An Analysis of Threads with No Responses in Online Asynchronous Discussions

Takashi Nakazawa
Faculty of Informatics
Shizuoka University
Hamamatsu, Japan
0000-0003-2107-2735

Tomoyuki Tatsumi
Faculty of Informatics
Shizuoka University
Hamamatsu, Japan
0000-0003-1249-8283

***Abstract*—This study examines the causes of "no-response threads" in thread-based online discussion support systems. We investigated the factors contributing to the occurrence of no-response threads by using a logistic regression analysis of discussions on the topic of Nagoya City's attractions in the online discussion system COLLAGREE. Results showed that the content quality of a post did not significantly affect the occurrence of no-response threads in the divergence phase, where non-response was found to be more dependent on environmental factors such as the displayed time on the first page and the number of entries and unique viewers after the thread was posted. On the other hand, in the convergence phase, content quality greatly affected whether a new post would receive a response.**

Keywords—collaboration, collective intelligence, consensus building, online discussion

I. Introduction

There has been increasing interest in online discussion support systems in recent years. The popularity of the Internet and social media platforms has spawned new discussion spaces where large numbers of people, though separated by time and space, can discuss problems and issues of common interest. Some argue for the potential of online communication to increase political participation and promote democratic deliberation [1], while others contend that online communication tends to be careless and chaotic, and often cause cyber cascades and group polarization [2].

New technical developments are needed to support and facilitate high-quality online discussion. To enhance such technical developments, it is necessary to identify problems in the existing online discussion forums, examine the causes of these problems, and formulate effective solutions.

II. Objective

The objective of this paper is to examine the causes of "lonely threads" on thread message boards. In discussion systems with a thread structure, the first post in a thread has the most important position, since it raises a new topic and becomes a potential starting point for discussion. For a good discussion to develop, other participants must respond to the original post and begin interacting with each other in their comments [3]. However, some initial posts become no-response threads, left behind in the flow of the overall discussion.

The occurrence of no-response threads is a major issue for online discussions, as new threads posted by participants are an important resource, and the occurrence of no-response threads is a waste of such resources. Furthermore, when a post receives few or no responses, the poster is likely to feel frustrated and lose motivation to participate in other discussions [4]. To stimulate improved online discussion, it is important to examine the factors underlying no-response threads and seek solutions that can prevent or reduce their occurrence.

This paper examines the hypothesis that the content quality of a post is not a major cause of no-response threads—that is, no-response threads occur regardless of their content quality. Indeed, if high quality content is what attracts replies, then the occurrence of no-response threads is not really much of a problem. However, it is quite possible that even high quality posts will be ignored and receive no reply because of other factors. For instance, in a system in which a newly posted thread appears at the top of the page, even a high quality post will suffer from a downward push and become less likely to be read if many new threads are posted in a short period of time [5]. Login-lag on asynchronous thread-based message boards [6] may make some good quality messages less likely to be read if they are posted at a time when few users are logged into the board. If it is indeed the case that many good quality threads are being wasted, improving the system and design of the discussion board to save such threads would be highly desirable. Accordingly, this paper examines the factors underlying the occurrence of no-response threads.

III. Data and Method

A. Data

The data used in this study were taken from records of an online discussion board on COLLAGREE hosted by research groups from the Nagoya Institute of Technology and Nagoya City. COLLAGREE is an asynchronous discussion support and opinion collection system produced by a research project led by the Takayuki Ito Laboratory at the Nagoya Institute of Technology.

To examine the causes of no-response threads, we used data from a social experiment conducted on COLLAGREE from December 2016 to January 2017. The theme of this online discussion was "Nagoya City attractions." The participants were ordinary citizens recruited through public communications, advertising, and word of mouth. All applicants were welcome, and a total of 820 people participated. In addition, a facilitator participated in the online discussion and took a lead role in advancing the dialog. Four well-known persons also took part as a way of attracting more participants and stimulating the discussion. The participants discussed Nagoya's selling points and ways to improve the city's attractions. The discussion process was separated into two phases: a divergence phase in which new ideas were encouraged and a convergence phase in which ideas were brought together and conclusion were reached. In total, the

This work was supported by the JST CREST JPMJCR15E1(15656580) "Intelligent Information Processing Systems Creating Co-Experience Knowledge and Wisdom with Human-Machine Harmonious Collaboration," Japan.

discussion generated 1,327 posts, including initial posts that started new threads, as well as replies. Ideas offered by participants were organized into 30 proposals, which were then evaluated by virtual currency voting [7].

