Journalism and Communications 新闻传播科学, 2026, 14(9), 101-109
Published Online September 2026 in Hans. https://www.hanspub.org/journal/jc
https://doi.org/10.12677/jc.2026.149241

文化产品网络评论的跨平台与多场景特征比较

——以电影《哪吒之魔童闹海》为例

张嘉耀

上海工程技术大学管理学院，上海

收稿日期：2026年8月4日；录用日期：2026年9月1日；发布日期：2026年9月11日

摘　要

为比较文化产品网络评论在不同平台与讨论场景中的分布特征，本文以电影《哪吒之魔童闹海》为例，对微博、知乎和Bilibili样本进行三分类情感分布及高频词比较。研究使用两组用途不同的数据：跨平台样本共4657条，包括微博310条、知乎3178条和Bilibili 1169条；Bilibili场景样本来自三个主题视频，共4482条。随稿材料仅保留正向、中性和负向标签及汇总结果，未保留原分类模型、训练语料和验证记录，因此本文将标签作为既有编码开展描述性统计，不推断模型性能。结果显示，微博、知乎和Bilibili样本的中性评论占比分别为54.8%、63.4%和62.8%，正向评论占比分别为34.8%、28.0%和31.7%，负向评论占比分别为10.3%、8.6%和5.6%。在三个Bilibili视频样本中，“隐喻分析”视频的正向评论占81.7%，而“好不好看”和“电影解说”视频的负向评论分别占56.7%和61.7%。研究表明，平台语境、视频主题、受众选择与观察窗口可能共同影响样本中的情感分布；在样本非随机且标签生成过程不可复核的条件下，相关差异应解释为探索性的样本特征。

关键词

文化产品，网络评论，情感分析，跨平台比较，多场景比较

Cross-Platform and Multi-Scenario Comparison of Online Reviews for a Cultural Product

—A Case Study of “Ne Zha 2”

Jiayao Zhang

School of Management, Shanghai University of Engineering Science, Shanghai

Received: August 4, 2026; accepted: September 1, 2026; published: September 11, 2026

文章引用: 张嘉耀. 文化产品网络评论的跨平台与多场景特征比较[J]. 新闻传播科学, 2026, 14(9): 101-109.
DOI: 10.12677/jc.2026.149241

Abstract

This study compares the distributional characteristics of online reviews for a cultural product across platforms and discussion scenarios, using the film "Ne Zha 2" as a case. It examines three-class sentiment distributions and high-frequency terms in samples from Weibo, Zhihu, and Bilibili. Two datasets serve distinct purposes: a cross-platform sample of 4657 comments (310 from Weibo, 3178 from Zhihu, and 1169 from Bilibili) and a within-Bilibili sample of 4482 comments drawn from three topic-oriented videos. The available research materials retain only positive, neutral, and negative labels and aggregate results; the original classifier, training corpus, and validation records were not archived. The labels are therefore treated as pre-existing codes for descriptive statistics, and no model-performance claim is made. Neutral comments account for 54.8% of the Weibo sample, 63.4% of the Zhihu sample, and 62.8% of the Bilibili sample. The corresponding positive shares are 34.8%, 28.0%, and 31.7%, while the negative shares are 10.3%, 8.6%, and 5.6%. In the three-video comparison, positive comments account for 81.7% of the metaphor-analysis sample, whereas negative comments account for 56.7% of the "Is it good?" sample and 61.7% of the film-commentary sample. The results suggest that platform context, video topic, audience selection, and observation window may jointly shape the observed distributions. Given non-random sampling and an untraceable labeling pipeline, the differences should be interpreted as exploratory sample characteristics.

