Do I Like This Movie? Movie Ratings, Box-Office Profit, and Factors Affecting Ratings

Asian Communication Research Vol.15 No.3, 20～46

DOI: 10.20879/acr.2018.15.3.20
https://www.comm.or.kr

Do I Like This Movie? Movie Ratings, Box-Office Profit, and Factors Affecting Ratings

Ji Youn Ryu
Graduate student, in the School of Media and Communication at Korea University

Hee Sun Park*
Professor in the School of Media and Communication at Korea University

Abstract

This study aimed to find out the influence of film ratings by three different groups—audience, netizens, and professional critics—on box-office revenue in South Korea. Study 1 examined how ratings by the three groups, as a form of word-of-mouth (WOM), are related to the total number of audience. Across 706 movies released in Korea from December 2013 to November 2016, ratings from all three groups were positively related to box-office record. Yet, a multiple regression analysis showed that audience rating was significantly related to the total number of audience for commercial films, but not so for independent/art films. To explore an explanation for the Study 1 finding, Study 2 hypothesized that efforts spent by individuals in watching films would lead them to grant higher ratings. However, it was discovered that it is not the effort but familiarity that brought about audience's favorable rating of a film. Therefore, we suggest that PR managers and advertisers focus on increasing the exposure of a film for positive reviews. More detailed findings and implications thereof are discussed in the article.

Key words: movie ratings, box-office revenue, word-of-mouth (WOM), effort justification, familiarity

* heesunpark321@gmail.com

Study 1

People make decisions after decisions every day. During decision-making processes, people not only rely on their own sources of information but also what others say. Especially when they are barely familiar with a product or have never performed a certain task, they may be more dependent on others' evaluations. For instance, if some individuals try to purchase a laptop computer for the first time in their lives, they may ask around for information from their friends or browse the Internet to look over others' comments on different laptop models. One of many possible reasons for such behaviors can be people's natural inclination to avoid risk-taking (Ha, 2002). By referring to others' previous experiences, people may prevent themselves from making hasty decisions.

As time passes, information gets accumulated, leading to creation of a sizable database loaded with various comments and experiential testimonies. By sharing their knowledge and experiences with others, individuals engage in word-of-mouth (WOM). The power of WOM is massive. Studies have found that it exerts considerable influence in shaping consumers' beliefs and attitudes (Brown & Reingen, 1987). Moreover, it affects consumer decision across different areas including automobile, sports viewership, online game, and tourism (Asada & Ko, 2016; Huang & Liao, 2017; Murtiasih & Siringoringo, 2013; Park, Kim, & Kim, 2016).

Positive WOM is considered as the ultimate factor that accounts for a product success (Day, 1971). The advent of the Internet further allows anyone to become a consumer and a producer of WOM at the same time. Through web-based platforms, individuals now have easier and broadened access to information online by engaging in electronic word-of-mouth

(eWOM) (Henning-Thurau, 2004). Research have shown that eWOM has become more influential than traditional WOM due to broad coverage and fast diffusion of information (Sung, Kim, Kwon, & Moon, 2012). In particular, online review, a popular form of eWOM, has been proven useful in guiding consumer decisions on various platforms including Amazon and eBay (Forman, Ghose, & Wiesenfeld, 2008; Pavlou & Dimoka, 2006).

Literature Review: Film Rating as a Form of Word-of-Mouth

Movie consumption and the power of WOM are often interrelated with one another. Dozens of movies are released every month, and it is inefficient for people to search every single movie to select one that they want to spend time on. Moreover, movies are experience goods of which product quality cannot be judged before consumers actually attend them (Baek, Oh, Yang, & Ahn, 2017; Koschat, 2012). This is when ratings come in handy as a form of WOM. By scanning ratings given by viewers who have already watched a film, an individual may decrease the burden of having to attain the information (Hu, Koh, & Reddy, 2014). For example, "Rotten Tomatoes (https://www.rottentomatoes.com)" is an online website in which tens of thousands of people share their evaluation of movies. The website provides "Tomatometer," which illustrates the percentage of critic and audience reviews that are positive for a given film. Movies that get 60 or higher percentage of positive reviews are labeled 'fresh' while those that get 59 or lower percentage of positive reviews are labeled as 'rotten.' Taking a look at this index greatly helps users decide on whether or not to watch a film at a glance. In short, by providing others' previous experiences, ratings work as a shortcut that enables people to save time and money.

