

# 基于变异系数的游戏用户体验度量方法研究

李晓娟, 涂俊, 章灏  
(南开大学, 天津 300071)

**摘要:** **目的** 对角色扮演游戏过程中被试的行为进行分析。探索和建立一种符合不同需求的游戏用户体验的研究方法, 为角色扮演游戏设计和评价提供客观依据。**方法** 通过预实验确定具体研究内容, 通过眼动技术研究游戏提示的可用性。采用基于行为观察和自我报告的度量研究方法, 通过新手—专家范式对实验结果进行分析。眼动数据结果基本符合大众认知, 同时也受游戏设置的影响。对于游戏的评价很大程度上受游戏的战斗效果和难易度设置影响。变异系数反映了被试组内和组间的差异, 新手与专家间表现的差异, 反映了游戏的难度特征及游戏本身特性。**结论** 利用变异系数研究新手组和专家组在游戏中的不同表现, 可以综合体现难易度和用户表现之间的关系。

**关键词:** 游戏; 用户投入量表; 眼动; 用户体验度量; 变异系数

**中图分类号:** TB472 **文献标识码:** A **文章编号:** 1001-3563(2019)08-0136-05

**DOI:** 10.19554/j.cnki.1001-3563.2019.08.023

## Game User Experience Measurement Method Based on CV

LI Xiao-juan, TU Jun, ZHANG Hao  
(NanKai University, Tianjin 300071, China)

**ABSTRACT:** The paper aims to analyze the user's behavior of subjects in the role playing game (RPG), explore and establish a research method of game user experience that meets different needs, so as to provide an objective basis for design and evaluation of RPG. The specific content of the study was determined through pre-experiment. The usability of the game cue was studied through the eye-track technology. The experimental result was analyzed through novice - expert paradigm based on behavior observation and self-report. The eye-track data were basically consistent with the public cognition, but were also influenced by the game design itself. Evaluation on a game was largely influenced by the game's combat effect and difficulty setting effect. CV (coefficient of variability) was a good reflection of the differences between group and the differences in group. The difference between the novice and the expert reflects the difficulty level characteristics and other design characteristics of the game. Studying different behaviors of expert group and novice group in the game through CV can reflect the relationship between the difficulty level and user performance.

**KEY WORDS:** games; user engagement scale; eye tracking; user experience measurement; coefficient of variability

游戏作为 20 世纪末的新兴娱乐媒介, 在最近 20 年的发展可谓是突飞猛进。新的游戏引擎技术使得游戏产业更加流程化、工业化<sup>[1]</sup>。随着游戏产业工业化的发展, 从用户的角度对游戏的评价——用户体验度量的需求越来越急迫。游戏产业越来越需要一个较为完善的评价方法, 指导游戏的游戏用户体验度量。我国游

戏产业多以大型多人角色扮演游戏作为主要产品<sup>[2]</sup>。这些产品由于更多的借鉴日本、韩国等国 10 年前的游戏生产经营模式, 竞争更加激烈, 急需再创新。游戏的游戏用户体验度量将帮助改善游戏产品, 创造新的增长点<sup>[3]</sup>。

国内对游戏的研究大多集中在游戏叙事和结构

收稿日期: 2018-12-21

基金项目: 国家自然科学基金资助项目 (71702080); 天津市艺术科学规划项目 (A16020)