Keywords

Cultural Products, Online Reviews, Sentiment Analysis, Cross-Platform Comparison, Multi-Scenario Comparison

1. 绪论

1.1. 研究背景与问题提出

电影、动漫、游戏等文化产品的网络传播具有可见的社会影响和从众效应。实验研究表明，社会影响信息会改变文化产品的受欢迎程度与在线评价，使结果呈现更高的不均衡性和不确定性[1] [2]。因此，网络评论既是观众表达体验的文本，也是观察文化产品接受过程的重要资料。

不同评论载体的文本长度、词汇、互动方式和评价对象并不相同。针对电影评论的研究发现，用户评论、论坛讨论、博客与专业影评在情感表达和内容覆盖上存在差异[3]，这提示跨平台比较不能把所有评论视为同质文本。

现有跨平台研究进一步表明，同一内容在不同评论平台上的参与规模、表达长度、主题和情绪可能不同[4]-[6]。平台的生产、分发和使用规则也会影响哪些内容更易被看见和回应[7]。因此，平台差异既可能反映社区机制，也可能来自采样对象、时间窗口和内容主题。

《哪吒之魔童闹海》上映后，相关讨论同时出现在微博、知乎和 Bilibili。三者分别偏向开放式社交传播、知识问答和视频社区互动，为比较同一文化产品在不同平台及不同视频主题下的评论特征提供了案例。

本文采用“网络评论”和“公众讨论”描述研究对象，避免把评论预设为需要治理的对象。文中的

“文化产品”仅指可被生产、传播和消费的影视文化内容，不包含规范性价值立场；“评价”和“文化意义”等术语也只用于描述观众表达。

基于现有材料，本文聚焦两个层面：一是不同平台样本的三类情感分布和高频词差异；二是 Bilibili 不同主题视频的评论结构差异。研究不进行平台总体推断、因果识别或纵向变化检验。

1.2. 研究问题与分析边界

1.2.1. 研究问题

围绕现有样本和能够复核的汇总结果，本文提出以下三个研究问题：

第一，微博、知乎和 Bilibili 三个平台样本的正向、中性与负向评论分布有何差异？

第二，Bilibili 三个不同主题视频样本的情感结构有何差异，这些差异与视频主题和观察窗口可能存在何种联系？

第三，在样本非随机、观察窗口重叠且情感标签生成过程不可复核的条件下，上述结果能够支持哪些结论，又受到哪些解释边界的约束？

1.2.2. 研究目标与分析思路

本文的直接目标是统一统计口径，描述不同平台和不同视频样本的情感结构，并以高频词辅助说明评论语境；解释目标是在不超出现有证据的前提下，讨论平台语境、视频主题和观察窗口与样本分布之间的可能联系。

相应地，本文采用探索性案例研究而非因果识别设计。所有比例均以对应分析单元的有效评论数为分母；相关表述使用“可能”“样本中显示”等限定语，不将描述性差异直接解释为平台用户属性或普遍规律。

1.3. 数据与方法

本文设置跨平台比较和 Bilibili 视频场景比较两个分析层次。前者观察微博、知乎和 Bilibili 样本之间的情感分布与高频词，后者比较三个不同主题视频样本的评论结构。

1.3.1. 数据来源与样本构成

跨平台比较样本包括微博 310 条、知乎 3178 条和 Bilibili 1169 条，共 4657 条。Bilibili 场景比较样本来自三个视频：“隐喻分析”1169 条、“好不好看”1068 条和“电影解说”2245 条，共 4482 条。两组数据服务于不同分析目的；其中 1169 条 Bilibili 评论可能同时进入两组，因此不将两组样本数相加为去重后的独立样本量。

据现有研究记录，原始评论由 Python 程序通过平台公开页面或可用接口采集，并经过重复项与无关内容剔除、中文分词和停用词过滤。随稿材料未保留完整采样关键词、采集字段、去重键、缺失值处理规则及视频链接，本文因此只报告能够核对的有效评论数与观察窗口，并不呈现用户名等可识别信息。

1.3.2. 情感标签、统计方法与质量控制

情感分析可采用词典规则、传统机器学习或深度学习等不同路线[8]-[10]。原数据处理将每条评论编码为正向、中性或负向，但随稿材料仅保留三分类标签及汇总图表，未归档分类算法或情感词典名称、软件版本、训练语料、标签阈值、人工标注规范和验证集。因此，本文无法确认标签的具体算法来源，也不能如实报告准确率、精确率、召回率或 F1 值；这些指标需要带真值标签的验证集，不能由类别数量反推[11]。为避免虚构模型信息，本文不生成重新训练或验证分类器，而将现有标签作为既有编码，仅用于样本内描述性比较。