As ratings play as a decent indicator of movie quality, they

have a significant impact on a film's box-office record. According to Dellarocas, Zhang, and Awad (2007), volume, valence, and dispersion of user ratings in the opening weekend significantly affect future national box-office revenue. Inconsistency exists among studies on which dimension of WOM most powerfully influences consumer decision. On one side, studies have reported that it is valence that matters the most, regardless of volume of WOM (Chevalier & Mayzlin, 2006; Chintagunta et al., 2010). For example, Chevalier and Mayzlin (2006) state that market sales are more likely to be affected by negative reviews than positive reviews. In contrast, East, Hammond, and Lomax (2008) argue that overall, the influence of positive WOM is greater than that of negative WOM. Meanwhile, Forman et al. (2008) report that extreme ratings, whether positive or negative, are considered useful in consumer decision making. These findings suggest that brand managers should take care of both positive and negative WOM for successful marketing strategies. It is not to mention that those who produce negative WOM are likely to produce positive WOM as well (East, Hammond, & Wright, 2007), implying the necessity of comprehensive monitoring of consumer evaluation.

On the other hand, other studies have emphasized the significance of the volume of WOM. For instance, Khare, Labrecque, and Asare (2011) suggest that WOM volume is more critical than valence because high volume enhances the impact of both positive and negative WOM. On a similar note, Kim, Park, and Park (2013) elaborate that the frequency of online WOM—that is, how many times a film is mentioned in online space—by the public is the most important predictor of box-office revenue. Although a disagreement exists on whether WOM valence or volume more strongly affects consumer behavior, there is no doubt that film ratings and reviews serve as guides for people's selection of movies.

In addition, professional critics can exert a considerable amount of impact on box-office record. To be specific, Basuroy, Chatterjee, and Ravid (2003) report that experts play a dual role of predictor and influencer of box-office revenue. In the U.S., for instance, positive reviews by well-known movie critics have a large influence on demand for newly-released movies (Reinstein & Sydner, 2005). In fact, moviegoers are more influenced by expert critiques than a cast of a film (Suárez-Vázquez, 2011). Nonetheless, Tsao (2014) states that the influence of expert reviews is relatively small compared to that of consumer reviews.

The current research aims to examine movie ratings as a form of WOM in South Korea. The powerful impact of WOM on box-office sales has been well-documented in several studies conducted in the Western countries. However, relatively speaking, empirical studies examining movie ratings and their related factors are scarce in South Korea. By examining the relationship between WOM and box-office revenue on 706 movies released in South Korea from December 2013 to November 2016, this study seeks to uncover if the findings from the West are applicable to South Korean contexts and if there are findings unique to South Korea. The current research specifically focuses on film rating as a representation of eWOM, since it is popularly used among moviegoers to search films on the Internet.

Hypothesis and Research Question

Based on previous research findings on the influence of WOM, we expect that film ratings will be positively related to box-office revenue in South Korea. Following Korean Film Council's method of calculating box-office record, we will use a total number of audience as an indicator of box-office profit.

Hypothesis 1: Film ratings will be positively related to a total number of audience.

Three types of film ratings are frequently available in online sites in South Korea; ratings by professional critics, audience (those who actually went to a theater), and netizen (those who leave ratings online). It is possible that any of the ratings can be more influential than others for box-office profit. Thus, the following research question is posed;

Research question 1: Which of the three rating types is significantly related to a total number of audience?

Figure 1 below illustrates the conceptual relationships among the proposed variables regarding H1 and RQ1.

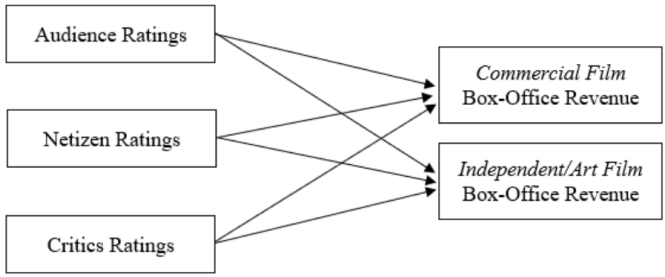

Figure 1. The conceptual model of the relationship among the variables in Study 1

Method

Data Collection Procedure

Information regarding 706 movies released in Korea from

December, 2013 to November, 2016 was gathered for analysis. The data set included the following: ratings by three different groups of people (professional critics, audience, netizens), film category (commercial vs. independent and art films), and total number of audience (i.e., an indicator of box-office profit record). All data were collected primarily from "Naver Movie (http://movie.naver.com)[1]" service and additionally from Korean Film Council (KOFIC)[2]. This study used Naver Movie as the main source of data because it holds the biggest number of users in South Korea compared to other movie ratings services. For instance, 10,777 people registered film ratings on *La La Land* (2016, dir. Damien Chazelle) on Naver Movie. Yet, on Daum Movie (http://movie.daum.net), which is the second biggest movie ratings service in Korea, only 1,344 people did. Thus, we concluded that Naver Movie would be the most suitable and credible source in studying the relationship between ratings and total number of audience.

