Proceedings Paper:
Banerjee, Snehasish orcid.org/0000-0001-6355-0470 and Chua, Alton Y.K. (2018) Tracing the Growth of IMDb Reviewers in terms of Rating, Readability and Usefulness. In: 2018 4th International Conference on Information Management (ICIM). International Conference on Information Management, 25-27 May 2018 IEEE, GBR, pp. 57-61.

https://doi.org/10.1109/INFOMAN.2018.8392809

Tracing the Growth of IMDb Reviewers in terms of Rating, Readability and Usefulness

Snehasish Banerjee
The York Management School
University of York
York, UK
e-mail: snehasish.banerjee@york.ac.uk

Alton Y. K. Chua
School of Communication and Information
Nanyang Technological University
Singapore
e-mail: altonchua@ntu.edu.sg

Abstract—**This paper seeks to investigate how reviewers in the IMDb online community grow from novices to experts in terms of rating, readability and usefulness. Reviewers' growth is conceptualized along two dimensions. One dimension includes three stages of the reviewer lifecycle, namely, novice, intermediate, and expert. The other dimension captures the rate of maturity. Data analysis involved both quantitative and qualitative approaches. Results indicated that reviewers in the novice stage contributed significantly higher ratings than in either the intermediate stage or the expert stage. As reviewers made the transition from novices to experts, the use of language became increasingly sophisticated, thereby taking a toll on readability. Moreover, in the novice and the intermediate stages of the reviewer lifecycle, entries submitted by slow-maturing reviewers were voted as being more useful than those contributed by fast-maturing individuals. The findings bear implications for both theory and practice.**

Keywords-IMDb, online review, rating, reviewer, usefulness

I. INTRODUCTION

A. *Background*

Popular review websites such as Amazon, TripAdvisor and IMDb empower the online community by serving as free outlets for the public's post-purchase experiences. In particular, IMDb allows its users to write movie reviews that comprise numerical ratings on a scale of one to 10 coupled with textual comments. When submitted online, each of these reviews comes with a close-ended question, "Was the above review useful to you?" with two response options: 'Yes' and 'No.' Based on the responses of registered IMDb users toward the question for a given review, the usefulness of the entry is indicated through an annotation of the form, "x out of y people found the following review useful."

When people are interested to know about a movie, they have the option to browse user-generated reviews on IMDb. They often look for reviews with either extreme or moderate ratings [1]. They can seek entries that are generally easy to read at a glance [2]. In addition, they are able to sort the entries based on review usefulness [3]. Conceivably, rating, readability and usefulness constitute crucial aspects that have the potential to shape the review reading experience of the online community.

B. *Problem Statement and Research Questions*

The popularity of reviews notwithstanding, research suggests that only a small fraction of reviewers remain active contributors in an online community [4]. Most reviewers leave the community after posting only a handful of entries due to lack of motivation and incentive. Commonly referred as the Funnel Model, the proportion of reviewers who actually remain passionate contributors could be as low as two percent [5].

This small proportion of reviewers, who often go on to attain influential status in the long run, are extremely valuable. After all, entries submitted by these reviewers exert immense influence on the purchase decisions of the online community that in turn has implications for firms' sales and profits [1]. Perhaps expectedly, scholars have tried to understand how novice reviewers mature into experts. According to research in this area [4,6], the progress in participation is made in stages. Furthermore, different reviewers can mature through the stages at different rates [7]. As reviewers grow from one stage of contribution to another, their intrinsic motivation to contribute is expected to evolve. As a result, their online behaviours are likely to change [8].

Building on these works, this paper argues that novice reviewers do not suddenly grow to become experts but grow through a maturing process. In addition, the intrinsic motivation to write reviews for a reviewer is not necessarily the same across the novice, the intermediate, and the expert stages. For the purpose of this paper, a reviewer who has just started contributing entries is said to be in the novice stage. One who has contributed more than a handful entries is said to be in the intermediate stage. Finally, a reviewer who has contributed at least 100 reviews is said to be in the expert stage [9,10].