B. Unit of analysis

The unit of analysis in this research was a "post" written by participants in COLLAGREE. Posts by participants on the system were categorized into two types: starting posts that created a new thread and reply posts. Since the purpose of this paper was to examine the causes of no-response threads, we focused on the posts that started a new thread. Posts by the facilitator and the four publicly prominent persons were removed from the analysis as we were interested in trends among the general participants. As a result, 334 thread-initiating posts were left to be analyzed.

C. Variables

The explained variable in this study was a binary dummy variable indicating whether a post was a no-response thread. No-response threads were defined as an initial post that had no replies regardless of the number of "likes" it received. Initial posts that did not receive a reply were coded as "1," while initial posts with at least one reply were coded as "0" regardless of the total number of replies the post received.

The study considered two types of variables to explain the occurrence of no-response threads: the content of the post (content factors) and factors constituting the conditions in which the new thread was posted and displayed (environmental factors).

The content factors included two variables: body size and content quality. The body size of a post refers to the number of characters used in the post. This was considered because the length of a post can affect the likelihood that the post is noticed and that participants decide to read it. Content quality was introduced to establish whether a high-quality post would be more likely to attract replies.

The content quality of a post was evaluated by two coders and was based on the extent to which the post's content would likely contribute to the development of the discussion by providing good ideas, views, proposals, supporting evidence or questions presented in a logical and easily understandable manner. Each post was rated on a five-point scale (1=lowest, 5=highest). The coding proceeded in four steps. First, the coders jointly evaluated 15 posts randomly selected from the set of 334 posts to tune the coding level. Second, each of the coders independently coded all of the posts. Third, the coders re-evaluated 63 posts for which the gap in the score between them was two points or wider, re-reading the posts and discussing with one another their reasons and their evaluation criteria. The average of the two scores after re-evaluation was then used as the indicator of the quality of the post. Following the re-evaluation procedure, the Intraclass Correlation Coefficient (ICC) was 0.832, with a 95% confidence interval from 0.796 to 0.862 ($p < 0.001$), a result that was considered "excellent" according to the criteria in [8].

As noted above, our hypothesis was that the content quality of a post does not play a major role in explaining the occurrence of a no-response thread; rather, such an occurrence is more dependent on the environment in which the new thread was posted—specifically, the degree of user activity around the time of posting and how (and where) the thread is displayed on the discussion board.

The degree of user activity includes the number of entries and the number of unique viewers by time unit after a new thread has been posted. The former refers to the number of new threads and replies that were posted during a particular time period following the initial posting. The latter is defined as the number of registered users who logged in to the system during that subsequent time period. We adopted these factors in order to explore whether the degree of user activity after a new thread is posted affects the likelihood of a no-response thread. By introducing these variables, we were able to assess the influence of the timing of the initial post. It was assumed that at times when there are many active participants, a newly posted thread would be more likely to receive replies. On the other hand, it is quite possible that when many replies and threads are being posted, a newly posted thread may become less noticeable and therefore draw less attention. As the time unit in our study, we adopted 6 hours and 24 hours after the initial thread is posted in order to examine the effect of user activity in both the short and long terms.

The other environmental factor concerns the system, focusing on the way COLLAGREE displays a thread. Specifically, we measured the total time that a thread was displayed on the first page of the board. In COLLAGREE, only 10 threads are displayed on the first page. A newly posted thread or a thread that receives a new reply comes to the top of the list by pushing down the other threads. It was assumed that the longer a thread is displayed on the first page, the more noticeable it becomes by participants and the more likely it will receive a reply.

D. Method

We began by summarizing the occurrence of no-response threads, the content quality of the initial posts, and the degree of user activity using descriptive statistics in order to identify the overall tendencies of the discussion. We then conducted a statistical analysis using logistic regression to determine the main factors and their relative influence on the occurrence of no-response threads. The logistic regression was conducted using R version 3.4.3 (2017-11-30). We defined statistical significance as $p < 0.05$. Since the purpose of the discussion differed between phases, as noted above, the analysis examined the divergence and convergence phases separately.