作者简介: 李晓娟 (1984—), 女, 河北人, 硕士, 南开大学实验师, 主要研究方向为图形图像处理、实验教学管理与应用。

通信作者: 涂俊 (1963—), 女, 湖北人, 硕士, 南开大学教授, 主要研究方向为新媒体艺术与应用、实验教学管理与应用。

学、游戏的心理效应、教育游戏的开发和可用性评价等方面。国外对游戏的研究除了上述方面以外，还采用眼动技术分析用户在游戏时的视觉行为，用户游戏时的浸入感、投入度等。

## 1 研究背景

### 1.1 游戏的评价方法

游戏的评价方法主要有游戏媒体评论、问卷调查方法、用户行为分析法、生理信号采集法等。游戏媒体评论是比较权威的评价方法，游戏媒体将游戏玩家、媒体编辑、专业访谈人员3个角色结合在一起，提升了关于玩家对于游戏产品意见的收集效率。以著名游戏媒体IGN(Imagine Games Network)为例，他们拥有庞大的编辑队伍和自有的打分体系。编辑们进行试玩游戏后，对游戏的个人感受撰写评论，并提炼出一些优点和不足，最后打出游戏的分数。

除了利用传统而且高效的问卷调查方法获得用户的个人信息、想法态度等数据以外，随着用户研究手段的发展，对用户行为的记录开始变得可能。用户的行为种类繁多，为了更便捷地收集和解释用户行为，眼动手段逐渐成为用户行为观察中常用的方法。另外，脑电图、皮电反应、常规的心电图甚至是脑部核磁共振，能够揭示被试的诸多心理或是情绪过程。

### 1.2 眼动技术与心流理论

眼动技术在游戏研究中的应用已经比较广泛，Andrea<sup>[4]</sup>等在2011年通过眼动手段分析了80 Days游戏。实验招募了12名13岁左右的奥地利儿童，游戏期间记录了他们的眼动数据，并在实验结束后进行知识的测验。研究表明，在测验中表现更好的被试，在游戏中拥有更多的注视次数，且视觉注意的分配更加均匀。

心流，是心理学家米哈里·契克森米哈赖在70年代末逐渐形成的一种理论<sup>[5]</sup>。心流被描述成一种心理上沉浸忘我的过程。“一种人们如此痴迷于一项活动中去的状态，以至于其他事仿佛不再存在。”用户投入量表，简称UES(User Engagement Scale)，最早来自于Jacques在1996年的成果<sup>[6]</sup>。他在博士论文中详细讨论了用户投入的定义以及使用用户投入度评价超媒体(Hypermedia)的方法。用户投入(Engagement)，从字面意思讲，即是用户投入到一项活动中去，类似于可用性。

### 1.3 变异系数

变异系数可以反映一组数据的相对变化程度，是反映样本数据离散程度的重要测度指标，常常用在相同类型事物之间的相对量的差异比较。变异系数越大说明比较的主体间的差异越大，变异系数的变化也体现了主体间差异的变化趋势。王文森<sup>[7]</sup>从经济统计的

角度概括了变异系数的作用。吴晓琪<sup>[8]</sup>在研究基于行业工资的居民收入差距时，利用变异系数考察行业间工资收入差异，并得出我国城乡居民之间工资差距，城镇行业工资差距均呈现出相对缩小的整体变化趋势。

## 2 眼动实验设计

### 2.1 问卷的准备

测前问卷，包括被试的游戏习惯、接触游戏的时间、游玩过的游戏类型、游玩游戏的动机。UES为五级李克特式量表来自文献，翻译成中文并删去了负载因子小于0.5的项目。

### 2.2 实验被试的选取

实验被试均来自南开大学的学生，男11名、女15名，平均年龄24.3岁，77%的玩家有5年以上的玩游戏经历。为了解玩家的动机，特别询问了玩家玩游戏的动机<sup>[9]</sup>。65%的玩家玩游戏动机均为消磨时间，19%的玩家是为了获得成就感。

### 2.3 实验设备

实验设备为两台计算机，其中一台用以被试玩游戏(呈现眼动刺激)，另一台连接眼动仪，用以实时监视玩家的眼动数据。架设摄像头，用于拍摄游戏画面，使Tobii Studio分析软件在Scene Camera模式运行。

实验素材为古墓丽影与古剑奇谭，所有游戏均以最高画质，全屏运行(1920×1080)。所有被试均被要求以普通难度进行游戏。古墓丽影为古墓丽影9版本，古剑奇谭为古剑奇谭2版本。