Furthermore, the intrinsic motivation to write reviews can also vary between fast-maturing and slow-maturing reviewers. If so, the three crucial aspects of reviews—rating, readability as well as usefulness—could differ with respect to the stages of the reviewer lifecycle and the rate of reviewers' maturity. Yet, little research has shed light on this area. In consequence, the scholarly understanding of how reviewers' contributions change as they grow from novices to experts in an online community is relatively limited.

To plug the gap in the literature, this paper seeks to address the following research questions (RQs) by drawing data from IMDb:

RQ 1: How does rating of reviews differ with respect to the stages of the reviewer lifecycle and the rate of reviewers' maturity?

RQ 2: How does readability of reviews differ with respect to the stages of the reviewer lifecycle and the rate of reviewers' maturity?

RQ 3: How does usefulness of reviews differ with respect to the stages of the reviewer lifecycle and the rate of reviewers' maturity?

The remainder of this paper is organized as follows: The next section describes the research methods employed. It is followed by a description of the results, which are discussed thereafter. The final section highlights the significance of the paper along with its limitations.

II. METHODS

A. Data Collection

Data were collected from the movie review website IMDb for two reasons. First, the website has been in existence for close to three decades. Hence, it boasts of a huge user base with several prolific reviewers. Second, a previous work has empirically shown that several reviewers on IMDb have a tendency to change their style of contribution as they go along [9]. This makes the platform particularly apt for this paper. The data were retrieved using a web crawler tool called import.io.

Identification of expert reviewers was a crucial decision in the data collection process. Since IMDb uses the voting of its most prolific reviewers to identify the best 250 movies of all time, reviewers for these movies with the highest volume of contributions were deemed as experts.

Therefore, the best 250 movies at the point of data collection were identified. For each movie, the top 10 reviewers in terms of their volume of contributions were identified, thus yielding a list of 2,500 reviewers (250 movies x 10 reviewers). Interestingly, most of these reviewers were not unique. After eliminating duplicates, a total of 106 unique reviewers was identified. Of these, 100 reviewers with the highest number of reviews were selected for investigation.

The first 100 reviews posted by each of these reviewers were retrieved. After all, the initial 100 contributions are sufficient to trace the growth of a reviewer in an online community [9,10]. Thus, the dataset comprised 10,000 reviews altogether (100 reviewers x 100 initial reviews).

B. Variable Operationalization

The research questions in this paper call for the operationalization of five concepts. These are as follows: (1) stage of the reviewer lifecycle, (2) rate of reviewers' maturity, (3) rating, (4) readability, and (5) usefulness.

With respect to the stages of the reviewer lifecycle, this paper groups reviewers' contributions into sets of 10. In particular, for a given reviewer, reviews #1 to #10, reviews #45 to #54, and reviews #91 to #100 were deemed to represent the novice stage, the intermediate stage, and the expert stage respectively. Such an approach of operationalization is consistent with the convention of handling timed data in human performance [9,10,11].

To distinguish between fast-maturing and slow-maturing reviewers, a median-split was applied on the time elapsed between the contribution of review #1 and that of review #100. Specifically, the median time in the dataset was 315.50 days. Therefore, any reviewer who took less than the median time to post 100 reviews was considered to be a fast-maturing individual. Conversely, any reviewer who took more than or equal to the median time to contribute 100 entries was treated as a slow-maturing individual. After median-split, there were 50 fast-maturing and 50 slow-maturing reviewers.

Rating of a given review was operationalized as the reviewer-assigned star value in the entry. It could range from one to 10.

Readability is commonly measured using metrics such as Gunning Fog Index, Coleman Liau Index, Automated Readability Index, and Flesch Kincaid Grade Level. A mean of these metrics was calculated to create a composite readability measure. The smaller the value of the composite measure, the better is the readability of a given review.

Usefulness of reviews was calculated as the ratio of useful votes to the total number of votes for a particular review. Reviews that are voted as useful by many would have a higher usefulness ratio than those that are mostly designated as not-useful.

