

Figure 4: Transition probabilities in 4-person groups

3.2.3 5-person Groups' Discussion

(1) First Practice

SNS condition: The exchanges of remarks tended to start from the member with the lowest skill score (i.e., E). The role of the member with a medium-level skill score (i.e., C) was not clear.

FTF condition: Most of the members maintained lively contact with each other, except for D, who had the second-lowest discussion skill score.

(2) Second Practice

SNS condition: No significant transition pattern of remarks was observed.

FTF condition: The exchange of utterances between the leader (i.e., A) and the member with the second-highest discussion skill score (i.e., B) disappeared. Additionally, the pattern of transition was similar to that observed in the first practice.

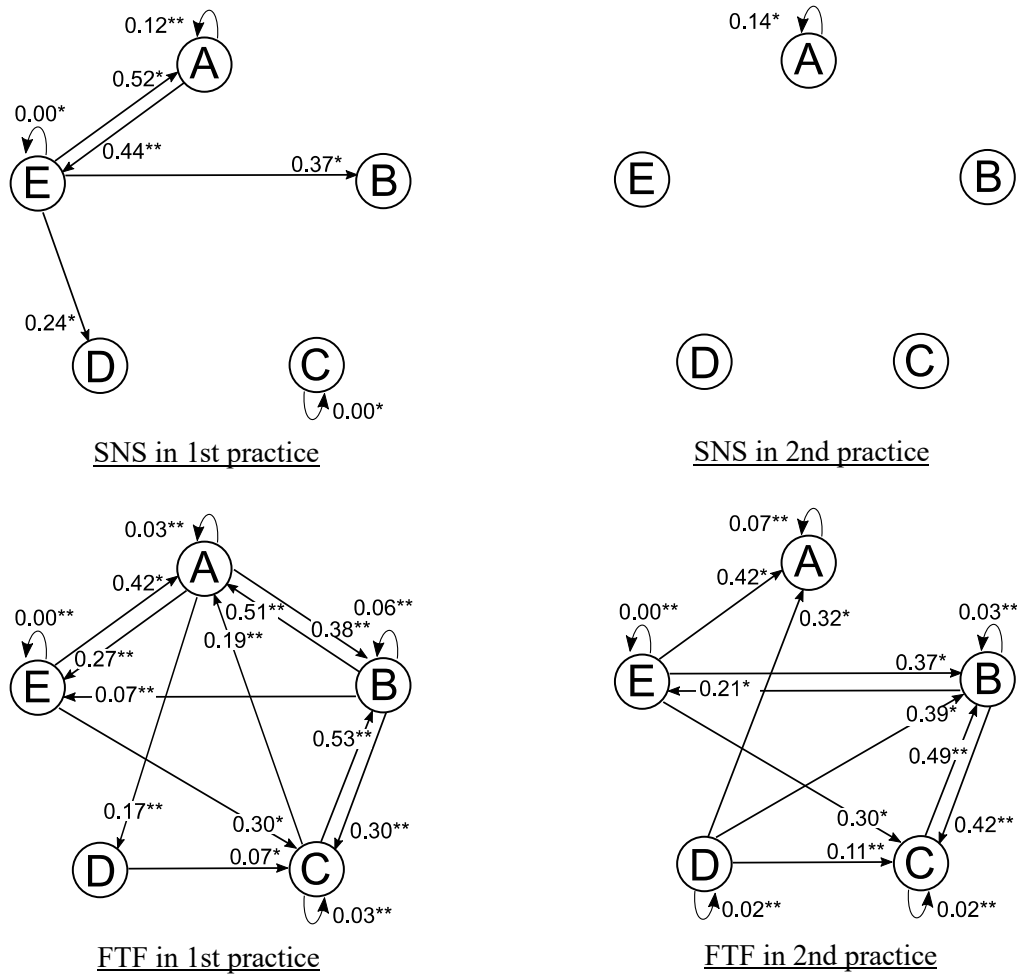


Figure 5: Transition probabilities in 5-person groups

3.3 Number of Utterances

Table 3 presents the means and standard deviations of the number of utterances for each condition.