IV. Results

A. Description of the occurrence of no-response threads

Fig. 1 shows the daily changes in the number of posts and unique viewers. As indicated, these numbers began to rise at the start of the discussion and reached a peak four days later, on December 16 (a Friday). A decrease over the weekend was followed by another peak on December 19 and 20, and then by another decrease. The number of posts and viewers continued to decline after the convergence phase began on December 21; the number of visitors increased again at the point of transition into the evaluation phase on January 4.

The overall trends in the occurrence of no-response threads and participant activity are shown in Table I and Table II. In total, 334 new threads were posted throughout the divergence and convergence phases. Among the 820 registered participants, 199 participants (24.3%) made at least one post either as a reply or a new thread, while 621 participants (75.7%) posted neither. Out of the 199 users who made a contribution, 84 posted only new threads and did not

reply to the comments of other users, while 29 made only replies and posted no new threads.

Fig. 2 shows the number of no-response threads and their percentage among all newly posted threads, excluding those posted by the facilitator and the four guests. No-response threads continued to appear throughout the two phases. Their percentage increased at the beginning of the convergence phase, as the overall number of new threads decreased. During the divergence phase (which lasted 9 days and 7 hours), 101 of 269 initial posts (37.5%) resulted in no-response threads; in the convergence phase (12 days and 17 hours), 31 of 65 initial posts (47.7%) became no-response threads.

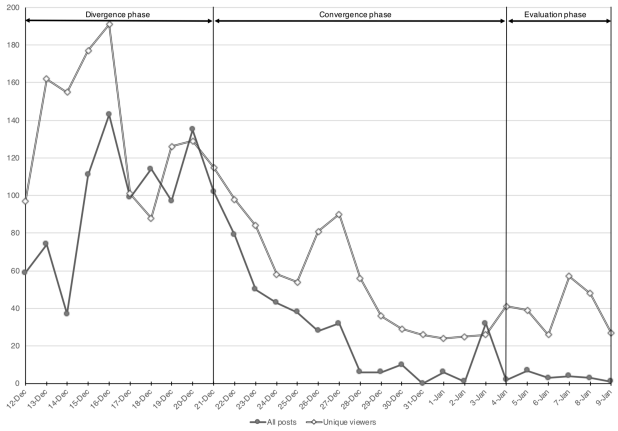

Fig. 1. Daily changes in the number of posts and unique viewers

TABLE I. OCCURRENCE OF NO-RESPONSE THREADS

Post types	Divergence phase		Convergence phase	
New threads	269	100.0%	65	100.0%
No-response threads	101	37.5%	31	47.7%
Replied threads	168	62.5%	34	52.3%
Reply posts	600		237	
Total posts	869		302	

TABLE II. USER TYPES BY CONTRIBUTION

	Users	Composition rate
Posted reply or new thread	199	24.3%
Posted reply and new thread	86	10.5%
Posted only new thread	84	10.2%
Posted only reply	29	3.5%
Posted neither reply nor new thread	621	75.7%
Registered participants	820	100.0%
Facilitator & guests	6	

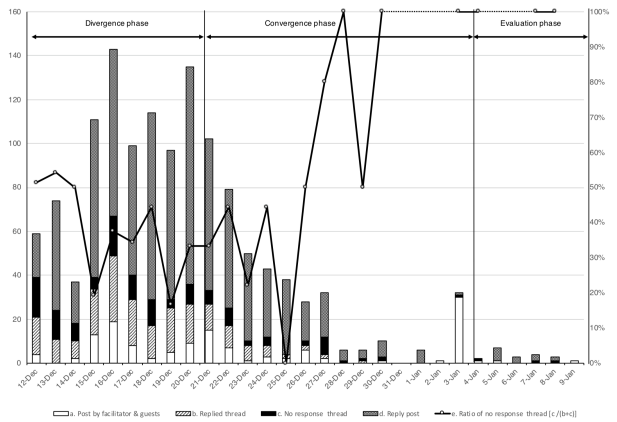

Fig. 2. Number of posts by types and percentage of no-response threads among all newly posted threads

The results shown in Table III indicate that the average rating of the no-response threads was substantially lower than the average for the posts with replies. At the same time, however, the ratings of the no-response threads showed greater variance; some of them, in both the divergence and convergence phases, were rated very high, whereas others were judged to be of low quality. As shown in Table III, among the no-response threads, there were 19 threads in the divergence phase and 9 threads in the convergence phase that were rated 4.0 or higher. This result indicates that some posts were wasted despite their good quality and that the occurrence of no-response threads cannot be explained solely by the quality of their content. In the next sub-section, we show the results of the logistic regression analysis to illuminate the relative influence of the various factors on the occurrence of no-response threads.