Data Characteristics

In this study, professional critics refer to journalists and film critics who work at Korean film magazine "Cine21". Audience refers to those who purchase tickets via Naver Movie service, visit a theater to watch a movie, and register ratings on Naver Movie service. The netizen category differs from the audience category in that it refers to those who register ratings on Naver

---

1 Naver Movie is a web-based movie ratings service in South Korea. It not only provides film ratings by audience, netizen, and professional critics, but also reviews and comments.

2 Korean Film Council (KOFIC) was founded in 1973 to promote qualitative improvement of films and support the growth of Korean movie industry. Their official website (http://www.kofic.or.kr) provides diverse information on films including box-office ranking, sales figure, total number of audience, screen share, etc.

Movie but *did not* purchase tickets via Naver Movie. Thus, it should be noted that netizen may include those who watch movies in ways other than visiting a theater (e.g., using a desktop or laptop computer, mobile phone, television, etc.) and even those who *have not* actually watched the film. It has been reported that Naver Movie separates the two groups to filter out the ratings from mindless or malicious users and ensure the credibility of audience ratings.

Movies were divided into two general categories—commercial film vs. independent and art film—based on the size and characteristics of distributing agencies. Commercial films refer to movies distributed by large-scale agencies such as CJ Entertainment, Showbox, Lotte Entertainment, SONY Pictures, and Universal Pictures Korea. Independent and art films refer to those distributed by small-scale agencies such as Jin-Jin, Big Pictures, and NEW.

Results

Pearson correlation analyses were conducted to assess the relationships among ratings by audience, netizens, and professional critics (independent variables), and total number of audience (dependent variable). The analysis was done separately for commercial vs. independent and art films. Then, a multiple regression analysis was conducted to identify which rating type explained a significant amount of the variation in the total number of audience.

Hypothesis 1

It was hypothesized that film ratings by three different groups—audience, netizens, and professional critics—would be positively

related to the total number of audience, which is an indicator of box-office record. In the case of *commercial films*, the correlations between ratings of all three groups and the total number of audience were significant. To be specific, audience ratings, r (264) = .42, p < .01, netizen ratings, r (267) = .35, p < .01, and critics ratings, r (264) = .28, p < .01, all had positive relationships with the total number of audience. Although there were some differences in the sizes of correlations, the data were consistent with H1 that WOM would be positively related to box-office profit. The results are illustrated in Table 1.

Table 1. Correlations between Film Ratings and Total Number of Audience: Commercial Films

	Audience Ratings	Netizen Ratings	Professional Critics Ratings	Total Number of Audience
Netizen Ratings	.88**	--		
Professional Critics Ratings	.44**	.50**	--	
Total Number of Audience	.42**	.35**	.28**	--
M	8.00	7.71	6.00	1,953,648[a]
SD	0.84	1.09	1.18	2,742,121[a]

*p < .05, **p < .01

Note. The degrees of freedom range from 262 to 267.
a. The mean and the standard deviation of total number of audience were rounded to the nearest integer.

Meanwhile, in the case of *independent and art films*, there was almost no correlation between the ratings and the total number of audience. That is, the ratings by audience, netizens, and professional critics respectively had correlations of r (418) = .08, p = .12, r (424) = .09, p = .08, r (417) = .02, p = .75 with the total number of audience, showing no statistical significance. The results are shown in Table 2.

Table 2. Correlations between Film Ratings and Total Number of Audience: Independent and Art Films

	Audience Ratings	Netizen Ratings	Professional Critics Ratings	Total Number of Audience
Netizen Ratings	.76**	--		
Professional Critics Ratings	.45**	.41**	--	
Total Number of Audience	.08	.09	.02	--
M	8.11	7.81	6.30	198,686[a]
SD	0.89	1.01	1.16	747,244[a]

*$p < .05$, **$p < .01$

Note. The degrees of freedom range from 417 to 427.
a. The mean and the standard deviation of total number of audience were rounded to the nearest integer.

An additional finding was that for both commercial and independent/art films, professional critic ratings were moderately related to audience ratings and netizen ratings, whereas audience and netizen ratings were highly related to one another. That is, the general public and professional critics seemed to have different criteria in evaluating a movie.

Research Question 1

To address research question 1, a multiple regression analysis was conducted to detect which group was important for the relationship between the ratings and the total number of audience. The result showed that for *commercial films*, audience ratings were the only significant variable, $\beta = .48$, $t = 4.10$, $p < .001$. On the other hand, for *independent and art films*, none of the variable was able to explain the variance in the total number of audience substantially. The results are shown in table 3.

