Decision Support Systems 149 (2021) 113610

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

How learning effects influence knowledge contribution in online Q&A community? A social cognitive perspective

Chencheng Shi [a], Ping Hu [a], Weiguo Fan [b], Liangfei Qiu [c,*]

[a] School of Management, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China
[b] Tippie College of Business, University of Iowa, Iowa City, Iowa 52242, USA
[c] Warrington College of Business, University of Florida, Gainesville, Florida 32611, USA

ARTICLE INFO

Keywords:
Online Q&A community
Knowledge contribution
Learning effect
Social cognitive theory

ABSTRACT

Informative contributions are critical for the healthy development of online Q&A communities, which have gained increasing popularity in solving personalized open-ended problems. However, little is known about whether past contribution behaviors and corresponding community feedbacks received affect the characteristics of subsequent contributions. Drawing upon the social cognitive theory, we examine the learning effects on users' knowledge contribution behaviors. Specifically, we focus on two types of learning effects: enactive learning from one's past contribution experience and vicarious learning from observation of others' performances in a question thread. Using a dataset collected from one of the largest online Q&A communities in China, we find that the length feature of past user contributions that garner highly positive feedback, no matter through enactive or vicarious learning, would influence the informativeness of subsequent contributions in the community. These learning effects are more effective for users with higher social status. The enactive learning effect is stronger for contributors with higher social status. For the vicarious learning on higher-status contributors, the influence of high-vote long answers is stronger, but the high-vote short answers show a weaker effect. Our research provides a deeper understanding of knowledge contribution behaviors in online knowledge communities and guides for establishing a healthy knowledge contribution environment.

1. Introduction

Online Q&A communities have grown rapidly and made up for the deficiencies in web search engines for acquiring customized information and knowledge [1–3]. For example, Quora, one of the most popular online Q&A communities, has 300 million monthly active users in 2018.[1] Through these online platforms, people could raise concerns, share personal experiences, or discuss others' answers to contribute knowledge to entire communities. The sustainability of such communities largely depends on high-quality and informative contributions from their participants, i.e., user-generated content (UGC) [4,5]. As such, studying the user's knowledge contribution behaviors in online Q&A communities has been a major research concern.

Through knowledge contributions in online communities, though without direct economic benefits, a contributor could harvest positive reputation, peer recognition, and social status due to past voluntary contributions [6–8]. Prior research on online knowledge contributions indicates that gaining social benefits is the main motivation for the provision of public goods in online social media [9,10]. Many online Q&A communities, such as Quora and Zhihu, possess the social mechanism of the "follow" function and the appraisal mechanism like the "vote" button that each answer is evaluated by all the users in the platform. Past contributions that receiving a higher level of recognition would attract the interests of other users and may draw them to follow the contributor. A follower could receive information about a contributor's latest participation in the community. In this way, the contributor becomes more influential and enjoys a higher level of social status. Accordingly, the community feedback of social recognition on each contribution could be treated as behavior consequences by which contributors could learn their performance of past contributions. According to the learning paradigm in the social cognitive theory (SCT) [11], learning occurs either through the consequence of actual doing (enactive learning) or by observing others' performance (vicarious learning) [12,13]. The latter effect is also known as "observational learning" in

* Corresponding author.
E-mail addresses: chencheng-shi@qq.com (C. Shi), helenhu@xjtu.edu.cn (P. Hu), weiguo-fan@uiowa.edu (W. Fan), liangfeiqiu@ufl.edu (L. Qiu).

[1] https://foundationinc.co/lab/quora-statistics/ (last accessed: February 21, 2020)

https://doi.org/10.1016/j.dss.2021.113610
Received 31 July 2020; Received in revised form 28 May 2021; Accepted 1 June 2021
Available online 5 June 2021
0167-9236/

both economics and Information Systems (IS) literature [14–17]. Similarly, the feedback of past contributions (both from past experience and by observing others) could help users infer the community preference and perform more quality contributions.

In this paper, we focus on the dimension of informativeness, which indicates the amount of information contained in an answer post [18–20]. Prior research demonstrates that the informativeness of online contributions is positively related to their quality or helpfulness [18,20–22] and generally reflects the contributor's effort or diligence in provision [23,24]. Different from the answer quality that requires the judgments of other users or experts, the content of each answer is completely determined by the contributors. It is the knowledge voluntarily shared by each contributor that constitutes the main content of these knowledge-oriented websites. Thus, motivating informative contributions is crucial for the sustainable development of online knowledge communities.

This study explores a new cue of learning effect that may impact informative contributions in online Q&A communities. Prior studies on knowledge contribution in online Q&A communities have focused on user motivations and continuance contribution [3,9,25–29]. The influence of past contributions is restricted to whether a user will make provision in the future or how many they might. Though the latest research has drawn attention to the order effect on the quality of contributions in the same question thread in online Q&A health forums [30], whether the feedback of past contributions would affect the content feature of future knowledge contributions (aka learning effect) is underexplored. Therefore, we ask:

1. *Does the learning effect derived from the feedback of past contributions (both enactive learning and vicarious learning) influence the informativeness of future contributions?*
2. *Does this learning effect proposed in the previous question impact distinctively on contributors with different levels of social status?*

The SCT holds that those actions bringing positive results are regarded as modeled behaviors, from which people can derive symbolic representations to guide their future actions [11,12]. Accordingly, to answer the research question, we consider both the length feature and feedbacks of past contributions and propose a research framework that considers the modeled contribution behaviors originating from two sides: one's historical contribution experience and the others' prior contributions in the same question thread. We also explore the corresponding interaction effects of a contributor's social status. Moreover, we pay specific attention to the distinct effect of modeled long and modeled short answers. Lastly, we discuss how our research contributes to the existing literature of online knowledge contribution and social cognitive theory, and give several design guidances for the website managers.

2. Theoretical backgrounds

2.1. Knowledge contribution in online Q&a community

The knowledge contribution issue has long been regarded as an essential question in online community research [9,28,29,31]. Researchers have applied various theories, such as social capital theory [9,32], self-determination theory [2,29], social exchange theory [26,31,32], social cognitive theory [33,34], to reveal the individual differences and social mechanisms that impact user contributions. They find that knowledge contribution behaviors are generally affected by the personal motivational status or conditions of social interaction at the time of contribution.

However, as suggested by the SCT, much human learning takes place in the social environment [11,12]. In the context of online Q&A communities, knowledge contributors do not provide knowledge in isolation or at one time. The feedbacks of past contributions are clearly displayed on the personal homepage or question thread page. A popular contribution could bring the contributor much social recognition (through others' votes) and increase personal social status (through the gaining number of followers) in the community. Through these mechanisms, when choosing the question to be answered, contributors do not have to pay much attention to the status or capacity of knowledge seekers due to the expectation of reciprocity benefits from them as suggested by prior research [9,32] but focus on finding an interesting question and knowledge provision itself. This design of social recognition provides knowledge contributors with opportunities to learn conveniently, and users' knowledge and belief upon contribution might evolve gradually based on the observations or experiences of prior contribution behaviors and corresponding outcomes in an online community. That is, through enactive learning by actual doing or vicarious learning by observing others' performances, users in online Q&A communities could learn what types of contributions are more likely to be recognized by other users on the platform. This may lead to the change of their future contribution behavior, i.e., the learning effect in online communities. To the best of our knowledge, this learning effect behind the knowledge contribution behaviors in online Q&A communities has not been investigated. Furthermore, numerous studies on the knowledge contribution behaviors are from the angle of quality or quantity in a certain period [2,29,33,35] rather than directly observing the particular content feature of contribution behavior. To address these important research gaps, we apply the social cognitive theory to explore the learning effect behind individual online knowledge contribution behavior.

In this study, we analyze the informativeness of knowledge contribution, which indicates the amount of information contained in an answer post [18–20]. This concept is a critical feature studying contributions in online UGC communities and could also be expressed in other similar words like information richness [21,36], comprehensiveness [37], and information depth [21,38]. In prior studies, the informativeness of online contributions, such as online reviews and answers, are usually treated as independent variables and are significantly related to the judgment of content quality and helpfulness [18,22,38] or the user's further decision-making (e.g., consumers' purchase decisions) [19,36]. Actually, these subsequent judgments or purchase behaviors are, to some extent, out of the control of the contributors when contributing, and largely depend on the preference of other users in communities. Only the content itself can be fully controlled by the contributors. Since we are studying specific knowledge contribution behavior, we take the textual content of the answer as our main research object. Among the characteristics of online knowledge contribution in online Q&A communities, the informativeness indicated by, for example, text length [18,38,39], is more straightforward than other text features such as readability and objectivity. It also reflects the contributor's effort in general [23,24]. Therefore, we choose informativeness as our research dimension for knowledge contributions in online Q&A communities.

2.2. Social cognitive theory and knowledge contribution

Developed from the social learning theory [40], social cognitive theory (SCT) explains human behavior using a triadic reciprocal model, in which the environments, personal variables (such as cognition), and behaviors interact with one another [11]. The personal cognitions of self-efficacy and outcome expectations are the main personal determinants shaping one's behavior. Self-efficacy indicates "the belief in one's capability to organize and execute the courses of action required to manage prospective situations" [41], while the outcome expectations refer to the expectancy of outcome produced by certain behaviors [42]. The environments indicate the social or physical environments that impact individual behaviors [43]. Researchers in IS, though admitting the richness of SCT, usually concentrate on the unidirectional relationships inside the model when using SCT to discuss the impacts of interest [28,44]. SCT regards learning as an information processing activity, where the information concerning the behaviors and environment could serve as symbolic representations and provide guidance for future behaviors [11]. SCT outlines two methods of learning: (1) enactive learning by the consequence of actual doing, and (2) vicarious learning by observing the performance of others [12,13]. In online Q&A communities, in which the relationships among users are mainly weak ties to obtain useful information [45], getting social recognition is an essential motivation to contribute knowledge [9,27]. The contributors, therefore, can learn the appropriateness of different contribution behaviors through the feedback they received from the community and might further adjust their future contribution behaviors to cater to the community preference.

