

社会化媒体的用户产品评价影响因素研究： 基于文本挖掘的方法

司格 张伦 张增一

摘要

社会化媒体区别于传统媒体的重要特点为用户自创内容（UGC，User-Generated Content），这一特点也为研究在线电子商务的信息传播行为注入了新的活力。本研究基于“使用与满足理论”，探究社会化媒体中用户产品评价的影响因素。本文选择“在线电子商务网站”这一形式的社会化媒体，将产品评分作为用户产品评价的量度，并对产品评论进行文本挖掘。研究发现，用户对产品的评价取决于用户对产品的使用经验、产品对用户需求的满足、以及在线社交影响（即其他用户的影响）。产品已售时间的长短、用户发表评论时间（如上午、下午等）以及用户所在地区也对用户评价有着显著影响。本研究还发现，上述影响因素的作用在不同产品类型中表现不同，即对于“搜索型产品”而言，其他用户的评论对用户产品评价无影响；而对于“体验型产品”而言，社交因素则显著影响用户产品评价。

关键词

社会化媒体、用户创建内容、使用与满足理论

作者简介

司格，中国科学院大学人文学院新闻与传播学系在读硕士研究生，研究方向为在线社交媒体的文本挖掘与行为研究。电邮：sige12@mailsucas.ac.cn，邮编：100049。

张伦（通讯作者），中国科学院大学人文学院新闻与传播学系讲师，研究方向为社会网络、用户在线行为研究。电邮：zhanglun@ucas.ac.cn。

张增一，中国科学院大学人文学院新闻与传播学系教授，研究方向为科学传播。电邮：zhzy@ucas.ac.cn。

Factors Influencing Online Products Rating on Social Media: Employing Text Mining Approach

SI Ge, ZHANG Lun, ZHANG Zengyi

Abstract

User-Generated Content (UGC) is an important characteristic of social media, which advances the development of study of online information processing. Based on Uses and Gratifications Theory, this study investigates the factors that influence customers' evaluation of products on social media. We extracted data from an online e-commerce platform (i.e., JD.com). This study found that user's evaluation of product is determined by their using experiences of the product, user satisfaction and online social influence (i.e., other users' comments on the product). In addition, the time period between comment and launch date of the product, the comment time (i.e., morning, afternoon, and evening), and user's geographical location also contribute to their evaluation of the product. Moreover, the relationship between these factors and the users' evaluation is conditional to the product type. The relationship between the factors of the users' evaluation is significant for the search goods while insignificant for the experience goods.

Keywords

social media, user-generated content, uses and gratifications theory

Author

Si Ge is a graduate student at the Dept. of Journalism & Communication, University of Chinese Academy of Sciences. Email: sigel2@mailsucas.ac.cn.

Zhang Lun is an assistant professor at the Dept. of Journalism & Communication, University of Chinese Academy of Sciences. Email: zhanglun@ucas.ac.cn.

Zhang Zengyi is a professor at the Dept. of Journalism & Communication, University of Chinese Academy of Sciences. Email: zhzy@ucas.ac.cn.

一、引言

集成了社会化媒体 (social media) 社会交往功能的电子商务网站一定程度上改变了消费者获取产品信息的渠道和产品购买行为, 打破了传统媒介时代消费者获取产品信息信源单一、信息匮乏的桎梏。社会化媒体的一个重要特点为用户创建内容 (user generated content, UGC), 电子商务网站用户在使用产品后能够创建在线评价。用户在浏览产品页面信息时, 同时能看到其他用户创建的评分和评论, 即不同用户创建的同一产品的信息可以在用户之间传播。Tuten与Solomon (2013:4)将社会化媒体分为社会化社区、社会化出版、社会化商务、社会化娱乐四类。其中社会化商务平台是指包含用户评分和评论的电子商务网站, 比如亚马逊、淘宝等。换言之, 电子商务网站是社会化媒体的形式之一。

用户在做出购买决策前, 除了依靠传统的大众媒体广告 (例如电视、报纸、户

外广告等)(Nelson, 1974), 用户在线创建的内容也能够补充产品的信息(Weathers et al., 2007)。从这个意义上说, 电子商务网站中其他用户创建的产品评价在较大程度上丰富了大众媒体广告信息, 从而成为用户做出购买决策的重要信息来源。

然而, 在线用户创建的产品评价是否客观反映了产品的性能? 本研究以“电子商务网站”为例, 以使用与满足理论为理论框架, 探讨用户产品评价的主要影响因素, 即产品使用时间、需求满足、社交影响, 以及用户自身特征。

二、用户在线消费行为研究

(一) 电子商务平台中用户产品评价的基本规律与特点

既往研究通过对用户产品评价进行的描述性研究发现, 用户产品评价展现出一定的规律性。第一, 用户偏向于发布极端正面的评价(Hu, Pavlou, & Zhang, 2006; Wu & Huberman, 2010), 即用户更喜欢给自己使用过的产品(例如看过的电影和书籍)更高的评价。例如, Hu et al. (2006)发现, 书、VCD、DVD的用户评分分布呈两极分化分布, 即高分和低分出现的频率较高, 特别是高分, 而中间分数段频率低。第二, 产品评分均值随时间呈递减趋势(Wu & Huberman, 2010)。最初的产品评价往往由该产品粉丝发布, 中后期的用户趋向于对产品做出理性客观的评价。第三, 用户意见的异质性反映出用户的“情感倾向混杂度”(郝媛媛, 2010)。如果某产品的用户评价具有较高的异质性(即评分方差较高), 则说明该产品是一个“非主流产品”。用户评价的异质性与产品的受欢迎程度相关。当且仅当评分的均值较低时, 较高的方差可能会伴随着较高的需求度以及较高的销量(Sun, 2012; Martin, Barron, and Norton, 2007)。

(二) 电子商务网站产品类型

产品按照用户使用体验的差异, 可以分为“搜索型产品”(search goods)和“体验型产品”(experience goods)(Nelson, 1974)。“搜索型产品”, 即用户主要通过信息搜索而不需要太多体验就能做出购买决策的产品; “体验型产品”, 即用户不能仅凭信息搜索也需要一定的使用经验才能判断其质量并做出购买决策的产品(Klein, 1998; Nelson, 1974)。例如, 手机、衣服、家具等属于“搜索型产品”; 旅游度假产品、餐饮类服务则属于“体验型产品”(Zeithaml, 1988)。有研究指出, 产品类型会调节(moderate)产品评价与用户购买决策之间的关系(Sundaram & Webster, 1999; Bone, 1995), 相比于“搜索型产品”的用户, “体验型产品”的用户更需要依靠其他用户的使用经验来做出购买决策(Park & Lee, 2009)。