By studying the differences between CMC (Computer-Mediated Communication) and FTF communication, Bordia pointed out that in a given time period, CMC produced fewer utterances than FTF communication [22]. Thus, we could not simply compare the number of utterances in the SNS condition to the number of utterances in the FTF condition. Therefore, we analyzed the SNS condition and the FTF condition separately.

3.3.1 Comparison by Group Size in the SNS Condition

We performed a two-way repeated-measures ANOVA to determine the effects of group size (3-person/4-person/5-person; as between-subject) and phase of practice (first practice/second practice; as within-subject). The ANOVA revealed that the main effects for group size were ($F(2, 35)=3.34, p<0.05, \eta_p^2=0.16$); the main effects for the phase of practice was observed to be ($F(1, 35)=4.57, p<0.05, \eta_p^2=0.12$). In addition, a significant interaction was observed between group size and phase of practice ($F(2, 35)=16.42, p<0.001, \eta_p^2=0.48$). A post hoc Holm's test demonstrated that 4-person groups had significantly fewer utterances

Table 3: Means and Standard Deviations of the Number of Utterances per Person

Types of Discussion	SNS				FTF			
	1st		2nd		1st		2nd	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
3-person groups	10.3	5.6	11.6	4.8	49.2	18.5	78.8	35.6
4-person groups	4.1	3.1	9.8	4.1	57.6	30.9	75.0	36.5
5-person groups	9.9	5.6	6.9	3.3	44.4	26.4	46.2	25.9

1st: first practice

2nd: second practice

than the other groups during the first practice. In addition, 5-person groups had significantly fewer utterances compared with 3-person groups during the second practice.

3.3.2 Comparison by Group Size in the FTF Condition

We performed a two-way repeated-measures ANOVA to determine the effects of group size (3-person/4-person/5-person; as between-subject) and phase of practice (first practice/second practice; as within-subject). The ANOVA revealed that the main effects for the phase of practice were ($F(1, 35)=13.15, p<0.01, \eta_p^2=0.27$). No significant interaction was observed between group size and the phase of practice. The second practice had the same number of utterances as the first practice.

3.4 Task Performances in the First Practice

Table 4 presents the means and standard deviations of the scores of the consensus exercise before and after discussions in the first practice.

3.4.1 Comparison by Group Size in the SNS Condition

We performed a two-way repeated-measures ANOVA to determine the effects of group size (3-person/4-person/5-person; as between-subject) and score (before discussion/after discussion; as within-subject). The ANOVA revealed that the main effects for group size were ($F(2, 35)=10.66, p <0.001, \eta_p^2=0.38$), and the main effects for the score were ($F(1, 35)=287.17, p <0.001, \eta_p^2=0.89$). In addition, a significant interaction was observed between group size and score ($F(2, 35)=3.64, p<0.05, \eta_p^2=0.17$). A post hoc Holm's test demonstrated that 4-person groups had significantly lower scores compared with 3-person groups and 5-person groups.

3.4.2 Comparison by Group Size in the FTF Condition

We performed a two-way repeated-measures ANOVA to determine the effects of group size (3-person/4-person/5-person; as between-subject) and score (before discussion/after discussion; as within-subject). The ANOVA revealed that the main effects for group size were ($F(2, 35)=5.34, p <0.01, \eta_p^2=0.23$).

