

# 基于用户体验的满意度评估

## ——以B站为例

葛 纓<sup>1\*</sup>, 谭云靳<sup>1\*</sup>, 刘 杰<sup>2</sup>, 何吉波<sup>3#</sup>, 罗 杰<sup>4</sup>

<sup>1</sup>重庆文理学院重庆市情绪与心理健康重点实验室暨用户体验与人机交互技术联合研究所, 重庆

<sup>2</sup>重庆第二师范学院教师教育学院暨西部儿童脑与认知科学研究院, 重庆

<sup>3</sup>南京师范大学心理学院, 江苏 南京

<sup>4</sup>贵州师范大学心理学院, 贵州 贵阳

收稿日期: 2024年7月19日; 录用日期: 2024年8月10日; 发布日期: 2024年8月21日

### 摘 要

目的: 本研究旨在探讨经典国外用户体验量表在中国本土环境下的适用性, 以丰富用户体验领域的测量工具。通过收集哔哩哔哩(B站)大学生用户的数据, 本研究旨在评估B站的用户满意度, 为其优化和改进提供有效建议, 确保用户需求能准确反映在APP的改进方向上, 从而为B站的发展提供实际的用户意见。方法: 本研究采用UMUX量表和WAMMI量表对1011名B站大学生的用户体验进行调查, 并以SUS量表作为效标。首先, 我们对UMUX量表进行了信效度检验; 接着, 对WAMMI量表的数据进行了探索性因素分析、验证性因素分析和信效度检验。我们使用经过修订的这两个量表的中文版来评估大学生对B站的用户体验情况。结论: 1) 修订后的UMUX量表显示出良好的信效度。对于WAMMI量表, 在删除第5题、第10题、第12题和第16题后, 剩余的16个题项保持了较高的信效度。两个量表的指标符合测量学标准, 表明它们在中国文化环境中具有一定的适用性。2) 总体而言, 中国大学生在使用B站时的用户体验较好, 且性别因素对用户体验无显著影响。然而, 使用时间越长、使用频率越高的用户体验越好。

### 关键词

用户体验测评, 适用性检验, 满意度评估, 哔哩哔哩(B站)

# Satisfaction Evaluation Based on User Experience

## —A Case Study of Bilibili

Ying Ge<sup>1\*</sup>, Yunjin Tan<sup>1\*</sup>, Jie Liu<sup>2</sup>, Jibo He<sup>3#</sup>, Jie Luo<sup>4</sup>

\*共同第一作者。

#通讯作者。

<sup>1</sup>Key Laboratory of Emotion and Emotion and Mental Health in Chongqing, User Experience & Human-Computer Interaction Technology Institute, Chongqing University of Arts and Sciences, Chongqing

<sup>2</sup>School of Teacher Development, West China Institute of Children's Brain and Cognition, Chongqing University of Education, Chongqing

<sup>3</sup>School of Psychology, Nanjing Normal University, Nanjing Jiangsu

<sup>4</sup>School of Psychology, Guizhou Normal University, Guiyang Guizhou

Received: Jul. 19<sup>th</sup>, 2024; accepted: Aug. 10<sup>th</sup>, 2024; published: Aug. 21<sup>st</sup>, 2024

## Abstract

**Purpose:** This study is to explore the applicability of the classical foreign user experience scale within the local context of China. This aims to enrich the measurement tools available in the field of user experience. By collecting data on college students' usage of Bilibili (Site B), the study seeks to evaluate user satisfaction with Site B, provide actionable suggestions for its optimization and improvement, ensure that user needs are accurately reflected in the direction of the APP's enhancements, and thereby contribute practical user feedback to the development of Site B. **Methods:** The UMUX scale and WAMMI scale were utilized to assess the user experience of 1011 college students on Site B, with the SUS scale serving as the criterion. Initially, we assessed the reliability and validity of the UMUX scale. Subsequently, we conducted exploratory factor analysis, confirmatory factor analysis, and validity testing on the WAMMI scale data. We employed the revised Chinese translations of these two scales to evaluate the user experience of college students on Site B. **Conclusion:** 1) The revised UMUX scale demonstrated good reliability and validity. Regarding the WAMMI scale, after eliminating questions 5, 10, 12, and 16, the remaining 16 items retained high reliability and validity. The indicators of the two scales fulfill the criteria for measurement, suggesting they possess a degree of applicability within the Chinese cultural context. 2) Overall, Chinese college students report a favorable user experience with Site B, and gender does not appear to significantly influence user experience. Nonetheless, increased duration and frequency of use are associated with improved user experience.

