

How Discussions Develop Online: The Relationship between Comments and Replies in News Portal Sites

Jung Lee

Hankuk University of Foreign Studies

jung.lee@hufs.ac.kr

Jinyoung Min

Chung-Ang University

jymin@cau.ac.kr

ABSTRACT

Although the online commenting system is generally regarded as a place for discussion, there should be empirical evidence that its components, such as comments, responses, and replies, are not unrelated fragments of the system but interrelated elements of the discussion. Using Habermas' public sphere theory and data from Korean news portal sites, we demonstrate that comments, responses, and replies all contribute to an online discussion by being in the discussion stages. Overall, the analysis of their relationship illustrates that comments that are likely to receive responses are more likely to receive replies. Furthermore, more responses and replies are followed when people disagree with the comments. Therefore, our findings show that comments, responses, and replies build the online discussion and that people discuss more when they disagree, suggesting that the commenting system becomes the place for online discussion.

Keywords: online discussion, online comments, news portal, deliberation, discussion stages.

1. INTRODUCTION

News portal sites provide engaging platforms among all the useful online communication platforms [1]. Particularly, they become a hot spot for building and spreading public opinion when the news concerns popular social issues, such as presidential elections and international wars [2]. The news portal sites spread public opinion through comments and replies attached to news articles, which are presented and read as public opinions [3,4]. We can observe how public opinions emerge through the comments and replies after the news. These comments are gathered and accumulated to form public opinions with highly valued collective intelligence [5].

Despite the popularity of news portal comment systems as public opinion incubators, how people comment on and reply to discussions has not been thoroughly studied. Although numerous studies on online comment systems exist, the majority of them treat comments as a single group of opinions without considering the structure and variety

within them. It is necessary to understand comment systems based on their structure because they demonstrate how discussions are initiated and progress [6,7].

Hence, in this study, we examine the comment systems of online news portals from a structural perspective. We consider people's comments on articles and the responses and replies to these comments. We estimate the possibility of quality discussions on online news portals by analyzing various measures, such as response and reply rates. Furthermore, we examine whether these platforms provide an adequate medium for quality discussions and demonstrate the value of comments as the outcome of discussions and as public opinion.

To define discussion, we introduce Habermas' public sphere [8]. This theory provides the foundations for communication, discussion, and social consensus. Furthermore, it defines discussion, emphasizes the importance of social consensus among participants, and discusses the implication of public spheres as a discussion platform. This theory benefits numerous studies on online platforms because it provides useful and comprehensive frameworks for online communications [9]. The concept of public spheres is used in this study to structure the comment systems of online news portal sites.

This paper is organized as follows. We begin by introducing Habermas' public sphere and analyzing the comment systems of online portal sites based on communication stages. Next, using the number of comments and replies as a parameter, we formulate hypotheses to show how online discussions progress. The difference between "likes" and "dislikes" will be used to determine the level of social agreement. To validate our hypotheses, we collected and analyzed over 300,000 comments from 150 news articles on a major Korean news portal site. Finally, we discuss the results and implications of the study.

2. THEORETICAL BACKGROUND

2.1 Online public spheres: The role of social consensus

A public sphere, as proposed by Jurgen Habermas in his 1962 book "Structural Change of the Public Sphere," involves a space for creating public opinions that are formed solely through individuals' free interaction, excluding any instrumental actions, such as labor [8]. Examples include newspapers, salons, and coffee shops. In these public spheres, social consensus on members' common interests is drawn through rational discussion [10].

Achieving social consensus is one of the objectives of the public sphere [11]. Debates are an example of a deliberative process for achieving social consensus [12]. Deliberation refers to the process of making decisions based on accurate information and fairly considering social diversity [13]. It includes listening to others and attempting to reach a consensus based on empathy. Social consensus is distinguished from a simple agreement of opinion. To understand social consensus, concepts such as "contestant" and "conflict" have been jointly studied for comparison with "consensus" [14,15].

Because online environments, such as SNS (social network sites or services), have structural similarities with public spheres, such as ease of entry and voluntary participation, Habermas' public sphere theory has been widely cited in research on platforms and online communities [16]. Furthermore, online environments generate various measures that aid in understanding the public sphere [8]. For example, the number of comments represents the enthusiasm for discussion, and the gap between "likes" and "dislikes" demonstrates the disparity between opinions in society [17]. Furthermore, it provides theoretical foundations for online discussion by showing how it achieves social consensus through conversations. Our study is based on the theory's core assumption that the public sphere also exists in the online environment

2.2 Discussion stages on online portals

Figure 1 illustrates the structure of typical online commenting systems on platforms such as YouTube. It shows how the platform supports people-building discussions in stages—first, Stage 0 starts when an article or a video is uploaded to the platform. Second, Stage 1 begins when some users who read the article or watch the video leave comments to express their opinions. Third, Stage 2 involves the users responding to the comments. People who read the comments may respond by clicking the like or dislike buttons to express their ideas. Finally, Stage 3 starts when people who read the comments reply to add their thoughts to the conversation. Stage 0 involves preparing the ground for discussion on the platform. The actual discussion is initiated in Stage 1 by people leaving comments. Table 1 summarizes these stages and activities with stage specifications.