Modeling, a general term in SCT, indicates a psychological matching process that takes place on learners when observing one or more models [11]. This process gives the observers the information to conduct behaviors in the observers' context and helps them latterly code rules in their mind for the exhibition of this behavior. People could also regard themselves as their models and observe their own performances as suggested by “self-modeling” [12]. Bandura (1986) holds that the utility of modeled behaviors to the observer is realized either by actual enacting the behavior and viewing its usefulness or by vicariously examining the outcomes of others' actions and receiving functional value. SCT illustrates that behaviors similar to the rewarded modeled behaviors are generally rehearsed more often due to the expectation of positive outcomes [11]. As explained above, this modeling process also occurs in online Q&A communities, and we treat the contributions receiving a higher level of recognition (i.e., the number of votes) as modeled answers. We would explore how these answers serve as antecedents for the knowledge contribution in the community.

Prior studies using SCT in online Q&A communities pay close attention to the impact of the cognitive forces (i.e., self-efficacy) and outcome expectation on subsequent behaviors [26,28,33]. However, the antecedent learning process facilitated by the community response [11] in an online Q&A community remains underexplored, even though this learning process has been examined in other territories in IS research, such as IT training [13,46,47] and e-learning [44]. Therefore, we treat the community feedback as behavior consequences serving as an incentive (social reinforcement as suggested in SCT) that attracts the contributors to perform what they have learned from past experience (enactive learning) or by observing others (vicarious learning).

On the other hand, contributors in the online Q&A community may not definitely perform according to what they have learned from social learning. The consequences of existing behavior indicate the appropriateness or accuracy of behavior and acts as sources of motivation and information [12]. The SCT demonstrates that learning and performance are two distinct processes and holds that whether one will perform what (s)he has learned depends on environmental variables like social pressure and personal impacts such as motivation and perceived need [12]. This conforms to the triadic reciprocal relationship from SCT in deciding personal behavior. Despite this, prior IS research related to computer self-efficacy also employs SCT and reveals that the consequences of prior performances could either exert direct impacts on future performances or first affect one's cognitions of self-efficacy and outcome expectations and then influence future performance [47,48]. The consequences of prior performances in these studies contain both individual historical behaviors and those by observing others' performances (environmental influence). The relationships are illustrated below in Fig. 1.

Hence, the feedback in prior contributions in online Q&A communities could contribute to both environmental and personal factors that affect future contributions. We adopt the paradigm of enactive and vicarious learning in SCT and investigate how prior behaviors impact the informativeness of future knowledge contribution behaviors.

3. Research model and hypotheses

Fig. 2 shows the conceptual model of this research. We begin by exploring the main effects of enactive learning and vicarious learning. Both these learning effects are originated from modeled behaviors featured by answer length. Then, we investigate the moderation effects between the contributor's social status and the main effects, respectively. Moreover, we compare the effects of modeled (high-vote) and unmodeled (low-vote) answers as one's social status grows. We also dig deeper into the length feature of prior contributions and pay specific attention to the modeled long answers and modeled short answers among prior contributions.

3.1. The influence of one's historical contributions

Enactive learning indicates learning by direct experience in which people formulate conceptions of behaviors from observing their response outcomes [11]. In online Q&A communities, the function of the voting mechanism enables contributors to evaluate the social acceptance and trustworthiness of community members to their past contributions. These increasing numbers of “votes” or “likes” one received could be internalized as positive psychological outcomes, thereby enhancing a user's perceived self-efficacy to subsequent contribution behaviors [11,29]. SCT posits that success in performing an action could initiate high self-efficacy while failure decreases it [11]. In addition, the past contributions with favorable outcomes could also increase a user's outcome expectation for performing a similar type of behavior [11,48]. Several studies in the IS research also suggest that individuals could adjust their future behaviors based on past experience and related evaluations or feedbacks [3,49]. Past contributions receiving high community evaluation scores could reinforce and act as a modeled behavior impacting future contributions. Several studies find that long answers are generally more informative than short answers [21,22,37]. Since we use the text length feature to characterize the informativeness of knowledge contributions in online Q&A communities, we expect, accordingly, that the length feature of one's past modeled contributions in an online Q&A community would impact the informativeness of this contributor's future contributions. We hence propose:

H1. *The informativeness of user contributions is positively associated with the length of modeled answers from the contributor's past contribution experience.*

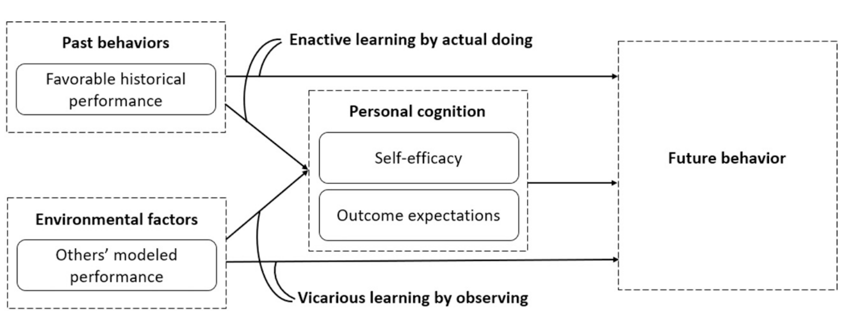

Fig. 1. A theoretical model based on SCT.

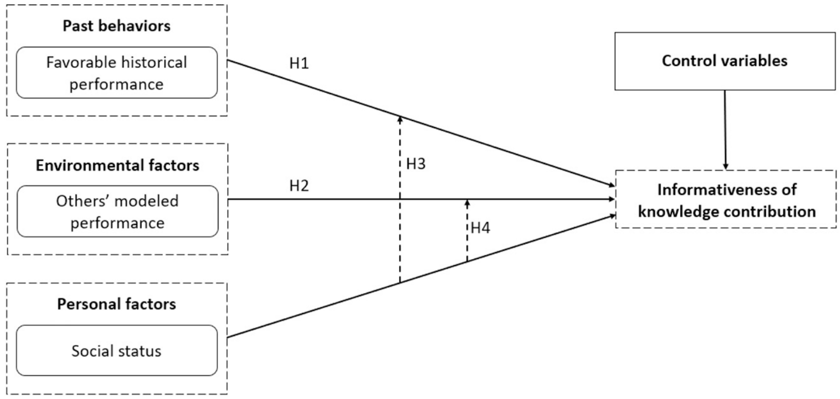

Fig. 2. Conceptual framework.

3.2. The influence of others' contributions in the thread

Apart from enactive learning, learning could also occur vicariously without overtly performing the behavior by the learner [11,12]. Vicarious learning (e.g., learning through electronic media or by observing others' actions) accelerates learning while eliminates the need for people to learn through experiencing the possible negative results [11–13]. This concept is similar to observational learning in the literature [14,15,50]. According to these theories, people in the community could infer the community preference and speculate which behavior is superior by observing others' actions and corresponding outcomes. SCT also suggests that observing similar others' successes might enhance the observer's self-efficacy and confidence in that they can accomplish it as well [11,12]. Thus, in the context of online knowledge contribution, we can expect that answers harvesting high community recognitions in online Q&A communities might be imitated by other contributors for the sake of obtaining positive consequences as well.

Similar to the part of enactive learning, we also focus on the dimension of informativeness. To control for the question heterogeneity in our research, we restrict the scope of vicarious learning to the contributions previously existing in the same question thread. Therefore, before making contributions to a question, potential contributors could observe the existing answers and their outcomes in advance, and speculate the preference of possible readers in the community. If the user observes that lengthy answers have been recognized, he/she can learn that many readers of this question prefer informational answers. To gain positive outcomes, the subsequent contributors could be more willing to express their ideas in detail. We hence raise that:

H2. *The informativeness of subsequent contributions is positively associated with the length of modeled answers already in the same question thread.*

3.3. Moderating effect of social status on the impacts of one's historical contributions

As explained in the introduction, the appraisal mechanism of "vote" and the social mechanism of "follow" function in online Q&A communities lead to various levels of social status for the contributors. Past popular contributions could bring contributors numerous social recognitions and social benefits. Research has shown that past successful experiences increase a user's self-efficacy and the expectation of outcome towards that behavior that further influence future contributions [9,11,26,29,47,48]. Users with higher status in online Q&A communities have more favorable experiences of past contributions and might possess a higher level of self-efficacy and expectation concerning their own feature of knowledge contribution due to the past successful experience. We thus believe that users with higher social status would regard more to their past successful contributions. Therefore, we propose:

H3. *The impact of modeled answers in the contributor's past contribution experience on the informativeness of knowledge contributions is stronger for contributors with a higher level of social status.*

3.4. Moderating effect of social status on the impacts of others' contributions in the thread

Following the SCT [11,47,48], the past successful contributions of users increase their expectation of future contributions. More favorable outcomes are expected for users with a higher social status. That is, they are more likely to show self-regulatory behavior [12] due to past exhibitions of popularity or expertise. Researchers in online product reviews have indicated that users with higher status are prone to regulate their behavior and exhibit less deviated behaviors [51]. In online Q&A communities, there exist competitions between answers in the same question thread, even if they might be weaker in questions asking for personal experience or opinion sharing. Once decided to contribute, subsequent contributors, especially those with a higher social status, may wish not to be in an inferior informative position for attracting potential readers, especially comparing with prior contributions that have received a certain level of recognition. We thus have:

H4. *The impact of modeled answers in the question thread on the informativeness of knowledge contributions is stronger for contributors with a higher level of social status.*

4. Research methodology

4.1. Research setting and data collection

We collect the research data from Zhihu, one of the most popular online Q&A communities in China. As of the end of November 2018, Zhihu has officially announced to have over 220 million users, 30 million questions, and 130 million related answers.[2] Registered participants in Zhihu can search or ask interesting questions, provide answers, vote or comment on others' answers, and follow or be followed by other users on this platform. Unlike other online knowledge communities, there is no time limit for answering questions in Zhihu. Users are encouraged to share knowledge at any time once they are prone to contribute after observing the existing answer conditions. In addition, no daily attendance or answer provision is required to obtain somewhat point rewards for user status promotion. In Zhihu, all the rewards or incentive mechanisms are solely relevant to the contributions themselves and rendered by other users, like recognitions from others and the growth of followers. This provides us with a relatively pure research

[2] https://www.sohu.com/a/281596300_100191066 (last accessed: April 26, 2020)

environment to study knowledge contribution from a learning perspective. Besides, unlike discussion forums, every answer in a specific question is independently arranged either chronologically or according to a certain ranking method. The comments are separated from the answers, and the vote numbers are clearly shown below each answer in the question thread, which is ideal for exploring vicarious learning in our research setting. For these reasons, we choose Zhihu as the target system for exploring the possible learning effect in knowledge contribution in online Q&A sites. Furthermore, Zhihu also issues regulations to guarantee the originality of answers and regulates citations in a certain format.[3] The administrator would "fold" answers with susceptible plagiarism and warn the contributors to modify.

Some questions in online Q&A sites originate from current hot events, which could gain numerous attention for a short while but remain silent after the trend passes. This time-related exogenous impulse could greatly affect users' contribution behaviors [52]. Moreover, the attributes of the question topic, like controversial sensitivity and specialization, might also exert influence on the knowledge contributions [53,54]. To keep the relative independence of the topic and eliminate the latent periodic impact, we focus on a subset of questions that belong to a particular topic: "English Learning." This topic presumably undergoes a relatively stable growth rate and is less affected by other topics, potentially avoiding the unexpected temporal endogeneity problem.

Not all the questions in "English Learning" are specifically related to technical or academic issues. Numerous questions are life-oriented, e.g., asking for others' opinions or recommendations. The users in Zhihu are mainly Chinese-speaking, and English is a compulsory course for nearly all the students in China from primary school (since 2001 according to the Chinese Ministry of Education[4]) to high school. According to a website survey in Zhihu conducted in February 2020, 97% of 5,518 respondents are born after 1980, and 89% of the respondents are born after 1990.[5] Thus, we believe that most of the users in Zhihu may have the experience of learning English. Even if an individual does not learn English to a high level, he/she could still provide answers to some questions under the topic of English learning. In other words, "English Learning" is not a very special topic that could only be contributed by users with unique characteristics. Although it is not a hot topic, it encourages the continuous sharing of skills, opinions, and recommendations in the community.

Our collected data from Zhihu contains all the 106,692 answers from 37,950 questions generated from January 9th, 2011 to August 1st, 2017. For every question and answer, we gather all the details like content, posting time, and feedback. For every contributor, we gather their profile information, such as demographics, participations, and feedback information. We first remove the questions without answers. As the contributors' personal features are considered in our study, we exclude anonymous answers and questions. Through descriptive statistics (shown in Table 1) of the remaining 69,124 answers belonging to 14,372 questions, we observe the number of answers in questions follows the long tail distribution, which is consistent with previous research [55]. 45.3% of the questions receive only one answer, and 1% of questions containing over 50 answers. For questions with too few answers, the question may not attract much attention, or it may not be an open-end question and has a definite answer. For questions with too many answers, the subsequent contributors might not be bothered to read all the prior answers before contribution, which spoils the vicarious learning effect. We hence restrict our samples to the questions containing 4 to 50 answers. Besides, answers are also excluded if they are the first one in a question thread or in a user's contribution history to make the vicarious and enactive learning meaningful. To control for the personal historical contributions, for every contributor in the dataset, we collect all the information on their historical answers, including content, vote number, and posting time. After these processings, we end up with a final corpus containing 20,152 answers in 3569 questions from 14,374 corresponding contributors with a total of 1,824,873 historical answers.

4.2. The influence of prior contributions on the informativeness of online knowledge contribution

In this section, we explore how prior contributions and corresponding outcomes in one's historical contributions and in the question thread influence the informativeness of subsequent contributions. We calculate the average length of high- and low-vote answers to compare their separate impacts.

4.2.1. Variables

We employ the absolute length and non-stop length to measure the informativeness of each contribution. As mentioned above, the length of contribution is a rather straightforward dimension indicating its informativeness and the codified effort of its contributor in general [31,37]. In online Q&A communities like Zhihu, the answers are mainly in the form of text, so we remove all the punctuation, special symbols, emoticons, pictures in the contributions, and calculate the length of the remaining text. Since our research topic is not specifically related to pictures, the pictures existing in contributions in our research act only as additional notes, similar to the role of emoticons. We thus do not use the number of pictures as a measure of informativeness like other studies [21,56]. In addition, we also use the number of nonstop words for the robustness of our measurement. The number of nonstop words quantifies the level of concrete information, i.e., noun, verb, and adjective, in each contribution [23]. Table 2 details the variables used in the research. Table 3 shows the descriptive statistics of these variables. We apply the log form to these two measurements due to their skewness shown in Table 3.

For our independent variables of modeled contributions, we need to define what is the high-vote answer. In Zhihu, as there is no specific time stamp for each vote, we could not directly acquire the exact number of votes at the time of subsequent contributor's provision. The vote function in Zhihu notifies every voter's follower of the vote issue, which is similar to the retweet mechanism on Twitter. It has been found that 75% of retweets happened within 24 h, and 10% occurred over a month [57]. Since we are only interested in figuring out the high-vote answers but not their specific number of votes at the moment of subsequent contributions, we exploit the final vote number as its approximation. Following the distribution of vote number (displayed in Table 4), we treat answers with over 5 (top 20%) and 10 (top 12%) votes as high-vote answers, respectively, to verify the robustness of the results. We then calculate the average length of high-vote answers in one's past contribution experience (enactive learning) and in the question thread (vicarious learning), separately. We also include the average length of low-vote answers in the control variables to make a comparison. These variables are zero if there is no such type of answers in one's historical contributions or in the question thread. Due to the skewness of these variables, we apply the natural log transformation to them.

We control for various factors that could potentially affect specific knowledge contribution, as described in Table 2, in three levels: question, thread, and personal level. In the question level, we control for the question features of length, label number, and question type. We pay special attention to the "how-to" questions, which are non-factoid and possess more uncertainty than factoid questions and more information than social or opinion questions [58]. From the thread level, we control for the answer's chronological rank and the average length of the low-vote answers in the thread. As for the contributor's personal characteristics, we

[3] https://www.zhihu.com/question/20258015 (last accessed: December 31, 2020)

[4] http://www.gov.cn/gongbao/content/2001/content_61196.htm (last accessed: December 19, 2020)

[5] https://www.zhihu.com/club/1156966703501107200/post/1216739470986543104 (last accessed: December 19, 2020)

Table 1
Distribution of the number of answers in questions.

N	Mean	S.D.	Skewness	Min	Max	Percentile													
						10	20	30	40	50	60	70	80	90	91	93	95	97	99
14,372	4.81	24.42	31.42	1	1569	1	1	1	1	2	2	3	4	7	8	10	13	20	51

Table 2
Description of variables.

Variable type	Variable	Description
Dependent variable	Knowledge contribution	Number of words in the answer/ Number of nonstop words in the answer
Independent variables	High-vote answers from past experience	Average length of high-vote answers in the contributor's past contribution experience
	High-vote answers from question thread	Average length of high-vote answers from prior contributions in the question thread
Control variables	Question length	Number of words in the question
	Question label	Number of the question label
	Question type	Whether it is "how to" question
	Low-vote answers from question thread	Average length of low-vote answers from prior contributions in the question thread
	Answer rank	Answer's chronological rank in the thread
	Contributor's self-presentation	Number of demographic items revealed by the contributor
	Contributor's social status	Weighted number of followers of a contributor
	Contributor's contribution experience	Number of past answers a contributor posted
	Low-vote answers from past experience	Average length of low-vote answers in the contributor's past contribution experience
	Contributor's contribution habits[a]	Average length of a user's recent contributions in 30 days.

[a] Though we have contained both the high- and low-vote historical answers in the regression, this habitual variable is additionally included for the robustness of the results. We have applied 12 variations for this variable (detailed in the last paragraph of Section 4.2.1). Here we use the average length of user's answers in recent 30 days at the time of contribution as an example in Tables 2 and 3. This variable is zero if there is no answer within 30 days which accounts for 15% of our sample.

Table 3
Descriptive statistics of variables ($N = 20{,}152$).