### 2.4 预实验

为了选择更有代表性的游戏，熟悉实验流程，寻找实验中的可能问题，特别进行了预实验。根据IGN 2011—2014年度游戏提名及游戏累计销量排名，在预实验中选取了5个游戏，选取3名被试(均有5年以上游戏经历)进行游戏并完成UES，结果见表1。

表1 预实验UES平均得分  
Tab.1 UES average score of Pre experiment

组别	游戏名称	UES得分
国外组	上古卷轴V:天际	51
	古墓丽影	67
	宠物小精灵白	52
国内组	仙剑奇侠传五前传	60
	古剑奇谭	70

选择国外组最高分古墓丽影和国内组最高分古剑奇谭进行正式实验。

### 2.5 实验流程

1) 被试到达实验场所时，填写关于过往游戏经

历的问卷。

2) 被试阅读关于两个游戏操作说明的文档。

3) 校准眼动仪。

4) 被试按照要求分别游玩两个游戏至规定的进度(古墓丽影需要游玩至谜题部分,古剑奇谭需要游玩至快速反应事件)。

5) 被试需要填写 UES。

6) 如果被试在游戏过程中被发现曾有难以完成任务、很快完成任务、频繁切换场景、长时间或多次尝试与任务无关的行为,访谈并询问原因。

### 3 结果

玩游戏经历在 5 年以上,且有玩过动作冒险类游戏的被试,定义为古墓丽影组的专家玩者。玩游戏经历在 5 年以上,且有玩过角色扮演类的被试,定义为古剑奇谭组的专家玩者。其中古墓丽影游戏新手 13 人、专家 9 人,古剑奇谭游戏新手 13 人、专家 9 人<sup>[10]</sup>。

#### 3.1 眼动实验

由于两个游戏的流程相差较大,在游戏内进行了新手和专家任务时间的对比。古墓丽影中任务时间是指被试拔出钢筋所需时间,古剑奇谭中任务时间是走出房间至任务触发地点所需时间。

古墓丽影专家组完成任务平均用时 24.12 s,新手组平均用时 24.87 s,任务完成时间没有显著区别。考虑到手型提示和文字提示是被试了解到按键进度的主要提示,特别考察了被试对提示的首次注视时间和注视持续时间。古墓丽影提示组首次注视时间均值为 0.343 s,注视持续时间均值为 1.651 s。新手组提示首次注视时间均值为 1.382 s,注视持续时间均值为 2.613 s。专家比新手更快注意到提示,注视时间更短。古墓丽影手型提示见图 1。



图 1 古墓丽影手型提示  
Fig.1 Cue of Tomb Raider

在古剑奇谭中,被试需要按照屏幕上的地图上所标定的任务地点,找到并进入任务地点。值得一提的是,这是关于任务地点的唯一提示。古剑奇谭专家组平均用时 56.73 s。新手组平均用时 51.28 s,任务完成时间上没有显著区别。古剑奇谭地图叹号提示见图 2。



图 2 古剑奇谭地图叹号提示  
Fig.2 Cue of Gujian

基于同样的理由,同样考察了被试对地图叹号提示的眼动指标。古剑奇谭专家组首次注视时间均值为 1.267 s,注视持续时间均值为 1.369 s。新手组首次注视时间均值为 1.473 s,注视持续时间均值为 1.171 s,专家组使用了更长的时间注视地图上的提示<sup>[11]</sup>。

从任务完成时间上来说,两款游戏的新手专家组之间都没有显著区别。从首次注视时间来说,两款游戏都是专家组首次注视时间短于新手组,这符合平常认知。从注视持续时间来说,古墓丽影游戏专家组注视时间短于新手组,这说明专家组很快明白了手型提示的意思。