Understanding online comment systems (Figure 1 and Table 1) has two implications. First, it describes online discussion activities from a stage perspective. This stage perspective informs users that not all comments develop into discussions and that only comments that receive responses and replies "survive" and develop into discussions [18]. If most of the comments are read adequately and developed into discussions, it may imply that the comment system adequately supports deliberative discussion processes. However, the system can be considered inactive discussion support if most comments do not often develop into online discussions and remain as a single comment with no response or reply.

Second, it highlights the quantifiable parameters in online discussions. Most portal systems provide conspicuous numerical parameters, such as the number of views, likes, and dislikes. These are significant measures that may represent the quality of discussions [19]. For example, comments with a high number of likes and dislikes are perceived as representing the opinions of a large audience [4, 20, 21]. The sum of the likes and dislikes represents the audience's overall response levels, while the difference between the two represents the opinion disparity or the overall agreement level [22]. The portal system may display the number of people who comment on the article and reply to the comments. These figures correspond to the overall participation level in discussions. These implications provide a basis for developing the hypothesis in subsequent sections.

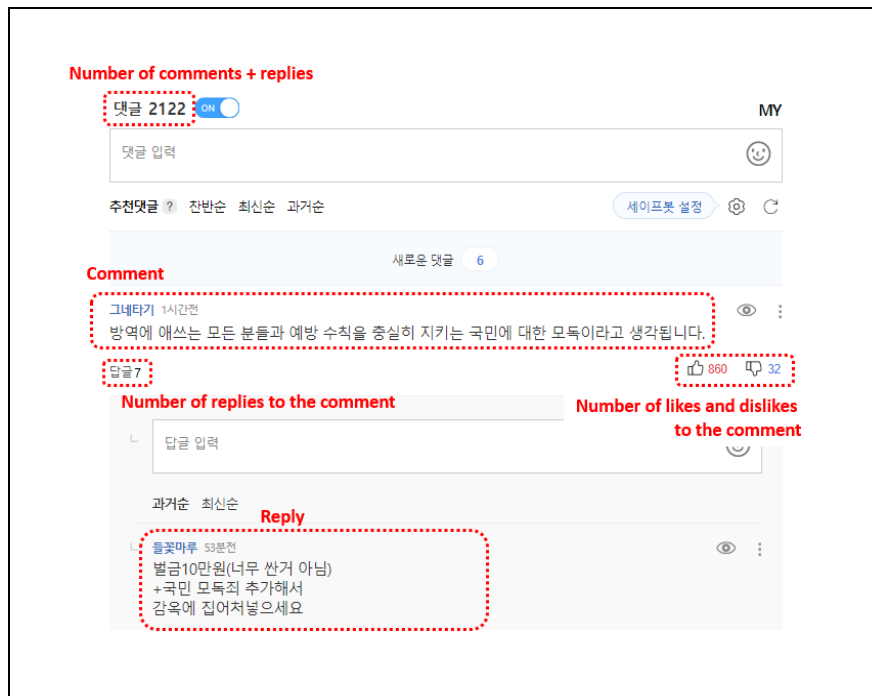


Figure 1. How comments, responses, and replies are displayed at the end of a news article

Table 1. Discussion stages in an online news portal

Stages	Terms	Actions	Features	Behavioral perspective	Provided parameters
Stage 0	News article and video upload	Topic specification	Each newspaper company has their characteristic	Gathering of an interesting topic	Total number of comments and replies, showing the overall participation level
Stage 1	Comment	Topic-related opinion postulation	Free access, id-trackable, multiple writing possible	Response to news; Personal opinion on the article	
Stage 2	Respond	Response to comment	Free access, non-id-trackable, like-dislike selective	Agree or disagree with the comment	The sum of likes and dislikes shows the overall response level; the difference between likes and dislikes shows the overall agreement level
Stage 3	Reply	Reply to comment	Free access, id-trackable, multiple writing possible (No possible response to a comment)	Stronger response; another opinion with higher motivation	

3. HYPOTHESES DEVELOPMENT

3.1 Stage Development: More responses lead to more replies

An online discussion is conceived when one user comments on an article on a portal. However, not all comments develop into online discussions because not all comments receive responses. Only “some” comments that attract people’s interest receive responses and acquire replies from others [23]. Responses and replies are the two representative forms of online discussion. Replies take the linguistic form and require writing opinions for the discussion. Responses take the form of an aggregated number that represents the people's overall agreement/disagreement level and require clicking the like/dislike buttons. In this sense, responses are called click speech, which expresses opinions simply and effectively in online discussions [24]. When those responses and replies are accumulated and woven into a thread of public opinions, it becomes an online discussion based on the interactions and communications among the participants.