图 1 至图 3 的高频词图来自既有预处理结果。现有记录能够确认其经过中文分词、停用词过滤和词频统计，但未保留分词词典与停用词表的版本。因此，高频词仅用作评论语境的辅助材料，不据此建立主题模型或显著性结论。

跨平台比较以各平台有效评论数为分母，视频场景比较以各视频有效评论数为分母，报告频数和百分比。本文不开展显著性检验、因果推断或传播网络分析，所有结果均限定于所选样本及其观察窗口。

1.4. 研究贡献与局限

本文的主要贡献是同时呈现跨平台样本和 Bilibili 视频场景样本，并对各组统计分母进行复核，使平台差异与内容场景差异能够在同一案例中得到比较。

在方法表述上，本文把情感标签、高频词统计与推断结论的功能边界分开，明确说明模型信息和原始数据档案的缺失，不以无法验证的方法指标支撑结论。

本文仍存在三方面局限：样本来自三个平台和少数视频，不能代表平台全体用户；原始采样记录、情感分类器及验证集不完整，分类误差无法评估；三个视频的主题、受众和时间窗口均不一致，组间差异不能解释为同一批评论随时间发生的变化。

2. 文献综述与分析框架

2.1. 文化产品网络评论与跨平台比较

文化产品的在线接受并非只由作品属性决定。人工文化市场实验和在线评分实验显示，先前可见的流行度或评价会改变后续选择与评分[1] [2]。电影评论研究还发现，不同评论类型在长度、词汇和讨论侧重点上存在差异[3]。这些发现说明，评论数据既反映观众体验，也受到信息环境和表达载体的影响。

跨平台研究强调比较对象与统计口径的可比性。针对相同新闻或事件的研究发现，不同平台在评论数量、长度、主题与情绪表达上存在差异[4] [5]；面向消费者评论的多平台研究也表明，平台之间的主题特征和情感倾向并不完全一致[6]。因此，平台比较需要同时控制内容对象、采样时间和分析分母。

网络媒体逻辑研究指出，平台在内容生产、信息分发和使用方式上具有不同规则[7]。微博的公开转发结构、知乎的问题组织方式和 Bilibili 的视频–评论关系，为用户提供了不同的表达入口；但描述性差异不能被直接归因于技术机制，仍需考虑受众自选择与样本构成。

Bilibili 评论和弹幕通常具有短文本、网络用语和语境依赖等特征。相关研究常结合分词、高频词、词云和情感分析处理这类文本[12]，但讽刺、反问和混合评价会增加三分类标签的误差风险，因而需要报告分类器来源与人工复核结果。

综上，现有研究为文化产品评论的跨平台比较提供了理论和方法基础，但也提示三个约束：平台样本未必同质，视频主题可能选择不同受众，自动情感标签必须有可追溯的质量控制。本文据此采用“平台–场景–情感分布”的描述性框架。

2.2. 情感分析与方法透明度

情感分析旨在识别文本中的主观评价与情感极性，常见路线包括情感词典、监督机器学习和深度学习[8]-[10]。无论采用哪种路线，研究都应说明算法或词典来源、软件与模型版本、类别规则、训练或适配数据、阈值设置及人工标注流程。

对于正向、中性和负向三分类任务，准确率只能反映总体正确比例，精确率、召回率和 F1 值则用于考察不同类别的识别质量[11]。在类别分布不均衡时，仅报告准确率可能掩盖少数类别的误差。本文因缺少验证集而不补造性能指标，并把这一缺失纳入结果解释。

3. 跨平台样本的评论特征

3.1. 三平台情感分布与高频词比较

3.1.1. 微博样本

微博样本共 310 条，其中中性评论 170 条，占 54.8%；正向评论 108 条，占 34.8%；负向评论 32 条，占 10.3%。正向评论约为负向评论的 3.4 倍。该样本整体以中性和正向表达为主，但由于样本量较小且采样规则未完整归档，结果不能用于推断微博全体用户的态度。具体分布及高频词见图 1。