在古剑奇谭游戏中,专家组对任务提示的注视持续时间长于新手组,这并不符合一般认知。就本游戏来说,这说明专家组更好地了解叹号提示的意义,并通过眼睛长时间注视小地图上叹号提示的同时,用鼠标控制主角走到任务触发地点。这种现象是由于不同游戏对于提示如何起作用的不同设置造成的。

通过比较首次注视时间(TTF, Time to First Fixation)和注视持续时间(FD, Fixation Duration)新手和专家组的变异系数(CV, Coefficient of Variation),见表 2 和表 3,发现古墓丽影专家组与新手组的组间差异较古剑奇谭更大。这说明古墓丽影游戏针对不同类型玩家更有区分度。另外,古墓丽影游戏新手组内差异大于专家组,古剑奇谭游戏新手组内差异小于专家组。这说明古墓丽影游戏上手较难,古剑奇谭游戏上手较容易。

表 2 古墓丽影 TTF,FD 变异系数对比  
Tab.2 CV comparison of TTF and FD in Tomb Raider

	TTF_QTE	FD_QTE
变异系数—新手/%	124.36	78.98
变异系数—专家/%	75.3	73.31

表 3 古剑奇谭 TTF,FD 变异系数对比  
Tab.3 CV comparison of TTF and FD in GuJian

	TTF_QTE	FD_QTE
变异系数—新手/%	43.99	62.26
变异系数—专家/%	46.14	81.00

### 3.2 主观评价

用户投入量表主要用于了解被试的投入度，分为强化注意 FA(Focused Attention)、易用性 PU(Perceived Usability)、审美性 AE(Aesthetics)、满意度 SA(Satisfaction)4 个维度，UES 总分即为 4 项和。UES 总分表示总体的投入度。UES 总分的及格分数为 63，当有分项或总分低于及格分数时，说明被试对此有负性的态度。本研究中两个游戏的 UES 总分和各分项均高于及格分数。其中，古墓丽影 UES 总分均值 72.19 分，古剑奇谭 UES 总分均值 68.62 分。UES 变异系数比较见图 3。

通过比较变异系数，可以看到古墓新手在审美性 AE、满意度 SA 上的分歧非常大，古墓丽影较为血腥，ESRB 评价<sup>[12]</sup>中含有血腥成分的游戏，被试在审美性 AE 上的意见不一致，审美上的分歧进一步导致满意度上的分歧。古剑专家组在审美性 AE 上的变异系数

也很大，经访谈，部分专家认为古剑奇谭的战斗动画比较重复，产生审美疲劳，这可能是整体审美的评价差异的原因。UES 配对样本 *t* 检验见表 4。

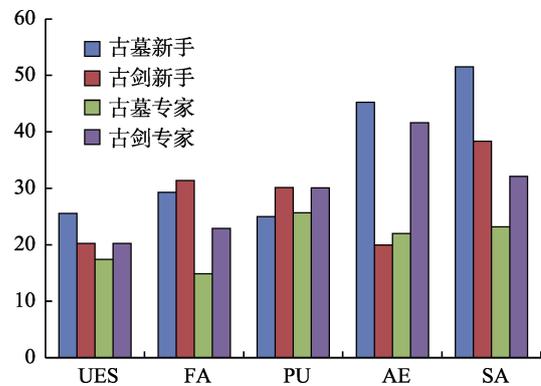


图 3 UES 变异系数比较  
Fig.3 CV Comparison of UES

表 4 UES 配对样本 *t* 检验  
Tab.4 UES paired sample *t* test

	成对差分					Sig.(双侧)
	均值	标准差	均值的标准误	差分的 95%置信区间		
				下限	上限	
古墓 UES—古剑 UES	3.577	17.569	3.445	-3.519	10.673	0.309
古墓 FA—古剑 FA	0.808	4.899	0.961	-1.171	2.786	0.409
古墓 PU—古剑 PU	0.308	9.768	1.916	-4.253	3.638	0.874
古墓 AE—古剑 AE	1.462	7.355	1.442	-1.509	4.432	0.321
古墓 SA—古剑 SA	1.615	3.981	0.781	0.008	3.223	<b>0.049</b>

