺少思路的任务先形成原型或方案。

协作过程

使用贯穿目标与范围设定、代码生成或修改、阅读和运行检查、反馈调整与项目接入。开发者把创作目标整理成任务，按项目周期和已有结构限定范围，再通过阅读、运行、逻辑检查或试玩发现偏差。

顺利与受阻

基础功能、原型形成和重复编写推进较顺利；要求遗漏、实现方式偏离项目习惯、界面或效果不符、运行错误，以及复杂任务缺少参考材料会造成受阻。结果继续进入项目仍需任务拆分、方案确认、运行验收和项目接入。

共同过程：创作目标 → 具体程序任务 → 生成或修改 → 运行/逻辑检查/试玩 → 反馈调整 → 决定项目接入

平台材料围绕可测试结果和持续修改展开，TapTap制造只承担案例锚点

6

份大学生平台创作记录

包含 TapTap 制造、造化工坊等多个平台

TapTap制造案例《炉釜猪》

呈现从整体方案、首轮偏离，到补充排除项和目标效果，再到7轮以上筛选、修改和玩法测试。

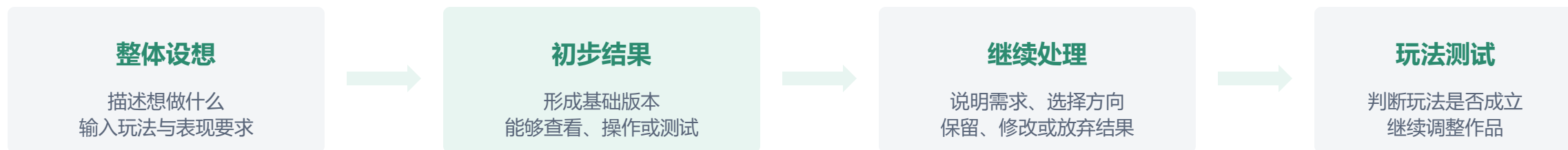
它用于解释可试玩结果如何形成，不用于代表 TapTap制造全部用户。

材料边界

没有一站式平台创作者访谈
不能比较不同平台差异
不能形成TapTap制造产品专属结论

问卷记录定位共同过程，案例记录解释首轮偏离与多轮修改

平台把想法推进到实际检验阶段，生成之后人的工作才有了具体对象



材料中较集中的平台作用

从想法生成原型

5 / 6

形成可玩逻辑

4 / 6

减少编码负担

4 / 6

快速验证玩法

4 / 6

三个对应问题

需求表达 → 目标说明

修改精度 → 结果调整

玩法深度 → 玩法完善

这些人数只用于描述当前6份记录；关键推理来自“初步结果”和“生成后工作”同时出现。

平台使用首先发生在原型形成和玩法验证阶段，用户获得的结果已经包含可供检查的作品形态和玩法逻辑。生成后的工作在初步结果形成后的调整阶段。

《炉釜猪》从首轮偏离到7轮以上调整，说明可试玩与持续修改同时存在

《炉釜猪》调整记录



“呈现出来的结果明显偏向市面上的玩法，必须得详细地说明哪些地方是我不要的，我要做到什么样的效果。”（第一次生成之后发生偏移）

大学生用户 | 样本 31305bf4

造化工坊补充观察

电子桌宠先考虑原生H5，随后因表现力不足转向Electron；作品已可运行，但创作者仍认为玩法逻辑较弱。

材料共同支持的判断

问卷同时记录了可运行或可试玩状态与修改不精确、玩法深度不足等困难。

《炉釜猪》在可试玩状态下继续完善美术。可运行或可试玩之后，作品仍有准确修改、玩法深化和表现统一等工作，作品状态和后续工作在同一记录中同时出现

平台主要支持创作启动、原型形成和玩法验证，完成仍需持续修改

直接作用

以TapTap制造为案例锚点的一站式平台把想法推进为能够查看、操作和测试的初步结果，降低创作启动阶段的编码负担。作品进入可运行或可试玩状态后，仍需持续修改才能接近创作目标。

推进过程

记录呈现从整体设想、首轮结果，到补充排除项和目标效果，再到多轮筛选、修改和玩法测试。创作者继续说明需求、选择方向、保留或放弃生成内容，并通过测试判断玩法是否成立。