TABLE III. DISTRIBUTION OF CONTENT QUALITY OF POSTS PER PHASE

	Divergence phase		Convergence phase	
content quality (point)	no-response threads	replied threads	no-response threads	replied threads
1	6	4	2	0
1.5	17	5	0	0
2	0	4	2	0
2.5	10	18	8	1
3	11	11	2	2
3.5	38	66	8	11
4	11	28	5	7
4.5	7	22	3	9
5	1	10	1	4
Number	101	168	31	34
Mean	3.00	3.51	3.18	3.99
SD	1.02	0.87	0.98	0.63

B. *Logistic regression analysis*

The logistic regression analysis proceeded as follows. First, we explored how the degree of user activity influenced the occurrence of no-response threads, using as the explanatory variables the number of entries and the number of unique viewers 6 hours and 24 hours after the initial post (Model 1). Second, the content factors (body size and content quality of the post) were added to examine the influence of the content factors relative to user activity (Model 2). Third, we added the total time that the post was displayed on the top page and examined the relative influence of all of the variables in our model (Model 3). For both the divergence and convergence phases, AIC (Akaike's Information Criterion) was lowest in Model 3. Descriptive statistics for all the variables are shown in Table IV.

TABLE IV. DESCRIPTIVE STATISTICS OF VARIABLES PER PHASE

factors	Divergence phase n=269		Convergence phase n=65	
	Mean	SD	Mean	SD
content				
body size (chars)	221.7	175.7	307.9	234.0
content quality (1 to 5 point)	3.3	1.0	3.6	0.9
environmental (user activity)				
number of entries in 6 hours(posts)	29.4	15.1	14.9	8.4
number of entries in 24 hours (posts)	105.5	29.3	44.4	20.6
number of viewers in 6 hours(persons)	58.9	25.0	36.4	12.5
number of viewers in 24 hours(persons)	142.3	32.3	78.6	21.6
environmental (system)				
displayed time on the first page (hours)	21.6	26.0	59.3	57.7

Table V shows the results of logistic regression analysis in the divergence phase: coefficient estimates, standard errors,

adjusted odds ratios, 95% confidence intervals for the odds ratios, and p-values.

TABLE V. RESULTS OF LOGISTIC REGRESSION ANALYSIS IN THE DIVERGENCE PHASE

	Coefficient	SE	adj.OR	95% CI	p-value	
Model 1						
Intercept	1.729	0.804			0.03	*
number of entries in 6 hours	0.029	0.012	1.03	1.00-1.05	0.02	*
number of entries in 24 hours	-0.023	0.006	0.98	0.97-0.99	<0.001	***
number of viewers in 6 hours	-0.007	0.007	0.99	0.98-1.01	0.32	
number of viewers in 24 hours	-0.002	0.005	1.00	0.99-1.01	0.68	
Model 2						
Intercept	3.803	1.010			<0.001	***
number of entries in 6 hours	0.023	0.013	1.02	1.00-1.05	0.08	.
number of entries in 24 hours	-0.025	0.006	0.98	0.96-0.99	<0.001	***
number of viewers in 6 hours	-0.008	0.007	0.99	0.98-1.01	0.28	
number of viewers in 24 hours	0.000	0.005	1.00	0.99-1.01	1.00	
content quality	-0.645	0.152	0.52	0.39-0.71	<0.001	***
body size	0.001	0.001	1.00	1.00-1.00	0.47	
Model 3						
Intercept	10.740	1.839			<0.001	***
number of entries in 6 hours	-0.015	0.020	0.99	0.95-1.02	0.46	
number of entries in 24 hours	-0.055	0.010	0.95	0.93-0.97	<0.001	***
number of viewers in 6 hours	-0.039	0.012	0.96	0.94-0.98	<0.001	***
number of viewers in 24 hours	0.018	0.007	1.02	1.00-1.03	0.02	*
content quality	-0.022	0.217	0.98	0.64-1.50	0.92	
body size	0.001	0.001	1.00	1.00-1.00	0.43	
displayed time on the first page	-0.440	0.070	0.64	0.56-0.74	<0.001	***