	Min	Max	Mean	S.D.	Skewness
Knowledge contribution (text length)	1.00	34,334.00	326.67	872.21	11.83
Knowledge contribution (nonstop text length)	0.00	3500.00	65.83	155.17	7.41
High-vote answers from past experience	0.00	23,112.00	433.97	806.60	7.00
High-vote answers from question thread	0.00	24,391.00	842.97	842.97	7.67
Question length	5.00	98.00	23.01	11.37	1.39
Question label	1.00	12.00	3.82	1.32	−0.42
Question type	0.00	1.00	0.35	0.48	0.63
Low-vote answers from question thread	0.00	5784.00	129.18	211.73	11.18
Answer rank	2.00	50.00	9.46	8.69	1.83
Contributor's self-presentation	0.00	6.00	2.53	1.90	0.40
Contributor's social status	0.00	566,124.70	2247.54	16,385.88	17.16
Contributor's contribution experience	1.00	15,667.00	175.35	930.76	11.60
Low-vote answers from past experience	0.00	8671.00	168.11	310.11	9.55
Contributor's contribution habits	0.00	18,002.00	268.62	613.44	7.98

control for four factors, including (1) the completeness of self-presentation; (2) social status in the community; (3) experience of past contribution; and (4) the average length of the low-vote answers in the contributor's past contribution experience. The number of followers is usually employed to indicate a user's social status in online communities [5,32,59]. As Zhihu does not display the following time of each follower, we have to estimate one's social status at the time of each contribution. Prior research in online Q&A communities has suggested the positive relationship between one's followers and the total number of votes received of past contributions [55]. We thus use estimated received vote at the time of contribution, total vote at the time of crawling, and the total number of followers at the time of crawling to gauge the weighted number of followers when he/she offers an answer showing below:

$$\text{Total follower number*Estimated vote number of the contributor} \\ /\text{Total vote number}$$

Similar to the above data processing, we take the logarithm to the highly skewed variables (skewness>2). We also control for the answer time dummies (monthly dummies).

For the robustness of the results, we have further controlled the personal habitual contribution behavior in recent and initial contributions. Specifically, we control for four types of habitual behavior separately: (1) the average length of recent personal answers in 30, 90, and 180 days before each contribution, (2) the average length of personal recent 3, 6, and 9 answers before each contribution, (3) the average length of answers in one's initial 30, 90, and 180 days of contribution history before each contribution, and (4) the average length of initial 3, 6, and 9 answers in one's contribution history before each contribution. The log form is applied to all these habitual variables to avoid overdispersion.

4.2.2. Model specification

For the structure of our data, answers are clustered by question or contributor. We, therefore, use the hierarchical linear model (HLM) [60,61]. The HLM accommodates unbalanced observations and could be used to separate the model variance into various levels. We find that our dataset is not a pure nest structure as observations are clustered within both question and contributor, but question and contributor are not nested within one another. Following the prior research [32], we conduct the HLM with cross-classified random effects controlling multiple sources of unobserved heterogeneity (controlling both contributor random effect and question random effect at the same time). The estimation equations are:

$$\begin{aligned} ln(Answer.length)_{ij} = {} & Q.Controls_i + Trd.Controls_{ij} + Ctor.Controls_{ij} + Ctor.high \\ & - vote.length_{ij} + Trd.high - vote.length_{ij} + Month.dummies_{ij} \\ & + \mu_i + \omega_j + \varepsilon_{ij} \end{aligned} \tag{1}$$

Table 4
Distribution of vote number, answer length, and social status.

	N	Mean	S.D.	Skewness	Min	Max	Percentile								
							10	20	30	40	50	60	70	80	90
Vote number	20,152	10.79	70.657	20.43	0	3073	0	0	0	0	1	2	3	5	13
Answer length	20,152	326.67	872.211	11.834	1	34,334	14	25	39	60	90	136	209	356	762
Social status	20,152	2247.54	16,385.88	17.16	0	566,125	0	1	3	7	16	37	86	245	1338

$$\begin{aligned} \ln\left(Answer\_nonstop\_length\right)_{ij} = & Q\_Controls_i + Trd\_Controls_{ij} + Ctor\_Controls_{ij} \\ & + Ctor\_high - vote\_length_{ij} + Trd\_high - vote\_length_{ij} + Month\_dummies_{ij} + \mu_i + \omega_j + \varepsilon_{ij} \end{aligned} \tag{2}$$

The subscript "*i*" refers to the Question_ID and "*j*" refers to the Contributor_ID. "*Ctor_high-vote_length$_{ij}$*" indicates the average length of high-vote long answers from one's past contribution experience. "*Trd_high-vote_length$_{ij}$*" suggests the average length of high-vote answers in the question thread. "*Q_Controls$_i$*" indicates the characteristics of *Questions*, such as length, type, and the number of labels. "*Trd_Controls$_{ij}$*" and "*Ctor_Controls$_{ij}$*" suggest the thread features and contributor's personal features of *Answer$_{ij}$* as shown in Table 2. We also examine these models with different standards of high-vote answers (5 and 10) to show the robustness of the results. For the moderation effects of contributor's social status, apart from investigating their moderating impacts with modeled (high-vote) answers that lead to enactive learning and vicarious learning, we also explore their moderating impacts with unmodeled (low-vote) answers in one's historical contributions and the same question thread, respectively. This could provide a comprehensive understanding of the influence of the prior contributions.

4.2.3. Empirical results

Table 5 lists the regression results of our research models. We employ the dependent variable of ln*(Answer_length)* in models 1, 2, and 3, and *ln (Ans_nonstop_length)* in models 4, 5, and 6.

For the enactive learning effect, the coefficients of *Ctor_high-vote_length* in models 1 to 6 are significantly positive. This validates **H1** that declares the enactive learning effect exists in the informativeness of online knowledge contribution. The significant positive coefficients of *Ctor_low-vote_length* in these models also suggest personal informative habit of contributions, no matter receive a high level of social recognition or not, would impact future contributions. Interestingly, for the moderating effect of contributor's social status, The coefficients of *Ctor_high-vote_length* (log) *X Ctor_socialstatus (log)* in models 2, 3, 5, and 6 are significantly positive, while the coefficients of *Ctor_low-vote_length (log) X Ctor_socialstatus (log)* are significantly negative. These results demonstrate that users with a higher social status in the community are more prone to be influenced by their high-vote answers in past contribution experience and less impacted by historical low-vote answers. This proves the stronger enactive learning effect on the informativeness of knowledge contributions when contributors have a higher social status level as illustrated in **H3**.

For the vicarious learning effect in the question thread, similar to the enactive learning effect, the coefficients of *Trd_high-vote_length* in models 1 to 6 are significantly positive, validating the existence of vicarious learning in online knowledge contributions as suggested in **H2**. Both significant positive coefficients of *Trd_high-vote_length* and *Trd_low-vote_length* in these models highlight the social influence (based on the relationship of answering the same question) of all the prior answers. Differently, the positive coefficient of *Trd_high-vote_length* (log) *X Ctor_socialstatus (log)* in models 2, 3, 5 and 6 demonstrates that contributors with higher social status tend to draw more attention to prior high-vote answers in the question thread to make their contributions competitive from the informative perspective, which confirms **H4**. Besides, by comparing the coefficients of independent variables, we find the impacts of one's historical contributions are generally larger than the influence of prior contributions in the thread.

The control for variables of question type, contributor's level of self-presentation, and contribution experience are significantly related to the informativeness of knowledge contribution. The control variable of the contributor's social status is consistently significantly positive in models 1 to 6. This also shows that contributors with higher social status generally provide more informative contributions. When adding the personal habitual variables to the regression model (shown in models 3 and 6), the effects of both *Ctor_high-vote_length* and *Ctor_low-vote_length* as well as the moderating impacts of contributor's social status are decreased, but still significant. This suggests the learning effect (modeling effect) in online knowledge contribution after controlling the personal contribution habits.

5. Robustness checks and additional analyses

5.1. Comparing the main effects with a series of samples

To compare the effects of unmodeled and modeled contributions more thoroughly, we further conduct a series of empirical analyses employing samples containing contributors with different levels of social status. Table 4 demonstrates that in our research sample, over 10% of the contributors have zero followers at the time of contribution, and 30% of the contributors harvest less than three followers when contributing. Though our sample contains contributors that have at least one historical answer when contributing, many of them may have little high-vote answers.

Next, we employ the regression model that only focuses on the direct impacts of modeled/unmodeled answers as in Column 1 of Table 5. We employ samples containing contributors with equal or over X historical followers. We first set X equals 0, 1, 2, 3, …, 22, 23, 24, and examine the coefficients of *Ctor_low-vote_length$_{ij}$*, *Ctor_high-vote_length$_{ij}$*, *Trd_low-vote_length$_{ij}$*, and *Trd_high-vote_length$_{ij}$*, respectively, in these 25 regressions. The results of the four variables in these regressions are all significant at the level of 0.01. The dynamics of results are illustrated in Fig. 3. The blue lines represent the influence of one's historical answers, while the orange indicates the effect of answers in the same question thread. The solid lines stand for the impact of high-vote answers, while the dash lines are for low-vote answers.

The two blue curves in Fig. 3 cross when X is around 3, suggesting that when a contributor's number of followers is over 3, the high-vote answers in one's historical contributions exert a bigger influence than the low-vote answers. As for the two orange curves in Fig. 3, when the contributor's number of followers is over 11, the high-vote answers in the thread are more influential than the low-vote answers.

After that, we convert the interval of X from 1 to 20 and set X equals

Table 5
Empirical results on the informativeness of online knowledge contributions.