## Keywords

User Experience Evaluation, Suitability Test, Satisfaction Evaluation, Bilibili (Site B)

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



## 1. 引言

根据 CNNIC 第 53 次《中国互联网网络发展状况统计报告》，截至 2023 年 12 月，我国手机网民已达 10.91 亿，网民使用手机上网的比例为 99.9%。我国网络视频(含短视频)用户规模达 10.67 亿，占整体网民的 97.7% [1]。国内市场上监测到活跃的 APP 数量为 260 万款，人均 APP 安装总量上升至 68 款[2]。为了推动 APP 市场持续发展，用户体验作为用户对 APP 最直观、最主观的反馈，日益受到各方关注。

随着社会生产力的提升，消费者对产品的需求逐渐从数量需求转为可用性需求，当基本需求得到满足之后，用户体验的重要性日益凸显。根据 ISO 9241-11:2018 的定义，用户体验涵盖了用户在使用或预期使用系统、产品或服务时所产生的感知和反应[3]。产品可用性是用户体验的核心，而用户体验则是产

品可用性的扩展和深化[4]。目前,产品可用性的研究已广泛应用于移动设备[5] [6]、智能家居系统[7] [8]、网络应用程序[9]以及公共服务平台[10] [11]等多个领域的产品设计和性能优化。

用户体验的量化评估是对用户体验质量进行具体化的测量过程,包括主观与客观的测量手段。访谈和问卷调查由于成本效益高、实施便利,成为评估用户体验最常用的手段之一。从已有文献可见,陈淑红和程春通过自制问卷调查了网易云音乐 APP 用户体验对用户忠诚度的影响,并发现感官体验、互动体验和社会体验对用户忠诚度具有显著的正向影响[12]; Lee 等利用问卷调查和访谈法探究了老年人群使用手机的用户体验,结果表明,老年用户普遍反映手机操作复杂,能够熟练使用的功能也十分有限[13]; 丁群和卢星运用访谈和问卷法评估了老年家庭康养产品设计的用户体验[14]。

目前,系统可用性量表(System Usability Scale, SUS)是被广泛认可的用研测量工具,可用于各类产品、系统与服务的可用性测评[15]。在我国,该量表已经取得可靠的信效度验证[16] (详见方法部分)。但现有大多数用户体验评估量表均源自国外,未经过中国本土环境的信效度检验。用户体验的可用性量表(Usability Metric for User Experience, UMUX) [17]作为 SUS 量表的一个变式,融合了 ISO 的相关标准,旨在评估用户的感知可用性,且在英语和意大利语的环境中表现出良好的信效度指标[18] [19]。除了传统的可用性属性测评,网站分析与测量调查表(Website Analysis and Measurement Inventory, WAMMI)是为评价网站在线服务而设计的量化评估工具[20],有英语、法语、德语等大部分欧洲语言测试版本可用[21]。UMUX 量表和 WAMMI 量表目前还需要中文文化环境下的信效度验证。

哔哩哔哩(简称 B 站)是一个以视频弹幕功能著称的网站。据 B 站 2022 年四季报及全年财报指出,其日均活跃用户数达到 9280 万,月均活跃用户 3.26 亿,用户日均使用平台时间长达 96 分钟[22]。虽然 B 站在视频领域中以其独特的内容、形式和设计展现出一定的创新性和新颖性,但与抖音、快手等竞争对手相比,其发展仍显不足。面临这样一个庞大的用户基础, B 站需要解决的关键问题是如何将新用户培养成为忠实用户,并进一步转化为付费用户,以增强运营商的经济效益。用户转化过程中,用户体验扮演着至关重要的角色——它是推动用户留存和付费转化的关键因素[23]。因此, B 站当前的核心追求在于更精准地满足用户的显性和隐性需求,以此提升用户对产品的忠诚度和满意度[23]。

B 站的用户体验,根据已有研究,主要得益于其独特的播放机制——弹幕、内容丰富的原创视频以及相对较少的广告。例如,李堂辉基于“使用与满足”理论指出, B 站的弹幕功能作为一种新兴的互联网信息交流和传播方式,有效满足了用户的多样化需求[24]。黄炜等的研究发现,用户对信息的选择关注和对视频的满意度会直接影响其对 Vlog 的使用决策,同时,对视频的整体感受会直接影响用户持续使用的意愿[25]。赵瑜的研究则表明, B 站短视频的文本内容和互动体验呈现出两种叙事类型和三种交互模式,这些不同的模式直接影响用户的体验[26]。秦紫函的分析进一步指出, B 站注重文化和用户的运营策略,以确保为用户提供更好的体验[27]。然而,尽管前人研究涉及 B 站的多个方面,但尚未有研究对 B 站 APP 的整体满意度进行全面的评估。

B 站的用户群体中, 18~35 岁占比 78% [28], 大学生是其中重要的活跃人群。鉴于此,本研究旨在评估大学生对 B 站的整体满意度,并进一步验证及提升中文版 UMUX 量表和 WAMMI 量表在中国文化环境应用的有效性。