Figure 1 and Table 1 depict three stages of online communication: commenting, responding, and replying. This structure provides a basis for the systematic development of online discussions. If people do not “discuss” online but instead express themselves without listening to others, the relationships between the comments, responses, and replies will be chaotic and amorphous, without any pattern. Conversely, if people discuss online, the comments that receive more responses are likely to receive more replies based on the stages of the online discussion. Consequently, we propose the following hypotheses.

H1a: For an article, the probability of a comment receiving at least one reply increases as the probability of the comment receiving at least one response increases.

The number of responses and replies pinned to a specific comment represents the intensity of the online discussion [25]. Comments with a high potential for developing into online discussions gain more responses and replies. H1a articulates the probability of a comment developing into a discussion by comparing the comments that develop and those that do not develop into discussions. The intensity of the discussion is articulated by focusing on the average number of responses and replies. Comments with high discussion potential will elicit more responses and replies than comments with low potential. As a result, we propose the following hypothesis:

H1b: For an article, the average number of replies to a comment increases as the average number of responses to the comment increases.

3.2 Distribution of comments, responses, and replies

People who reply to comments by writing words are usually more motivated than those who only respond by clicking like or dislike [26]. Replying is more time-consuming, effortful, and informative than responding. One click is enough to respond; however, words and thoughts must be prepared to reply. When people who have read the comments intend to express their opinions, they often give one click first to respond—

to like or dislike—and then some proceed to reply in case they have anything further to add. Repliers are distinguished from responders based on their motivations. When we compare the people who respond and reply to comments with consumers, the responders are ordinary consumers, whereas the repliers are loyal consumers with higher motivations.

This subset relationship between the responders and the repliers gives the distribution of responses and replies an interesting shape. Because repliers are more motivated and loyal to the discussion than responders, the replier group will exhibit more stable participation attitudes than responders. In other words, the replier group will be more consistent than the responder group. Loyal customers are hard to get but are more faithful and stable in their behaviors [27].

The distribution's variance is reduced because of the repliers' stability. The possibility that comments will receive replies is more stable than the possibility that they will receive responses, meaning that people who reply initiate their behaviors more regularly than those who respond. Thus, we have the following hypothesis:

H2a: The distribution of the probability that a comment receives at least one reply exhibits a larger kurtosis than the distribution of the probability that a comment receives at least one response.

The average number of replies and responses per article will exhibit similar distribution patterns. H2a articulates the “initiation” of the responding and replying behaviors, while H2b discusses the “intensity” of the behaviors. People who reply more regularly and routinely are more likely to initiate their behavior than those who respond because of their superior loyalties. Thus, the average number of responses and replies will exhibit different patterns. From this, we have the following hypothesis:

H2b: The distribution of the average number of replies per article exhibits a larger kurtosis than the distribution of the average number of responses per article.

3.3 Consensus level as a discussion accelerator

People discuss in public spheres in order to achieve social consensus [28]. The public sphere participants share their ideas and hold discussions because they may have different opinions but want to reach a consensus by persuading others. Historically, social confusion occurred frequently when the discussion medium did not function properly and a social consensus was not attained [29]. When opinions differ, discussion can help to avoid violence and coercion [30].

A discussion commences when people respond and reply to others' comments. When people in a group have a low agreement level on a certain issue, the motivation to start a discussion is relatively high. When people have diverse opinions and reaching a social consensus seems difficult, they are strongly motivated to engage in a discussion to reach a social consensus. Conversely, people will be less motivated to add their opinions and reply to comments if they have similar opinions.

Such high motivation for discussion is quantified in online environments by the

frequency of responses and replies. A low agreement level and a high disagreement level will promote discussion. Based on the stage development perspective presented in Table 1, comments at the initial stage will have a high probability of receiving responses and replies from readers if people generally have different opinions. Accordingly, the following hypotheses are proposed:

H3a: The lower the overall agreement level, the higher the probability that a comment will receive at least one response.

H3b: The lower the overall agreement level, the higher the probability that a comment will receive at least one reply.

When participants realize their opinions differ, they tend to insist on them or add more. Offline, when opinions differ, people tend to raise their voices and talk more and faster than when they are similar. In other words, when the social consensus level is low, discussion participants add their opinions, increasing the number of added opinions. From these, the following hypotheses are proposed.

H4a: The lower the overall agreement level is, the higher the average number of responses per article will be.

H4b: The lower the overall agreement level is, the higher the average number of replies per article will be.

4. METHODOLOGY AND DATA COLLECTION

We collected data from the Daum news portal (<http://media.daum.net>) in Korea to validate the hypotheses. Daum is selected for the following reasons. First, it is one of Korea's two largest news portals and has been providing news portal services to Korean internet users since 1999. It is famous for its active comment system, which is supported by enthusiastic users [31]; its news section page recorded 87,896,000 views per month as of June 2020 [32]. Second, until January 2022, it provided the most commented news list. Sampling was based on this list because it strengthens the data's reliability and validity by maintaining the large scales and the popularity of the sources.