Table 4: Means and Standard Deviations of Scores of the Consensus Exercise per Person

Types of Discussion group size	SNS				FTF			
	Before		After		Before		After	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
3-person groups	18.0	4.3	47.6	7.5	17.3	5.4	18.8	5.2
4-person groups	13.9	6.0	36.7	9.9	15.5	4.3	13.5	2.7
5-person groups	13.1	5.0	46.6	3.2	15.4	6.5	12.0	4.2

Before: Before discussion

After: After discussion

3.5 Variables Likely to Correlate with Typing Speed

Typing is indispensable for discussions using SNS. Bordia reported that typing speed was related to the number of utterances and the ability to reach consensus [22]. Based on this, we expected correlations between typing speed and such variables in this study. Therefore, the correlation between typing speed and number of utterances, and the correlation between typing speed and task performances in the first practice were analyzed.

3.5.1 Correlation between Typing Speed and Number of Utterances

To analyze whether typing speed affected the number of utterances, the correlation between the typing speed and the number of utterances was analyzed. A positive relationship was observed in 3-person groups in the second practice. The correlation coefficients between typing speed and number of utterances in the SNS condition are presented in Table 5.

Table 5: Correlation Coefficients (*r*) between Typing Speed and Number of Utterances in the SNS Condition

Group size	1st practice	2nd practice
3-person groups	0.28 <i>n.s.</i>	0.69*
4-person groups	-0.37 <i>n.s.</i>	-0.09 <i>n.s.</i>
5-person groups	-0.13 <i>n.s.</i>	-0.18 <i>n.s.</i>

* $p < .05$

3.5.2 Correlation between Typing Speed and Task Performances in First Practice

To examine whether typing speed affected the task performances, the correlation between typing speed and the score after the discussion was analyzed. No correlation was observed between the typing speed and the score after the discussion. The correlation coefficients between typing speed and the score after the discussion in the SNS condition are presented in Table 6.

Table 6: Correlation Coefficients (*r*) between Typing Speed and Score of after Discussion in the SNS Condition

Group size	After discussion
3-person groups	-0.20 <i>n.s.</i>
4-person groups	0.11 <i>n.s.</i>
5-person groups	-0.41 <i>n.s.</i>

4 Discussion

4.1 General Characteristics of the Collected Data

4.1.1 *Typing Speed*

This study deals with synchronous communication; thus, it is regarded not only as a CSCA study but also as a Synchronous Computer-Mediated Communication (SCMC) study. Lin noted that participants needed to be equipped with good/fast typing skill, because they needed to respond immediately to conversational partners in SCMC [23]. On the other hand, Sykes noted that variables such as typing speed could easily affect the outcome of a study [24]. In the results of the correlation analysis, no correlation was observed between the typing speed and the number of utterances, except for the 3-person groups in the second practice. Additionally, no correlation was observed between the typing speed and the score. In the 3-person groups, the number of times of turn-taking for the utterance is heavier than that of the other group sizes. Individuals with good typing skills can keep up with a conversation, whereas individuals with poor typing skills may be left out of a conversation and miss opportunities to remark.

Because the pace of turn-taking of utterances slows as the number of people increases, it is considered that 4-person group and 5-person group have a lower likelihood of losing the opportunity to remark than 3-person group. Based on these considerations, in terms of typing speed, 3-person group is not an appropriate group size for collaborative argumentation.

4.1.2 *Transition Probabilities*

In 3-person and 4-person groups, the number of transitions of conversations between two specific talkers tended to increase during the second practice. Moreover, in the second practice, we surmised that the participants may have attempted to promote a smooth discussion based on the discussion during the first practice. In both the SNS condition and the FTF condition, the tendency to talk more frequently after a specific person tended to be lower in 5-person groups in the second practice than in the first practice.

4.1.3 *Mean of the Number of Utterances per Person*

The mean of the number of utterances per person was significantly less in the SNS condition than in the FTF condition. This finding may be because typing was indispensable for the SNS condition, that is, participants assigned to the SNS condition used more time to input their thoughts. In addition, unlike a face-to-face discussion, it was not necessary to return a quick response to keep pace with others. This seemed to afford them enough time to form their opinions in their pace. Therefore, their remarks tended to be delayed [25].