三、理论框架与研究假设

使用与满足理论认为用户是有着特定“需求”的个人,用户通过“使用”某产品或技术,使其某些需求得到“满足”。Katz, Blumler和Gurevitch (1974: 20) 对使用与满足理论进行了经典论述:“基于社会影响和个人心理的需求,激发了用户对媒介和其他来源的期望,这一期望导致了不同的媒介接触,从而造成了用户需求的满足以及其他后果。”该论述涉及到使用与满足理论的七个核心概念——“社会和心理起源”、“需求”、“媒介和其他来源”、“期望”、“媒介接触”、“满足”和“其他后果”。

经典的使用与满足理论研究往往只关注媒介采纳,特别是不同类型的媒介满足了用户的哪些社交和心理需求(Chua, Goh, & Lee, 2012), 并常与动机理论密切结合(Chen, 2011)。本研究拟将使用与满足理论扩展到产品采纳领域。在互联网时代的电商平台中,产品采纳涉及的要素包括产品的使用经验、产品对用户需求的满足以及用户评价。用户在“产品接触”后会创建“评价”,该“评价”即基于需求满足而产生的“其他后果”。已有研究表明,用户评价与产品对用户需求的满足呈正相关关系(Wiebe, Bruce, & O'Hara, 1999)。本研究认为,“用户评价”这一概念的加入扩展了使用与满足理论的适用范畴(generalizability)。

本研究基于使用与满足理论框架,旨在探讨用户对某种产品的使用经验(using experience)如何影响了用户评价。此外,产品对用户功能需求的满足、在线用户之间的“社交影响”也是解释用户产品评价的重要影响因素,因此,本研究加入了上述两个因素,旨在扩展使用与满足理论的解释范畴(explanatory power)。具体来说,本研究以电子商务网站为平台,探讨用户产品评价的主要影响因素,即产品使用时间、用户的需求满足、社交影响,以及用户自身特征(如图1所示)。

(一) 产品使用时间

使用与满足理论认为,“使用经验”为影响产品评价的一个重要维度。“人类的时间是一项资源”(Dimmick, 2002: 31)。在面临多种媒介选择时,用户使用何种媒介与他们使用该种媒介的时长有关(Dimmick, 2002; Lo & Leung, 2009)。事实上,产品之间的竞争也是对用户时间的争夺。与媒介选择类似,用户对产品的评价也可能与其使用该产品的时间有关。此外,有研究发现使用习惯会影响媒介采纳(Rubin, 1984),用户使用某类型媒介的时间越长,由于习惯性,用户会对该类型媒介的评价越高。Li与Hitt(2008)发现用户使用产品的时间越长,越不愿意对其做出负面评价。基于此,提出以下假设:

假设1: 用户使用产品的时间正向影响了产品评价。

(二) 需求满足

研究发现, 用户对产品的评价取决于该产品对用户需求的满足程度。使用与满足理论认为, 用户对媒介类型的选择建立在自身需求以及需求被满足程度的基础之上。用户在使用产品的过程中, 对产品的各项性能做出判断, 这些判断通过用户自创内容反映出来, 表达了产品对用户需求的满足程度。用户在其自创内容中正面提及产品的特征, 说明用户的功能需求得到了满足, 用户对产品的评价就可能很高; 而用户在其创建内容中没有提及特征或者提及负面特征时, 说明用户的功能并未得到满足, 甚至非常不满足, 用户对产品的评价就可能较差。基于此, 提出以下假设:

假设2: 产品对用户需求的满足程度正向影响了用户评价。

(三) 社交影响

用户对产品的评价还可能受到电子商务平台中其他用户的影响。为检验社交影响, 本研究引入四个变量: 其他用户对产品的评分, 其他用户评论的可读性, 其他用户评论的情感倾向, 其他用户对产品特征的提及。

产品评分是用户对产品最直接的评价。电子商务网站, 以及影评、书评网站均将用户评分置于醒目位置, 且这些网站通常按用户评价倒序排列产品, 使得用户通过评分能够迅速对产品、电影和书籍等产生基本印象。评分较低的对象, 用户可能直接拒绝购买、观赏或阅读。研究发现, 用户对产品的评价受到在线社交关系的影响, 即存在“选举者困境(voters' paradox)”(Garman & Kamien, 1968)。具体而言, 用户倾向于做出与他人不同、甚至相反的评价(Wu & Huberman, 2010; Godes & Silva, 2006)。用户为了吸引他人注意, 会最大程度上改变既有评价(Dichter, 1966; Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004; Sundaram & Webster, 1999)。因而, 在前期用户的评分较高时, 后期用户很可能倾向于给出较低的评价。基于此, 本研究提出以下假设:

假设3: 其他用户的产品评分负向影响了用户对其所购买产品的评价。

Petty与Cacioppo (1986)指出, 在一定条件下, 用户对质量较高的信息会采纳积极的信息处理模式(central route)。具体而言, 信息论点越是强有力, 用户的态度越倾向于发生改变(Metzger, 2007)。高质量的评论具有更强的说服力, 因为其包含的产品信息是相关的、可靠的、具体的、容易理解的, 而低质量的用户评论则相反(Lee, Park, & Han, 2008)。可读性是信息质量的一个重要衡量标准, 指文本被用户理

解并留有印象的容易程度(Hu, Bose, Koh & Liu, 2012; Lee et al., 2008)。前期用户评论的可读性较高时,便于后期用户的理解,并对后期用户的产品评价产生影响。基于此,本研究提出以下假设:

假设4:其他用户评论的可读性正向影响了用户对产品的评价。

评论质量的另一个方面为情感倾向,即文本表现出的情感极性(Hu et al., 2012; Lee et al., 2008)。有研究表明,人的主观态度与评论中使用的形容词有显著的正相关关系(Wiebe, 2000; Wiebe, Bruce, & O'Hara, 1999)。因而,评论的情感倾向反映了用户对于产品体验优劣的认知。前期用户的负面评论比正面评论更容易引起用户的注意,从而对该用户创建的评价产生影响(Chevalier & Mayzlin, 2006; Dichter, 1966; Hennig Thureau et al., 2004)。Wu and Huberman (2010)则指出用户更倾向于表达出与之前用户相反的观点,当前期用户的情感倾向多为正面时,后期用户可能倾向于发表负面评价。基于此,提出以下假设:

假设5:其他用户评论的情感倾向负向影响了用户产品评价。

在文本分析中,其他用户评论所提及的产品特征也是影响当前用户产品评价的一个重要的变量(Hu & Liu, 2004; Shi & Chang, 2006)。Hu and Liu (2004)指出,用户描述产品特征有两种方式,一是直接提及的显性特征(例如“做工优良”);二是间接提及的隐性特征(例如“容易放进包中”是形容尺寸,但未直接提及)。

用户在评论中提及某一特征,说明该特征给用户留下了深刻印象,用户提及的特征既可能是与预期一致的令人满意的特征,也可能是意外的令人满意(或不满意)的特征。这些用户创建内容,区别于产品广告,丰富了产品信息,也体现了产品对于用户功能需求的满足程度,且可以为后期用户判断产品质量提供参考(Hu & Liu, 2004)。如果用户并未提及特征,那么该用户的评论就可能不会对后期用户产生影响。

另外,用户对产品特征的描述存在方向性(正/负),但本研究并未选择“用户正面(或负面)描述产品特征”作为影响用户评价的自变量之一,这主要是基于以下两点原因。一是不少用户对特征的描述是中性的,或者是信息性的,判断其情感倾向的意义不大;二是某些特征描述的正负倾向是相对的,前期用户给出的负面描述,对于后期用户而言可能是正面的。比如前期用户指出“手机屏幕尺寸小,阅读体验差”,这样的描述偏向负面,但是某后期用户则认为这样的尺寸合适,可以单手操作,即,前期用户对功能的负面描述,对后期用户而言是正面的。综上所述,相对于“用户正面描述产品特征”,本研究认为“是否提及产品特征”作为影

响用户评价的自变量之一更有实际意义。基于此,提出以下假设:

假设6: 其他用户评论是否提及产品特征正向影响了产品评价。

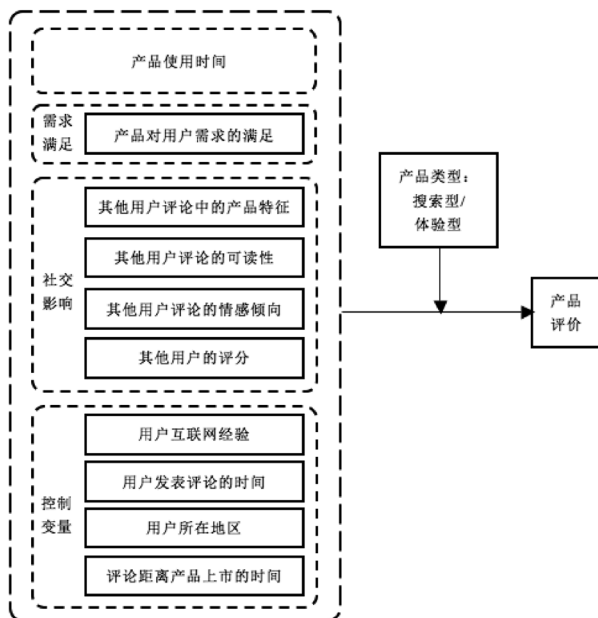


图1 理论框架

三、研究方法 with 测量

(一) 数据

根据第三方市场研究公司艾瑞咨询的数据,京东商城(JD.com)是中国最大的自营式电商企业。¹2010年,京东跃升为中国首家规模超过百亿的网络零售企业。²因而,京东商城能够在一定程度上反映出电子商务网站的普遍规律。本研究采用的数据即为2010年1月从京东商城(JD.com)上抓取的iPhone4手机(以下简称“iPhone4”)和“诺基亚(NOKIA)BL-5C原装锂电手机电池”(以下简称“手机电池”)的用户评论数据。iPhone4的用户评论数目为4452条,手机电池的评论数目为5424条,数据包括用户名(考虑到隐私,已将用户名做了隐私处理)、用户级别、用户所在城市、用户产品评分(5分、4分、3分、2分、1分)、对产品优点和不足的评价、使用心得、评论日期和购买日期。

该数据有三大特点。一是数据量大,信息丰富,评论和评分起始于产品上市;二是包含具体的评论日期和购买日期,精确记录了用户的购买和评论行为;三是评论中包含“优点”、“不足”和“使用心得”三部分,研究者能够对这些内容进行

文本语义层面的挖掘与分析。

iPhone4在网络上信息详尽,用户通过信息搜索可以对其质量有着充分的估计,因而,本研究将iPhone4定义为“搜索型产品”。手机电池品牌多样,质量参差不齐,用户仅凭手机电池的基本信息很难获知产品质量,并做出购买决策。用户需要在获得他人使用经验之后才能判断某一种手机电池的质量。因而,本研究将手机电池定义为“体验型产品”。

(二) 测量

“用户的产品评价”利用产品评分来测量。在本研究使用的数据中,评分的等级划分为5级,5分为最高分,1分为最低分。其中,1、2分属于用户对产品“差评”,用户的产品评价最低;3分表明用户对产品给予“中评”,即用户对产品基本满意,用户的产品评价中等;4、5分为表明用户对产品给予“好评”,即用户的产品评价最高。

“用户产品使用时间”利用数据集中用户“购买时间”和“评论时间”的差值来测量。

“评论中的产品特征”运用自然语言处理领域文本挖掘的手段,对用户评论进行分词和词性标注,当名词和形容词相连时,认为该名词为产品的特征(Hu & Liu, 2004)。本研究抽取500条评论,经过人工编码,发现计算机文本挖掘的正确率在65%左右。

“产品对用户需求的满足程度”利用“用户是否提及产品的特征”与“评论的情感倾向”的乘积来测量。若用户评论提及产品特征,则“用户是否提及产品的特征”值为1;若用户评论未提及产品特征,“用户是否提及产品的特征”值为0。若情感分数为正,情感倾向值为1;若情感分数为负,情感倾向值为-1。总的来说,当用户提及特征并给出正面评价时,需求满足的值为1,即产品满足用户需求的程度最高;当用户未提及特征或评价中性时,需求满足为0,即产品满足用户需求的程度中等;当用户提及特征并给出负面评价时,需求满足为-1,即产品满足用户需求的程度最低。