## 2. 研究方法

### 2.1. 研究对象

本次研究以大学生 B 站用户为调查对象,采取方便取样方法,共进行了三次样本的发放,剔除无效及规律性作答的,最后有效样本依次为:第一次 1011 份有效问卷,有效回收率 98.73%,其中男性 432 名,女性 579 名,用于两个量表的有效性检验,我们又在一个月后对其中 80 名相同被试进行了两个量表

的重复测量。第二次有效问卷仍然是 1011 份, 用于两个量表的用户体验现状调查。第三次我们选择 41 名大学生被试, 男性 22 名, 女性 19 名, 目的在于通过定性访谈的方式, 深入了解他们对于 B 站的整体印象以及不满意之处。

本研究采用 SPSS 26.0 进行描述性统计分析和探索性因素分析, 使用 Amos 24.0 软件对数据做验证性分析。另外, 利用 NVivo 11.0 软件对访谈资料进行整理和分析, 提取文本内容中的关键词频。

## 2.2. 研究工具

### 2.2.1. 用户体验的可用性量表(Usability Metric for User Experience, UMUX)

UMUX 量表被看作是 SUS 量表的一个变式。最初它包含 12 个问题, 后来经过修订, 缩减为一个包含四个问题的单维量表。这个量表包括综合体验(第 1 题)、满意度(第 2 题)、有用性(第 3 题)和高效性(第 4 题)四个部分, 其中后三个方面用来测量 ISO 可用性定义[17]。在本研究中, 我们使用这个四题版的 UMUX 量表, 采用七点计分, 分数从 1 (非常不同意)到 7 (非常同意)。量表中包含正向和反向的描述, 转化分值范围在 0~6。UMUX 总分通过将四个问题的得分相加后除以 24, 再乘以 100, 计算得到与 SUS 量表相同的 0 到 100 的分数区间。分数越高, 表明用户的可用性体验越佳[17]。

### 2.2.2. 网站分析和测量调查表(Website Analysis and Measurement Inventory, WAMMI)

爱尔兰科克大学的人因研究组(HFRG)于 20 世纪 90 年代早期研发了针对评估软件应用的软件可用性测试调查表(Software Usability Measurement Inventory, SUMI) [29], 随后该研究组又与斯德哥尔摩的 Nomos management AB 合作开发出来 WAMMI 量表用于网站评估。WAMMI 已经被用来对世界范围内的上百个网站进行了评估。当前版本的 WAMMI 由 20 个项目构成, 采用 5 点 Likert 评分, 包括肯定与否定陈述, 从 1 (强烈同意)到 5 (强烈不同意), 一共包含了五个维度: 吸引力(Attractiveness)、可控性(Controllability)、效率(Efficiency)、帮助性(Helpfulness)、易学性(Learnability) [20]。标准化的整体 WAMMI 分数的平均值是 50, 标准差是 10,100, 是最高值。WAMMI 量表的总分越低, 代表着 B 站用户体验越好。

## 2.3. 效标选择

系统可用性量表(System Usability Scale, SUS)由 John Brooke 创建[15], 目的是对系统或产品的感知可用性进行评估。它目前是应用最为广泛的用户体验测量工具之一。2020 年, Wang 及其同事对中文版 SUS 量表进行了修订, 并验证了其效度的可靠性[16]。在本研究中, 我们采用了这个中文版的 SUS 量表作为效标工具, 与另外两个量表一同发放。量表由十个采用李克特 5 点计分法的问题组成, 评分范围从 1 (代表“非常不同意”)到 5 (代表“非常同意”)。其中, 偶数题项是反向描述问题, 每题的得分转换范围为 0~4。正向描述题目的转换分值为原始分数减 1, 而反向描述题目的转换分值为 5 减去原始分数。所有题目的转换分值总和乘以 2.5 后, 即可得到 SUS 量表的总分, 其分数范围在 0~100 之间[16]。在本项研究中, SUS 量表的 Cronbach's  $\alpha$  系数为 0.870。

## 2.4. 翻译与回译

为确保翻译质量, 我们采取了严格的双重校对流程。首先, 我们邀请两位语言学教授和一位具有英语教育背景的心理教授共同翻译了 UMUX 量表和 WAMMI 量表的初稿。随后, 一位心理学教授对初稿进行了细致的审阅和归纳, 并根据实际应用场景对题项表述进行了调整。为了确保翻译的准确性, 我们随后邀请了两位英语专业的教授对这两个量表进行回译。在此过程中, 我们还特别邀请了一位心理学专家对回译稿与原稿之间的一致性进行了细致的比较和研究。最后, 为了确保量表翻译的适应性和实用性, 我们选取了 10 名 B 站用户进行调查, 并根据用户的反馈对存在问题的题项进行再次修订, 从而形成