*** significant at <0.001; ** significant at <0.01; * significant at <0.05.
N=269, AIC: Model 1=347.86, Model 2=332.17, Model 3=172.74

To assess the effect of user activity, Model 1 included as the explanatory variables the number of participants and the number of entries 6 and 24 hours after the initial posting. Results show that both the number of entries for the 6- and 24-hour intervals influenced the occurrence of no-response threads. However, the direction of the effect appears to be opposite; more entries during the first 6 hours increased the rate of no-response threads, while no-response threads were less likely to occur when there were more entries during the 24-hour period. On the other hand, the number of unique viewers, i.e., logged on users, during both the 6- and 24-hour intervals appeared to have no influence.

In Model 2, the two content factors (content quality and body size) were added to the factors in Model 1. Content quality turned out to be statistically significant, with a strong effect on the occurrence of no-response threads, while body size appeared not to be an influencing factor. The number of entries in 24 hours remained significant; however, the p-value for the number of entries in 6 hours was above 0.05.

Tellingly, when the duration of a thread being shown on the first page was added to Model 2 (producing Model 3), content quality became insignificant, while the duration of first page display exerted a strong influence with an odds ratio of 35.6%; the longer a thread was displayed on the first screen, the less likely it was that it would end up as a no-response thread. Results also showed that the number of entries in the subsequent 24 hours remained significant. However, this was not the case for the 6-hour interval. On the other hand, the number of unique viewers in both the 6- and 24-hour intervals was shown to be significant, although the effect was in opposite directions: having more users logged in during the 24 hours following the initial posting raised the likelihood that the thread would have no responses, while for the 6-hour interval, having more participants lowered the likelihood of no responses.

As shown in Table VI results in the convergence phase showed different tendencies from those in the divergence phase, especially in the case of Model 3. In Model 1, in both phases, the ratio of no-response threads increased as the number of entries during the 6-hour interval increased, while no-response threads were less likely to occur when there were more entries during the 24-hour period following the initial posting. Similar to what occurred in the divergence phase, when content quality was added (Model 2), its effect was statistically significant, with an odds ratio of 72.7%. However, unlike what happened in the divergence phase, content quality remained significant even after thread display duration was added (Model 3).

TABLE VI. RESULTS OF LOGISTIC REGRESSION ANALYSIS IN THE CONVERGENCE PHASE

	Coefficient	SE	adj.OR	95% CI	p-value	
Model 1						
Intercept	1.729	0.804			0.03	*
number of entries in 6 hours	0.029	0.012	1.03	1.00-1.05	0.02	*
number of entries in 24 hours	-0.023	0.006	0.98	0.97-0.99	<0.001	***
number of viewers in 6 hours	-0.007	0.007	0.99	0.98-1.01	0.32	
number of viewers in 24 hours	-0.002	0.005	1.00	0.99-1.01	0.68	
Model 2						
Intercept	4.707	1.826			0.010	**
number of entries in 6 hours	0.095	0.075	1.10	0.95-1.27	0.20	
number of entries in 24 hours	-0.076	0.035	0.93	0.87-0.99	0.03	*
number of viewers in 6 hours	-0.047	0.050	0.95	0.86-1.05	0.34	
number of viewers in 24 hours	0.047	0.032	1.05	0.98-1.12	0.14	
content quality	-1.297	0.471	0.27	0.11-0.69	0.006	**
body size	0.000	0.002	1.00	1.00-1.00	0.79	
Model 3						
Intercept	7.488	2.485			0.003	**
number of entries in 6 hours	0.192	0.110	1.21	0.98-1.50	0.08	.
number of entries in 24 hours	-0.194	0.074	0.82	0.71-0.95	0.008	**
number of viewers in 6 hours	-0.110	0.079	0.90	0.77-1.05	0.16	
number of viewers in 24 hours	0.104	0.053	1.11	1.00-1.23	0.05	*
content quality	-1.124	0.530	0.32	0.12-0.92	0.03	*
body size	0.000	0.002	1.00	1.00-1.00	0.94	
displayed time on the first page	-0.035	0.014	0.97	0.94-0.99	0.02	*