“可读性”是指文本被用户理解并留有印象的容易程度(Hu et al., 2012)。该文章认为评论长度对可读性有着重要的影响,并以字母数、单词数、句子数构成的“自动可读性指标”(Smith & Senter, 1967)来测量可读性。英文为字母文字,单词之间存在空格,能够清楚地算出字、词、句的数目;而就中文而言,词的分界不清晰,网络语言使用不规范,标点随意,难以判断词、句的数目。因而,本研究将评

论的长度（即字数）作为可读性的量度。

“情感倾向”是指文本的情感极性。Hu et al. (2012)以情感分数为量度。本研究利用ROST软件³直接计算每一条评论的情感分数。ROST软件能够对评论进行分词和词性标注，找出表达情感的词语，根据强弱情感词库，判断这些词语的正面、负面倾向及其强弱程度，计算出评论的情感分数。⁴该分数为正，表明用户情感倾向为正面；分数为负，则表明用户情感倾向为负面。分数的绝对值越大，表明用户情感越强烈。该软件目前下载量超过7000次，100多所院校使用该软件进行学术研究 (Shen et al., 2009a; Shen et al., 2009b; 袁红, 2010; 许梅华, 2010)。本研究邀请两位非本研究小组人员，随机抽取100条评论分别进行人工分析。将其结果与ROST软件的结果相比较后，发现ROST软件的正确率平均在70%以上。

此外，本研究将用户的互联网使用经验、用户所在地区、用户发布评论的时间、评论距离产品上市的时间作为控制变量。“用户互联网经验”利用会员级别⁵测量；为便于比较，本研究将“用户所在地区”划分为东部地区、中部地区和西部地区；“用户评论时间”则被划分为上午、下午、晚上和凌晨；“评论距离产品上市的时间”为评论时间与产品发售时间之差，以天为单位。

四、研究结果

（一）总体描述

表1为用户产品评价、所在地区和评论发布时间的分布。从中可以看出，大多数用户评价为最高。在iPhone4的4231条用户评论中，评价最高占总评论数的97.8%，评价中等和最低的评论数分别占0.9%和1.4%；在手机电池的5341条用户评论中，评价最高的评论比例高达98.6%，“中评”和“差评”的评论数分别占1.2%和0.2%。这与此前研究中的发现相符——用户偏向于发布极端正面的评价(Godes & Silva, 2006; Wu & Huberman, 2010)。

表1 用户评价、产品对用户需求满足程度、所在地区、评论时间分布				
	iPhone4		手机电池	
	N	%	N	%
用户评价				
最低	37	0.9%	66	1.2%
中等	58	1.4%	11	0.2%
最高	4136	97.8%	5264	98.6%
产品对用户需求满足程度				

最低	37	0.9%		157	2.9%
中等	3708	87.6%		3976	74.4%
最高	486	11.5%		1208	22.6%
用户所在地区					
东部	3704	87.5%		4626	86.6%
中部	210	5.0%		415	7.8%
西部	238	5.6%		300	5.6%
缺失值	79	1.9%			
评论时间					
凌晨	119	2.8%		139	2.6%
上午	1334	31.5%		1519	28.4%
下午	1548	36.6%		1931	36.2%
晚上	1230	29.1%		1752	32.8%
总计	4231	100.0%		5341	100.0%

图2为iPhone4的用户评分均值随时间变化的曲线，从中可以看出，产品上市初期，用户评分集中于5分，评分均值较高，随后较快降至4.75左右，再小幅上升，渐趋平稳，整体保持在较高水平上。图3为手机电池的用户评分均值随时间变化的曲线，与iPhone4不同的是，手机电池的评分均值初期波动很大，在几乎稳定于4.55分左右后，逐渐上升，小幅下降，最后趋向平稳。从图2可以看出，两类商品的评分均值均保持在较高水平。其次，相对于iPhone4来说，手机电池的评分均值波动较大，那么其评分的波动性则会更大；尤其是初期，结合用户评价分布情况，手机电池的用户评价更偏“U型”——“好评”、“差评”均高于“中评”比例。

图2 iPhone4评分均值分布

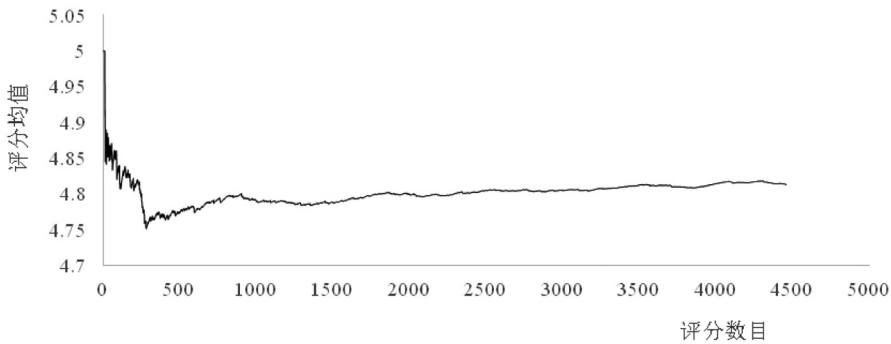


图3 手机电池评分均值分布

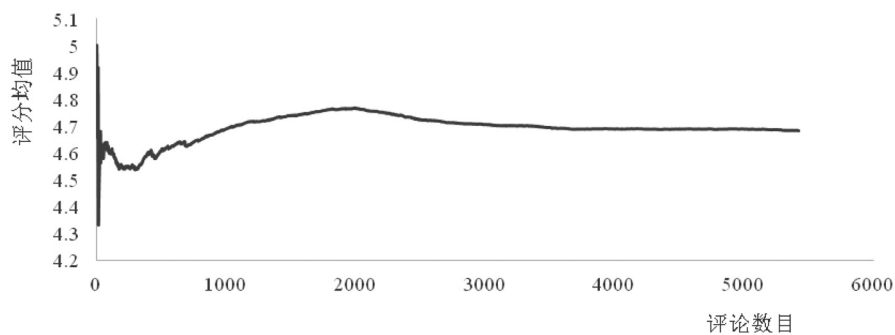


表1中“产品对用户需求满足程度”的分布显示，iPhone4和手机电池中“中等”程度的评论数所占比例分别为87.6%和74.4%，远高于“最低”和“最高”的评论数，说明用户给出的评价与实际的功能满足情况并不相符，用户给出高分并不等价于该产品满足了其功能需求。“用户所在地区”的分布显示，用户主要集中于东部地区，两种产品的东部用户比例均在85%以上。东部地区的经济发展水平较高，用户消费力强，互联网普及率高，且电商物流网点密集，在线购物方便快捷。两类产品的“用户评论时间”分布相似，凌晨较少，上午、下午、晚上的分布相差无几，各占三成左右。