	Ans_length			Ans_nonstop_length		
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept (level 0)	3.11***	2.92***	2.86***	1.93***	1.76***	1.71***
	0.072	0.085	0.085	0.068	0.080	0.079
Question (level 1)						
Q_length	0.000063	0.00013	0.00016	−0.0012	−0.0012	−0.0011
	0.0011	0.0011	0.0011	0.0011	0.0010	0.0010
Q_label	−0.0024	−0.004	−0.0044	−0.0051	−0.0064	−0.0067
	0.0094	0.0093	0.0093	0.0091	0.0090	0.0090
Q_type	0.16***	0.16***	0.16***	0.14***	0.14***	0.14***
	0.026	0.026	0.026	0.026	0.025	0.025
Contributor (level 1)						
Ctor_self-presentation	0.020***	0.022***	0.020***	0.016***	0.017***	0.016***
	0.0061	0.0060	0.0059	0.0055	0.0054	0.0054
Ctor_contributionexperience (log)	−0.36***	−0.32***	−0.32***	−0.32***	−0.28***	−0.29***
	0.0087	0.0092	0.0091	0.008	0.0084	0.0083
Ctor_socialstatus (log)	0.19***	0.083***	0.085***	0.16***	0.062***	0.063***
	0.0070	0.022	0.022	0.0064	0.020	0.020
Ctor_low-vote_length (log)	0.20***	0.28***	0.25***	0.18***	0.25***	0.22***
	0.0078	0.011	0.011	0.0071	0.010	0.010
Ctor_recent_habit_30days (log)			0.064***			0.056***
			0.0051			0.0046
Ctor_high-vote_length (log)	**0.12*****	**0.089*****	**0.082*****	**0.11*****	**0.080*****	**0.074*****
	0.0049	0.0055	0.0055	0.0045	0.0050	0.0050
Ctor_low-vote_length (log) X Ctor_socialstatus (log)		**−0.031*****	**−0.026*****		**−0.028*****	**−0.024*****
		0.0027	0.0027		0.0025	0.0025
Ctor_high-vote_length (log) X Ctor_socialstatus (log)		**0.035*****	**0.031*****		**0.033*****	**0.029*****
		0.0025	0.0025		0.0023	0.0023
Thread (level 0)						
Trd_answerrank	0.0010	0.00055	0.00081	0.0023***	0.0018	0.0021
	0.0015	0.0015	0.0015	0.0014	0.0014	0.0014
Trd_low-vote_length (log)	0.045***	0.048***	0.047***	0.034***	0.035***	0.034***
	0.0068	0.010	0.010	0.0063	0.0092	0.0092
Trd_high-vote_length (log)	**0.042*****	**0.027*****	**0.029*****	**0.030*****	**0.016*****	**0.018*****
	0.0042	0.0057	0.0057	0.0039	0.0053	0.0053
Trd_low-vote_length (log) X Ctor_socialstatus (log)		−0.0011	−0.00089		−0.00054	−0.00038
		0.0020	0.0020		0.0018	0.0018
Trd_high-vote_length (log) X Ctor_socialstatus (log)		**0.0044*****	**0.0040*****		**0.0042*****	**0.0037*****
		0.0011	0.0011		0.0010	0.0010
Month fix effect	Yes	Yes	Yes	Yes	Yes	Yes
Q_random effect	Yes	Yes	Yes	Yes	Yes	Yes
Ctor_random effect	Yes	Yes	Yes	Yes	Yes	Yes
AIC	68,422.1	68,159.3	67,145.7	64,801.8	64,525.2	64,389.0
BIC	68,627.8	68,396.6	67,390.9	65,007.5	64,762.5	64,634.3
Log-likelihood	−34,185.1	−34,049.7	−33,541.8	−32,374.9	−32,232.6	−32,163.5
N	20,152	20,152	20,152	20,152	20,152	20,152
N_Question (Group i)	3569	3569	3569	3569	3569	3569
N_Ctor (Group j)	14,374	14,374	14,374	14,374	14,374	14,374

Notes. The results are robust when we change the standard of high-vote answers to 10 votes. The variance inflation factors (VIFs) of variables show no serious multicollinearity problem in these regressions (below 5). *Ctor_recent_habit_30days (log)* refers to the variable of contributor's contribution habits and the results are robust when we change it to other personal habitual variables mentioned in Section 4.2.1.

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

10, 30, 50, …, 290, 310, 330, and conduct regressions, respectively. The coefficients of the four variables in these regressions are still significant at the level of 0.01. The dynamics of results are shown in Fig. 4.

Similar to Fig. 3, the solid blue curve in Fig. 4 reveals that as user's social status grows, the impact of the modeled answers in one's historical contributions increases continually. It further shows an inverted U-shaped growth curve, which suggests the influence of modeled answers in one's historical contributions (enactive learning) grows dramatically as the number of followers grows initially, and when the number of followers is over 110, this effect is still gradually increasing, but the growth trend becomes relatively stable. This indicates that as social status grows, users in online Q&A communities would gradually learn from past high-vote contributions and eventually form a rather stable writing style in terms of informativeness. The decreasing blue dash curve in Fig. 4 suggests as the social status increases, the influence of low-vote answers in one's historical contribution decreases steadily.

The two orange curves in Fig. 4 are comparatively flat, though the solid line rises gently and is consistently higher than the orange dash line. This demonstrates that for the impact of prior answers in the thread, the contributors with higher social status (over 10 followers) draw more attention to the high-vote answers in the thread, which is consistent with our hypothesis.

Interestingly, by contrasting the curves of two colors in Fig. 4, we find the blue curve is generally higher than the orange curve. This

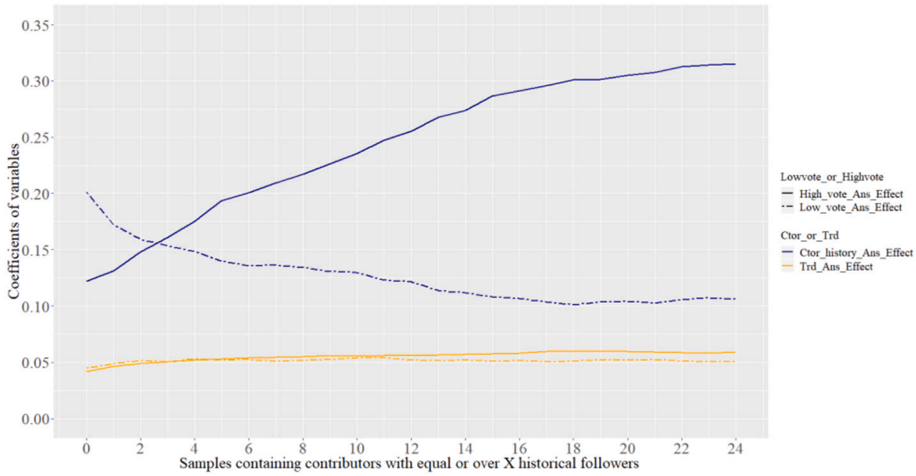

Fig. 3. Dynamics of the effects of high- and low-vote answers as social status grows (internal = 1).

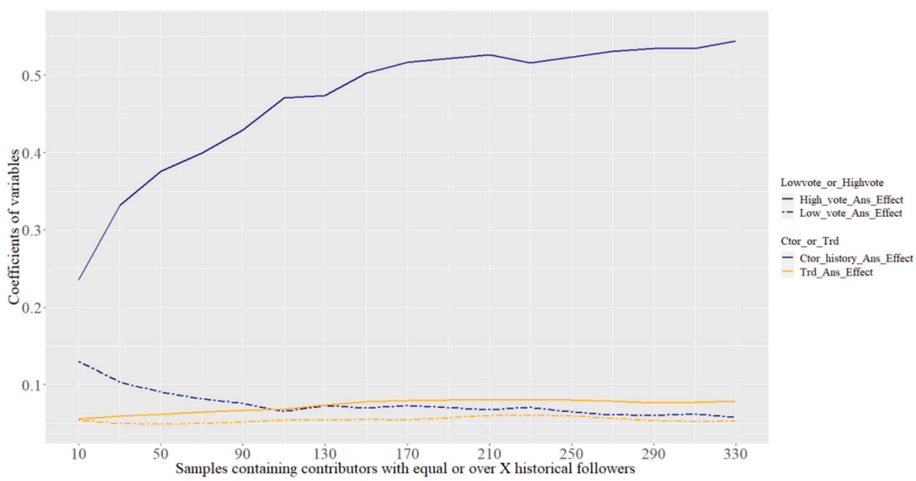

Fig. 4. Dynamics of the effects of high- and low-vote answers as social status grows (internal = 20).

suggests in the context of online knowledge contribution, the impacts of historical personal behaviors are higher than others' influence. This finding is consistent with previous studies on online rating behavior [62]. The solid orange curve in Fig. 4 surpasses the blue dash curve when X is over 130, highlighting the distinct impacts between modeled and unmodeled behaviors.

5.2. Further exploration of modeled contributions

Among these highly-voted contributions in online Q&A communities, truly informative contributions could receive numerous recognitions (up-votes, likes) due to their detailed and profound knowledge contribution. Nevertheless, not every highly-voted contribution is genuinely informative [50]. Like other social media, online Q&A communities could also discourage attention-grabbing contents [14], which tend to have inferior quality but could gain tremendous popularity through sensuous arousing or stimulus. The occurrences of these attention-grabbing contributions might impair the enthusiasm of content contributors for rigorous contributions, leading to a less conscientious atmosphere of contribution in the community.