在组间变异系数比较中，满意度 SA 项的 *P* 值小于 0.05，两游戏之间在满意度上有显著性的差异，古墓丽影给玩家带来了更大的满足感。古剑奇谭在强化注意 FA、易用性 PU 上的变异系数更大，新手与专家对于游戏的易用性评级有较大差异。结合表 3 的结果，古剑奇谭游戏确实对新手来说上手容易，但专家组觉得难度不够，所以易用性方面组内差异较大。

### 4 结果分析

变异系数在经济研究中十分常用，用于评价区域经济的差异。在本实验中，变异系数很好地反映了被试组内和组间的差异。新手与专家间表现的差异，反映了游戏的难度特征及游戏本身特性。心流的体验会带来极大的喜悦感、控制感和成就感。这种体验发生的条件是玩家所面临的挑战和需要的技能都呈适中状态。能力高、挑战低会感到无聊；能力低，挑战高则会感到焦虑。

将被试的表现（评分、绩效、眼动等统计量的均值）和被试的变异系数结合起来，可以划出 4 个象限。其中，高表现、高变异，代表着区分度较高，均值较

高的情况。这时游戏被更多人肯定，同时玩家的表现差异最大；正如雅达利创始人所说“好的游戏应是易于上手，难于掌握”。高表现、低变异，说明游戏表现较低，区分度不高，好比愤怒的小鸟一样，更像是休闲小游戏。低表现、高变异，说明游戏表现较低，区分度很高，类似的游戏有 I wanna；低表现低变异，说明游戏表现低，区分度也低，这时玩家的评价最低，类似的游戏有 E.T。表现与变异系数的 4 个象限见图 4。

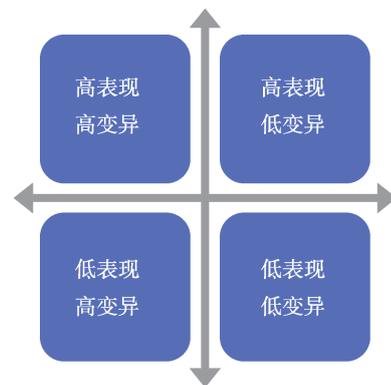


图 4 表现与变异系数的 4 个象限  
Fig.4 Four quadrants of CV and expression

影响变异系数的因素包括玩家情感和游戏本身的可用性,变异系数代表了游戏所导致的玩家间的行为差异。游戏的可用性影响因素包括关卡设计、外观设计和提示的可见性;玩家的情感由玩家自身沉浸感和游戏共情因素来决定<sup>[13]</sup>。

## 5 结语

整个实验过程中,需要精细的实验设计,否则由于游戏时间较长会造成海量数据。游戏中的提示种类很多,快速反应事件 QTE 也是常出现的游戏环节<sup>[14]</sup>。在未来研究中,对 QTE 的研究能更好地丰富对游戏可见性的研究。另外,玩家自身的共情能力也影响整个研究,未来通过测试被试的焦虑性和共情能力更好把握被试的差异性,从而更好地专注研究游戏本身的特性<sup>[15]</sup>。