*** significant at <0.001; ** significant at <0.01; * significant at <0.05.
N=65, AIC: Model 1=92.57, Model 2=83.42, Model 3=74.12

V. DISCUSSION

Four findings are deserving of particular note. First, we found that, in the divergence phase, content quality and body size were statistically insignificant factors in explaining the occurrence of no-response threads when first page thread display duration was included in the model. This is consistent with our hypothesis that the content quality of a post is not a major cause of no-response threads. However, we should temper any conclusion by recognizing that the occurrence of a no-response thread might be related not just to the quality, but also to the type, of content. For instance, Reference [9] showed that responses can be triggered by including controversy, unexpectedness, personalization, and uncertainty in postings and by avoiding incomprehensibility and negativity. Taking the type of content into account may well produce a different result.

Second, the major factor affecting the ratio of no-response threads in our study was how long the thread was displayed on the first page of the message board. This implies that even high quality posts are likely to end up not being found by other users when the post is shown for only short time on the first page, most likely the result of being pushed down the board by other entries.

Third, the degree of user activity had a significant effect on the occurrence of no-response threads. In all three models, the results demonstrated that more entries during the 24-hour period after a thread is posted decreased the likelihood that that thread would not be responded to. On the other hand, in Model 1 and Model 2, it was found that more entries during the 6-hour period immediately following the posting led to a higher rate of no-response threads. This is possibly because,

in a time span as short as six hours, a new thread is more likely to be buried and overlooked among the relatively large number of new entries, while the likelihood that a new thread attracts a reply increases in an environment in which users participate in the discussion more actively over the longer time span. This result is consistent with the aforementioned effect of the duration of a thread displayed on the first page. Furthermore, Model 1 showed that no-response threads are less likely to occur when there are more unique viewers during the six hours after a thread is posted. This seems logical, as a new thread is more likely to be found and read when there are more participants. However, counter-intuitively, the results of Model 3 found that the rate of no-response threads increased when more users are logged on within the 24-hours period after the thread is posted. This result is difficult to explain logically; it might be somehow related to the fact that, as noted above, most users were not active in the discussion even though they had logged in; only 199 out of the 820 registered participants contributed either by posting a new thread or a reply.

Fourth, the convergence phase differed from the divergence phase in that, in the convergence phase, content quality was found to be a significant factor even after adding the thread's display duration to the model. Although thread duration was shown to be statistically significant, its effect was much smaller than in the divergence phase. This might be because the purpose of the convergence phase was different from that of the divergence phase. Whereas various ideas and opinions were sought in the divergence phase, the convergence phase sought to bring ideas together and make proposals. For that reason, participants may have been more attentive to the content quality of posts and thus making that a more significant factor. In addition, the number of new posts was much lower than in the divergence phase, allowing a newly posted thread to remain on the first page for a relatively longer time period (on average, 21.6 hours for a thread posted in the divergence phase versus 59.3 hours for a thread posted in the convergence phase). This might serve to weaken the effect of thread display duration on the occurrence of no-response threads.

VI. Conclusion

Results of the study imply that the degree of user activity and how a thread is displayed on the discussion board−rather than the quality of the post-matter most in explaining the occurrence of no-response threads. Many asynchronous thread-based online discussion systems are designed so that threads with numerous replies appear closer to the top of the display, on the assumption that users have evaluated them as important and relevant. However, as the study showed, posts with high quality evaluations do not necessarily attract many replies; some of them end up as no-response threads, especially if posted at a time when they are more likely to be overlooked among a relatively large number of entries over short time period. Although lively discussion is desirable, it brings with it the potential risk that beneficial threads may become buried. Therefore, there is a need for a system that reactivates buried threads.