（二）用户评价的影响因素

本研究采用逻辑回归模型（Logistic Regression Model）检验用户产品评价的影响因素（分析结果见表2）。“搜索型产品”（即iPhone4）和“体验型产品”（即手机电池）两个模型均显著（ $-2LL_{iPhone4} = 981.918, p < .01$ ； $-2LL_{手机电池} = 729.026, p < .001$ ），可以认为在这两个模型中，自变量有效解释了因变量“用户产品评价”。

表2 用户评价的多元逻辑回归结果							
	iPhone4				手机电池		
	B (SE)	Wald	OR		B (SE)	Wald	OR
用户产品评价为最低							
（常数）	-7.159 (11.529)	0.386			0.847 (5.295)	.026	
用户使用							
使用时间	-0.008 (0.009)	0.799	0.992		-0.007★ (0.003)	4.735	0.993
用户需求满足的程度							

最低	-16.718 (0.000)		0.000		2.550*** (0.573)	19.787	12.810
中等	1.526 (1.018)	2.248	4.601		1.066** (0.475)	5.052	2.905
最高	0	.	.		0	.	.
社交影响							
其他用户的评分	-0.072 (2.293)	0.001	1.075		-0.021 (0.024)	.753	0.979
其他用户评论中的产品特征	-0.229 (3.019)	0.006	0.796		-0.105** (0.037)	8.166	0.901
其他用户评论的情感倾向	-0.123 (0.071)	3.014	0.884		0.000 (0.001)	.049	1.000
其他用户评论的长度	0.040 (0.049)	0.643	1.040		0.001 (0.000)	3.302	1.001
控制变量							
用户互联网经验	-0.039 (0.036)	1.204	0.962		-0.334*** (0.100)	11.154	0.716
评论距离上市的时间	0.012* (0.005)	4.905	1.012		0.000 (0.002)	.001	1.000
用户所在地区							
东部	0.580 (1.022)	0.323	1.787		0.287 (0.600)	.228	1.332
中部	1.980 (1.089)	3.305	7.244		0.485 (0.700)	.481	1.625
西部	0		.		0	.	.
评论时间							
凌晨	-17.603 (6448.207)	0.000	0.000		-0.347 (1.039)	.111	0.707
上午	0.212 (0.413)	0.263	1.236		0.336 (0.328)	1.047	1.399
下午	-0.083 (0.432)	0.037	0.920		0.289 (0.317)	.832	1.336
晚上	0	.	.		0	.	.
表2 用户评价的多元逻辑回归结果（续）							
	iPhone4				手机电池		
	B (SE)	Wald	OR		B (SE)	Wald	OR
用户产品评价为中等							
（常数）	5.239 (9.112)	0.331			-99.473*** (24.007)	17.168	
用户使用							
使用时间	-0.022 (0.012)	3.496	0.979		0.009 (0.005)	2.644	1.009
用户需求满足的程度							

最低	-1.839 (0.896)	4.218	6.293		0.472 (0.000)	.	1.603
中等	0.541 (0.526)	1.058	1.717		17.985 (0.000)		64675292.009
最高	0	.	.		0	.	.
社交影响							
其他用户的评分	-1.117 (1.843)	0.368	0.327		0.160 (0.091)	3.113	1.174
其他用户评论中的产品特征	-1.314 (2.490)	0.279	0.269		-0.131 (0.105)	1.566	0.877
其他用户评论的情感倾向	-0.007 (0.054)	0.016	0.993		0.009 (0.008)	1.374	1.009
其他用户评论的长度	-0.084 (0.044)	3.695	0.919		-0.007 (0.008)	.849	0.993
控制变量							
用户互联网经验	-0.001 (0.024)	0.002	0.999		-0.421 (0.273)	2.372	0.657
评论距离上市的时间	-0.007 (0.005)	2.087	0.993		0.040*** (0.010)	15.059	1.041
用户所在地区		2.095	3.355				
东部	-0.122 (0.605)	0.041	0.885		16.949*** (0.952)	316.734	22942815.791
中部	1.123 (0.690)	2.652	3.074		18.334 (0.000)	.	91680524.270
西部	0	.	.		0	.	.
评论时间							
凌晨	-0.805 (1.027)	0.615	0.447		1.537 (1.376)	1.248	4.650
上午	-0.690* (0.335)	4.246	0.502		1.926* (0.952)	4.095	6.865
下午	-0.739* (0.321)	5.301	0.477		0.069 (0.974)	.005	1.071
晚上	0	.	.		0	.	.

表2 用户评价的多元逻辑回归结果（续）

	iPhone4				手机电池		
χ^2	57.25**				139.879***		
-2LL	978.918				729.026		
Nagelkerke R2 (%)	6.2				17.2		

*p<.05; **p<.01;*** p<.001 , SE为标准误, OR为发生比率。

1. “搜索型产品”用户评价影响因素分析

由表2可以看出，对于iPhone4模型而言，“用户发表评论的时间”为上、下午

时,评价“中等”相较于“最高”的发生比分别下降约50%和52% (Wald = 4.246, $p < .05$; Wald = 5.301, $p < .05$)。用户在上下午而非凌晨和晚上发表评价时,产品评价更高。“评论距离产品上市的时间”一个单位的增加,导致评价“最低”相较于“最高”的发生比上升1.2% (Wald = 4.905, $p < .05$) ;换言之,评论距离产品上市的时间负向影响了用户评价。总的来说,用户对于iPhone4这样的“搜索型产品”的评价不受需求满足程度、产品使用时间、社交影响、以及用户自身特征的影响。