Table 3. Multiple Regression Analysis Results

Commercial Film		Independent and Art Films	
	β		β
Audience Ratings	.48**	Audience Ratings	.04
Netizen Ratings	-.14	Netizen Ratings	.07
Critics Ratings	.15	Critics Ratings	-.03
F (3, 259) = 21.02, $p < .001$, $adj.R^2 = .186$		F (3, 409) = 1.29, $p = .28$, $adj.R^2 = .002$	

*$p < .05$, **$p < .01$

Discussion

The main finding of this study is that ratings are positively related to the total number of audience (i.e., an indicator of box-office revenue) in the case of commercial films. In general, the findings are consistent with previous research findings on power of WOM. The findings lead to two different implications. First of all, it may be crucial for commercial film makers and promoters to work on improving the valence of WOM and keep an eye on ratings as predictors of box-office record. Although professionals' opinions do have a significant correlation with a movie's success, it may be more important to attract ordinary audience. This is due to the fact that ordinary audience and experts emphasize different standards in appreciating a movie (Holbrook, 1999). In addition, previous studies have reported that people rely on expert knowledge when there is "low information volume," but prefer WOM created by the ordinary citizens when there is high information volume (Flanagin & Metzger, 2013). These findings may account for the current finding that critic ratings were not as significantly correlated to box-office profit compared to audience ratings. Since Naver Movie holds a large number of users, it is possible that people credited user-generated WOM more than expert opinions. In sum, positive WOM in the form of audience movie ratings can be consequential for the PR

managers in the commercial film business.

In contrast, independent and art film makers may need to focus more on enlarging audience population. As revealed in Study 1, there was almost no correlation between the ratings and the total number of audience. No matter how high or low ratings were, independent and art films did not attract much audience from the first place. The average number of total audience for independent films was less than 200,000, while the average number was almost 2 million for commercial films. Thus, the first step that independent and art film promoters have to take is creating bigger WOM regardless of valence. Although valence matters in the long run, it may be more important to establish sizable WOM in the beginning. In other words, making the film noticeable would be the most effective way to draw potential customers (Hsu & Jane, 2016).

Limitations

One limitation of this study is that a causal relationship between the audience ratings and the total number of audience could not be examined. That is, the authors were not able to find out if WOM led to box-office profit or vice versa because temporal order could not be addressed. Another limitation was that it was not possible to assess the impact of other variables which could have affected box-office profits. For example, screen domination would force audience to watch certain movies regardless of their choices. Screen domination refers to large distributing agencies compelling theaters to show their movies more exclusively than other smaller agencies' movies. Lastly, there may be a questionable credibility issue in netizen ratings, since the netizen group included those who did not actually watch the film.

Study 2

The data collection of 706 movies from Study 1 displayed an interesting pattern. On average, audiences rated films 0.3 point higher than netizens did. For the majority of movies (535 out of 706), audience ratings were higher than netizen ratings. Thus, a question can be raised; what triggered the difference between audience and netizen ratings? In Study 2, an experiment was conducted to test if Aronson and Mills' (1959) effort justification theory would provide an explanation for higher audience ratings than netizen ratings.

Literature Review: Effort and Favorability

Effort justification refers to individuals' tendency to grant higher value on something they have strived to attain, compared to that requiring minimum effort (Aronson & Mills, 1959). According to the theory, perceived value on a task increases as the degree of effort put into the task increases. For instance, suppose person X spends 20 hours while person Y spends only 2 hours both to get an A mark on a math exam. Although X and Y equally earned A mark, X would place greater value on his or her score because s/he has put much effort. In other words, X tries to account for, or justify his or her endeavor by enhancing the value of outcome. What if X gets C+ mark after studying for 20 hours? Interestingly, the level of perceived value is likely to get even higher so as to reduce discomfort created by discrepancy between long study time (effort) and low mark. Despite unsatisfactory outcome (C+ mark), X would justify his effort by reminding himself of what he gained after 20 hours of studying (i.e. "At least I learned how to calculate properly. Scores do not matter.") (Zentall, 2010). As such, effort justification is an effective way to resolve cognitive

dissonance (Festinger, 1957) by emphasizing the positive side of an outcome, thus convincing oneself that his or her effort is worthwhile.