受阻与完成

主要受阻条件是需求难以准确表达、局部修改不够精确、玩法深度不足（《炉釜猪》个案中，首轮结果更接近已有常见玩法），以及素材和逻辑难以统一。初步结果解决创意能否进入实际检验的问题，作品完成还要继续处理玩法、操作反馈、素材和整体表现。

消费端只把经复核的实际体验材料放在主线，其他回答承担补充说明

6 名ai平台产出作品实际体验者 依据平台或作品线索、实际游玩深度和答卷一致性复核后纳入	访谈主要体验材料 开发者 B 与N=6重叠，具有开发者与玩家双重经历 玩家 D 独立直接体验个案，记录事先知情、玩法、操作和视觉评价	补充材料 未体验原因 44 人 AI标注态度 63 人 用于补充首次体验障碍和信息透明态度，不称为对照组。
主线回答实际体验后的感知、质量评价与后续选择		

44人和63人属于补充材料，不是实验对照，玩家D是独立访谈个案，不并入问卷侧的6人统计

AI参与先通过事先信息和作品直接呈现被理解，美术与文本更容易进入评价



“部分剧情语句句式模板化，人物对话缺少自然的情绪逻辑。”

开放回答|大学生实际体验者 | 样本 860630e2

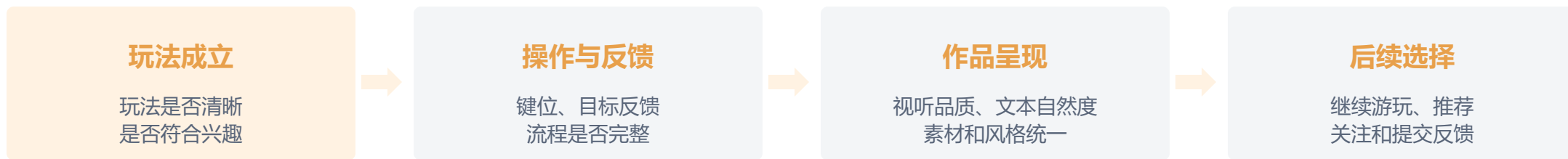
问卷与访谈怎样相互解释

量表记录美术质量、视觉统一和文本自然度；开放回答指出画面粗糙、文本模板化和对话生硬；访谈说明玩家直接接触的是画面、文字、音乐和操作结果，代码过程较少进入评价。

推理结果

AI参与因此更多通过作品呈现被注意，美术和文本的完成质量也更容易影响整体判断

玩家先判断玩法能否成立，再评价反馈、视听、文本与整体统一



量表与开放回答共同呈现

玩法清晰有4人给4—5分且无人低评；目标反馈无人给4—5分；美术质量和视觉统一各有2人给1—2分。材料呈现“基础玩法成立、局部完成质量（主要集中在视觉体验）不足”。

知道或感受到AI参与后：4人评价不变、1人更愿尝试、1人略减分。

“游戏的本质还是好玩.....应该是开发者主导AI的审美。”

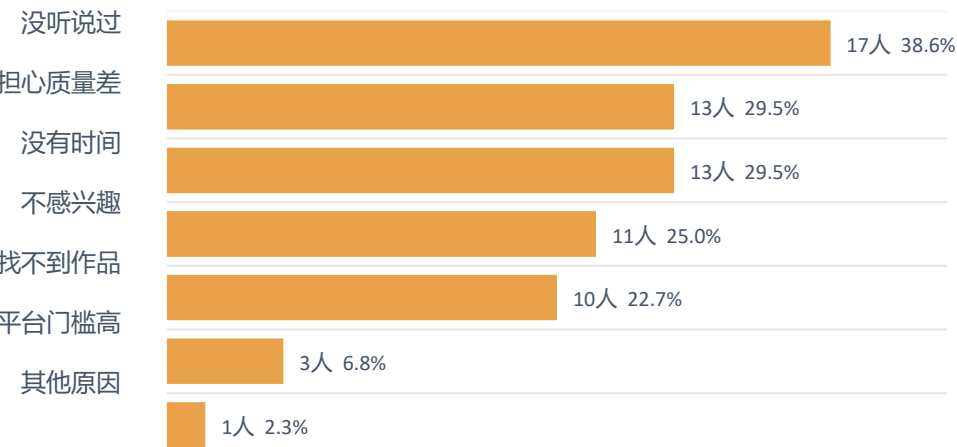
开发者 B | 访谈

“如何呈现作品，这个是主创团队需要思考的问题。”