中文版的 UMUX 量表和 WAMMI 量表的最终版本。

### 3. 数据分析及结果

#### 3.1. 中文版 UMUX 量表和 WAMMI 量表的适用性检验

##### 3.1.1. UMUX 量表研究结果与分析

###### 1) 项目分析

通过对 1011 名被试 UMUX 量表得分进行相关分析, 结果显示, 各题项与总分的相关系数在 0.679~0.752 之间, 均在 0.01 水平上达到显著。本研究以总分为依据, 对研究对象进行了排名, 将分数最高的个体(前 27%)分为高分组, 将分数最低的个体(后 27%)分为低分组, 并对高分组与低分组样本在 4 个题项的得分进行独立样本  $t$  检验。结果表明, 两组样本在每个题项的得分均差异显著( $p < 0.001$ )。这说明 UMUX 量表所有题项具有较好的区分度。

###### 2) 信度检验

问卷的信度分析主要是对问卷测量的可靠性与稳定性的考量, 其主要目的是检验测量所得结果的内部一致性程度。UMUX 量表的  $\alpha$  系数为 0.746, 重测信度为 0.792, UMUX 量表是单维度, 合成信度为 0.850, 说明 UMUX 量表在中国具有一定的可行性。

###### 3) 效度检验

效度分析用来评判测量结果的正确程度和准确程度。探究 UMUX 量表的四个因子之间及与总量表的相关分析结果发现, 四个因子与总量表之间的相关系数较大(0.679~0.752), 说明各因子都围绕一个共同特质; 四个因子之间存在较低程度的相关(0.216~0.538), 显示每个因素有各自独立的作用。这些说明 UMUX 量表具有良好的结构效度。

以 SUS 量表作为效标, 本研究测得 UMUX 量表得分与该效标的得分呈现显著的正相关( $r = 0.723, p < 0.01$ ), 说明该量表效标效度良好。

##### 3.1.2. WAMMI 量表研究结果与分析

###### 1) 项目分析

同理 UMUX 量表, 对 WAMMI 量表的各题目与总分做相关分析, 其中第 5 题“B 站需要更多的介绍性说明”题总相关系数为 0.147, 未达到 0.400 的标准, 鉴别能力较低, 予以删除。其余题目与总分的相关系数都在 0.600 以上。接着将研究对象按总分高低排序, 进行高低分组(前后各 27%), 对两组在余下的 19 个题项上的得分进行独立样本  $t$  检验, 结果显示存在显著差异( $p < 0.001$ )。这说明 WAMMI 量表的各个题项具有较好的区分度。

###### 2) 探索性因素

本研究将被试随机分成两部分: 一部分( $n = 504$ )用于探索性因素分析, 另外一部分( $n = 507$ )接受验证性因素分析。

本研究使用 KMO 和 Bartlett 球形检验样本的充足度和进行因素分析的适宜度。结果表明,  $KMO = 0.904$ , Bartlett 球形检验结果达到显著性水平( $\chi^2 = 3560.221, df = 190, p < 0.001$ ), 这说明样本数据适合进行探索性因素分析。

基于上述结果, 本研究使用主成分分析法(Principal Components)提取各因子, 并使用最大方差法(Varimax)进行旋转。在因子选择上, 删除因子下题项只有一个或两个的因子。探索性因素分析结果显示, 有两个因子包含题项在三个以下。因此, 再次删除 3 个题项(第 10 题、第 12 题、第 16 题), 保留余下 16 个题项, 并对剩下的 16 个题项进行第二次探索性因素分析, 提取因子方法如前。16 个题项的公因子方

差值范围为 0.451~0.732，最终抽取出四个因子(图 1)，四个因子之间成分相关矩阵的相关系数范围为 0.416~0.831，累积解释总变异的 61.58% (表 1)。对新因子进行重新命名，依次为：1) 可用性：指 B 站能否被正常使用，以及使用 B 站所需要的时间和努力程度。2) 易用性：指 B 站能够被多数用户在最短的时间内熟练使用的程度。3) 满意度：由用户对 B 站使用需求与期望所产生，是用户使用产品中的主观感受，包括视觉、时间、流畅度等带来的感受。4) 易学性：用户认为他们掌握 B 站系统或学会使用 B 站新功能的速度和容易性。