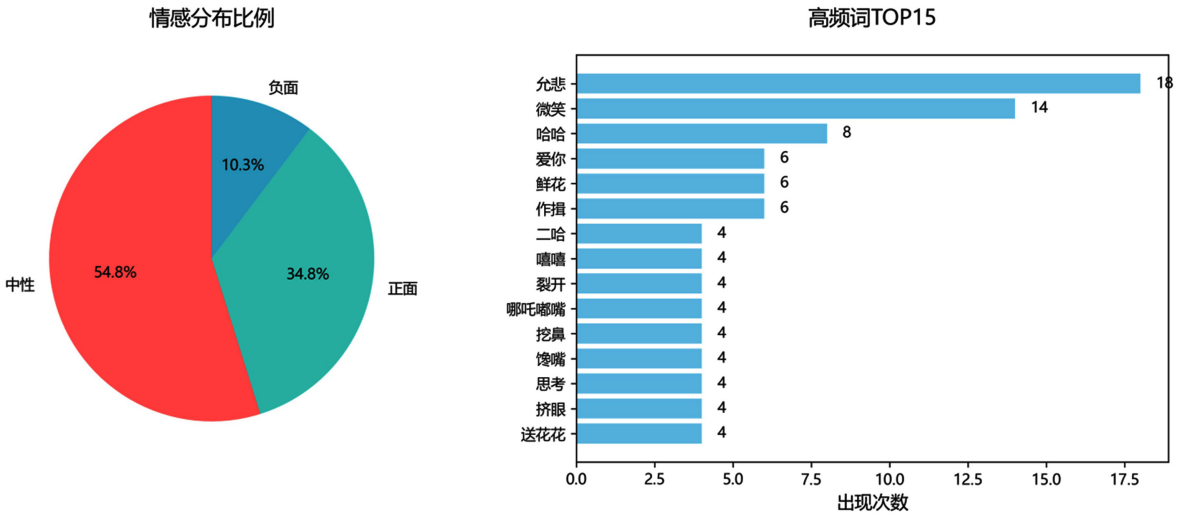

Figure 1. Sentiment distribution and high-frequency words in the Weibo sample
图 1. 微博样本情感分布与高频词

3.1.2. Bilibili 样本

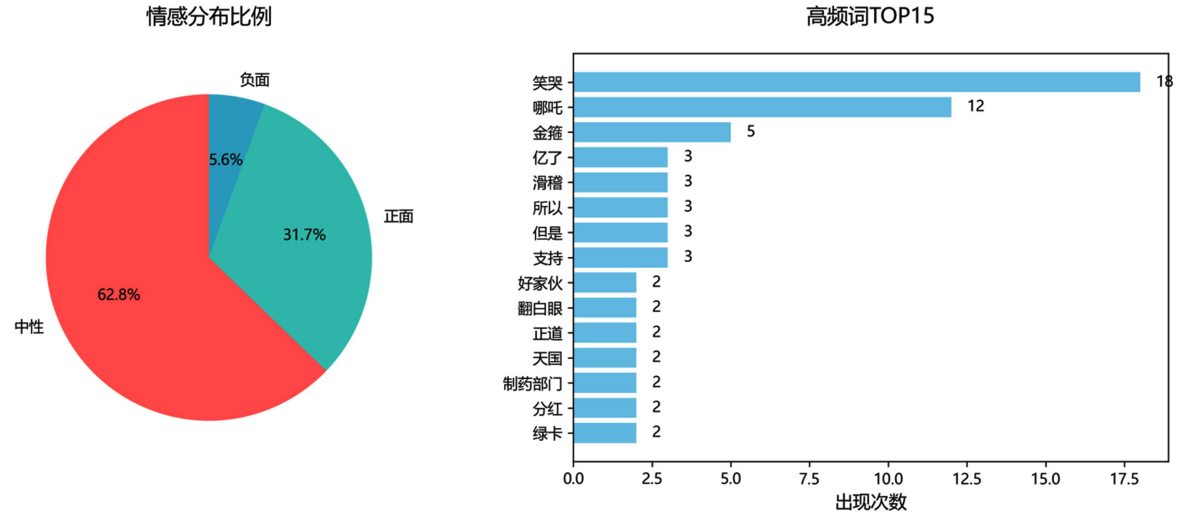

Figure 2. Sentiment distribution and high-frequency words in the Bilibili sample
图 2. Bilibili 样本情感分布与高频词