One solution to these problems is to offer users multiple display orders rather than simply offering a chronological arrangement. Displaying posts chronologically or according to the number of comments or likes focuses too much on the newness or popularity of a post, which is not necessarily representative of the post's quality. A sorting or filtering function that permits users to rediscover buried posts can help to address this problem. Another effective method is a re-posting system that shows buried threads separately from the main display. With this feature in place, when a user logs in again, newly posted threads and those with a low number of views since the previous log-out are selected and displayed. In addition, it may be effective to provide an opportunity for posts that initially received little attention to be displayed and re-evaluated at the point of transition from the divergence phase to the convergence phase. All these functions could mitigate the problem of posts not being viewed and becoming buried without response. Importantly, technologies are being developed that can automatically assess the quality of a post; for instance, by detecting arguments and reasons to support them in posted messages through argumentation mining [10], [11]. Such technologies might help the functions noted above to work more effectively, although we need to be careful about the potential risk of algorithm bias and discrimination that such a technological development might cause.

Results of the study also suggest non-technological solutions to the problem of no-response threads. As the study showed, content quality was a significant factor affecting the occurrence of no-response threads in the convergence phase in which the purpose of the discussion was different from that in the divergence phase. This implies that we may be able to manage participant behavior in an online discussion board and constrain the occurrence of no-response threads by clarifying the purpose of the discussion and instructing participants on how they might contribute.

With the development of natural language processing and artificial intelligence, discussion support systems are entering a new era. To address the challenges that accompany this form of interaction, further studies are needed on the problems in communication that arise when we apply such new technologies to the realm of discussion support. By focusing on no-response threads and examining the causes of their occurrence, this study makes a significant contribution to the refinement and future development of online discussion support technologies.

Acknowledgment

The authors are grateful to the NITech Center for Collective Intelligence, E. Hideshima, Takayuki Ito, Takanori Ito, S. Shiramatsu and their laboratory members for sharing their dataset with us.

References

[1] Z. Papacharissi, “Democracy online: civility, politeness, and the democratic potential of online political discussion groups,” *New Media Soc.*, vol. 6, no. 2, pp. 259–283, 2004.

[2] C. R. Sunstein, *# Republic: Divided Democracy in the Age of Social Media*. Princeton University Press, 2017.

[3] P. Aragón, V. Gómez, and A. Kaltenbrunner, “To thread or not to thread: the impact of conversation threading on online discussion.,” in *Proceedings of the Eleventh International AAAI Conference on Web and Social Media*, Montreal, Quebec, Canada, 2017, pp. 12–21.

[4] K. F. Hew, W. S. Cheung, and C. S. L. Ng, “Student contribution in asynchronous online discussion: A review of the research and empirical exploration,” *Instr. Sci.*, vol. 38, no. 6, pp. 571–606, 2010.

[5] M. S. Bernstein, A. Monroy-Hernández, D. Harry, P. André, K. Panovich, and G. G. Vargas, “4chan and/b/: an analysis of anonymity and ephemerality in a large online community.,” in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, Catalonia, Spain, 2011, pp. 50–57.

[6] D. Dufner, S. R. Hiltz, and M. Turoff, “Distributed group support: a preliminary analysis of the effects of the use of voting tools and sequential procedures,” in *Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Sciences*, Wailea, HI, USA, 1994, pp. 114–123.

[7] T. Nishida, T. Ito, and T. Ito, “Verification of consensus building support system on large-scale social experiment where celebrities participate in discussion,” in *Proceedings of the 3rd IEEE International Conference on Agents*, Singapore, Singapore, 2018, pp. 126–131.

[8] D. V. Cicchetti, “Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology.,” *Psychol. Assess.*, vol. 6, no. 4, pp. 284–290, 1994.

[9] 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,” *J. Commun.*, vol. 64, no. 6, pp. 1111–1138, 2014.

[10] G. Morio and K. Fujita, “End-to-end argument mining for discussion threads based on parallel constrained pointer architecture,” in *Proceedings of the 5th Workshop on Argument Mining*, Brussels, Belgium, 2018, pp. 11–21.

[11] M. Ida, G. Morio, K. Iwasa, T. Tatsumi, T. Yasui, and K. Fujita, “Can you give me a reason?: argument-inducing online forum by argument mining,” in *Proceedings of The World Wide Web Conference 2019*, San Francisco, CA, USA, 2019, pp. 3545–3549.