We collected data from the daily top 50 most-commented news rankings provided by the portal. We chose the weekend over the weekdays because people use the portal more on weekends than they do on weekdays. June 12-14, 2020, constituted the dates of the news publications, and data were collected in mid-July 2020. This one-month time gap between the generation and the collection of the data guarantees that the data are collected after the discussions are closed.

We checked for news that could generate significantly dominant topics, such as a global earthquake or a presidential demise. Such extreme and rare news should be avoided because it can overwhelm discussions on online portals. For example, a single article of such extreme news often receives over 20,000 comments from users, and it dominates the overall tone of the discussions. In such cases, strong words are used, and opinions are often one-sided rather than diverse.

During those three days, the most significant news was about a mother who killed her stepson and husband (i.e., the Ko Youjeong case), and it had a maximum of 13,000 comments for a single article. Based on this number, we concluded that it would not overwhelm the overall commenting system enough to cause bias. In total, 150 news articles and their comments were collected. These processes confirm that the resulting data are not biased in their popularity or topics.

Table 2 describes the categories of the 150 news articles provided by the portal. Among them, the largest category was social issues, including topics like coronavirus, home and work, and small business issues. The second-largest category was politics, which included congress issues and politicians' actions. Social issues and politics accounted for almost 90% of all the news collected.

Table 2. The categories of 150 news articles

Sections	Frequency	Percentage (%)	Cumulative percentage (%)
Social issues	75	50.0	50.5
Politics	56	37.3	87.3
International	9	6.0	93.3
Economics	6	4.0	97.3
Technology	1	.7	98
Sports	1	.7	98.7
Culture	1	.7	99.3
Series	1	.7	100.0
Total	150	100.0	

5. DATA ANALYSIS

5.1 Descriptive Study

We prepared various parameters to test the hypotheses (Table 3). Four statistics were directly obtained from the raw data: *com*, *rep*, *a*, and *b*. The remaining statistics were further manipulated based on these four numbers. The average number of comments per article was 2,105. Approximately 31% of the comments received at least one response from the users, whereas only 11% received at least one reply. These numbers reveal that 70% of the comments did not receive any response from others, and 90% did not receive any reply from other users.

Table 3. Descriptive study

Statistics	Description	Mean	SD	Min	Median	Max
com	Number of comments on the article	2,105.03	1,713.34	723	1,493	13,255
rep	Number of replies to comments on the article	295.493	241.233	44	217.5	1,801
res (=a+b)	Number of responses to comments on the article	19,516.9	14,115.2	1,913	15,815	101,716
a	Number of the agreements with all the comments on the article	17,916.0	13,193.9	1,847	14,084	98,431
b	Number of disagreements with all the comments on the article	1,600.93	2,432.13	50	1,018	20,051
res1	Number of comments that received at least one response on the article	602.367	454.802	110	449.5	2,655
rep1	Number of comments that received at least one reply on the article	65.3667	48.0897	11	51	312
M_res (=res/com)	The average number of responses to the comments on the article	10.2908	4.9810	1.4217	9.2665	27.722
M_rep (=rep/com)	The average number of replies to the comments of the article	0.1619	0.0974	0.0111	0.1361	0.5885
P_res1 (=res1/com)	The proportion of comments that received at least one response among all the comments on the article	0.3078	0.1183	0.0268	0.2961	0.6278
P_rep1 (=rep1/com)	The proportion of comments that received at least one reply among all the comments on the article	0.1146	0.0369	0.0416	0.1087	0.2455
P_a (=a/res)	The proportion of agreement on responses to comments on the article	0.9114	0.0846	0.6495	0.9426	0.9938

5.2 H1 test

We formulated and tested single-variate regression equations to validate H1 (Tables 4 and 5). The result reveals that as the average number of responses to the comments on an article (i.e., M_res) increases, so does the average number of replies to the comments (i.e., M_rep) (Sig. <0.01). The change in the average number of responses explained by the average number of replies is 37.5%. Therefore, H1a is supported. Moreover, it shows that as the proportion of comments that received at least one response (i.e., P_res1) increases, so does the proportion of comments that received at least one reply (i.e., P_rep1) (Sig. <0.01). The change in P_rep1 that P_res1 explains is 63%. Therefore,

H1b is supported.

