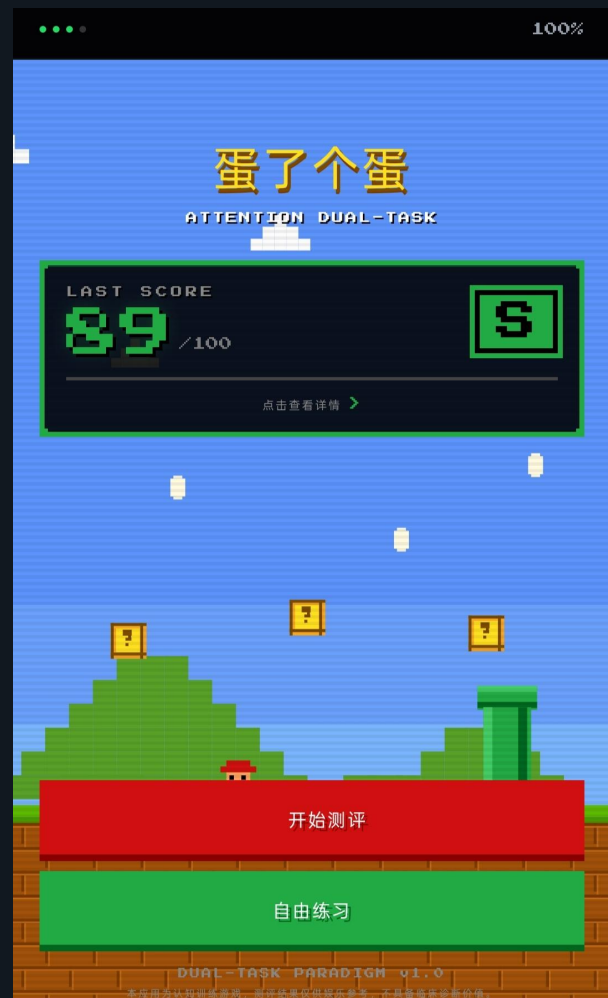


蛋了个蛋

注意力分配双任务测评游戏

认知心理学课程作业汇报

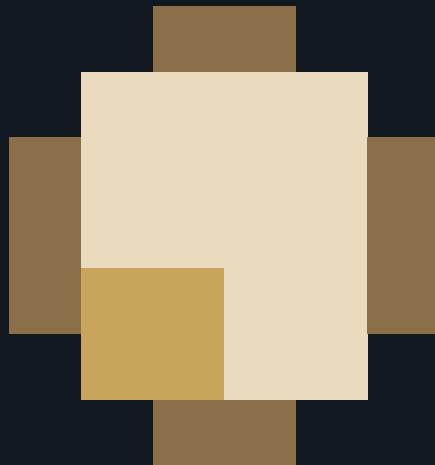
eggtask.bouncingpig.cn



.....

为什么要做这个项目？

想试试看，
能不能利用认知心理学的经典实验范式
做成一个有点意思的游戏



理论基础

为什么选择这个范式

- “忙不过来”的体验：类似节奏游戏里同时处理多个轨道的信息、经营游戏里处理紧急订单。
- 有机会在“实验任务”和“游戏体验”之间找到一个比较自然的接合点。至于这个接合点能否真正成立，就是这个项目想要探索的。

01

双任务范式

Dual-Task Paradigm

让被试同时执行两项任务，测量双任务条件下的表现衰减。若两项任务都需要注意力资源，双任务表现会显著低于单任务。衰减幅度 = 注意力资源损耗 = 注意力分配能力的反向指标。