Bilibili 跨平台比较样本共 1169 条，其中中性评论 734 条，占 62.8%；正向评论 370 条，占 31.7%；

负向评论 65 条，占 5.6%。正向评论约为负向评论的 5.7 倍，中性评论占比最高。该分布说明所选视频评论中直接表达明确褒贬的文本相对较少，但中性标签也可能包含混合评价或情感强度较弱的表达。具体分布及高频词见图 2。图中的“正道”为样本原始词项，其含义依赖具体语境；本文仅报告其频次，不把该词解释为规范性价值立场。

3.1.3. 知乎样本

知乎样本共 3178 条，其中中性评论 2015 条，占 63.4%；正向评论 889 条，占 28.0%；负向评论 274 条，占 8.6%。正向评论约为负向评论的 3.2 倍。与另外两个平台样本相比，知乎样本的中性比例最高、正向比例最低。该差异可能与评论长度、问题设置和采样对象有关，不能仅凭情感比例推断平台用户特征。具体分布及高频词见图 3。

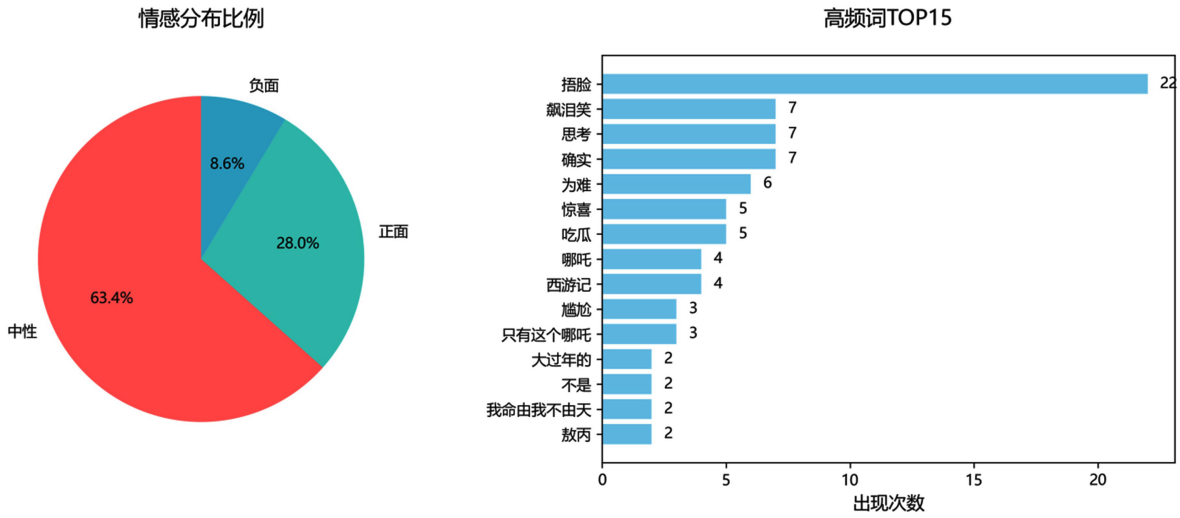

Figure 3. Sentiment distribution and high-frequency words in the Zhihu sample
图 3. 知乎样本情感分布与高频词

4. Bilibili 视频样本的多场景比较

4.1. 场景界定与比较口径

三个视频样本分别围绕“隐喻分析”“好不好看”和“电影解说”展开，评论窗口分别为 2025 年 2 月 2 日至 2 月 5 日、2025 年 9 月 6 日至 12 月 15 日以及 2025 年 1 月 31 日至 12 月 15 日。时间范围来自现有整理记录，仍需由原始评论时间戳复核。

三组样本的主题、视频发布者、受众选择和评论窗口均不相同，且部分时间区间相互重叠。因此，它们不构成同一评论池的连续追踪，也不能被排列为作品讨论随时间变化的三个阶段。