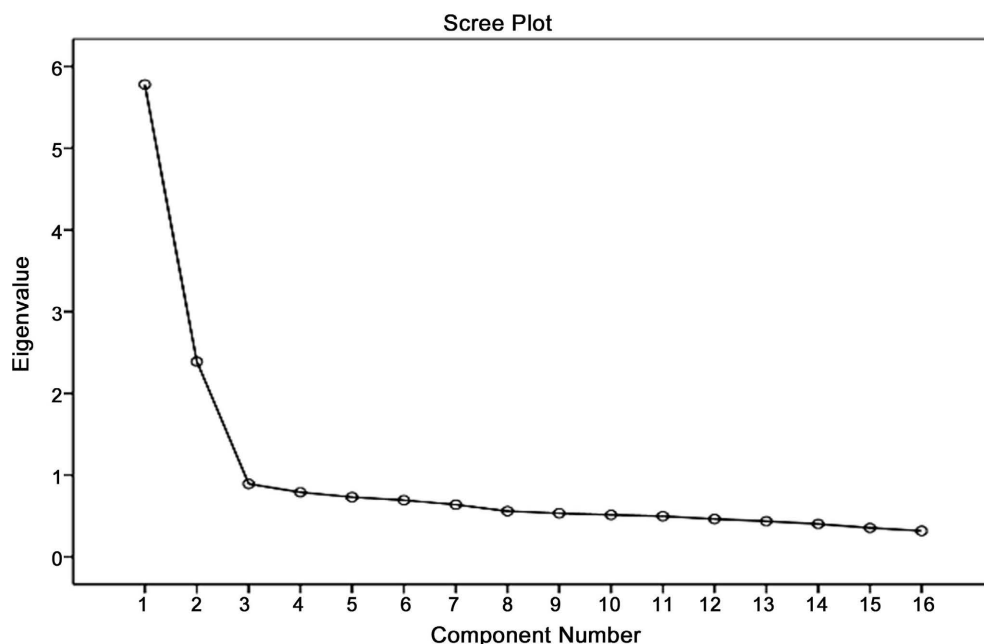
**Table 1.** Results of EFA of WAMMI scale (n = 504)

**表 1.** WAMMI 量表探索性因素分析结果(n = 504)

项目	共同度	因素负荷	特征值	贡献率(%)	累积贡献率(%)
因素 1	可用性		5.778	36.114	36.114
3	0.732	0.831	在 B 站上，我能快速找到我想要的东西		
4	0.657	0.747	在我看来，B 站是符合逻辑的		
9	0.624	0.744	B 站可以帮助我找到正在查找的内容		
13	0.577	0.698	当使用 B 站时，我感到效率高		
1	0.677	0.650	B 站有许多我感兴趣的内容		
因素 2	易用性		2.390	14.937	51.050
8	0.615	0.744	B 站太慢		
14	0.632	0.756	很难说 B 站是否有我想要的东西		
17	0.607	0.710	难以记住自己在 B 站上的位置		
18	0.612	0.691	使用 B 站就是浪费时间		
2	0.572	0.651	B 站浏览起来是困难的		
因素 3	满意度		0.895	5.594	56.645
7	0.658	0.67	当我使用 B 站时，我感觉可以掌控操作		
6	0.568	0.557	B 站的网页非常吸引人		
11	0.677	0.508	我不喜欢使用 B 站		
因素 4	易学性		0.790	4.935	61.580
20	0.701	0.790	B 站上的每件事情都容易理解		
15	0.451	0.469	第一次使用 B 站是容易的		
19	0.492	0.416	当点击 B 站时，能得到我期望的内容		

### 3) 验证性因素

根据探索性因素分析的结果，我们对另一部分样本数据(n = 507)进行了验证性因素分析，以检验 WAMMI 量表的四个因子的拟合程度。如表 2 所示，四个因子的拟合指数都达到了理想标准，说明模型对数据具有很好的拟合能力。此外，各项拟合指数都符合预期值要求，表明模型具有很高的可靠性。



**Figure 1.** Analysis scree plot of EFA of WAMMI scale

**图 1.** WAMMI 量表探索性因素分析碎石图

**Table 2.** Main fitness indexes of WAMMI scale

**表 2.** WAMMI 量表主要拟合指数

统计检验量	模型拟合指数					
	$\chi^2/df$	RMSEA 值	GFI 值	NFI 值	TLI (NNFI)值	CFI 值
适配标准值	<3	<0.08	>0.9	>0.9	>0.9	>0.9
模型值	3.011	0.066	0.928	0.907	0.918	0.933

#### 4) 信度检验

WAMMI 量表的内部一致性分析显示,其总体的 Cronbach's  $\alpha$  系数为 0.882,表明量表具有较高的可靠性。此外,各维度的  $\alpha$  系数介于 0.650 至 0.860 之间,进一步证实了其良好的内部一致性。重测信度达到 0.766,说明该量表在不同时间点测量的一致性也较为理想。

#### 5) 效度检验

分析 WAMMI 量表的四个因子及其与总量表之间的相关性显示,这四个因子与总量表的相关性较强(相关系数介于 0.738 至 0.800 之间),而因子之间的相关性则相对较弱(相关系数介于 0.319 至 0.501 之间),这表明 WAMMI 量表具备良好的结构效度。此外,WAMMI 量表与作为效标的 SUS 量表之间的相关系数达到 0.709 ( $p < 0.01$ ),这一结果说明两量表之间效标效度合乎标准。

### 3.2. UMUX 量表和 WAMMI 量表用户体验现状调查

#### 3.2.1. 量表人口学变量总体分布情况

由表 3 可见,本研究中男性 432 份(42.7%),女性 579 份(57.3%),男女大致平衡;使用时间小于 1 年的有 54 人,1 年到 3 年的有 381 人,3 年以上的有 576 人,大多数人使用时间较长;大部分人的使用频率都比较频繁,基本上每周都有使用。