Table 4. H1 model summary

Testing H	Model	R (=correlation coefficient)	R squared	Adjusted R squared
H1a	X=M_res, Y=M_rep	.616	.379	.375
H1b	X=P_res1, Y=P_rep1	.795	.632	.630

Table 5. H1 regression coefficients

Testing H	Model	Variables	Unstandardized coefficients		t	Sig	Result
			B	Std. error			
H1a	X=M_res, Y=M_rep	(constant)	.038	.014	2.620	.010	Supported
		M_res	.012	.001	9.512	.000	
H1b	X=P_res1, Y=P_rep1	(constant)	-.003	.003	-1.016	.311	Supported
		P_res1	.124	.008	15.950	.000	

5.3 H2 test

To validate H2, we checked several parameters defining distributions' shape, including kurtosis. Kurtosis indicates the thickness of the distribution's tails relative to a normal distribution. An increase in the numbers lightens the tails of the distributions. As shown in Table 6, the replies' kurtosis is higher than that of the responses (i.e., 3.009 > 2.043), meaning that replies have lighter tails and, accordingly, thicker centers than responses. This result is consistent with H2. However, other measures, such as coefficients of variance and interquartile range, showed controversial results. Therefore, we concluded that H2 is partially supported.

Table 6. Descriptive measures

Testing H	Variable	Coefficient of variance	Standardized			Skewness	Kurtosis	Result
			Max	Min	Inter-quartile range			
H2a	M_res	0.4840	3.500	-1.781	1.0184	1.277	2.043	Partially
H2a	M_rep	0.6018	4.379	-1.548	1.1701	1.509	3.009	Supported
H2b	P_res1	0.3843	2.706	-2.376	1.4042	0.455	-0.427	Partially
H2b	P_rep1	0.5187	3.980	-1.789	1.3721	0.908	1.135	Supported

5.4 H3 and H4 tests

To validate H3 and H4, we formulated and tested regression equations (Tables 7 and 8). First, as the proportion of likes among the total responses (i.e., P_a) increases, the proportion of comments that received at least one response (i.e., P_res1) decreases (Sig. <0.01). The change in P_a that P_res1 explains is 53.1%. Therefore, H3a is supported.

As the proportion of likes among the total responses (i.e., P_a) increases, the proportion of comments that received at least one reply (P_rep1) decreases (Sig. <0.01). The change in P_rep1 that P_a explains is 33.6%. Therefore, H3b is supported.

Second, as the proportion of likes among the total responses (i.e., P_a) increases, the average number of responses to comments on the article (i.e., M_res) does not decrease. Although the beta coefficient is positive, the significance level is less than 0.1. Therefore, H4a is not supported. Next, as the proportion of likes among the total responses (i.e., P_a) increases, the average number of replies to comments (M_rep) decreases (Sig. <0.01). The change in P_a explained by M_rep is 18.7%. Therefore, H4b is supported.

Table 7. H3 and H4 model summary

Testing H	model	R (=correlation coefficient)	R squared	Adjusted R squared
H3a	X=P_a, Y=P_res1	.729	.531	.528
H3b	X= P_a, Y=P_rep1	.580	.336	.332
H4a	X=P_a, Y=M_res	.144	.021	.014
H4b	X=P_a, Y=M_rep	.432	.187	.181

Table 8. H3 and H4 regression coefficients

Testing H	Model	Unstandardized coefficients		t	Sig	Result	
		B	Std. error				
H3a	X=P_a, Y=P_res1	(constant)	1.236	.072	17.170	.000	Supported
		P_a	-1.019	.079	-12.951	.000	
H3b	X=P_a, Y=P_rep1	(constant)	.151	.013	11.285	.000	Supported
		P_a	-.127	.015	-8.660	.000	
H4a	X=P_a, Y=M_res	(constant)	2.582	4.383	.589	.557	Not
		P_a	8.458	4.789	1.766	.079	Supported
H4b	X= P_a, Y=M_rep	(constant)	.615	.078	7.874	.000	Supported
		P_a	-.497	.085	-5.826	.000	

6. DISCUSSION

6.1 Summary of findings

Table 9. Hypotheses test results

Hs	Hypotheses	Result
H1a	H1a: For an article, the probability of a comment receiving at least one reply increases as the probability of the comment receiving at least one response increases.	Supported
H1b	H1b: For an article, the average number of replies to a comment increases as the average number of responses to the comment increases.	Supported
H2a	H2a: The distribution of the probability that a comment receives at least one reply exhibits a larger kurtosis than the distribution of the probability that a comment receives at least one response.	Partially supported
H2b	H2b: The distribution of the average number of replies per article exhibits a larger kurtosis than the distribution of the average number of responses per article.	Partially supported
H3a	H3a: The lower the overall agreement level, the higher the probability that a comment will receive at least one response.	Supported
H3b	H3b: The lower the overall agreement level, the higher the probability that a comment will receive at least one reply.	Supported
H4a	H4a: The lower the overall agreement level is, the higher the average number of responses per article will be.	Not Supported
H4b	H4b: The lower the overall agreement level is, the higher the average number of replies per article will be.	Supported

Of the eight hypotheses, five are fully supported, two are partially supported, and one is not supported. Based on these, we summarize the results as follows. First, online discussions progress in distinct stages. The combined results of H1 and H2 indicate that the creation of the replies and the responses are highly correlated, meaning that the intense responses lead to heated replies. Notably, the probability of receiving replies may be constant, while the probability of receiving responses may change if both actions are not correlated. We explain the reason for the correlation using a stage-based discussion mechanism. It demonstrates that people do not respond or reply routinely but instead carefully read the comments and rationally decide their subsequent reactions.