### 参考文献:

- [1] 赫伊津哈. 游戏的人[M]. 杭州: 中国美术学院出版社, 1996.  
JOHAN H. Study of the Play-element in Culture[M]. Hangzhou: The China Academy of Art Press, 1996.
- [2] 新华网. 2015 年全球电子游戏收入排行榜出炉: LOL 登顶 COC 居次[EB/OL]. [http://news.xinhuanet.com/info/2016-01/28/c\\_135050057.htm](http://news.xinhuanet.com/info/2016-01/28/c_135050057.htm)  
Xinhua Net. 2015 global video game revenue list: LOL top COC in[EB/OL]. [http://news.xinhuanet.com/info/2016-01/28/c\\_135050057.htm](http://news.xinhuanet.com/info/2016-01/28/c_135050057.htm)
- [3] 中国音像与数字出版协会游戏出版工作委员会. 2015 年中国游戏产业报告[M]. 北京: 中国书籍出版社, 2015.  
Chinese Video and Digital Publishing Association Game Publishing Committee. China Game Industry Report[M]. Beijing: China Book Publishing House, 2015.
- [4] MICHAEL D, KICKMEIER R. Tracking the UFO's Paths: Using Eye-Tracking for the Evaluation of Serious Games[M]. USA: Orlando, 2011.
- [5] 米哈里·契克森米哈赖. 生命的心流[M]. 北京: 中信出版社, 2009.  
MIHALY C. Finding Flow the Psychology of Engagement with Everyday Life[M]. Beijing: CITIC Press, 2009.
- [6] RICHARD D J. The Nature of Engagement and Its Role in Hypermedia Evaluation and Design[D]. London: South Bank University, 1996.
- [7] 王文森. 变异系数——一个衡量离散程度简单而有用的统计指标[J]. 中国统计, 2007(6): 41—42.  
WANG Wen-Sen. Coefficient of Variation: a Statistical Indicator of the Simple and Useful Degree of Dispersion[J]. China Statistics, 2007(6): 41—42.
- [8] 吴晓琪. 基于行业工资的居民收入差距研究[J]. 工业技术经济, 2018(3): 137—146  
WU Xiao-qi. Resident Income Difference Analysis Based on Industrial Salary[J]. Industrial Technology & Economy, 2018(3): 137—146.
- [9] 郑显亮, 顾海根. 大学生网络利他行为和网络行为偏好的关系: 班级环境的作用[J]. 心理与行为研究, 2013, 11(5): 690—696.  
ZHENG Xian-liang, GU Hai-gen. The Relationship between Online Altruistic Behavior and Behavioral Preference of College Students: the Role of Classroom Environment[J]. Psychological and Behavioral Research, 2013, 11(5): 690—696.
- [10] 崔运坤, 贾燕, 张英波, 等. 知觉预测领域专家——新手研究范式的元分析[J]. 北京体育大学学报, 2016, 39(4): 109—117.  
CUI Yun-kun, JIA Yan, ZHANG Ying-bo, et al. Analysis on the Meta-analysis of Novice Research Paradigm in the Field of Perceptual Prediction[J]. Journal of Beijing Sports University, 2016, 39(4): 109—117.
- [11] 闫国利, 熊建平. 阅读研究中的主要眼动指标评述[J]. 心理科学进展, 2013, 21(4): 589—605.  
YAN Guo-li, XIONG Jian-ping. Commented on the Main Indexes of Eye Movement of Reading Research[J]. Advances in Psychological Science, 2013, 21(4): 589—605.
- [12] ESRB. Rating Information[EB/OL]. <http://www.esrb.org/ratings/search.aspx?from=home&titleOrPublisher=tomb%20raider>.
- [13] 金瑜. 心理测量[M]. 上海: 华东师范大学出版社, 2005.  
JIN Yu. Psychometrics[M]. Shanghai: East China Normal University Press, 2005.
- [14] 汤晓静, 虞素馨, 孙丽君, 等. 加强医学生情绪智力和自我效能感培养的重要意义[J]. 卫生职业教育, 2017, 35(14): 63—64.  
TANG Xiao-jing, TUO Su-xin, SUN Li-jun, et al. Important Significance of Strengthening the Emotional Intelligence and Self-efficacy of Medical Students[J]. Health Profession Education, 2017, 35(14): 63—64.
- [15] 潘彦谷, 刘衍玲. 共情的神经生物基础[J]. 心理科学进展, 2012, 20(12): 2011—2021.  
PAN Yan-gu, LIU Yan-ling. Empathic Neurobiological Basis[J]. Progress in Psychological Science, 2012, 20(12): 2011—2021.