02

短时记忆容量

「7±2」理论

Miller (1956): 人类短时记忆容量为 7 ± 2 个信息单元。超出容量后，记忆准确率显著下降。

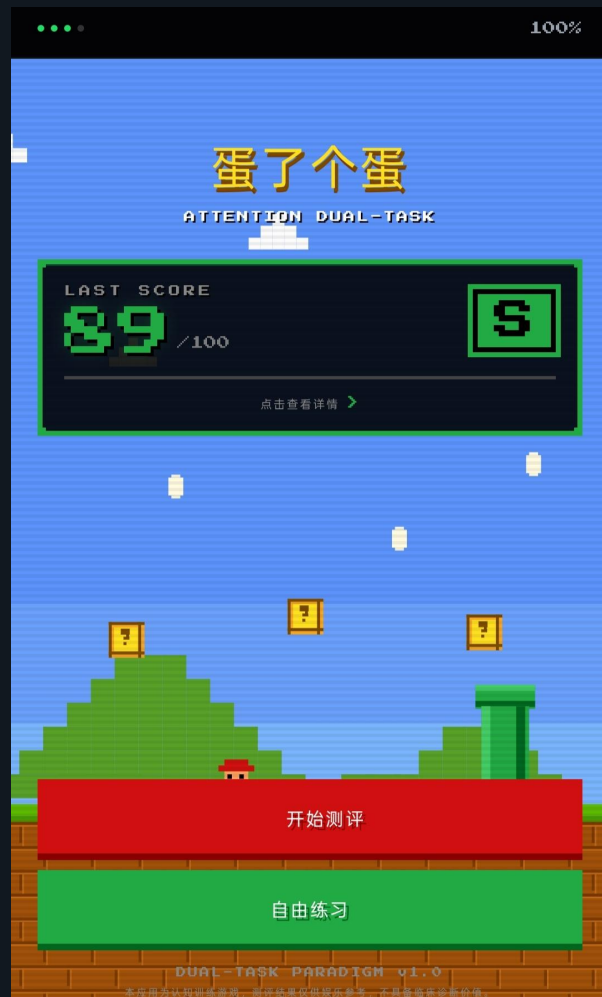
03

注意力分配理论

Kahneman (1973)

注意力是有限资源，多任务时需在任务间分配。双任务损耗越大 → 注意力分配能力越弱。

- 测评模式
- 练习模式

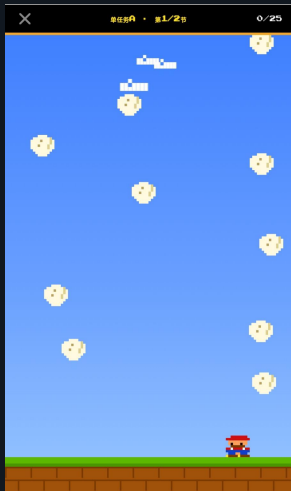


测评模式

任务 A：追踪任务（接鸡蛋）

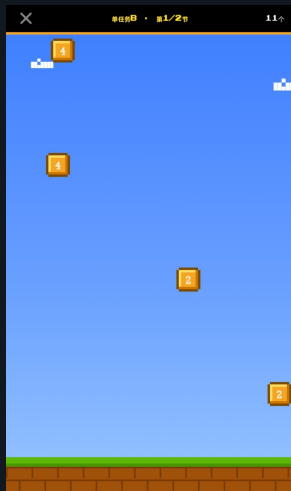
玩家操控角色左右移动，接住下落的鸡蛋
对应认知心理学中的追踪任务（Tracking Task）

测量：手眼协调 + 持续注意力



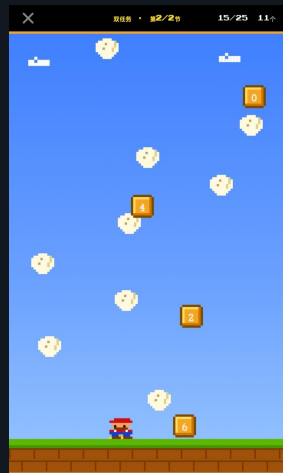
任务 B：记忆任务（数字回忆）

屏幕中随机下落数字方块，玩家需记住数字
小节结束后用数字按钮盘回忆输入
对应自由回忆范式（Free Recall）
测量：短时记忆 + 编码能力



双任务条件

双任务条件：鸡蛋和数字同步下落，玩家需同时完成接球和记忆。对比单任务基线，计算注意力损耗。



测评模式

单任务 A (2节) → 休息 → 单任务 B (2节) → 休息 → 双任务 (2节)

阶段	内容	鸡蛋数	数字数	时长/节
单任务 A	仅接鸡蛋	25	0	12秒
单任务 B	仅记数字	0	11	12秒
双任务	接鸡蛋+记数字	25	11	12秒

设计要点

- ▶ 固定顺序：先建立单任务基线，再测双任务
- ▶ 每节固定时长 12 秒，隐藏计时器避免时间压力干扰
- ▶ 每个任务间有休息界面，玩家自主控制节奏，避免疲劳累积
- ▶ 每个条件重复 2 节取均值，减少随机误差

分数计算

Step 1

单任务基线

S_A = 单任务A接球正确率 (2节均值)

S_B = 单任务B记忆正确率 (2节均值)

Step 2

双任务实测

D_A = 双任务接球正确率 (2节均值)

D_B = 双任务记忆正确率 (2节均值)