Second, the overall agreement level moderates the discussion activeness levels. The overall vitality of the discussion is depicted by the summation of the likes and dislikes. The difference between likes and dislikes shows the disparity in opinions. It represents the overall agreement level and moderates the discussion temperature. This moderation implies that people rationally and systematically participate in online discussions to express their opinions and voices. In that sense, the online portal's comment system provides a modest discussion platform for people.

6.2 Academic contribution

This study contributes to the academic field in the following aspects. First, it analyzes the features of online discussions from a stage perspective and shows that online discussions develop in stages. Most prior studies on online communication assumed that the comments attached to articles were the products of deliberative discussions instead of an amorphous pile of unstructured words [33]. However, this assumption has not often been empirically validated. Hence, this study proved that people actually “discuss” things online by showing how the discussion stages are correlated. Using various empirical data analyses to enhance its academic contribution, it theorizes the three stages (i.e., commenting, responding, and replying) and validates their relationships.

Second, by parameterizing the number of likes and dislikes, it validates the role of the social agreement in online discussions. The overall agreement level in a discussion group has long been considered critical to achieving social consensus [34]. Numerous studies have used the number of likes and dislikes to measure online behaviors [35]. However, few studies have combined these two, used the gap between likes and dislikes as a measure of agreement in a discussion, and tested its role. This study contributes to the research by equating these two essential concepts (i.e., the level of the social agreement in a discussion and the gap between the number of likes and dislikes) and verifying them using empirical analysis.

Last, this study presents the role of platforms as an online discussion medium. Numerous studies have discussed the role of platforms from various perspectives [36]. However, few studies have highlighted the role of platforms as discussion advocates. This study demonstrates how platforms support discussions and how they become active as they develop. Furthermore, it verifies several quantifiable measures of online discussions by ramifying the discussion stages into three stages (i.e., commenting, responding, and replying). While some portals seemingly transform themselves from discussion accelerators to information curators, platforms still need to advocate discussions in order to build and protect public opinions [37].

6.3 Practical implications

This study presents the following practical implications. First, it shows how users induce discussions with different weights on each communication stage on platforms. People often believe that all the words from users are true and meaningful. It was difficult to comprehend how each communication stage differs regarding their roles and significance in the discussion processes. Hence, this study informs people that comments are the outcomes of user discussions. Moreover, comments and replies should be comprehended differently based on the discussion stages. The number of comments represents the quantity and quality of discussions. Understanding comments from a stage perspective is important because it defines the objectives and future directions of major portal sites. Adopting this stage understanding for the strategy for commenting systems, for example, portal sites may consider offering information on how the stages are established, which can be executed by showing how a particular

reply is linked to a certain response. Distinguishing between the replies generated by the like response and those generated by the dislike response may provide the background information of the replies and encourage more and better discussions. Second, this study helps us understand how opinion disparity influences discussion vitality. Most organizations agree that many user comments are assets containing consumer opinions [38]. However, the opinion gaps within comments have not been comprehensively investigated, even though they can be discussion drivers to explain the different influences of having numerous comments. Discussion drivers are important because they increase the number of comments and enhance the possibility of reaching social consensus. This study addresses the question of whether people do meaningless things online or communicate and discuss. The latter is likely. According to this finding, platforms may analyze the most disagreed-upon comments and generate new content that demonstrates how they are argued. Because natural language processing techniques are rapidly evolving, embedding those techniques into commenting system to automatically analyze and show what topics people disagree on and how those topics are discussed can enrich the discussion. This can also be the platforms' strategy to create new services and attract more users. Third, this study provides strategic ideas for portals' content display management. Portals' content display strategy has always been business managers' primary interest [39]. For example, one of the major portals in Korea recently removed the news ranking lists to avoid the possible bias it could portray to the readers [40]. As a syndicator of the content and not the producer, they randomize the news displays to be politically fair, maintain variety, and guarantee slots to every newspaper company instead of giving incentives to large companies. This study can be helpful to those company managers because it provides a detailed view of the structure of comments and their relationships. In addition, portal sites can use our findings to use various strategies for displaying comments. The commenting systems on portal sites provide several options for sorting comments: by the most responded to, by the most liked, by the most disliked, and by the most recently uploaded. Portal sites may consider using not only responses but also replies as the sorting standard to encourage more discussion. If the comments with many replies are exposed more to people, people with opinions will be drawn to those comments, and intense discussion may be encouraged.