本文据此采用多场景横向比较：把每个视频视为一个独立讨论场景，以该视频有效评论数作为分母，比较正向、中性和负向评论的比例，并在解释中同时考虑视频主题与采样窗口。

4.2. 三类视频样本的情感分布

Bilibili 场景比较样本共 4482 条，来自三个视频：“隐喻分析”1169 条、“好不好看”1068 条和“电影解说”2245 条。为保证组间口径一致，表 1 中的占比均以各视频自身的有效评论数为分母，而不是以 2245 条或全部 4482 条为共同分母，结果见表 1。

“隐喻分析”视频样本中，正向评论 955 条，占 81.7%；中性评论 115 条，占 9.8%；负向评论 99 条，占 8.5%。该组以正向表达为主，但结果只能说明所选视频及其评论样本的特征，不能代表影片总体评价。

“好不好看”视频样本中，负向评论 606 条，占 56.7%；中性评论 260 条，占 24.3%；正向评论 202 条，占 18.9%。相较于“隐喻分析”样本，该组负向比例更高，可能与视频问题设置直接要求总体评价有关。

“电影解说”视频样本中，负向评论 1386 条，占 61.7%；中性评论 507 条，占 22.6%；正向评论 352 条，占 15.7%。该组覆盖时间较长，且内容对象与前两组不同，因此较高的负向比例既可能反映评论争议，也可能受到视频立场、受众选择和评论语境的影响。

Table 1. Sentiment distribution across three Bilibili video samples
表 1. Bilibili 三个视频样本的情感分布

视频内容	情感分类	评论数	组内占比(%)	主要情感
哪吒 2 隐喻分析	中性	115	9.8	正向
哪吒 2 隐喻分析	正向	955	81.7	正向
哪吒 2 隐喻分析	负向	99	8.5	正向
哪吒 2 好不好看	中性	260	24.3	负向
哪吒 2 好不好看	正向	202	18.9	负向
哪吒 2 好不好看	负向	606	56.7	负向
哪吒 2 电影解说	中性	507	22.6	负向
哪吒 2 电影解说	正向	352	15.7	负向
哪吒 2 电影解说	负向	1386	61.7	负向

注：各组占比以该视频有效评论数为分母；时间范围依据现有整理记录，仍需由原始数据时间戳复核。

4.3. 分组比较结果与解释边界

三组样本呈现出明显不同的情感结构：“隐喻分析”样本以正向评论为主，另外两组以负向评论为主。这一结果支持“不同视频主题和讨论场景对应不同情感分布”的描述性判断，但不能证明评论必然从正向转为负向。由于未获得完整高频词清单、原始评论文本和可复核的分类流程，本节不进行语义变化或因果机制推断。

5. 结论与展望

5.1. 主要研究结论

第一，三平台样本均以中性评论为主，但比例存在差异。微博、Bilibili 和知乎样本的中性评论分别占 54.8%、62.8%和 63.4%；正向评论分别占 34.8%、31.7%和 28.0%；负向评论分别占 10.3%、5.6%和 8.6%。这些结果描述所选样本的情感结构，不构成对平台全体用户的总体推断。

第二，Bilibili 三个视频样本之间的情感分布差异较大。按各组自身分母计算，“隐喻分析”样本的正向评论占 81.7%，“好不好看”和“电影解说”样本的负向评论分别占 56.7%和 61.7%。视频主题、观察窗口和受众选择可能共同造成差异。

第三，现有证据只支持横向场景比较。三个视频并非同一评论队列的连续观察，时间窗口又相互重叠，因而组间差异不能解释为评论态度随时间发生的阶段性变化。

5.2. 理论与实践启示

第一，跨平台评论研究应同时报告样本来源、内容对象、采样窗口和统计分母。

平台间比例差异可能来自社区生态，也可能来自视频主题和采样方式。只有在口径一致的条件下，跨平台比较才具有较好的解释力；实践中宜把不同平台视为互补的观察窗口，而不是用单一平台代表全部公众评价。