2. “体验型产品”用户评价影响因素分析

从手机电池的回归模型中可以看到,“使用时间”一个单位的增加,与评价“最高”相比,“最低”的发生比下降0.7% (Wald = 4.735, $p < .05$) ,即用户使用产品时间正向影响了用户评价,假设1成立。“产品对用户需求的满足程度”为“最低”和“中等”而非“最高”时,与评价“最高”相比,“最低”的发生比分别上升11.8倍和1.9倍 (Wald = 19.787, $p < .001$; Wald = 5.052, $p < .01$) ,即产品对用户需求的满足程度正向影响了用户评价,假设2成立。“其他用户的评分”、“其他用户评论的情感倾向”和“其他用户评论的可读性”对用户评价都没有显著的影响。假设3、4、5均不成立。“其他用户评论中的产品特征”一个单位的增加,与评价“最高”相比,“最低”的发生比下降约10% (Wald = 8.165, $p < .01$) ,即提及产品特征的其他用户的评论正向影响了用户评价,假设6成立。

“用户互联网经验”一个单位的增加,导致评价“最低”相比于“最高”的发生比下降约28% (Wald = 11.154, $p < .001$) ,即,用户本身使用互联网的经验正向影响了用户评价。“评论距离上市的时间”一个单位的增加,导致与评价“最高”相比,“中等”的发生比上升约4% (Wald = 15.059, $p < .001$) ,即,评论距离上市的时间负向影响了用户评价。

“用户所在地区”为东部地区时,评价“中等”而非“最高”的发生比大幅上升 (Wald = 316.734, $p < .001$) ,东部地区用户给出的评价最可能是中等。“用户发表评论的时间”为上午时,与评价“最高”相比,“中等”的发生比上升5倍多 (Wald = 4.095, $p < .05$) ,用户在上午发表评价最可能是“中等”的。用户在上午时可能最为客观,下午、晚上更容易给出极端评论。

总的来说,手机电池比较能反映出“体验型产品”的特点。用户对产品的评价不仅取决于用户的使用经验、产品对用户需求的满足,还取决于其他在线用户的影响。同时,评论时间、用户本身的特点(互联网经验、地域、评论习惯)等变量也会对用户评价产生显著影响。

五、讨论

(一) 用户评价的影响因素

本研究以使用与满足理论为研究框架,探究在线电子商务平台中用户评价的影响因素。“搜索型产品”和“体验型产品”两类产品的评分均值随时间的变化趋势相似;两类产品的评价均多为极端正面评价;两类产品评论距离产品上市时间均负向影响评价。这些与Wu和Huberman (2010), Godes和Silva (2006)研究中的发现相符,具有普适性。初期用户通常是粉丝,或是产品经销商雇佣的“水军”,他们判断产品的标准具有强烈的主观色彩,后期用户则有更多时间“搜索”产品信息,产品评价也更为客观。

其他用户的评分没有对用户评价产生显著的影响,说明用户在进行产品评价时不注重其他用户的评分,而是更注重评论,这印证了其他研究中指出的用户不会“随大流”(Dichter, 1966; Engel et al., 1993; Sundaram et al., 1998; Hennig-Thurau et al., 2004)。我们认为,其他用户的评分无显著影响,主要有两个原因。一是早期评分的质量不高,不能反映产品的质量,具体来说,评分绝大部分是高分;有些用户虽然打了高分,但其评论是负面的。另一个原因可能在于,“其他用户的评分”这一变量的量度区间比较窄,变化范围较小,从统计分析的角度来看,易出现“不显著”的结果,未来的研究应注重对该变量测量方式的改进。

评论质量包括可读性和情感倾向(Hu et al., 2012),本研究发现,其他用户评论的可读性和情感倾向对用户评价没有显著影响,说明当前用户的评价并不取决于前期用户的评论质量。这与既往研究中“信息论点越是强有力,用户态度越倾向于发生改变”(Metzger, 2007)的发现相悖,在电子商务网站的语境下,其他用户评论质量对用户态度形成或改变的影响力降低。

更重要的是,本研究发现,其他用户评论中提及产品特征正向影响了用户评价。也就是说,电子商务网站用户在做出产品评价时存在社交影响。Web2.0时代,用户利用社交媒体结交朋友和维持友谊。“社交”并非电子商务平台的主要功能,但不同用户能够对同一产品做出评价,并彼此影响,从而产生虚拟的社交群体。用户浏览前期用户的评价后(例如前期用户评价中对于产品特征的提及频次等),其信息处理和决策过程均受影响,其评价又可能进一步影响其后出现的用户。该发现颠覆了人们对于电子商务网站中“产品评价客观性”的假设。用户对于产品的评价并非单纯客观地反映该产品的质量、品牌等特征,而是受到在线其他用户既存信息的影响。这种在线社交影响使得产品评论呈现出一定的时序性演化规律。

（二）对使用与满足理论的再思考

本研究进一步丰富和拓展了使用与满足理论。使用与满足理论作为传播学的经典理论，已经广泛应用于媒介的研究。而在新媒介环境中，该理论有着更大的发展空间。本研究将其置于电子商务网站这一特定的社会化媒体平台，不再仅仅关注媒介采纳，而立足于新媒体平台中的产品采纳。本研究发现，用户对于产品的评价不仅仅取决于该产品满足用户需求的程度，更取决于其他在线用户的影响，后者是对使用与满足理论的深入阐释，丰富了使用与满足理论的基本观点。

此外，本研究还丰富了使用与满足理论的阐释语境与范畴。本研究结果表明，使用与满足理论不仅仅适于探讨不同媒介使用的比较(LaRose, Mastro, & Eastin, 2001)，其对于解释互联网、手机等新技术的采纳时也有强大的生命力(Chua et al., 2012; LaRose et al., 2001; Stafford, Stafford, & Schkade, 2004)。本研究引入“产品类型”这一变量，将该理论应用于不同类型的产品采纳，出现了显著的对比效果，说明使用与满足理论也非常适于解释不同类型产品的采纳行为；这是对使用与满足理论在语境层面的拓展。

（三）文本挖掘对研究在线用户行为的方法论借鉴意义

本研究采用多元逻辑回归模型及相关的统计方法，辅以计算机手段进行文本分析。本研究利用文本挖掘方法，自动提取了9000多条评论中的情感倾向以及产品特征。本研究的文本分析主要分为两个方面，一是判断评论的情感倾向和强烈程度，二是判断评论是否提及特征。