玩家 D | 访谈

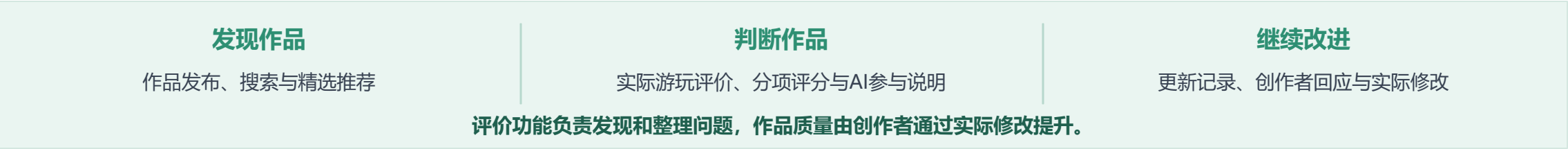
作品质量决定平台天花板，突破口在发现、判断与持续改进

首次体验障碍 | N=44，多选

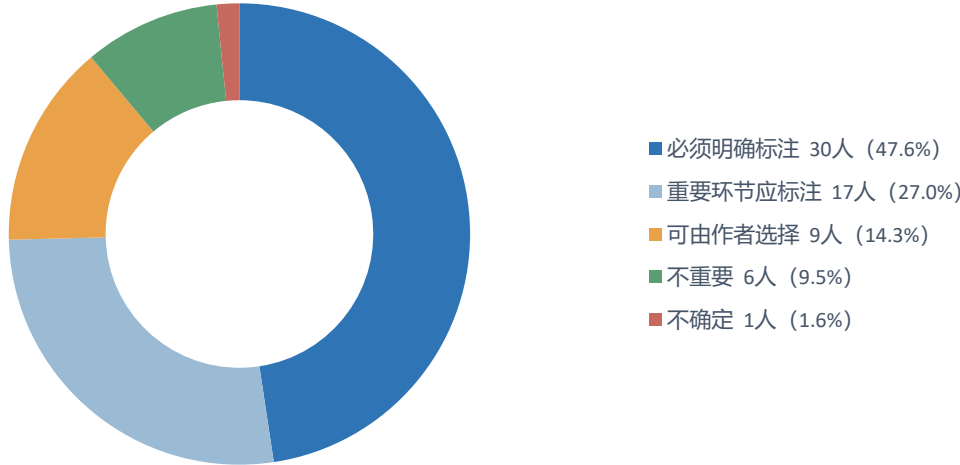


知名度、质量担忧和获取渠道是最集中的首次体验障碍。

平台突破口：让作品被发现、被判断并持续改进



AI参与标注态度 | N=63



47人（74.6%）希望作品至少说明关键AI参与情况

N=6：5人对标注透明给4—5分；知晓或感受到AI参与后，4人评价不变。

N=44与N=63为补充材料；44人均回答标注题，N=63另含19人

AI参与本身没有形成一致加分或减分，作品质量仍是评价中心

评价方向

在经平台或作品线索、实际游玩深度和答卷一致性复核后纳入的6名大学生实际体验者中，AI参与本身没有形成一致的加分或减分。评价主要围绕玩法是否清晰、操作与目标反馈、视觉与文本表现以及整体完成质量形成。

评价依据

体验者主要通过事先信息和作品直接呈现理解AI参与，美术、画面和文本比程序实现方式更容易进入评价。玩家先判断玩法能否成立，再结合操作、反馈、视听品质、文本自然度和整体统一评价作品。