**Table 3.** Distribution of demographic variables  
**表 3.** 人口学变量分布情况

		频数	百分比(%)
性别	男	432	42.7
	女	579	57.3
使用时间	小于 1 年	54	5.3
	1 年到 3 年	381	37.7
	3 年以上	576	57.0
使用频率	每天都使用	455	45.0
	每周都使用	429	42.4
	每月都使用	127	12.6

### 3.2.2. UMUX 量表的现状分析

UMUX 量表得分越高,代表用户体验越好。对不同性别、使用时间、使用频率的 B 站用户进行 UMUX 量表得分的多因素方差分析,结果发现,四个维度和量表总分均不存在交互作用。使用时间在综合体验、有用性以及量表总分上存在着主效应差异( $F_{\text{综合体验}(3, 1010)} = 21.132, p < 0.001$ ;  $F_{\text{有用性}(3, 1010)} = 20.687, p < 0.001$ ;  $F_{\text{总分}(3, 1010)} = 16.833, p < 0.001$ ),使用时间越长的用户体验感越好。使用频率表现出主效应,频率高的用户有更好的有用性体验( $F_{\text{有用性}(3, 1010)} = 14.420, p < 0.001$ )。

### 3.2.3. WAMMI 量表的现状分析

总体而言, WAMMI 量表的分数越低,意味着用户体验越佳。同样采用性别、使用时间和使用频率三个自变量进行多因素方差分析,结果显示这四个维度与量表总分上没有交互作用。使用时间对于可用性、易用性、用户满意度以及量表总分等方面具有显著的主效应影响( $F_{\text{可用性}(3, 1010)} = 17.550, p < 0.001$ ;  $F_{\text{易用性}(3, 1010)} = 13.120, p < 0.01$ ;  $F_{\text{满意度}(3, 1010)} = 31.040, p < 0.001$ ;  $F_{\text{总分}(3, 1010)} = 27.040, p < 0.001$ ),使用频率在满意度上存在主效应差异( $F_{(3, 1010)} = 25.116, p < 0.001$ )。使用时间越长,使用效率越频繁,用户的使用体验感越好,满意度越高。

## 3.3. 用户体验访谈结果分析

根据访谈结果可知,大学生用户对于 B 站用户体验的使用体验较好。根据词云图(图 2)发现,“使用”被提及的次数最多,其次“符合”、“功能”、“操作”、“容易”、“学习”、“吸引力”、“效率”、“帮助”等也被提及较多。



**Figure 2.** Word cloud map of interview results  
**图 2.** 访谈结果词云图



## 4. 讨论

### 4.1. 关于中文版 UMUX 量表和 WAMMI 量表的修订与检验

在中文版 UMUX 量表上,本研究通过对 1011 份问卷进行数据分析后发现:1) 项目结果显示,UMUX 量表的题总相关系数在 0.679~0.752 之间,题与题之间的相关系数在 0.216~0.655 之间,表明量表的每一个题项均指向了一个共同的主题,但题项之间本身的重合度不高,具有较好的区分度,反映了量表可从 B 站用户体验的不同方面进行评估。2) UMUX 量表的  $\alpha$  系数为 0.746,重测信度为 0.792,合成信度为 0.850,与 SUS 量表的效标效度为 0.723,说明量表在中文语境下具有一定的可施测性。

在中文版 WAMMI 量表上:1) 项目分析结果显示, WAMMI 量表各题项与总分的相关系数在 0.147~0.660 之间,均在 0.01 水平上达到显著;其中第 5 题与总分间的相关系数为 0.147,题目的鉴别能力较低,应当剔除。2) 数据高低分组后,两组样本在题项的得分均差异显著( $p < 0.001$ ),说明 WAMMI 量表的题项具有较好的区分度,可从不同维度对 B 站用户体验进行多方面的评估。3) 在进行探索性因素分析中发现,其中两个因子下分别只有两个题项,不符合每个因子下至少三个题项的要求,因此将这四个题项(第 5 题、第 10 题、第 12 题、第 16 题)予以剔除,然后进行验证性因素分析,发现各种拟合指数达到标准,说明剔除题项后的模型对数据拟合较好。4) WAMMI 量表总分  $\alpha$  系数为 0.882,每个维度的  $\alpha$  系数在 0.650~0.860 之间,重测信度达到 0.766。四个因子(可用性、易用性、满意度、易学性)与总量表之间的相关系数较大,说明各因子都围绕一个共同特质,四个因子之间存在较低程度的相关,每个因素有各自独立的作用,进而说明 WAMMI 量表具有良好的结构效度。与作为效标的 SUS 量表之间的相关系数达到 0.709 ( $p < 0.01$ ),效标效度也较理想。