C. Data Analysis

Prior to analysis, the dataset was processed to create reviewer-centric data points. For this purpose, reviews for each of the three stages in a given reviewer's lifecycle were grouped together. This helped calculate the reviewer's mean rating, mean readability, and mean usefulness separately in the novice, the intermediate, and the expert stages. All reviewers were further dummy-coded to indicate whether they were fast-maturing or slow-maturing reviewers.

The RQs were addressed using split-plot analysis of variance (ANOVA). This statistical procedure is suitable to analyse data in mixed-design settings that include a combination of within-participants and between-participants variables. In the context of this paper, stage of the reviewer lifecycle represents a within-participants variable. This is because each reviewer has to grow through the novice stage, the intermediate stage, and the expert stage. In contrast, rate of reviewers' maturity represents a between-participants variable. Obviously, a slow-maturing reviewer cannot be a fast-maturing individual at the same time, and vice-versa.

Split-plot ANOVA was conducted thrice with rating (RQ 1), readability (RQ 2), and usefulness (RQ 3) as the three separate dependent variables. The analysis primarily focused on the presence of any statistically significant interaction. When absent, the simple effects were checked.

Split-plot ANOVA is susceptible to the assumption of sphericity. Therefore, violation of sphericity was checked using Mauchly's test. When the assumption was violated, Greenhouse-Geisser corrections were employed.

Besides, this paper also offers a qualitative treatment of the reviews. Specifically, to delve deeper, a manual content analysis was done on the reviews submitted by the five fastest-maturing reviewers and the five slowest-maturing reviewers. The purpose of this exploratory analysis was to identify any interesting trends or patterns.

III. RESULTS

Table I presents the descriptive statistics for the 100 reviewers, who were categorised as either fast-maturing or slow-maturing based on a median-split. The trends in terms of rating, readability and usefulness are pictorially depicted in Figure 1, Figure 2, and Figure 3 respectively.

As the fast-maturing reviewers grew from novices to experts, three trends could be identified: (1) Rating of reviews falls from the novice stage to the intermediate stage before rising again in the expert stage. (2) Readability scores of reviews rise continuously across the three stages. In other words, reviews become less readable as the reviewers grow. (3) Usefulness of reviews seem to grow.

Among the slow-maturing reviewers, the three trends are as follows: (1) Rating of reviews falls consistently across the three stages. (2) Reviews become less readable as the reviewers grow. (3) Usefulness of reviews seem to fall steadily.

With respect to rating (RQ 1), the split-plot ANOVA did not reveal any statistically significant interaction between the stages of the reviewer lifecycle and the rate of reviewers' maturity. Nonetheless, the simple effect of the stages was statistically significant, $F = 2.62$, $p < 0.01$. Post-hoc analysis revealed that reviewers in the novice stage contributed significantly higher ratings than in either the intermediate stage or the expert stage.

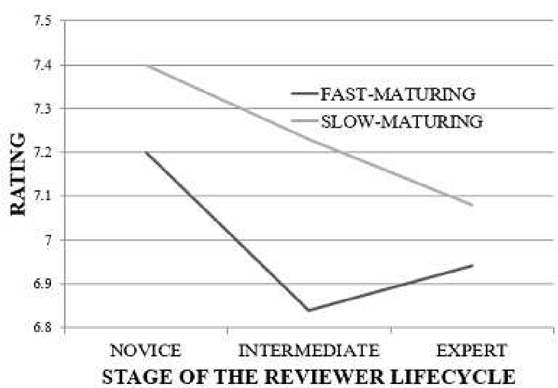

Figure 1. Trends of rating.

With respect to readability (RQ 2), no statistically significant interaction was identified. However, the simple effect of the stages was statistically significant, $F = 4.92$, $p < 0.001$. Reviewers particularly contributed reviews with the lowest readability scores in the initial stage. In other words, individuals in the initial stage of the reviewer lifecycle were most likely to contribute easy-to-read reviews. As they grew through the stages, their use of language became more sophisticated, thereby taking a toll on readability.