Justification of effort is commonly observed in human behavior. Series of studies have showed how it takes place in different settings. Aronson and Mills (1959) first conducted an experiment in which college students were required to speak sexually arousing words out loud to join group discussion. The study found that those who endured the most severe embarrassment test thought most highly of the discussion although it was intentionally designed to be "dull and banal" (p. 179). Participants who underwent tough initiation chose to overestimate the attractiveness of the dull discussion to reduce dissonance, thus justifying their effort. Gerard and Mathewson (1966) conducted a replication study and found that those who received severe electrical shock (severe initiation group) expressed high satisfaction on following discussion compared to those who received mild shock or no shock.

Other studies also have consistently demonstrated the existence of effort justification. Klein, Bhatt, and Zentall (2005) report how people favored shapes that followed high effort (clicking on a mouse 20 to 30 times). The finding is intriguing in that a sheer presence of effort drove people to value certain shapes, which are usually meaningless and are not subject to value judgment. The input of effort transformed something random into ones that are special and likable.

Benozio and Diesendruck (2015) further discover that justification of effort occurs even among preschool children, suggesting that it is a natural instinct which people are born with. Even for socially-desirable values (i.e., outcome is not the only thing that matters) to which preschool children have not been exposed or educated to conform, children naturally exhibit effort

justification. That is, children appraise the value of the unattractive reward after effortful task. Even more surprising is that non-human species such as pigeons and rats also demonstrated preference towards rewards that were harder to obtain (Lydall, Gilmour, & Dwyer, 2010; Zentall, 2010).

As a frequently occurring phenomenon, effort justification could account for the higher audience ratings than netizen ratings. Study 1 explained that audience and netizen groups were not homogenous. The audience group was guaranteed by Naver Movie to have actually paid for movie tickets and visited a brick-and-mortar theater to watch films. The netizen group was not guaranteed so, implying that the group may have included those who did not watch films at a theater or those who did not officially pay for the movie (e.g., downloading pirated movie at home). It is possible that the main difference between the two groups is the amount of effort that they have put in watching a movie. Compared to the netizen group, the audience group had to spend considerable time and money (effort) to purchase the tickets and visit a theater. Therefore, it may be that the audience group rated movies more positively to compensate for the invested effort. Study 2 tested the idea that liking toward a film would increase in presence of effort, compared to absence of effort.

Hypothesis

The primary objective of Study 2 is to examine the impact of effort justification in movie ratings. An experiment was designed to test whether the investment of effort would directly affect perceived value of film, thus influencing the ratings. After presenting two different movie clips, we asked the experimental group to write comments about the movies (i.e., requirement of

effort) while the control group did not write any comments. Then, we compared how the control group and experiment group evaluated the films.

Hypothesis 2: People who write comments after watching a movie will rate the film more highly than those who do not.

Figure 2 below illustrates the conceptual relationships between the proposed variable regarding H2.

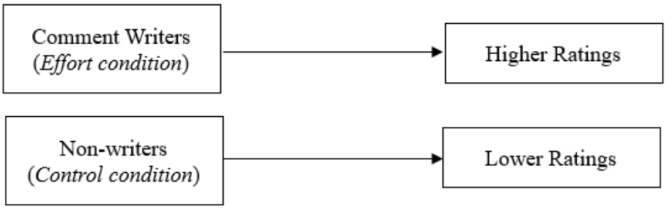

Figure 2. The conceptual model of the relationships among the variables in Study 2

Method

Data Collection Procedure

We selected two movies of different genres and presented them to the participants. The first film (Movie 1) was an independent animation *El Empleo* (2008, dir. Santiago Bou Grasso), which is generally gloomy and calm. The second clip was a commercial action film *The Hunger Games* (2012, dir. Gary Ross) (Movie 2), which is tense and vibrant. The reason choosing movies different in type and style was to reduce confounding influence of personal taste. The two films were presented to participants consecutively. For *the Hunger Games,* because of time constraint for the study, we were not able to play the entire film. Instead, we extracted a well-known scene where the protagonists start to play a survival game. In the case of *El Empleo*, we were able to show the entire film because its running time is less than 6 minutes. For

both movie clips, there was no dialogue. This was intentionally designed to prevent a language barrier which can potentially affecting film watching experience.

Participants were 126 undergraduates and 3 graduate students attending Korea University at Seoul. Of the participants, 90 were women, 32 were men, and 7 did not indicate their gender. Sixty nine of them identified themselves as Koreans while 60 said they were non-Koreans (e.g., international students from China, the U.S., Denmark, Hong Kong, Indonesia, etc.) The participants' age ranged from 20 to 32, with the median age being 23.

The participants were randomly assigned into two groups. In the control condition, 56 students were asked to give ratings on two films right after they watched the films. On the other hand, 73 students in the experimental group were instructed to write reviews and comments on the films and rate them afterwards. To be specific, we asked the participants to write about their favorite scenes, lines, and cinematography so that they would have to think about the films as carefully as possible. There was no interaction between participants in the control group and those in the experimental group.