From this view, modeled long and short answers will have distinct effects on subsequent knowledge contributions and may be moderated differently by the contributor's social status. It would be interesting to portray the prior modeled behaviors in more detail in addition to calculating the average value like prior studies usually did [63,64]. As such, we draw separate attention to the influence of two modeled behaviors: modeled "one-liner" (high-vote short answers) and modeled "real stuff" (high-vote long answers), and investigate their impacts on knowledge contribution, respectively. We adopt the number of highly voted answers as a proxy of the degree of each effect based on social impact theory (SIT) [65,66] in the psychology literature. Studies adopting SIT in the field of marketing and information systems have proved that the number of sources with certain influence exerts positive effects on the level of social impact [67–69].[6]

According to the distribution of answer length shown in Table 4, we select the answers whose length is at the top 20% (over 356 words) as long answers and the last 20% (less than 25 words) as short answers. The criteria of high-vote answers (5 and 10) are the same as the prior section. Based on these classifications and SIT, we count the number of each type of answers in the contributor's past contribution experience and in the question thread to set up the new independent variables. Due to their skewness, we apply the natural log transformation of ln(1 + X) to these variables.

[6] Besides source number, SIT also considers the impact of source strength and immediacy on the social impact [65]. Here in an online Q&A community, we assume that the strength and immediacy of former answers in specific question threads are the same to the subsequent contributors once an answer is viewed as a high-vote answer in this weak tie [45] community.

The HLM are also employed and the equations are as follows:

$$ln\left(Answer\_length\right)_{ij} = Q\_Controls_i + Trd\_Controls_{ij} + Ctor\_Controls_{ij} + \left(Ctor\_high - vote\_longNum_{ij} + Ctor\_high - vote\_shortNum_{ij} + Trd\_high - vote\_longNum_{ij} + Trd\_high - vote\_shortNum_{ij}\right) \cdot \left(1 + Ctor\_socialstatus_{ij}\right) + Month\_dummies_{ij} + \mu_i + \omega_j + \varepsilon_{ij} \tag{3}$$

Table 6
The learning effect on the informativeness of online knowledge contributions. (modeled long and short answers)

	Ans_length		Ans_nonstop_length	
	(7)	(8)	(9)	(10)
Intercept (level 0)	1.25***	1.26***	0.27***	0.29***
	0.091	0.092	0.086	0.087
Question (level 1)				
Q_length	0.00052	0.00053	−0.00078	−0.00077
	0.00094	0.00094	0.00094	0.00094
Q_label	−0.0074	−0.0073	−0.0095	−0.0094
	0.0081	0.0081	0.0081	0.0081
Q_type	0.13***	0.13***	0.12***	0.12***
	0.023	0.023	0.023	0.023
Contributor (level 1)				
Ctor_self-presentation	0.016***	0.016***	0.012**	0.012**
	0.0057	0.0057	0.0052	0.0052
Ctor_contributionexperience (log)	−0.17***	−0.17***	−0.15***	−0.16***
	0.0094	0.010	0.0087	0.0092
Ctor_socialstatus (log)	0.066***	0.053***	0.054***	0.040***
	0.0079	0.0089	0.0073	0.0081
Ctor_habit_avg_length (log)	0.47***	0.48***	0.42***	0.43***
	0.011	0.012	0.010	0.011
Ctor_high-vote_longNum (log)	**0.15*****	**0.11*****	**0.15*****	**0.11*****
	0.017	0.030	0.015	0.028
Ctor_high-vote_shortNum (log)	**−0.10*****	−0.019	**−0.097*****	−0.0011
	0.016	0.036	0.015	0.033
Ctor_high-vote_longNum X Ctor_socialstatus		**0.0098****		**0.0097****
		0.0042		0.0038
Ctor_high-vote_shortNum X Ctor_socialstatus		**−0.015*****		**−0.017*****
		0.0054		0.0050
Thread (level 0)				
Trd_answerrank	0.00026	0.00036	0.0018	0.0019
	0.0017	0.0017	0.0016	0.0016
Trd_avg_length (log)	0.20***	0.20***	0.17***	0.17***
	0.011	0.011	0.011	0.011
Trd_high-vote_longNum (log)	**0.062****	0.020	0.012	−0.031
	0.028	0.032	0.027	0.031
Trd_high-vote_shortNum (log)	**−0.12*****	**−0.17*****	**−0.085****	**−0.14*****
	0.040	0.051	0.039	0.049
Trd_high-vote_longNum X Ctor_socialstatus		**0.013****		**0.014****
		0.0054		0.0049
Trd_high-vote_shortNum X Ctor_socialstatus		0.017		**0.019***
		0.011		0.010
Month fix effect	Yes	Yes	Yes	Yes
Q_random effect	Yes	Yes	Yes	Yes
Ctor_random effect	Yes	Yes	Yes	Yes
AIC	66,841.79	66,866.61	63,294.44	63,312.71
BIC	67,063.3	67,119.76	63,515.95	63,565.87
Log-likelihood	−33,392.89	−33,401.3	−31,619.22	−31,624.36
N	20,152	20,152	20,152	20,152
N_Question (Group i)	3569	3569	3569	3569
N_Ctor (Group j)	14,374	14,374	14,374	14,374

Notes. The results are robust when we change the standard of high-vote answers to 10 votes. The variance inflation factors (VIFs) of variables show no serious multicollinearity problem in these regressions (below 5).

* $p < 0.1$.

** $p < 0.05$.

*** $p < 0.01$.

$$ln(Answer\_nonstop\_length)_{ij} = Q\_Controls_i + Trd\_Controls_{ij} + Ctor\_Controls_{ij} + \left(Ctor\_high - vote\_longNum_{ij} + Ctor\_high - vote\_shortNum_{ij} + Trd\_high - vote\_longNum_{ij} + Trd\_high - vote\_shortNum_{ij}\right) \cdot \left(1 + Ctor\_socialstatus_{ij}\right) + Month\_dummies_{ij} + \mu_i + \omega_j + \varepsilon_{ij} \tag{4}$$

"$Ctor\_high\text{-}vote\_longNum_{ij}$" and "$Ctor\_high\text{-}vote\_shortNum_{ij}$" indicate the enactive learning effect of high-vote long and short answers from one's past contribution experience. "$Trd\_high\text{-}vote\_longNum_{ij}$" and "$Trd\_high\text{-}vote\_shortNum_{ij}$" suggest the vicarious learning effect of high-vote long and short answers in the question thread. As we pay specific attention to the high-vote answers in this section, different from the control variables of eqs. 1 and 2, we further control the average length of both contributor's historical contributions and prior answers in the same question thread. The results are shown in Table 6.

Similar to the results in Section **4.2.3**, the results of coefficients *Ctor_high-vote_longNum*, *Ctor_high-vote_shortNum*, *Trd_high-vote_longNum*, and *Trd_high-vote_shortNum* in models 7 and 9 also validate both the enactive learning and vicarious learning effect on the informativeness of online knowledge contributions. The modeled long answers and modeled short answers have opposite effects on the informativeness of future contributions. That modeled long answers, no matter from historical experience or prior contributions in the thread, would lead to more informative contributions and vice versa. Differently, the coefficients *Ctor_high-vote_longNum* are higher than *Ctor_high-vote_shortNum* in models 7 to 10, but the coefficients *Trd_high-vote_longNum* are lower than *Trd_high-vote_shortNum* in these models. This indicates modeled long answers in one's historical contributions generally have a higher impact than modeled short answers, while the influence of modeled short answers in the question thread is stronger than modeled long answers.

For the moderating effects of social status on enactive learning, the results of coefficients *Ctor_high-vote_longNum X Ctor_socialstatus* and *Ctor_high-vote_shortNum X Ctor_socialstatus* in models 8 and 10 are significant and are in the same direction as the main enactive learning effects. This indicates that users with higher social status are more influenced by their successful historical experience when making provisions in the future. The impacts of historical modeled behaviors, including both modeled long and short answers, are stronger on contributors with higher social status.

Interestingly, as for the moderating effects on vicarious learning, both the coefficients of *Trd_high-vote_longNum X Ctor_socialstatus* and *Trd_high-vote_shortNum X Ctor_socialstatus* in models 8 and 10 are positive, though the variable of *Trd_high-vote_shortNum X Ctor_socialstatus* in models 8 is not significant. That is, the moderating effect of contributor's social status on *Trd_high-vote_shortNum* is in the opposite direction comparing with the direct impact of *Trd_high-vote_shortNum*. This shows that although all the contributions of subsequent users are influenced by the prior modeled short answers in the thread, users with higher social status are less affected by these less informative answers. On the contrary, they are influenced more by the prior high-vote long answers in the question thread. This reveals that once users with higher social status decide to provide answers, they are inclined to regulate their behaviors and regard more informative contributions.

6. Discussion and implications

Motivated by the diverse level of informativeness in online knowledge contributions under the same question or by the same contributor, we investigate whether the length feature of past contributions and corresponding feedback (social incentives received) can motivate contributors to contribute more or less content in subsequent contributions. We find the informativeness of future contributions is significantly influenced by the past contributions and corresponding feedback, no matter in one's past experience or in the same question thread.[7] For users with higher social status, the enactive learning effects drawn from past successful experiences are the most influential factor among the four determinants. Interestingly, the impacts of modeled answers are more effective for users with higher social status, while users with lower social status are prone to value the unmodeled answers. We further pay attention to two specific types of modeled answers: high-vote long answers and high-vote short answers, and validate their opposite impacts on future informative contributions. The subsequent contributors with higher social status regard more on the high-vote long answers in the same question and are less affected by the high-vote short answers.