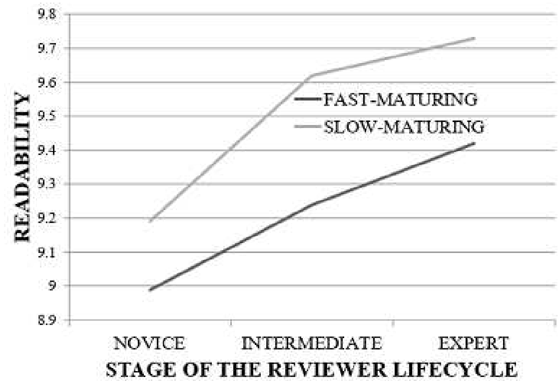

Figure 2. Trends of readability.

With respect to usefulness (RQ 3), the split-plot ANOVA detected a statistically significant interaction between the stages of the reviewer lifecycle and the rate of reviewers' maturity, $F = 3.70$, $p < 0.001$. The difference in the usefulness of reviews between fast-maturing and slow-maturing reviewers that existed in the novice stage was reduced in the intermediate stage, and almost disappeared in the expert stage.

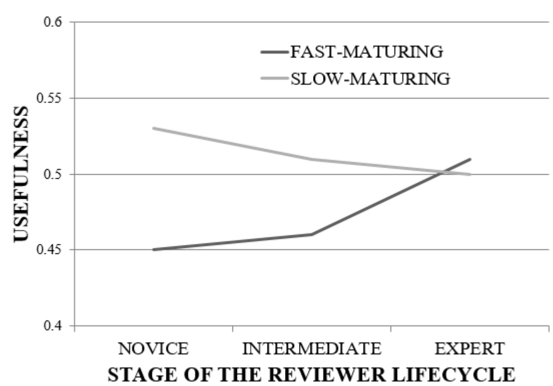

Figure 3. Trends of usefulness.

TABLE I. DESCRIPTIVE STATISTICS (MEAN ± SD)

	Fast-Maturing Reviewers (N = 50)			**Slow-Maturing Reviewers (N = 50)**		
	Novice Stage (Review #1 - #10)	**Intermediate Stage (Review #45 - #54)**	**Expert Stage (Review #91 - #100)**	**Novice Stage (Review #1 - #10)**	**Intermediate Stage (Review #45 - #54)**	**Expert Stage (Review #91 - #100)**
Rating	7.20 ± 1.43	6.84 ± 1.22	6.94 ± 1.56	7.40 ± 1.25	7.23 ± 1.32	7.08 ± 1.57
Readability	8.99 ± 2.56	9.24 ± 2.28	9.42 ± 2.63	9.19 ± 2.00	9.62 ± 1.90	9.73 ± 1.95
Usefulness	0.45 ± 0.13	0.46 ± 0.15	0.51 ± 0.16	0.53 ± 0.21	0.51 ± 0.18	0.50 ± 0.19

Based on the content analysis of the reviews submitted by the five fastest-maturing reviewers and the five slowest-maturing reviewers, two patterns emerged. First, the fast-maturing reviewers were more expressive in their use of language compared with the slow-maturing individuals. The former used phrases such as “Mad Mad Mad Mad Mad” and “One word=F-L-A-W-L-E-S-S” while the latter was more subdued as evident from milder words such as “disappointing” and “beautiful”. The excessive use of punctuations such as “?” and “!” in review titles by the fast-maturing reviewers was particularly conspicuous. Examples of such titles include “So inspiring!” and “Best Film of all time?”

Second, the fast-maturing reviewers focused on the movies per se whereas the slow-maturing reviewers emphasized on their personal experiences of movies. The former seemed keen to highlight plots as well as the quality of script and direction. Interestingly, for one of the five fastest-maturing reviewers, the titles of all the 100 reviews were simply the respective movie names. In contrast, the slow-maturing reviewers offered more personalised views with substantial use of pronouns. They often used phrases such as “I did not enjoy” and “one of my favourite movies”.

Besides the two patterns, a serendipitous finding is worth pointing. The five fastest-maturing reviewers in the dataset took less than three weeks to submit reviews #1 to #100. Even by the standards of super-movie fanatics, it seems daunting for any individual to view 100 movies and review them within such a short span. This raises the question of whether such reviewers were spammers.