后续选择

AI参与只有通过这些具体呈现进入体验，作品质量仍是主要依据。知道或感受到AI参与后，实际体验者的评价方向并不一致，继续游玩、推荐、关注和提交反馈也分别形成选择。

三端核心结论回顾

Codex

可继续处理的程序结果



关键限定

提效与能力补足对应不同任务
进入项目仍需任务拆分、检查
反馈调整与项目接入

一站式平台

支持创作启动与玩法验证



关键限定

可运行或可试玩不等于完成
仍需说明需求、筛选结果
测试玩法并持续修改

相关作品实际体验者

作品质量仍是评价中心



关键限定

在6名实际体验者中
AI参与未形成一致加减分
评价仍围绕玩法、反馈与作品呈现

三项对象结论

分别回答工具如何使用、作品如何推进、玩家如何评价。

AI先形成可检查、修改或体验的结果，随后进入人的判断与处理

Codex

关注程序结果能否检查、
修改和接入项目



人的工作

设定任务范围
阅读与运行检查
决定项目接入

一站式平台

关注想法能否转化为可测
试作品



人的工作

说明需求与选择方向
筛选结果
测试玩法

实际体验者

依据作品呈现判断玩法、
反馈、视觉、文本和完成
质量



人的工作

判断玩法与反馈
评价视觉和文本
形成后续选择

初步结果提供继续处理的对象

最终质量仍由具体任务中的检查、修改和选择形成。

前期预设 在材料中得到三点修正

	前期预设	材料暴露的问题	修正后的判断	研究最终得到的内容
01	“提效”与“能力补足”二选一	不同任务中的作用并不互斥	二者随任务并存	核心是快速形成可继续处理的结果
02	一体化流程直接推进作品完成	可运行与持续修改同时出现	主要支持创作启动	完成仍需持续修改
03	AI标签形成稳定负面预期	实际体验后的反应方向不一	不形成统一加减分	评价仍主要围绕作品质量形成

修正预设比把材料压成“支持/不支持”更能说明研究实际得到了什么。

这些材料能够回答具体过程，但不能推断总体比例或建立因果关系

- | | | |
|----|---------|---|
| 01 | 样本来源与规模 | 主要来源于开发者群组 and 光核共创营内部招募，69名大学生主样本不能推断总体比例，也不支持稳定的分组比较。 |
| 02 | 平台材料范围 | 6份平台记录包含多个平台（TapTap制造，造化工坊）且没有平台创作者访谈，不能形成TapTap制造产品专属结论。 |
| 03 | 对照条件 | 没有AI标注/不标注或AI参与/非AI参与的实验对照，不能判断AI参与本身造成评价差异。 |
| 04 | 作品过程追踪 | 没有连续追踪同一作品从创作到玩家评价，三端只能比较问题层面的共同点。 |
| 05 | 访谈使用边界 | 5名访谈对象只解释具体过程；相近判断重复出现也不承担群体比例推断。 |

按研究方法分工

问卷组	定性访谈组	桌面研究组	数据处理组	H5搭建与 统筹整合 1人
6人	3人	2人	2人	H5系统 研究方案、报告 汇总与对齐 整体统筹
设计定量问卷问题	执行访谈并编写摘要	提供背景信息与初始假设	处理与分析问卷数据	

前三周的协作方式	各组先内部讨论负责区域，再通过例会汇报、交流和对齐。第二周导师参与例会后，研究重点开始发生实质调整。
第三周后 持续执行	随着人员可投入时间下降，协作逐渐转为由少数成员推进；继续完成问卷投放、样本招募与筛选、H5维护、访谈执行和方案修订。
后期 材料整合与成稿	完成数据清洗、访谈摘要和桌面研究材料整理；持续维护数据口径、重组证据结构，并根据导师反馈重写报告。

初始分工解决了启动问题，但没有覆盖后续不断变化的研究对象和执行任务。

研究重点和对象范围的两次收束，都来自“原问题无法被现有条件严谨回答”