Step 3

注意力损耗

$Loss_A = (S_A - D_A) \div S_A$

$Loss_B = (S_B - D_B) \div S_B$

Step 4

综合分数

$FinalScore = 100 \times (1 - (Loss_A + Loss_B) \div 2)$

解读：分数越高 → 双任务下损耗越小 → 注意力分配能力越强

分数解读

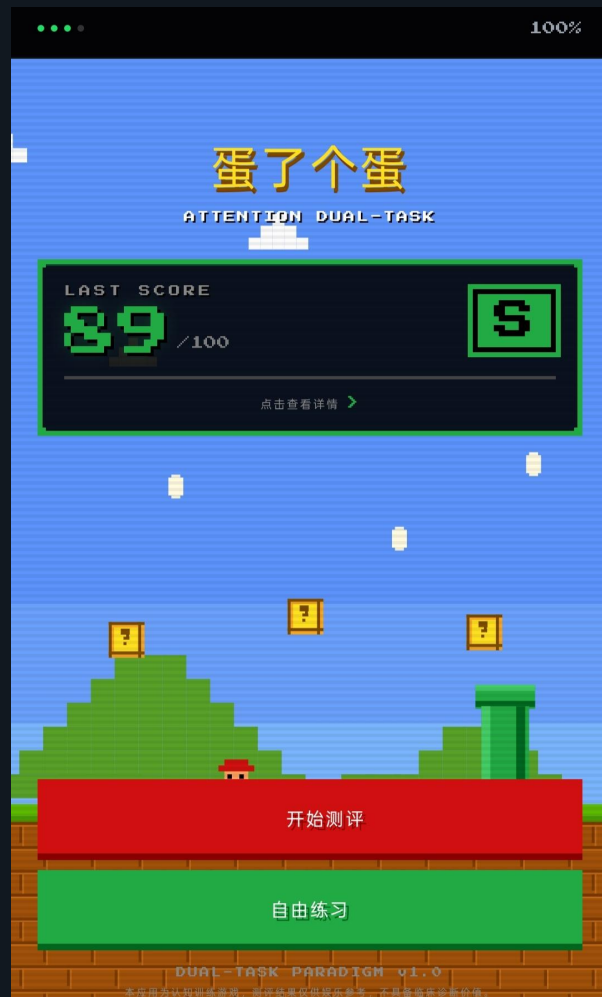
等级	分数范围	含义
S	≥85	注意力分配能力优秀，双任务下几乎无衰减
A	≥70	良好，双任务下有轻微衰减
B	≥55	一般，双任务下有明显衰减
C	<55	待提升，双任务下衰减严重

设计说明

- ▶ 等级划分参考常见游戏评级 (S/A/B/C)，便于用户理解和传播
- ▶ 分数阈值为经验设定，非临床诊断标准
- ▶ 结果页明确标注"仅供娱乐参考，不具备临床诊断价值"



- 测评模式
- 练习模式



自由练习模块

▶ 定位：

测评模式固定为11个数字，这是为了分数的唯一性和可比性。但代价是玩家只能体验一种难度。

▶ 练习模式补上了这个缺口：

轻量：选好难度打一局12秒就走，没有测评那么重的流程

多档：5/7/9/11/13/15个数字，体验记忆负荷从低到高的变化

数字数	难度	说明
5	入门	试玩熟悉操作
7	简单	在 7 ± 2 范围内
9	中等	接近记忆容量上限
11	标准	与测评一致
13	困难	超出记忆容量
15	挑战	远超记忆容量



▶ 鸡蛋数固定 25 个，只调节数字数量



局限与反思



局限 — 关于定位

这个产品最底层的困惑是：它到底是什么？

如果是游戏

用户体验围绕「好玩」展开

界面语言：挑战、成就、进度感

目标：让玩家愿意再次打开

如果是认知测评工具

用户体验围绕「可靠」展开

界面语言：准确、标准、可解释

目标：让结果值得信任

当前版本在两者之间来回摇摆。两边都不彻底，这也是下面所有问题的根源。

后续推进需在两个方向中做选择：

方向 A：彻底转向游戏

接受严谨性上的妥协，用游戏语言重新包装体验（界面、反馈、激励）

方向 B：彻底转向工具

接受「用完即走」的定位，提升测评严谨性（增加试次、校准参数、验证效率度）

中间地带的模糊定位是最差的状态，因为它让每个决策都变得两难。



作为心理学实验的反思

作为心理学实验的反思

分配性注意和多任务处理



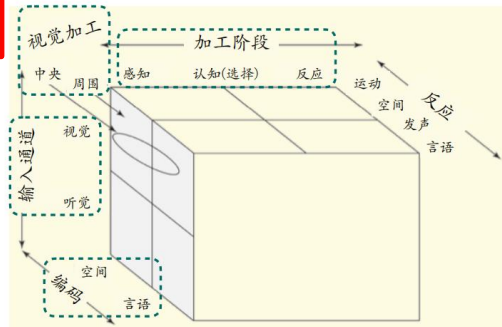
- 多重资源模型(维肯斯 Wickens, 1984; 2008)
 - 用来预测在多任务处理情景下，两个任务之间会产生多大的干扰；
 - 大脑具有不同的独立的注意资源(多重资源)，**当同时处理两个任务时，相互之间的干扰取决于两个任务消耗占据了**多少相同的资源。

分配性注意和多任务处理



- 多重资源模型(维肯斯 Wickens, 1984; 2008)