6.1. Theoretical contributions

Firstly, drawing on the learning paradigm in social cognitive theory, we provide a novel understanding of contribution behaviors in online Q&A communities, under the influence of content characteristics of past contribution performance. The SCT suggests that individuals would learn either by actual doing (enactive learning) or by observing others (vicarious learning), and highlights those received positive outcomes would be treated as modeled behaviors and impact future behaviors. This learning paradigm has been applied to contexts such as IT training [13,46,47] and e-learning [44]. Similarly, the social incentive generated from past knowledge contributions in online Q&A communities could also activate this learning paradigm. Our research extends the applicability of this theoretical perspective and validates both the enactive learning effect and vicarious learning effect on the informativeness of knowledge contributions in the online Q&A community. Besides, when probing the past contributions in more detail, a few distinct findings are also indicated by our empirical results. These include: (1) High-vote long answers in one's historical contributions generally have a higher impact than high-vote short answers, while the influence of high-vote short answers in the question thread is stronger than high-vote long answers. (2) Contributors with higher social status consider more on high-vote long answers in the question thread and are less affected by high-vote short answers when making knowledge provisions. In general, we suggest the learning effect to be a helpful theoretical lens to understand the contribution behaviors in online communities with incentives or feedback mechanisms.

Secondly, our research also contributes to the literature on online knowledge contribution. Prior research has identified the impact of past contributions to future knowledge contribution behaviors [3,30] or how community feedback influences users' continuous online knowledge contributions [29]. We complement these studies by combining past contributions with their corresponding feedback and investigate their compound influence on the informativeness of future contributions. The past contributions in the online Q&A community are considered in both one's historical contributions and prior contributions in the same

[7] In an untabulated empirical analysis, we also employ a new measurement of informativeness concerning whether answers containing pictures and the results support the learning effect of online knowledge contribution.

question thread. Our results suggest the learning effect in online knowledge contribution after controlling the personal contribution habit. Contributors, especially those with higher social status, are influenced by their past successful contributions, and the influence of high-vote long answers is stronger. The high-vote answers in the question thread, especially the high-vote short answers, also affect the informativeness of subsequent contributions, and users with higher social status regard more on the high-vote long answers and less on the high-vote short answers in the thread. Besides, we also illustrate the distinct impacts of modeled and unmodeled answers on contributors with various levels of social status. The finding that one's historical personal contributions are generally more influential than others' contributions in the thread is also consistent with research on online rating behavior [62]. These results provide a more comprehensive understanding of online knowledge contributions by shedding the relationships among specific contribution behaviors along with relating feedbacks and suggest the potential for motivating informative contributions.

For the research generalizability exclusively from China, as mentioned in the topic selection, the users in the online Q&A community of Zhihu are mainly young people and speaking Chinese, and English is a compulsory course for nearly all the young people in China. This assures the broad potential users on this topic. As Zhihu is not a website academically for learning English, the topic of English learning in Zhihu thus is a relatively casual platform for sharing techniques, opinions, and recommendations initiated by questions. This topic is also not related to hot events. Therefore, this research might be generalized in other settings out of China for casual personal sharing in themes not related to instant hot spots like learning foreign languages.

6.2. Practical implications

Our findings highlight the value of past contributions and feedback received on the informativeness of future contributions and guide the management and design of online communities for motivating informative contributions (more "real-stuff" answers). This might include: (1) leading the contributor's behaviors by showing them tips in the interface of contribution like "Following past performances, informative answers are easier to get more votes"; (2) providing more opportunities of positive feedback for informative answers; When a reader stays in a long answer for a long time, the website could remind them of voting by the tips like "It is not easy to contribute, don't grudge your appreciation." This could lead to more recognition for informative knowledge contributors. (3) providing informative answerers with titles such as "Outstanding Topic Answerer" to encourage their diligent and informative contributions.

On the other hand, if the online knowledge website decides to employ a developing strategy of encouraging a diversified style of contributions, a similar encouraging design method could also be applied to other types of contributions such as interesting "one-liner" answers in certain sections or topics in the community. This could attract users who prefer this category of online contributions.

6.3. Limitations and future research

Our study is not without limitations. First, although we provide rationales for choosing a rather stable topic of data to prevent unexpected exogenous impulses, it would be interesting to examine the existence of learning effects in other types of topics relating to such as hot events. There could be hundreds of thousands of answers to questions related to hot spots in a short period in online Q&A communities like Zhihu. The learning effect on content generation may behave differently in this instant background of online contributions. Future research in broader contexts could provide a more profound vision of contribution behaviors, IT artifacts, and learning effects in the online community.

Second, the dimension of the explained variable in this research is restricted to the informativeness of knowledge contribution. There are several other angles of text analysis, such as objectivity and sentiment. Besides the text-based online community, the learning effect might also affect personal styles in other types of UGC communities like YouTube or Instagram.

Third, our research focuses on the learning effects of high-vote answers after controlling the impacts of low-vote answers and personal contribution habits. These two types of controls also significantly influence future contributions. Though we have explored the distinct impacts of the high- and low-vote answers on contributors with various levels of social status, it would be interesting to dig deeper into the underlying mechanisms among these contribution behaviors. Other theories such as strategic behavior in online review [51] and the research on the early stage of social cognition [11] could also be employed to provide a more comprehensive understanding of individual knowledge contribution behaviors.

The fourth limitation is related to the research design. Although we control for comprehensive personal and contextual factors to demonstrate the enactive and vicarious learning effects on knowledge contribution, such settings may not sufficiently verify causal relationships of the theoretical model. Future research could adopt randomized experiments to strengthen causal identifications.

Funding

This research was supported by the National Social Science Foundation of China (17GBL035), and the China Scholarship Council (201806280325).

References

[1] C. Shah, S. Oh, J. Sun, Research agenda for social Q&A, Libr. Inf. Sci. Res. 31 (2009) 205–209, https://doi.org/10.1016/j.lisr.2009.07.006.
[2] J. Lou, Y. Fang, K.H. Lim, Contributing high quantity and quality knowledge to online Q&A communities, J. Am. Soc. Inf. Sci. Technol. 64 (2013) 356–371, https://doi.org/10.1002/asi.
[3] L. Khansa, X. Ma, D. Liginlal, S.S. Kim, Understanding members' active participation in online question-and-answer communities: a theory and empirical analysis, J. Manag. Inf. Syst. 32 (2015) 162–203, https://doi.org/10.1080/07421222.2015.1063293.
[4] J. Chen, H. Xu, A.B. Whinston, Online communities and quality of user-generated content moderated online communities and quality of user-generated content, J. Manag. Inf. Syst. 28 (2011) 237–268, https://doi.org/10.2753/MIS0742-1222280209.
[5] X. Zeng, L. Wei, Social ties and user content generation: evidence from flickr, Inf. Syst. Res. 24 (2013) 71–87.
[6] P.A. Pavlou, D. Gefen, Building effective online marketplaces with institution-based trust, Inf. Syst. Res. 15 (2004) 37–59, https://doi.org/10.1287/isre.1040.0015.
[7] R. Willer, Groups reward individual sacrifice: the status solution to the collective action problem, Am. Sociol. Rev. 74 (2009) 23–43.
[8] J. Lerner, J. Tirole, Some simple economics of open source, J. Ind. Econ. 50 (2002) 197–234.
[9] M.M. Wasko, S. Faraj, Why should I share? Examining social capital and knowledge contribution in electronic networks of practice, MIS Q. 29 (2005) 35–57.
[10] M.X. Zhang, F. Zhu, Group size and incentives to contribute: a natural experiment at Chinese wikipedia, Am. Econ. Rev. 101 (2011) 1601–1615.
[11] A. Bandura, Social Foundations of Thought and Action: A Social Cognitive Theory, Prentice Hall, Englewood Cliffs, NJ, 1986.
[12] D.H. Schunk, Learning Theories an Educational Perspective Sixth Edition, Pearson, 2012.
[13] S. Gupta, R. Bostrom, An investigation of the appropriation of technology-mediated training methods incorporating enactive and collaborative learning, Inf. Syst. Res. 24 (2013) 454–469.
[14] L. Qiu, Q. Tang, A.B. Whinston, Two formulas for success in social media: learning and network effects, J. Manag. Inf. Syst. 32 (2015) 78–108, https://doi.org/10.1080/07421222.2015.1138368.
[15] H. Cai, Y. Chen, H. Fang, Observational learning : evidence from a randomized natural field experiment, Am. Econ. Rev. 99 (2009) 864–882.
[16] Y. Chen, Q. Wang, J. Xie, Online social interactions: a natural experiment on word of mouth versus observational learning, J. Mark. Res. 48 (2011) 238–254, https://doi.org/10.1509/jmkr.48.2.238.