以往的在线行为研究主要依靠便于处理的结构性数据，文本语义层面的挖掘通常主要依靠人工编码分析。在线行为研究涉及的数据量大，人工分析的成本较高，此外，分析结果受到主观因素的影响，不同分析人员对于同一文本材料可能给出差异较大的结果，因此，人工分析文本语义从效率和质量上都存在严重的问题。事实上，文本中包含着内容、表达方式、态度等丰富的信息，能够反映出人的情感、性格、教育程度、所处环境等特征，使研究者更清晰地描述、归纳和解读在线行为。利用自然语言处理的方法进行文本挖掘，能够突破小样本和人工分析误差的束缚，从文本中挖掘信息，这对于在线行为研究有着革命性的意义，也将会扩展传播学研究的方法。

（四）本研究不足之处以及未来研究方向

第一，本研究对于“可读性”这一变量的测量比较简单。可读性，即一段语料容易被读者理解的程度(Smith & Senter, 1967)。Smith与Senter(1967)利用“自

动可读性指标”测量可读性。这一指标涉及的变量包括一段语料中的句子数、词数和字母数,从英文的角度来看这一指标考虑了字、词、句三个维度,对于“可读性”的测量具有较好的信度和效度。然而,在中文中,词的界限不易区分,且电商网站中评论的长度有限,句子数的差异不大,因而本研究将评论长度,即字数作为“可读性”的量度,但仅囊括一个维度,测量难免粗糙。未来的研究可以对分词进一步优化,同时关注词语本身的含义,甚至标点符号,并加入词组、句子等更多维度作为“可读性”度量,这或许可以使得对“可读性”的测量更具效度,也可能发现影响用户产品评价的更多评论文本特征(例如关键词和标点特征)。

第二,本研究选取了“iPhone4”和“手机电池”作为“搜索型产品”和“体验型产品”的代表。研究发现在电子类产品中,“搜索型产品”和“体验型产品”的评价影响因素不同。该发现还可以在其他类别(例如母婴产品、图书音像产品等)中加以验证。未来的研究可选择更多的“搜索型产品”和“体验型产品”,从而一般性地验证产品类型对社交影响、产品使用等因素与产品评价之间的调节作用。

第三,本研究使用的数据截止时间为2010年。经过几年的发展,电子商务网站中用户评价的格式和展示形式发生了一定的变化(例如加入了图片上传、“点赞”等功能,且对优质评论有一定的奖励措施),这些变化都可能会影响用户对产品的评价。未来的研究可关注这些变化,从而探究这些因素是否影响用户评价。

第四,有网文称京东会删除差评⁶,因而,本研究使用的数据也可能并非是完整数据,甚至可能是遭到一定程度“清洗”的数据。如果数据不完整,可能会影响到研究的结果。未来的研究为避免此风险,可在数据抓取方面有更多的改进。为了将网站“删改评论”等行为的影响降至最低,研究者应当实时抓取数据,延长数据收集时间,使得抓取的数据更接近原始数据,对用户真实行为的描述也就更严谨。此外,以同一时间段内的评分和评论为研究对象,在不同时间点抓取,比如一天后、一周后、一个月后、一年后,那么或许能够发现网站删改评论的规律,从而更好地解释用户的在线行为。但是,对于本文所采纳的非实验研究方法,这个问题或许将永久存在。研究者无法要求研究情境达到理想化设定,即完全避免人为干预与操作。

第五,本研究利用京东会员等级来测量用户的互联网经验,这一测量方式有可能存在一定的误差。比如,有些用户拥有较丰富的互联网经验,特别是网购经验,但可能由于注册京东会员的时间较短,在京东的会员级别不高。未来的研究可以在

测量用户互联网经验上做更多的优化;或利用传统的“使用经验”测量方式与用户级别进行对比,以验证该操作化定义的恰当性。

最后,本研究使用的数据包含的用户自身特征较少,且由于抓取时间较早,彼时的电子商务网站用户主要为年轻人,年龄范围小;用户地区也集中于东部沿海发达地区;用户互联网经验也明显不如目前的用户丰富。总之,当前的电子商务网站用户背景显然更为多样化。受到数据集和时间的限制,本研究所能囊括的控制变量有限,未来的研究可以多关注用户年龄、性别、受教育程度等自身特征,通过严格控制这些变量,得出更普遍的用户评价与使用、社交影响、需求满足等之间的关系。

(责任编辑:杨雅)

注释 [Notes]

1. 京东商城官网企业简介<http://www.jd.com/intro/about.aspx>.
2. 央广网京东简介http://tech.cnr.cn/techzt/jdipo/jj/201404/t20140418_515322146.shtml.
3. ROST软件地址 <http://hi.baidu.com/ietynxalzdjoye/item/77fa3d392e5f8b6f7c034bae>.
4. ROSTCM6版本说明<http://hi.baidu.com/ietynxalzdjoye/item/77fa3d392e5f8b6f7c034bae>.
5. 京东会员级别介绍 <http://help.jd.com/help/question-57.html>.
6. 京东删除用户差评<http://tech.sina.com.cn/i/2010-06-28/11244358537.shtml>.

参考文献 [Reference]

- 许梅华 (2010). 基于共词分析的近年国内发展心理学研究热点分析.《现代情报》, 30 (8), 171-175.
- [Xu, Meihua(2010). Hot spots analysis of China' s developmental psychology based on co-words analysis method. *Journal of Modern Information*, 30(8), 171-175.]
- 袁红 (2010). 基于网络内容分析的高校门户网站可用性测评——以江苏省为例.《现代图书情报技术》, 10, 70-75.
- [Yuan, hong(2010). Web usability evaluation of the university portal based on web content analysis: case study of Jiangsu province. *New Technology of Library and Information Service*, 10, 70-75.]
- 郝媛媛 (2010).《在线评论对消费者感知与购买行为影响的实证研究》. 哈尔滨工业大学博士论文. 哈尔滨.
- [Hao, Yuanyuan (2010). *An empirical research about the impact of user online reviews on consumer perception and purchase behaviors*. Doctoral dissertation, Harbin Institute of Technology. Harbin.]
- Bone, Paula Fitzgerald. (1995). Word-of-mouth effects on short-term and long-term product