6.4 Limitations and future study

This study has the following limitations and scopes for future study. First, it focuses on immediately quantifiable measures, such as the number of responses and replies. However, other important quantifiable measures exist, such as topics and sentiments [41]. In the future, qualitative measures should be included in the study to comprehend online discussion environments and user behaviors fully. Second, to observe how discussions continue online, this study only selects news articles with numerous comments. However, in the future, news articles with few comments should be included to avoid possible bias. Lastly, the factors that accelerate or deteriorate the speed of discussions should be investigated in future studies. This study tested only one factor, the agreement level. However, other factors should be investigated, such as a specific

word, tone, or discussion direction.

ACKNOWLEDGEMENT

This work was supported by 2023 Hankuk University of Foreign Studies Research Fund. Jinyoung Min is the corresponding author and can be contacted at: jymin@cau.ac.kr

REFERENCES

- [1] M. Ziegele, T. Breiner, and O. Quiring, "What creates interactivity in online news discussions? An exploratory analysis of discussion factors in user comments on news items," *Journal of Communication*, Vol. 64, No. 6, pp. 1111-1138, 2014.
- [2] P. Weber, "Discussions in the comments section: Factors influencing participation and interactivity in online newspapers' reader comments," *New Media & Society*, Vol. 16, No. 6, pp. 941-957. 2014.
- [3] M. J. Metzger, A. J. Flanagin, and R. B. Medders, "Social and heuristic approaches to credibility evaluation online," *Journal of Communication*, Vol. 60, No.3, pp. 413-439, 2010.
- [4] C. Peter, C. Rossmann, and T. Keyling, "Exemplification 2.0," *Journal of Media Psychology*, Vol. 26, No.1, pp. 19-28, 2014.
- [5] T. B. Ksiazek, "Commenting on the news: Explaining the degree and quality of user comments on news websites," *Journalism Studies*, Vol. 19, No. 5, pp. 650-673, 2018.
- [6] C. Budak, R. K. Garrett, P. Resnick, and J. Kamin, "Threading is sticky: How threaded conversations promote comment system user retention," *Proceedings of the ACM on Human-Computer Interaction*, Vol. 1, No. 27, pp. 1-20. 2017.
- [7] F. Toepfl and E. Piwoni, "Public spheres in interaction: Comment sections of news websites as counterpublic spaces," *Journal of Communication*, Vol. 65, No. 3, pp. 465-488, 2015.
- [8] Z. Papacharissi, "The virtual sphere 2.0: The Internet, the public sphere, and beyond," in *Routledge handbook of Internet politics*, Routledge, pp. 246-261, 2008.
- [9] R. Zamith and S. C. Lewis, "From public spaces to public sphere: Rethinking systems for reader comments on online news sites," *Digital Journalism*, Vol. 2, No. 4, pp. 558-574. 2014.
- [10] L. Dahlberg, "Rethinking the fragmentation of the cyberpublic: from consensus to contestation," *New Media & Society*, Vol. 9, No. 5, pp. 827-847, 2007.
- [11] K. R. Phillips, "The spaces of public dissension: Reconsidering the public sphere," *Communications Monographs*, Vol. 63, No. 3, pp. 231-248, 1996.
- [12] M. E. Del Valle, R. Sijtsma, H. Stegeman, and R. Borge, "Online deliberation and the public sphere: developing a coding manual to assess deliberation in Twitter political networks," *Javnost-The Public*, Vol. 27, No. 3, pp. 211-229, 2020.
- [13] J. Fishkin, *When the People Speak*, Oxford University Press, 2011.