总之,修订后的中文版 UMUX 量表和 WAMMI 量表具有良好的信效度指标,适用于评估大学生对 B 站的用户体验情况。

### 4.2. 两个量表的用户体验评价

#### 4.2.1. 总体用户体验评价

整体而言,大学生用户对于 B 站的用户体验较好,与前期访谈的结果一致,这可能与其用户本身的特殊性有密不可分的联系。B 站的用户大多为 Z 世代用户[28](Z 世代用户:指出生于 1990 年~2009 年的用户),这些用户大都是独生子女,在儿童成长时期缺乏兄弟姐妹的陪伴,所以他们希望能够得到认同,找到归属感,B 站的出现恰好为 Z 世代用户提供了他们所需要的认知需求和情感需求。除此之外,B 站会赋予用户广泛的自主权,包括发布弹幕、一键三连等,久而久之,用户在这种环境下与 B 站的连接会更加紧密,使用体验也会更加偏向积极[30]。

#### 4.2.2. 两个量表的现状分析

两个量表的结果均表现出,使用时间长和使用频率高的用户体验感更好,满意度也更高。首先,使用时间长和使用频率高,用户会与 B 站建立更深层次的联系,对 B 站分享的信息能进行更深层次的加工,产生强烈的参与感和社会存在感[31]。此外,B 站的用户只有通过高门槛的会员答题制度后才能成为 B 站的正式会员,这个行为本身就需要花费更多的时间与访问频率,而愿意这样做的用户就表明是认同或者喜欢 B 站的某些内容的,具有极高的忠诚度和归属感[30]。

但 B 站用户体验在不同性别上的差异不显著,赵阳佳紫的研究中发现性别对于 B 站跨年晚会的使用动机没有差异[32]。一方面是因为 B 站的分区十分丰富,包括电影、电视、动漫、资讯、运动、美食、汽车漫画、鬼畜等,不同的性别人群都能在 B 站里找到符合自己期望的内容,女生可以追星、刷刷、看吃播、学美妆,男生可以看汽车、动漫、运动等;另一方面是 B 站注重用户体验和用户反馈,通过持续

的优化网站界面布局、功能设置、推荐算法等来适应不同性别用户的需求，让用户感到使用起来很舒适和顺畅。

### 4.3. 用户体验访谈评价

在访谈过程中，我们发现“使用”这一词汇频繁地被提及，这无疑反映出 B 站用户在体验过程中的极高满意度。那些如“符合”、“功能”、“操作”、“容易”、“学习”、“吸引力”、“效率”、“帮助”等关键词，更是清晰地揭示了大学生用户群体对 B 站的普遍认可。他们认为 B 站不仅操作简便，使用起来轻松愉悦，而且能够高效满足他们的需求，提供极具吸引力的内容和功能，帮助他们更方便地获取信息和解决问题。

## 5. 结论

1) 修订后的 UMUX 量表的信效度较好，WAMMI 量表在删除第 5 题、第 10 题、第 12 题、第 16 题后，保留剩余的 16 个题项，具有较高的信效度。两个量表的指标符合测量学标准，在中国文化环境中具有一定的适用性。