- judgments. *Journal of Business Research*, 32(3), 213-223.
- Chen, Gina Masullo. (2011). Tweet this: A uses and gratifications perspective on how active Twitter use gratifies a need to connect with others. *Computers in Human Behavior*, 27(2), 755-762.
- Chen, Yubo, & Xie, Jinhong. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management Science*, 54(3), 477-491.
- Chevalier, Judith A, & Mayzlin, Dina. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3), 345-354.
- Chua, Alton YK, Goh, Dion Hoe-Lian, & Lee, Chei Sian. (2012). Mobile content contribution and retrieval: An exploratory study using the uses and gratifications paradigm. *Information Processing & Management*, 48(1), 13-22.
- Dellarocas, Chrysanthos, & Narayan, Ritu. (2006). *What motivates consumers to review a product online? A study of the product-specific antecedents of online movie reviews*. Paper presented at the WISE.
- Dichter, Ernest. (1966). {How word-of-mouth advertising works}. *Harvard business review*, 44(6), 147-160.
- Dimmick, John W. (2002). *Media competition and coexistence: The theory of the niche*: Routledge.
- Garman, M. B., & Kamien, M. I. (1968). The paradox of voting: Probability calculations. *Behavioral Science*, 13(4), 306-316.
- Godes, David, & Silva, Jose. (2006). The dynamics of online opinion: working paper.
- Hennig-Thurau, Thorsten, Gwinner, Kevin P, Walsh, Gianfranco, & Gremler, Dwayne D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of interactive marketing*, 18(1), 38-52.
- Hu, Mingqing, & Liu, Bing. (2004). *Mining and summarizing customer reviews*. Paper presented at the Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining.
- Hu, Nan, Bose, Indranil, Koh, Noi Sian, & Liu, Ling. (2012). Manipulation of online reviews: An analysis of ratings, readability, and sentiments. *Decision Support Systems*, 52(3), 674-684.
- Hu, Nan, Pavlou, Paul A, & Zhang, Jennifer. (2006). *Can online reviews reveal a product's true quality?: empirical findings and analytical modeling of Online word-of-mouth communication*. Paper presented at the Proceedings of the 7th ACM conference on Electronic commerce.
- Katz, E., & Blumler, J. G. en M. Gurevitch (1974). 'Utilization of Mass Communication by the Individual.' . JG Blumler en E. Katz (red.), *The Uses of Mass Communications. Current Perspectives on Gratifications Research*, 19-32.
- Klein, Lisa R. (1998). Evaluating the potential of interactive media through a new lens: search versus experience goods. *Journal of business research*, 41(3), 195-203.
- Korfiatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and

- helpfulness of online product reviews: the interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications*, 11, 3, 205-217.
- LaRose, Robert, Mastro, Dana, & Eastin, Matthew S. (2001). Understanding Internet usage a social-cognitive approach to uses and gratifications. *Social Science Computer Review*, 19(4), 395-413.
- Lee, Jumin, Park, Do-Hyung, & Han, Ingoo. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications*, 7(3), 341-352.
- Li, Xinxin, & Hitt, Lorin M. (2008). Self-Selection and Information Role of Online Product Reviews. *Information Systems Research*, 19(4), 456-474.
- Lo, Olivine Wai-Yu, & Leung, Louis. (2009). Effects of gratification-opportunities and gratifications-obtained on preferences of instant messaging and e-mail among college students. *Telematics and Informatics*, 26(2), 156-166.
- Metzger, Miriam J. (2007). Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078-2091.
- Nelson, Phillip. (1974). Advertising as information. *The Journal of Political Economy*, 729-754.
- Park, Cheol, & Lee, Thae Min. (2009). Information direction, website reputation and eWOM effect: A moderating role of product type. *Journal of Business Research*, 62(1), 61-67.
- Petty, Richard E, & Cacioppo, John T. (1986). The elaboration likelihood model of persuasion. *Advances in experimental social psychology*, 19, 123-205.
- Rubin, A. M. (1984). Ritualized and instrumental television viewing. *Journal of Communication*, 34, 3, 67-77.
- Ruggiero, Thomas E. (2000). Uses and gratifications theory in the 21st century. *Mass communication & society*, 3(1), 3-37.
- Schlosser, Ann E. (2005). Posting versus lurking: Communicating in a multiple audience context. *Journal of Consumer Research*, 32(2), 260-265.
- Shen, Y., Li, S., Zheng, L., Ren, X., & Cheng, X. (2009a, August). Emotion mining research on micro-blog. In Web Society, 2009. SWS'09. 1st IEEE Symposium on (pp. 71-75). IEEE.
- Shen, Y., Liu, Z., Luo, S., Fu, H., & Li, Y. (2009b, September). Empirical research on e-government based on content mining. In Management of e-Commerce and e-Government, 2009. ICMECG'09. International Conference on (pp. 91-94). IEEE.
- Shi, Bin, & Chang, Kuiyu. (2006). *Mining chinese reviews*. Paper presented at the Data Mining Workshops, 2006. ICDM Workshops 2006. Sixth IEEE International Conference on.
- Stafford, Thomas F, Stafford, Marla Royne, & Schkade, Lawrence L. (2004). Determining uses and gratifications for the Internet. *Decision Sciences*, 35(2), 259-288.

- Sun, Monic. (2012). How Does the Variance of Product Ratings Matter? *Management Science*, 58(4), 696-707.
- Sundaram, DS, & Webster, Cynthia. (1999). The role of brand familiarity on the impact of word-of-mouth communication on brand evaluations. *Advances in Consumer Research*, 26, 664-670.
- Tuten, T.L. & Soloman, M.R. (2013). *Social media marketing*. Upper Saddle River, New Jersey: Pearson Education, Inc.
- Weathers, D., Sharma, S., & Wood, S. L. (2007). Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods. *Journal of Retailing*, 83(4), 393-401.
- Wiebe, Janyce. (2000). *Learning subjective adjectives from corpora*. Paper presented at the AAAI/IAAI.
- Wiebe, Janyce M, Bruce, Rebecca F, & O'Hara, Thomas P. (1999). *Development and use of a gold-standard data set for subjectivity classifications*. Paper presented at the Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics.
- Wu, Fang, & Huberman, Bernardo A. (2010). Opinion formation under costly expression. *ACM Transactions on Intelligent Systems and Technology*, 1(1), 1-13.
- Zeithaml, Valarie A. (1988). Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. *The Journal of Marketing*, 2-22.