IV. DISCUSSION

Tracing the growth of IMDb reviewers, this paper gleans two major findings. First, rating and readability of reviews show a gradual convergence of patterns regardless of reviewers’ rate of maturity. Specifically, with respect to ratings, the journey from novices to experts seemed to make reviewers increasingly critical. For this reason, reviewers in the novice stage submitted higher ratings on average than either in the intermediate stage or in the expert stage. The leniency waned as they matured in the community.

As reviewers progressed through the stages in their lifecycle, the readability of their reviews gradually declined. This does not mean that they become less adept in expressing themselves. Instead, the longer they stayed in the online community, the greater was their level of sophistication in the use of language.

The concept of legitimate peripheral participation could be used as a theoretical lens to explain this finding [12]. It posits that when newcomers join a group, they mostly engage in low-risk activities. After staying in the community for a relatively longer period of time, they start to operate as matured individuals. In the case of IMDb, newcomers perhaps do not take the risk of submitting overly critical ratings with highly sophisticated language. This could be due to fear of isolation. With experience however, novices become more critical in their ratings, and more sophisticated in their use of language. This type of behavioural convergence is commonly referred as online herding [13].

Second, in the novice and the intermediate stages of the reviewer lifecycle, entries submitted by slow-maturing reviewers were voted as being more useful than those contributed by fast-maturing individuals. This finding is promising given that the behaviour of some of the fast-maturing reviewers resembled that of potential spammers.

Insights from prior research on online review spam could be brought to bear in this context. For one, works such as [14] suggested that submission of several reviews within a short duration is a sign of spamming. On a similar note, this paper found that the five fastest-maturing reviewers had submitted 100 reviews within a duration of three weeks.

Again, works such as [15] found that fake reviews were more likely to contain punctuations such as “!” compared with authentic entries. In a similar vein, this paper identified substantial use of punctuations in reviews contributed by the fastest-maturing reviewers. Overall, it might not always be wise for users to rely on reviews submitted by those who contribute a huge pool of entries in a short span of time.

V. CONCLUSION

This paper investigated how IMDb reviewers grow from novices to experts in terms of rating, readability and usefulness. Results indicated that reviewers in the novice stage contributed significantly higher ratings than in either the intermediate stage or the expert stage. As reviewers made the transition from novices to experts, the use of language became increasingly sophisticated. Moreover, in the novice and the intermediate stages, entries submitted by slow-maturing reviewers were voted as being more useful than those contributed by fast-maturing individuals.

The paper is significant for both theory and practice. On the theoretical front, it presents a new two-dimensional conceptualization of reviewers’ growth in an online community. One dimension included three stages of the reviewer lifecycle, namely, novice, intermediate, and expert. The other dimension captured the rate of maturity. Moreover, the paper augments the scholarly understanding of how rating, readability and usefulness of reviews change as a function of the stages in the reviewer lifecycle as well as the rate of reviewers’ maturity [6,9,13,16].

On the practical front, the findings of this paper have implications for review websites on how to retain users. According to [6], an understanding of expert reviewers helps devise strategies to attract potentially loyal reviewers. Review websites need to encourage newcomers, who have a tendency to leave the community after submitting only a handful of entries, to maintain a long-term association. The newcomers could be exposed to a short demonstration of how others have grown on to become experts from novices. Ways to become an expert could be suggested. Moreover, review websites need to closely monitor accounts from which several reviews are posted in short durations. This could be a small step toward weeding out review spam [14].

That said, two limitations inherent in this paper need to be acknowledged. First, it drew data from a single review website. Hence, the results are not necessarily generalizable to other platforms such as Amazon and TripAdvisor. Interested scholars could replicate the paper in the context of other review websites as well as social media platforms such as Facebook and Twitter [17].

Second, to trace the growth of the reviewers in the online community, it relied not on the actual reviewers but on their reviews. This could not be avoided because the nature of social media research is such that access to the real individuals is seldom available. Nonetheless, where feasible, future research could consider administering surveys or interviews among prolific reviewers. This would enable obtaining richer insights into online reviewers' journey from novices to experts.