2) 中国大学生用户的 B 站使用体验整体来说较好，在性别上不存在差异，但表现出使用时间越长、使用频率越高的用户体验更好。

## 基金项目

国家自然科学基金重点项目(No. 61936010); 重庆文理学院横向项目(No. WLHX-2020-0033)。

## 参考文献

- [1] 中国互联网络信息中心(CNNIC). 第 53 次中国互联网络发展状况统计报告[R/OL]. <https://www.cnnic.net.cn/NMediaFile/2024/0325/MAIN1711355296414FIQ9XKZV63.pdf>, 2024-08-05.
- [2] MoonFox. 2023 年 Q1 移动互联网行业数据研究报告[R/OL]. <https://moonfox.cn/insight/report/1281>, 2024-08-05.
- [3] 杨智宝, 袁小伟, 刘远. 用户体验在我国的发展概述[J]. 中国质量, 2020(6): 98-103.
- [4] 葛列众, 许为. 用户体验: 理论与实践[M]. 北京: 中国人民大学出版社, 2020.
- [5] 高蕾, 王洪阁. 基于眼动分析的触屏手机界面设计研究[J]. 艺术与设计(理论), 2014(5): 122-124.
- [6] 韦艳丽, 徐曦, 蒋小庆. 基于情境思维的权限获取页面设计方法研究[J]. 设计, 2022, 35(3): 36-39.
- [7] 雷尚仲, 张瑞秋. 智能电视的自然交互设计研究[J]. 沿海企业与科技, 2020(6): 27-30.
- [8] 经顺林, 陈红, 蓝雯静. 基于体验理论的智能洗衣机软件用户界面设计[J]. 中国高新科技, 2021, 85(1): 125-126.
- [9] 王萍. 文遗主题类 APP 的用户体验设计方法研究[J]. 包装工程, 2016, 37(8): 63-66.
- [10] 段欣园, 覃一崧, 刘巧, 等. 大学生二手书购买共享平台的建立及调查分析[J]. 科技与创新, 2022(4): 112-115.
- [11] 汪颖, 吕富强. 基于眼动数据的 ATM 机界面可用性测试[J]. 人类工效学, 2017, 23(1): 48-54.
- [12] 陈淑虹, 程春. 网易云音乐 APP 用户体验对用户忠诚的影响[J]. 北方经贸, 2021(12): 50-53.
- [13] Lee, Y.S., Smith-Jackson, T.L., Chair, P.D., et al. (2007) Older Adults' User Experiences with Mobile Phones: Identification of User Clusters and User Requirements. Virginia Polytechnic Institute and State University.
- [14] 丁群, 卢星. 基于用户体验的老年家庭康氧产品设计研究[J]. 工业设计, 2022(3): 149-151.
- [15] Brooke, J. (1996) SUS—A “Quick and Dirty” Usability Scale. In: Jordan, P., Thomas, B. and Weerdmeester, B., Eds., *Usability Evaluation in Industry*, Taylor & Francis, 4-9.
- [16] Wang, Y.H., Lei, T. and Liu, X.X. (2020) Chinese System Usability Scale: Translation, Revision, Psychological Measurement. *International Journal of Human-Computer Interaction*, 36, 953-963. <https://doi.org/10.1080/10447318.2019.1700644>
- [17] Finstad, K. (2011) The Usability Metric for User Experience. *Interacting with Computers*, 22, 323-327. <https://doi.org/10.1016/j.intcom.2010.04.004>

- [18] Lewis, J.R. (2013) Critical Review of “The Usability Metric for User Experience”. *Interacting with Computers*, **25**, 320-324. <https://doi.org/10.1093/iwc/iwt013>
- [19] Borsci, S., Federici, S., Bacci, S., Gnaldi, M. and Bartolucci, F. (2015) Assessing User Satisfaction in the Era of User Experience: Comparison of the SUS, UMUX and UMUX-LITE as a Function of Product Experience. *International Journal of Human-Computer Interaction*, **31**, 484-495. <https://doi.org/10.1080/10447318.2015.1064648>
- [20] Kirakowski, J. and Cierlik, B. (1998) Measuring the Usability of Web Sites. *Proceedings of the Human Factors & Ergonomics Society 42nd Annual Meeting*, Chicago, 5-9 October 1998, 424-428. <https://doi.org/10.1177/154193129804200405>
- [21] Tullis, T.S. and Albert, B. (2008) *Measuring the User Experience: Collecting, Analysis, and Presenting Usability Metrics*. Morgan Kaufmann.
- [22] Bilibili. B 站发布 2022 年 Q4 及全年财报[R/OL]. <https://www.bilibili.com/read/cv22153039>, 2024-08-05.
- [23] 胡佳漪. 用户体验对健身类 APP 用户付费课程购买意愿的影响研究[D]: [硕士学位论文]. 上海: 上海体育学院, 2021.
- [24] 李堂辉. “使用与满足”理论视角下弹幕视频网站用户分析——以 Bilibili 为例[J]. 戏剧之家, 2019(24): 210-211.
- [25] 黄炜, 沈欣芸, 李伟卿. 视频博客 Vlog 用户的持续使用行为影响因素研究——以 Bilibili 为例[J]. 现代情报, 2022, 42(6): 69-79.
- [26] 赵瑜. 叙事与沉浸: Bilibili “互动短视频”的交互类型与用户体验[J]. 西南民族大学学报(人文社会科学版), 2021, 42(2): 129-134.
- [27] 秦紫函. 用户视角下“B 站”的运营策略浅析[J]. 新媒体研究, 2019, 5(20): 48-49.
- [28] Bilibili. Bilibili 用户画像分析[R/OL]. <https://www.bilibili.com/read/cv18496411/>, 2024-08-05.
- [29] Kirakowski, J. (1996) The Software Usability Measurement Inventory: Background and Usage. In: Jordan, P. and Weerdmeester, B., Eds., *Usability Evaluation in Industry*, Taylor & Francis, 169-178.
- [30] 聂丽豆. 哔哩哔哩网站生活类短视频主流化趋向发展研究[D]: [硕士学位论文]. 呼和浩特: 内蒙古大学, 2022.
- [31] Ruiz-Mafe, C., Bigne-Alcaiz, E., Sanz-Blas, S., et al. (2018) Does Social Climate Influence Positive eWOM? A Study of Heavy-Users of Online Communities. *BRQ Business Research Quarterly*, **21**, 26-38. <https://doi.org/10.1016/j.brq.2017.12.001>
- [32] 赵阳佳紫. 哔哩哔哩弹幕网站跨年晚会传播效果研究[D]: [硕士学位论文]. 南昌: 江西财经大学, 2021.