Results

Hypothesis 2 predicted that those who wrote comments or reviews on films will grant higher ratings than those who did not. The result showed inconsistency; for Movie 2, the average rating of people who wrote comments (experimental group) was 1 point higher than that of people who did not (control group). To specify, the experimental group granted average rating of 7.08 (SD = 1.54) while the control group granted average rating of 6.08 (SD = 2.39), t (127) = 2.87, p = .005. On the other hand, for Movie 1,

there was almost no difference in average rating given by two groups. The experimental group gave 7.73 points on average (*SD* = 1.72) while the control group gave 7.65 points on average (*SD* = 1.96). Although the average rating by experimental group was slightly higher, t-test showed that 0.08 point difference is not significant at all, t (127) = 0.25, p = .80. Thus, for Movie 2, the data were consistent with hypothesis 2, but for Movie 1, the data were inconsistent with hypothesis 2.

The above analyses may appear to produce statistical results leading to a conclusion of partial support for the hypothesis 2. However, considering that this study used two different types of film as the stimulus materials, it is possible that the partial support for the hypothesis was not due to simply writing comments or not. To explore the effect of effort justification further, additional analyses were conducted as described below.

Researchers of this study assessed the length of comments written by participants in the experimental group and correlated that with their ratings. The length of film review comments, as an indicator of the degree of effort, however, was not related to ratings, r (71) = .15, p = .21 (Movie 1), r (71) = .16, p = .18 (Movie 2). The only significant correlation found was that between the lengths of comments on Movie 1 and Movie 2, r (71) = .64, p < .01. That is, those who wrote lengthy comments for Movie 1 also did for Movie 2, and vice versa. Table 4 illustrates the correlations among the variables.

Table 4. Correlations between Length of Comments and Film Ratings

	Comment Length (Movie 1)	Comment Length (Movie 2)	Ratings (Movie 1)	Ratings (Movie 1)
Comment Length (Movie 2)	.64**	--		
Ratings (Movie 1)	.15	.09	--	
Ratings (Movie 2)	.19	.16	.04	--
M	8.55	8.03	7.69	6.65
SD	2.05	2.67	1.82	2.01

*$p < 0.05$, ** $p < 0.01$

Note. The degrees of freedom range from 71 to 127.

To further explore the data, we conducted a series of additional analyses. First of all, we coded and divided the comments into three types—positive, negative, and neutral—reflecting the degree of favorability towards the movies. Then, for both movies, post hoc analyses examined type of comments and ratings to see if the process of writing positive or negative reviews polarized the ratings. For instance, it may be that some people felt moderately unfavorable towards a film at first but as they wrote comments, they had to recall what they did not like about the film, thus the degree of un-favorability increased. If this happened, it is possible that people who wrote negative comments granted particularly low ratings; thus it contributed to lowering the average film rating, nullifying the effect of effort. However, the post hoc analysis showed that positive or negative comments did not greatly affect how they rated the films afterwards. For both movies, positive comment writers and negative comment writers did not differ from one another in their ratings of the movies (1.27 point gap, 1.72 point gap respectively), F (2, 70) = 3.08, p = .052, η^2 = .08 (Movie 1), F (2, 70) = 11.51, p < .001, η^2 = .25 (Movie 2). In other words, participants' ratings were not influenced by the contents of their own reviews. Therefore,

the post hoc analysis did not explain why the effect of effort justification was shown only for Movie 2.

As the last attempt, we wondered if familiarity would affect evaluation of a film. This speculation arose because quite a few people in the experimental group reported in their comments that they have watched *The Hunger Games* before. It was interesting to see that the questionnaire only asked for general impressions and memorable scenes, but people voluntarily mentioned their prior exposure to the film. Then, it is possible that the familiarity towards *The Hunger Games* that comes from prior experience triggered favorability, thus increasing the average ratings. We categorized the participants into three groups: those in the experiment group who reported or implied in their written comments that they have already watched *The Hunger Games* before, those in the experiment group who did not indicate any clue regarding prior exposure, and the control group.

Testing the effect of familiarity, ANOVA analysis yielded a significant result, F (2, 124) = 6.07, p = .003, η^2 = .089. As shown in figure 3, post hoc comparisons using Tukey's HSD test at p = .05 showed that those who already watched the film ($M = 7.70_{b,}$ SD = 1.67) evaluated the film more highly than the control group ($M = 6.08_{a,}$ SD = 2.39). The presence of effort itself (writing reviews), however, did not make the ratings ($M = 6.68_{ab,}$ SD = 1.34) to be different from other conditions, rejecting hypothesis 2.