ACKNOWLEDGMENT

The authors would like to thank Phyu Mon Theint for her help in the data collection and analysis.

REFERENCES

[1] N. N. Ho-Dac, S. J. Carson, and W. L. Moore, "The effects of positive and negative online customer reviews: Do brand strength and category maturity matter?," Journal of Marketing, vol. 77, pp. 37-53, 2013.

[2] V. Woloszyn, H. D. P. dos Santos, and L. K. Wives, "The influence of readability aspects on the user's perception of helpfulness of online reviews," Salesian Journal on Information Systems, vol. 1, pp. 2-8, 2016.

[3] B. Li, F. Hou, Z. Guan, A. Y. L. Chong, and X. Pu, "Evaluating online review helpfulness based on elaboration likelihood model: The moderating role of readability," Proceedings of the Pacific Asia Conference on Information Systems, AIS, 2017, available at: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1016&context=pacis2017

[4] K. Crowston, and I. Fagnot, "Stages of motivation for contributing user-generated content: A theory and empirical test," International Journal of Human-Computer Studies, vol. 109, pp. 89-101, 2018.

[5] J. Porter, Designing for the Social Web. Berkeley, CA: New Riders, 2008.

[6] J. Preece, and B. Shneiderman, "The reader-to-leader framework: Motivating technology-mediated social participation," AIS Transactions on Human-Computer Interaction, vol. 1, pp. 13-32, 2009.

[7] J. J. McAuley, and J. Leskovec, "From amateurs to connoisseurs: Modeling the evolution of user expertise through online reviews," Proceedings of the ACM International Conference on World Wide Web, pp. 897-908, New York, NY: ACM, 2013.

[8] T. W. Gruen, T. Osmonbekov, and A. J. Czaplewski, "eWOM: The impact of customer-to-customer online know-how exchange on customer value and loyalty," Journal of Business Research, vol. 59, pp. 449-456, 2006.

[9] L. Hemphill, and J. Otterbacher, "Learning the lingo? Gender, prestige and linguistic adaptation in review communities," Proceedings of the ACM Conference on Computer Supported Cooperative Work, pp. 305-314, New York, NY: ACM, 2012.

[10] P. Victor, C. Cornelis, A. M. Teredesai, and M. De Cock, "Whom should I trust? The impact of key figures on cold start recommendations," Proceedings of the ACM Symposium on Applied Computing, pp. 2014-2018, New York, NY: ACM, 2008.

[11] T. Monk, D. Buysse, C. Reynolds III, S. Berga, D. Jarrett, A. Begley, and D. Kupfer, "Circadian rhythms in human performance and mood under constant conditions," Journal of Sleep Research, vol. 6, pp. 9-18, 1997.

[12] J. Lave, and E. Wenger, Situated Learning: Legitimate peripheral participation. Cambridge university press, 1991.

[13] L. Michael, and J. Otterbacher, "Write like I write: Herding in the language of online reviews," Proceedings of the International Conference on Weblogs and Social Media, AAAI, 2014, available at: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8046

[14] G. Fei, A, Mukherjee, B. Liu, M. Hsu, M. Castellanos, and R. Ghosh, "Exploiting burstiness in reviews for review spammer detection," Proceedings of the International Conference on Weblogs and Social Media, AAAI, 2013, available at: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewPaper/6069

[15] S. Banerjee, and A. Chua, "Theorizing the textual differences between authentic and fictitious reviews: Validation across positive, negative and moderate polarities," Internet Research, vol. 27, pp. 321-337, 2017.

[16] H. T. Tsai, and P. Pai, "Why do newcomers participate in virtual communities? An integration of self-determination and relationship management theories," Decision Support Systems, vol. 57, pp. 178-187, 2014.

[17] A. Pal, A. Chua, and D. Goh, "Does KFC sell rat? Analysis of tweets in the wake of a rumor outbreak," Aslib Journal of Information Management, vol. 69, pp. 660-673, 2017.