- 1. 加工的阶段
 - 感知、认知、反应执行等
- 2. 加工的编码方式
 - 空间编码、言语编码等
- 3. 信息输入的通道
 - 听觉、视觉等
- 4. 视觉加工
 - 中央、外周等



作为心理学实验的反思

1. 单任务A的天花板效应

单任务A是纯滑动接物任务，属于平滑追踪，对健康成年人几乎完全自动化。

调节物理参数（数量、速度、x/y轴间距）只能触及运动极限，触及不到认知负荷。正确率不会平滑下降——保持100%直到临界点，然后断崖式下跌。

根本原因：追踪任务缺乏认知决策成分。如果重做，应由追踪任务改为需要认知参与的任务（如视觉搜索、规则切换等），让任务有足够的资源竞争。

2. 构念效度问题

当接球被证明是自动化任务后，双任务范式名存实亡——两个任务没有形成真正的资源竞争。

最终测到的本质上是「记忆任务在低负荷追踪条件下的表现」，而不是「注意力分配能力」。这是构念效度问题——测的东西和声称要测的不是同一个。

这个偏差无法通过调整参数修复，只能通过更换任务类型解决。

作为心理学实验的反思

3. 缺少预实验

核心参数的确定缺乏实证支撑。数字数量为什么是11个？鸡蛋为什么是25个？每条件为什么是2节？每个参数都基于个人测试，未经过小规模预实验校准。

正确流程：先招募少量被试测试不同参数组合 → 分析难度梯度是否合理 → 用数据确定最终参数。这一步缺失意味着所有后续结果都建立在未经检验的假设之上。

5. 信度问题

每条件仅2小节，双任务共4小节，总测评时长不到3分钟。这么少的试次，分数在多大程度上反映真实能力、多大程度上是随机波动，是未知的。

如果重做，应该增加小节数，或采用多次测评取平均的方式提高信度。

4. 资源分配策略

双任务实验中，被试天然会策略性调整资源投向。实验者需要能识别这种策略。数据分析阶段应同时关注Cost_A和Cost_B两个指标，而非只看总分。如果一个玩家的Cost_A远大于Cost_B（接球严重下降、记忆几乎无损），说明他把注意力优先分配给了记忆任务，这个策略模式本身就应该被记录和分析。研究者收集完数据后，需要注意在后台分析时分开看两个成本。

6. 外部效度缺失

这个测评分和真实世界的注意力表现有多大关系？目前没有任何验证来建立映射关系。

作为严肃的认知测评工具，这是必要的一步；作为娱乐产品，这个问题可以暂时搁置。定位决定了对这个问题的重视程度。



作为游戏的反思

作为游戏的反思

1. 用户体验的「实验感」过重

双任务范式的核心体验「忙不过来」本身可以很有趣，类似烹饪游戏里同时处理多个订单。但当前界面语言是「单任务A」「双任务」「基线」，结果展示「损耗率」，这些词带有实验严肃感，而非游戏兴奋感。玩家看到这些词，心理预期自然变成「我在做测试」。如果想让玩家觉得有趣，用词、界面设计等方面需要调整。

3. 社交传播锚点

一个东西能传播，需要考虑：可以比（朋友间竞争）、可以聊（带出有意思的话题）、可以玩（分享本身让人想点进来）。

这个产品目前只做到了第一点。注意力分配的分数的缺乏社交属性。

“我的注意力分配能力比你强”，这件事有多少人在意？它不像审美、幽默感、知识储备那样带有社交价值标签，缺乏被讨论的话题度。

2. 留存机制缺失

测完第一次，拿到分数，分享出去，体验闭环就结束了。除非想看看进步、挑战更高分、和朋友竞争，否则用户没有动机再次打开。但当前版本没有历史趋势图、排行榜、成就系统。是否增加这些取决于定位：如果是体验型产品，用完即走正常；想做持续运营，必须回答「用户为什么回来」。

4. 盈利模式的思考？

常见路径：免费+广告（简单但对体验有损耗）、免费测评+付费报告（适合工具型）、一次性付费下载（适合单机体验型）。选择取决于目标用户画像和产品定位。

盈利模式不是先于产品存在的东西，它取决于用户是谁、他们为什么来。在用户画像清晰之前，这个问题无法被回答。



总结 — 核心矛盾

实验逻辑和游戏逻辑在底层存在根本性冲突：

	实验逻辑	游戏逻辑
试次	越多越好（信度）	越少越好（不无聊）
变量	严格控制	灵活可调
反馈	中性反馈	积极激励
流程	流程固定	自由探索

这个矛盾无法通过优化解决，只能通过选择解决。

当前版本同时承载了两种期待，结果两边都做不彻底。因此定位不解决，所有后续决策都会在两种逻辑之间摇摆。



谢谢观看

2026.7.2