[17] K. Kang, J. Hahn, P. De, Learning effects of domain, technology, and customer knowledge in information systems development: an empirical study, Inf. Syst. Res. 28 (2017) 797–811.
[18] X. Sun, M. Han, J. Feng, Helpfulness of online reviews: examining review informativeness and classification thresholds by search products and experience products, Decis. Support. Syst. 124 (2019) 113099, https://doi.org/10.1016/j.dss.2019.113099.
[19] K.Z.K. Zhang, S.J. Zhao, C.M.K. Cheung, M.K.O. Lee, Examining the influence of online reviews on consumers' decision-making: a heuristic–systematic model, Decis. Support. Syst. 67 (2014) 78–89, https://doi.org/10.1016/j.dss.2014.08.005.
[20] D. Park, J. Lee, I. Han, The effect of on-line consumer reviews on consumer purchasing intention: the moderating role of involvement, Int. J. Electron. Commer. 11 (2007) 125–148, https://doi.org/10.2753/JEC1086-4415110405.
[21] J. Fang, L. Chen, X. Wang, B. George, Not all posts are treated equal: an empirical investigation of post replying behavior in an online travel community, Inf. Manag. 55 (2018) 890–900, https://doi.org/10.1016/j.im.2018.04.003.
[22] X. Liu, M.A. Wang, W. Fan, Z. Zhang, Finding useful solutions in online knowledge communities: a theory-driven design and multilevel analysis, Inf. Syst. Res. 31 (3) (2020) 731–752.
[23] J. Pu, Y. Chen, L. Qiu, K. Cheng, Does identity disclosure help or hurt user content generation? social presence, inhibition, and displacement effects, Inf. Syst. Res. 31 (2) (2020) 297–322.
[24] Z. Ke, D. Liu, D.J. Brass, Do online friends bring out the best in us? the effect of friend contributions on online review provision, Inf. Syst. Res. 31 (4) (2020) 1322–1336.
[25] H.-T. Tsai, R.P. Bagozzi, Contribution behavior in virtual communities: cognitive, emotional, and social influences, MIS Q. 38 (2014) 143–163.
[26] J. Jin, Y. Li, X. Zhong, L. Zhai, Why users contribute knowledge to online communities : an empirical study of an online social Q&A community, Inf. Manag. 52 (2015) 840–849, https://doi.org/10.1016/j.im.2015.07.005.
[27] W. Chen, X. Wei, K.X. Zhu, Engaging voluntary contributions in online communities: a hidden markov model, MIS Q. 42 (2018) 83–100, https://doi.org/10.25300/MISQ/2018/14196.
[28] T. Guan, L. Wang, J. Jin, X. Song, Knowledge contribution behavior in online Q&a communities : an empirical investigation, Comput. Hum. Behav. 81 (2018) 137–147, https://doi.org/10.1016/j.chb.2017.12.023.
[29] L. Chen, A. Baird, D. Straub, Why do participants continue to contribute? Evaluation of usefulness voting and commenting motivational affordances within an online knowledge community, Decis. Support. Syst. 118 (2019) 21–32, https://doi.org/10.1016/j.dss.2018.12.008.
[30] R. Mousavi, T.S. Raghu, K. Frey, Harnessing artificial intelligence to improve the quality of answers in online question-answering health forums, J. Manag. Inf. Syst. 37 (2020) 1073–1098.
[31] A. Kankanhalli, B.C.Y. Tan, K.-K. Wei, Contributing knowledge to electronic knowledge repositories: an empirical investigation, MIS Q. 29 (2005) 113–143.
[32] R. Beck, I. Pahlke, C. Seebach, Knowledge exchange and symbolic action in social media-enabled electronic networks of practice: a multilevel perspective on knowledge seekers and contributors, MIS Q. 38 (2014) 1245–1270.
[33] C. Chiu, M. Hsu, E.T.G. Wang, Understanding knowledge sharing in virtual communities : an integration of social capital and social cognitive theories, Decis. Support. Syst. 42 (2006) 1872–1888, https://doi.org/10.1016/j.dss.2006.04.001.
[34] C. Chen, S. Hung, To give or to receive? Factors influencing members' knowledge sharing and community promotion in professional virtual communities, Inf. Manag. 47 (2010) 226–236, https://doi.org/10.1016/j.im.2010.03.001.
[35] W. Jabr, R. Mookerjee, Y. Tan, V.S. Mookerjee, Leveraging philanthropic behavior for customer support: the case of user support forums, MIS Q. 38 (2014) 187–208.
[36] K. Goh, C. Heng, Z. Lin, Social media brand community and consumer behavior: quantifying the relative impact of user- and marketer-generated content, Inf. Syst. Res. 24 (2013) 88–107.
[37] Y. Lu, K. Jerath, P.V. Singh, The emergence of opinion leaders in a networked online community: a dyadic model with time dynamics and a heuristic for fast estimation, Manag. Sci. 59 (2013) 1783–1799.
[38] S.M. Mudambi, D. Schuff, What makes a helpful online review? a study of customer reviews on amazon.com, MIS Q. 34 (2010) 185–200.
[39] K.K.Y. Kuan, K.-L. Hui, P. Prasarnphanich, H.-Y. Lai, What makes a review voted? An empirical investigation of review voting in online review systems, J. Assoc. Inf. Syst. 16 (2015) 48–71.
[40] A. Bandura, R.H. Walters, Social Learning Theory, Prentice-hall, Englewood Cliffs, NJ, 1977.
[41] A. Bandura, Self-Efficacy: The Exercise of Control, W.H. Freeman and Company, New York, 1997.
[42] A. Bandura, Self-efficacy mechanism in human agency, Am. Psychol. 37 (1982) 122–147.
[43] R. Wood, A. Bandura, Social cognitive theory of organizational management, Acad. Manag. Rev. 14 (1989) 361–384.
[44] Z. Wan, D. Compeau, N. Haggerty, The effects of self-regulated learning processes on e-learning outcomes in organizational settings, J. Manag. Inf. Syst. 29 (2012) 307–339, https://doi.org/10.2753/MIS0742-1222290109.
[45] M.S. Granovetter, The strength of weak ties, Am. J. Sociol. 78 (1973) 1360–1380.
[46] R. Santhanam, D. Liu, W.M. Shen, Gamification of technology-mediated training: not all competitions are the same, Inf. Syst. Res. 27 (2016) 453–465.
[47] M.A. Bolt, Testing the interaction effects of task complexity in computer training using the social cognitive model, Decis. Sci. 32 (2001) 1–20.
[48] D.R. Compeau, C.A. Higgins, Application of social cognitive theory to training for computer skills, Inf. Syst. Res. 6 (1995) 118–143.
[49] P.B. Lowry, J. Zhang, C. Wang, M. Siponen, Why do adults engage in cyberbullying on social media? An integration of online disinhibition and deindividuation effects with the social structure and social learning model, Inf. Syst. Res. 27 (2016) 962–986.
[50] C.C. Chen, Y. Tseng, Quality evaluation of product reviews using an information quality framework, Decis. Support. Syst. 50 (2011) 755–768, https://doi.org/10.1016/j.dss.2010.08.023.
[51] W. Shen, Y.J. Hu, J.R. Ulmer, Competing for attention: an empirical study of online reviewers' strategic behavior, MIS Q. 39 (2015) 683–696.
[52] A. Aaltonen, S. Seiler, Cumulative growth in user-generated content production: evidence from wikipedia, Manag. Sci. 62 (2016) 2054–2069.
[53] G.C. Kane, S. Ransbotham, G.C. Kane, S. Ransbotham, Content and collaboration: an affiliation network approach to information quality in online peer production communities, Inf. Syst. Res. 27 (2016) 424–439.
[54] H.S. Kim, Attracting views and going viral: how message features and news-sharing, J. Commun. 65 (2015) 512–534, https://doi.org/10.1111/jcom.12160.
[55] G. Wang, K. Gill, M. Mohanlal, H. Zheng, B.Y. Zhao, Wisdom in the social crowd: an analysis of quora, in: Proc. 22nd Int. Conf. World Wide Web, ACM, 2013, pp. 1341–1351.
[56] Y. Jin, J. Huang, X. Wang, What influences content popularity? an empirical investigation of voting in social q&a communities, in: Proc. 21st Pacific Asia Conf. Inf. Syst., Langkawi, 2017, p. 161.
[57] H. Kwak, C. Lee, H. Park, S. Moon, What is twitter, a social network or a news media?, in: Proc. 19th Int. Conf. World Wide Web ACM, 2010, pp. 591–600.
[58] M. Surdeanu, M. Ciaramita, H. Zaragoza, Learning to rank answers to non-factoid questions from web collections, Comput. Linguist. 37 (2011) 351–383.
[59] A. Susarla, J. Oh, Y. Tan, Social networks and the diffusion of user-generated content: evidence from YouTube, Inf. Syst. Res. 23 (2012) 23–41.
[60] K.J. Klein, S.W.J. Kozlowski, Multi-Level Theory, Research, and Methods in Organizations, Jossey-Bass, San Francisco, CA, 2000.
[61] S.W. Raudenbush, A.S. Bryk, Hierarchical Linear Models: Applications and Data Analysis Methods, Sage Publications, Newbury Park, CA, 2002.
[62] B. Gao, N. Hu, I. Bose, Follow the herd or be myself? An analysis of consistency in behavior of reviewers and helpfulness of their reviews, Decis. Support. Syst. 95 (2017) 1–11, https://doi.org/10.1016/j.dss.2016.11.005.
[63] B.Y. Chen, F.M. Harper, J. Konstan, S.X. Li, Social comparisons and contributions to online communities : a field experiment on movielens, Am. Econ. Rev. 100 (2010) 1358–1398.
[64] X. Ma, L. Khansa, Y. Deng, S.S. Kim, Impact of prior reviews on the subsequent review process in reputation systems, J. Manag. Inf. Syst. 30 (2014) 279–310, https://doi.org/10.2753/mis0742-1222300310.
[65] B. Latané, The psychology of social impact, Am. Psychol. 36 (1981) 343–356.
[66] A. Nowak, J. Szamrej, B. Latané, From private attitude to public opinion : a dynamic theory of social impact, Psychol. Rev. 97 (1990) 362–376.
[67] X. Zhang, S. Li, R.R. Burke, A. Leykin, An examination of social influence on shopper behavior using video tracking data, J. Mark. 78 (2014) 24–41.
[68] C. Ding, H.K. Cheng, Y. Duan, Y. Jin, The power of the "like" button: the impact of social media on box office, Decis. Support. Syst. 94 (2017) 77–84, https://doi.org/10.1016/j.dss.2016.11.002.
[69] T. Schlager, C. Hildebrand, G. Häubl, N. Franke, A. Herrmann, Social product-customization systems: peer input, conformity, and consumers' evaluation of customized products, J. Manag. Inf. Syst. 35 (2018) 319–349, https://doi.org/10.1080/07421222.2018.1440763.