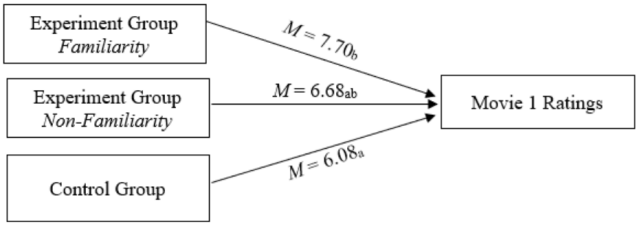

Figure 3. The relationship between familiarity and ratings on Movie 1 (*The Hunger Games*)

Discussion

The experiment revealed that investment of effort does not primarily affect how people evaluate a movie. Thus, the hypothesis that justification of effort would occur to people who write comments was not supported. Nonetheless, this study found that familiarity was significantly related to ratings. It was shown that the level of liking was higher for those who already watched the presented film before participating in the current study

Implication

The finding leads back to the implication of Study 1 that it is important for movie directors and advertisers to make their movies known to people. While Study 1 revealed how WOM are associated with positive sales, Study 2 discovered that simply being familiar with a film leads to higher evaluation. The result is consistent with preceding studies that stated the linkage between familiarity and liking. As people are continuously exposed to something, they start to favor it even when they are not conscious of the exposure (i.e., the mere exposure effect) (Ballard, Hennigan, & McClure, 2017; Bornstein & D'agostino, 1992;

Montoya, Horton, Vevea, Citkowicz, & Lauber, 2017). Then, a good way to attract audience may involve increasing exposure to a film. If one encounters a movie poster repetitively on the bus, subway, and billboard, the possibility that he chooses to watch that film would increase.

The findings from Study 1 and Study 2 together indicate that WOM may be used as a tool to create frequent exposure. Suppose a newly-released film gained popularity among active users of social networking services (SNS). The SNS users inspired by the film will post positive comments on their SNS. The information spreads and creates WOM, triggering more people to share their own opinions about the film with even a greater number of others. Consequently, a random person's likelihood of exposure to some kind of information about the movie goes up, leading to increase of familiarity unconsciously or consciously.

Limitations

The current study has at least two limitations. One has to do with individual differences in favorite types of movies. Although two different films were presented to the participants, there still exists a possibility that participants' personal taste in movies might have influenced the findings. There are numerous film genres, and everyone has his or her own preferences. For indie movie fans, action or thriller movies may be dissatisfying. Similarly, for action movie lovers, documentary films would be perceived as boring. Therefore, future research would have to consider people's movie preferences and present several types of films. Another limitation concerns the design of the questionnaire. Because we did not include a question asking whether the participants watched the presented films or not, we had to infer it from the participants' written comments. Even when they already

watched any of the films before, they might have not written anything indicative of it. For a future study, researchers should assess participants' previous knowledge and familiarity on each movie to be used as research stimulus.

Conclusion

Study 1 and 2 both underline the importance of utilizing WOM as a powerful promotional tool for a movie. As more and more people engage themselves in creating, sharing, and transforming information, the impact of WOM would continue to grow. The current research showed that online movie rating was positively associated with the number of movie goers. Although the current research did not clarify the reasons for audience ratings being more influential than other types of ratings, continuing this line of research has potentials to improve understandings of people's online behaviors and uncover effective marketing strategies.

References

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*, *59*, 177-181.

Asada, A., & Ko, Y. J. (2016). Determinants of word-of-mouth influence in sport viewership. *Journal of Sport Management*, *30*, 192-206.

Baek, H., Oh, S., Yang, H. D., & Ahn, J. (2017). Electronic word-of-mouth, box office revenue and social media. *Electronic Commerce Research and Application*, *22*, 13-23.

Ballard, I. C., Hennigan, K. & McClure, S. M. (2017). Mere exposure: Preference change for novel drinks reflected in human ventral tegmental area. *Journal of Cognitive Neuroscience, 29*(5), 793-804.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing*, *67*(4), 103-117.

Benozio, A., & Diesendruck, G. (2015). From effort to value: Preschool children's alternative to effort justification. *Psychological Science, 26*, 1423-1429.

Bornstein, R. F., & D'agostino, P. R. (1992). Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology*, *63*, 545-552.

Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, *14*, 350-462.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, *43*, 345-354.

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science*, *29*, 944-957.

Day, G. S. (1971). Attitude change, media and word of mouth. *Journal of Advertising Research, 11*(6), 31-40.

Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing, 21*, 23-45.