第二，内容场景应成为情感分析的重要分层变量。

不同视频样本的情感结构差异提示，研究不应只报告总体正负比例，还应按视频主题、内容类型和时间窗口分层展示，以避免总体平均值掩盖局部差异。

第三，方法透明度是结果可解释性的前提。

用于自动分类时，应保存模型或词典名称、版本、训练与验证数据、类别阈值和混淆矩阵，并报告各类别的精确率、召回率与 F1 值；否则，自动标签更适合作为待复核的辅助编码。

5.3. 研究不足与未来展望

第一，后续研究应扩大平台与视频覆盖，记录完整的采样关键词、时间、视频链接、采集字段、去重键和缺失值处理过程，并明确不同分析样本是否重合。

第二，应重新取得原始评论与分类程序，说明情感标签的生成方法，设置人工标注验证集并报告准确率、精确率、召回率、F1 值及混淆矩阵，以检验结论对分类误差的敏感性。

第三，可围绕同一评论池设置互不重叠的时间窗口，连续计算评论数量、情感比例和高频词，并结合影片上映、媒体报道和争议节点开展时间序列分析，从而区分时间变化与样本差异。

第四，未来若进行平台机制或视频主题的因果解释，应采用匹配样本、实验设计或准实验方法，并对平台自选择、视频立场和受众构成进行控制。

总体而言，本文提供了一个经过统计口径校正的探索性案例，并明确现有证据能够与不能够支持的结论。后续研究若能补齐原始数据、模型记录和可比的时间设计，可进一步检验文化产品网络评论的跨平台与多场景差异。

参考文献

[1] Salganik, M.J., Dodds, P.S. and Watts, D.J. (2006) Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, **311**, 854-856. https://doi.org/10.1126/science.1121066

[2] Muchnik, L., Aral, S. and Taylor, S.J. (2013) Social Influence Bias: A Randomized Experiment. *Science*, **341**, 647-651. https://doi.org/10.1126/science.1240466

[3] Na, J.C., Thet, T.T. and Khoo, C.S.G. (2010) Comparing Sentiment Expression in Movie Reviews from Four Online Genres. *Online Information Review*, **34**, 317-338. https://doi.org/10.1108/14684521011037016

[4] Ben-David, A. and Soffer, O. (2019) User Comments across Platforms and Journalistic Genres. *Information, Communication & Society*, **22**, 1810-1829. https://doi.org/10.1080/1369118x.2018.1468919

[5] Waterloo, S.F., Baumgartner, S.E., Peter, J. and Valkenburg, P.M. (2018) Norms of Online Expressions of Emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. *New Media & Society*, **20**, 1813-1831. https://doi.org/10.1177/1461444817707349

[6] 周婷玮. 基于共现网络与情感分析的多平台消费者评论主题比较研究[J]. 知识管理论坛(网络版), 2023, 8(2): 79-91.

[7] Klinger, U. and Svensson, J. (2015) The Emergence of Network Media Logic in Political Communication: A Theoretical Approach. *New Media & Society*, **17**, 1241-1257. https://doi.org/10.1177/1461444814522952

[8] Pang, B. and Lee, L. (2008) Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, **2**, 1-135. https://doi.org/10.1561/1500000011

[9] Medhat, W., Hassan, A. and Korashy, H. (2014) Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams*

Engineering Journal, **5**, 1093-1113. https://doi.org/10.1016/j.asej.2014.04.011

[10] Rodríguez-Ibánez, M., Casánez-Ventura, A., Castejón-Mateos, F. and Cuenca-Jiménez, P. (2023) A Review on Sentiment Analysis from Social Media Platforms. *Expert Systems with Applications*, **223**, Article 119862. https://doi.org/10.1016/j.eswa.2023.119862

[11] Sokolova, M. and Lapalme, G. (2009) A Systematic Analysis of Performance Measures for Classification Tasks. *Information Processing & Management*, **45**, 427-437. https://doi.org/10.1016/j.ipm.2009.03.002

[12] 刘宇婷, 杨燕. 弹幕文本挖掘与情感分析[J]. 人工智能与机器人研究, 2023, 12(4): 361-372.