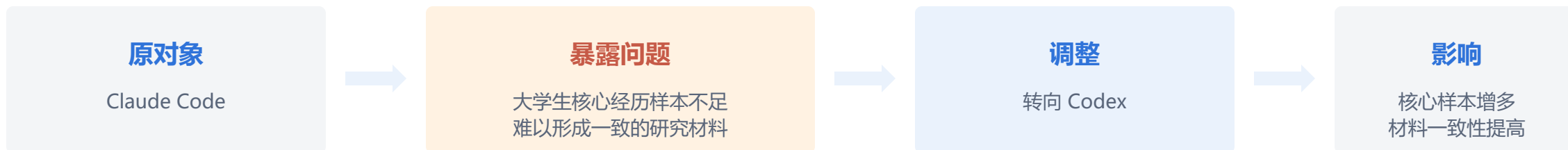
	原设想	暴露问题 / 导师反馈	实际调整	调整带来的影响
01	参考传统App分析，关注留存率与流失率	AI工具持续更新；单独追踪流失不能解释用户在具体任务中怎样使用	重点转向使用者	核心问题改为诉求、任务、协作和评价
02	以“AI工具”“AI平台”为宽泛主语	覆盖整个类别需要大量跨产品数据；单一产品又不能代表整个类别	限定三条具体路径	同时限定研究对象、问题和结论边界

这两次调整的主要目的

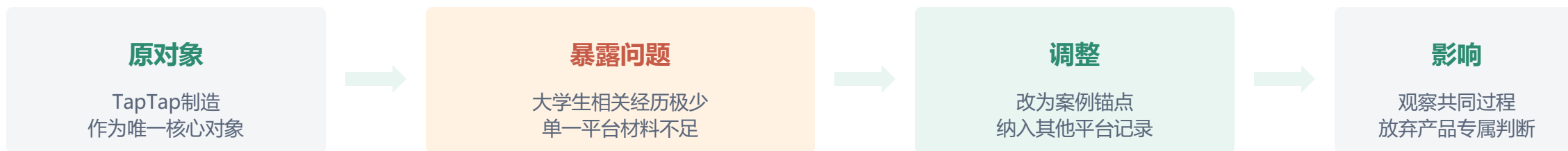
- 1、针对对象本身性质做出方案，不局限于传统App调研
- 2、把“想研究什么”改成“现有时间、人力和材料能够严谨回答什么”。

首轮投放缺少核心经历样本，工具和平台口径因此分别调整

程序工具路径



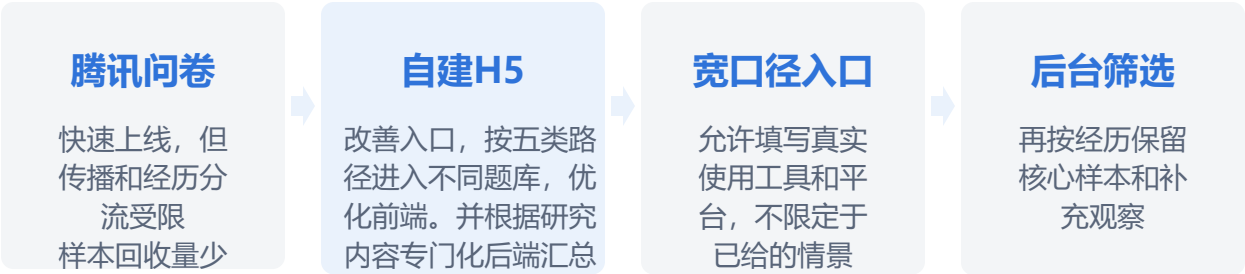
一站式平台路径



对象调整的原则：根据能够获得和核验的材料，重新限定结论能够说明到哪里。

入口调整解决招募和分流问题，报告重构解决证据重复与结论分散

问卷入口为什么改变



报告结构为什么改变



这次项目真正训练的是让研究对象、证据、结论和执行机制持续对齐

仍然存在的不足

- 01

招募偏向开发者

平台和消费端核心经历材料仍然有限
- 02

平台材料不对称

只能观察多平台共同过程，不能形成产品专属判断
- 03

动态分工不足

研究对象变化后没有同步重分工和设置阶段验收

我获得的具体认识

- 研究对象

必须同时限定对象、问题和结论边界
- 证据分工

定量描述现象，开放回答和访谈解释过程
- 报告组织

研究方法服务于对象，不能反过来决定报告结构
- 项目管理

研究设计变化后，分工和验收机制也要同步变化

最终完成的不是一份一次成型的报告，而是一次连续发现问题、修正设计并控制结论边界的用户研究实践。