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. *International Journal of Research in Marketing, 25*, 215-224.

East, R., Hammond, K., & Wright, M. (2007). The relative incidence of positive and negative word of mouth: A multi-category study. *International Journal of Research in Marketing, 24*, 175-184.

Festinger, L. (1962). *A theory of cognitive dissonance* (Vol. 2). Palo Alto, CA: Stanford University Press.

Flanagin, A. J., & Metzger, M. J. (2013). Trusting expert-versus user-generated ratings online: The role of information volume, valence, and consumer characteristics. *Computers in Human Behavior, 29*, 1626-1634.

Forman, C., Ghose, A. & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*(3), 291-313.

Gerard, H. B., & Mathewson, G. C. (1966). The effects of severity of initiation on liking for a group: A replication. *Journal of Experimental Social Psychology, 2*, 278-287.

Ha, H. Y. (2002). The effects of consumer risk perception on pre-purchase information in online auctions: Brand, word-of-mouth, and customized information. *Journal of Computer-Mediated Communication, 8*(1), JCMC813. https://doi.org/10.1111/j.1083-6101.2002.tb00160.x

Henning-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004).

Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, *18*, 38-52.

Holbrook, M. B. (1999). Popular appeal versus expert judgments of motion pictures. *Journal of Consumer Research*, *26*, 144-155.

Hu, N., Koh, N. S., & Reddy, S. K. (2014). Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. *Decision Support Systems, 57*, 42-53.

Hsu, Y. L., & Jane, W. J. (2016). Bidirectional causality for word of mouth and the movie box office: An empirical investigation of panel data. *Journal of Media Economics*, *29*, 139-152.

Huang, M., Ali, R., & Liao, J. (2017). The effect of user experience in online games on word of mouth: A pleasure-arousal-dominance (PAD) model perspective. *Computers in Human Behavior*, *75*, 329-338.

Khare, A., Labrecque, L. I., & Asare, A. K. (2011). The assimilative and contrastive effects of word-of-mouth volume: An experimental examination of online consumer ratings. *Journal of Retailing*, *87*, 111-126.

Kim, S. H., Park, N., & Park, S. H. (2013). Exploring the effects of online word of mouth and expert reviews on theatrical movies' box office success. *Journal of Media Economics*, *26*, 98-114.

Klein, E. D., Bhatt, R. S., & Zentall, T. R. (2005). Contrast and the justification of effort. *Psychonomic Bulletin & Review*, *12*, 335-339.

Koschat, M. A. (2012). The impact of movie reviews on box office: Media portfolios and the intermediation of genre. *Journal of Media Economics*, 25, 35-53.

Lydall, E. S., Gilmour, G., & Dwyer, D. M. (2010). Rats place greater value on rewards produced by high effort: An animal analogue of the "effort justification" effect. *Journal of Experimental Social Psychology*, *46*, 1134-1137.

Montoya, R. M., Horton, R. S., Vevea, J. L., Citkowicz, M., & Lauber, E. A. (2017). A re-examination of the mere exposure effect: The influence of repeated exposure on recognition, familiarity, and liking. *Psychological Bulletin, 143*(5), 459-498.

Murtiasih, S., & Siringoringo, H. (2013). How word of mouth influence brand equity for automotive products in Indonesia. *Procedia-Social and Behavioral Sciences*, *81*, 40-44.

Park, J. A., Sung, J., Kim, M. K., & Kim, S. K. (2016). Effect of service quality on revisit intention and word-of-mouth at the F-1 Chinese grand prix. *International Journal of Applied Sports Sciences*, *28*, 42-52.

Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information Systems Research, 17*(4), 392-414.

Reinstein, D. A., & Snyder C. M. (2005). The influence of expert reviews on consumer demand for experience goods: A case study of movie critics. *The Journal of Industrial Economics*, 53, 27-28.

Sung, Y., Kim, Y., Kwon, O., & Moon, J. (2012). An explorative study of Korean consumer participation in virtual brand communicates in social network sites. *Journal of Global Marketing, 23*, 430-445.

Suárez-Vázquez, A. (2011). Critic power or star power? The influence of hallmarks of quality of motion pictures: An experimental approach. *Journal of Cultural Economics*, *35*, 119-135.

Tsao, W. C. (2014). Which type of online review is more persuasive? The influence of consumer reviews and critic ratings on moviegoers. *Electronic Commerce Research*, *14*, 559-583.

Zentall, T. R. (2010). Justification of effort by humans and pigeons: Cognitive dissonance or contrast? *Current Directions in Psychological Science*, *19*, 296-300.

Submission: September 30, 2017
Review: October 1, 2018
Decision of publication: December 3, 2018