- [14] R. W. Martin, "Between consensus and conflict: Habermas, post-modern agonism and the early American public sphere," *Polity*, Vol. 37, No. 3, pp. 365-388, 2005.
- [15] P. Markell, "Contesting consensus: Rereading Habermas on the public sphere," *Constellations*, Vol. 3, No. 3, pp. 377-400, 1997.
- [16] N. Poor, "Mechanisms of an online public sphere: The website Slashdot," *Journal of Computer-Mediated Communication*, Vol. 10, No. 2, 2005.
- [17] J. Phua and S. J. Ahn, "Explicating the 'like' on Facebook brand pages: The effect of intensity of Facebook use, number of overall 'likes', and number of friends' 'likes' on consumers' brand outcomes," *Journal of Marketing Communications*, Vol. 22, No. 5, pp. 544-559, 2016.
- [18] M. A. Shapiro, and H. W. Park, "Climate change and YouTube: Deliberation potential in post-video discussions," *Environmental Communication*, Vol. 12, No. 1, pp. 115-131, 2018.
- [19] A. L. Burrow and N. Rainone, "How many likes did I get?: Purpose moderates links between positive social media feedback and self-esteem," *Journal of Experimental Social Psychology*, Vol. 69, pp. 232-236, 2017
- [20] T. K. Naab, D. Heinbach, M. Ziegele, and M. T. Grasberger, "Comments and credibility: How critical user comments decrease perceived news article credibility," *Journalism Studies*, Vol. 21, No. 6, pp. 783-801, 2020.
- [21] P. Porten-Che e and C. Eilders. "The effects of likes on public opinion perception and personal opinion," *Communications*, Vol. 45, No. 2, pp. 223-239, 2020
- [22] M. Walrave, K. Poels, M. L. Antheunis, E. Van den Broeck, and G. van Noort, "Like or dislike? Adolescents' responses to personalized social network site advertising," *Journal of Marketing Communications*, Vol. 24, No. 6, pp. 599-616, 2018.
- [23] K. Stsiampkouskaya, A. Joinson, L. Piwek, and C. P. Ahlbom, "Emotional responses to likes and comments regulate posting frequency and content change behaviour on social media: An experimental study and mediation model," *Computers in Human Behavior*, Vol. 124, pp. 1-14, 2021.
- [24] A. D. Sklan, "@ Socialmedia: speech with a click of a button-# socialsharingbuttons," *Cardozo Arts & Ent. LJ*, Vol. 32, p. 377, 2013.
- [25] S. Matook, A. R. Dennis, and Y. M. Wang, "User comments in social media firestorms: A mixed-method study of purpose, tone, and motivation," *Journal of Management Information Systems*, Vol. 39, No. 3, pp. 673-705, 2022.
- [26] R. Chugh, S. B. Patel, N. Patel, and U. Ruhi, "Likes, comments and shares on social media: Exploring user engagement with a state tourism Facebook page." *International Journal of Web Based Communities*, Vol. 15, No. 2, pp. 104-122, 2019.
- [27] H. S. Kim and C. H. Yoon, "Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market," *Telecommunications Policy*, Vol. 28, No. 9-10, pp. 751-765, 2004.
- [28] M. Samuelsson, "Education for deliberative democracy and the aim of consensus," *Democracy and Education*, Vol. 26, No. 1, pp. 1-9, 2018.
- [29] F. X. Kaufmann, "Normative conflicts in Germany: basic consensus, changing

- values, and social movements,” in *The Limits of Social Cohesion*, Routledge. 2018, pp. 84-112.
- [30] A. Nelson, R. Lewy, F. Ricardo, T. Dovydaitis, A. Hunter, A. Mitchell, and C. Kugel, “Eliciting behavior change in a US sexual violence and intimate partner violence prevention program through utilization of Freire and discussion facilitation,” *Health Promotion International*, Vol. 25, No. 3, pp. 299-308, 2010.
- [31] W. Park, (2021, January). “0.03% accounts for 30%’...Are portal news comments public opinion?”. SBS News. Available: https://news.sbs.co.kr/news/endPage.do?news_id=N1006153616&plink=COPYPASTE&cooper=SBSNEWSSEND
- [32] J. Kum, (2021, March)“NAVER News usage is decreasing rapidly”. Media Today. Available: <http://www.mediatoday.co.kr/news/articleView.html?idxno=212279>
- [33] M. Ziegele and O. Quiring, “Conceptualizing online discussion value: A multidimensional framework for analyzing user comments on mass-media websites,” *Annals of the International Communication Association*, Vol. 37, No. 1, pp. 125-153, 2013.
- [34] B. Schyns and D. Day, “Critique and review of leader–member exchange theory: Issues of agreement, consensus, and excellence,” *European Journal of Work and Organizational Psychology*, Vol. 19, No. 1, pp. 1-29, 2010.
- [35] C. Ding, H. K. Cheng, Y. Duan, and Y. Jin, “The power of the “like” button: The impact of social media on box office,” *Decision Support Systems*, Vol. 94, pp. 77-84, 2017.
- [36] S. Tadelis, “Reputation and feedback systems in online platform markets,” *Annual Review of Economics*, Vol. 8, pp. 321-340, 2016.
- [37] H. Baek, S. Lee, and S. Kim, “Are female users equally active? An empirical study of the gender imbalance in Korean online news commenting,” *Telematics and Informatics*, Vol. 62, pp. 1-12, 2021.
- [38] Y. Liu, W. X. Gan, and Q. Zhang, “Decision-making mechanism of online retailer based on additional online comments of consumers,” *Journal of Retailing and Consumer Services*, Vol. 59, pp. 1-9, 2021.
- [39] H. Wan, W. Ji, G. Wu, X. Jia, X. Zhan, M. Yuan, and R. Wang, “A novel webpage layout aesthetic evaluation model for quantifying webpage layout design,” *Information Sciences*, Vol. 576, pp. 589-608, 2021.
- [40] Chae, (2021, March) “Naver to shut down portal's trending search chart this month”. Yonhap News. Available: <https://en.yna.co.kr/view/AEN20210204010300320>
- [41] M. B. Pandur, J. Dobša, S. Beliga, and A. Meštrović, “Topic modelling and sentiment analysis of COVID-19 related news on Croatian Internet portal,” *Proceeding of Conference on Data Mining and Data Warehouses 2021*, Ljubljana, Slovenian, pp. 5-9, 2020.