4.1.4 *Mean of the Score of Consensus Exercise per Person*

Prior to the discussions during the first practice, participants ranked the ten items individually in the consensus exercise. We observed no significant difference in the mean scores of the consensus exercise between the SNS condition and the FTF condition. The mean scores of the consensus exercise, which were ranked based on the results of the discussions, had a significantly higher score in the SNS condition than in the FTF condition. In the rule of this consensus exercise, lower scores indicate better performance and greater consensus. In the SNS condition, the discussion

progressed rather slowly. A few groups did not complete the exercise. Therefore, to simply compare the SNS condition with the FTF condition was difficult. As further research, it is necessary to study the intergroup differences. In particular, it is important to determine whether consensus is obtained after a sufficient amount of time is provided for the discussion in the SNS condition.

4.2 Characteristics by Group Size

4.2.1 Characteristics of 3-person Groups

According to the transition probabilities, the tendency of members assigned to the FTF condition to talk with each other was high for both practices. In the SNS condition, participants who regarded themselves as poor communicators (coded as “C”) tended to be isolated from the other members.

In the SNS condition, the number of utterances in 3-person groups tended to be higher than the number of utterances in 4-person and 5-person groups. This finding may be due to the number of persons in these groups, that is, members in the 3-person groups may have talked more than members of the other groups. The members of the 3-person groups talked more frequently in the second practice than in the first practice. This finding was true for both discussion types. Because the participants had become accustomed to the types of discussions and group members, the number of utterances may have increased.

The scores of the consensus exercise in the first practice in the SNS condition showed that the mean of the scores of 3-person groups after the discussion was significantly higher than the scores obtained by the 4-person groups. No significant difference was observed between the scores of 3-person groups and the scores of 5-person groups. In the first practice, lower scores indicated better group performance. The performance achievement of the 3-person groups and the 5-person groups were lower than that of the 4-person groups.

Collaborative argumentation in 3-person groups in the SNS condition may have been dominated by the two members with the highest discussion skill scores. In addition, because of poor group performance achievement, these groups had difficulty reaching consensus.

4.2.2 Characteristics of 4-person Groups

In the FTF condition, from the standpoint of the transition probabilities, many participants conversed with each other during both practices. No exchanges with any specific person were observed during the first practice in the SNS condition. We observed that the number of interactions with a specific person increased during the second practice (e.g., C talked continuously; C talked after B; A talked after D). From the utterances of A after D, it can be deduced that A, the leader, supported D who was not particularly good at discussions.

In the SNS condition, the number of utterances during the first practice in the 4-person groups was significantly less than the number of utterances in 3-person and 5-person groups. No significant difference was observed in the second practice. Many groups had an equal number of male and female participants—typically, two each. Daibo et al. reported that females were more active than males in participation and utterances towards problem solving in small groups of university students [26]. As was the case with Daibo et al. [26], in the SNS condition of this study, it seemed there was a possibility that the conversation did not advance because of gender differences at the first practice.

The number of utterances of 4-person groups was lower than those of 3-person groups and 5-person groups in the first practice. Nevertheless, the mean scores of the 4-person groups were

significantly lower than those of the 3-person and 5-person groups. This indicates that the performance achievement in 4-person groups was good. In the SNS condition, 4-person groups tended to reach a consensus more easily compared with the other groups.

4.2.3 Characteristics of 5-person Groups

In 5-person groups, members from the SNS condition and the FTF condition tended to not talk after someone in particular. In the SNS condition, the number of utterances in 5-person groups was less than the number of utterances in 3-person groups.

Based on the scores of the consensus exercise, discussions in the FTF condition led to consensus more easily compared with the other groups. By contrast, in the SNS condition, there was no significant difference between 3-person groups and 5-person groups. The difference was significantly higher between 4-person groups and 5-person groups. Therefore, we posit that reaching consensus in the SNS condition was difficult.

According to Kreijns et al., anonymity and nonparticipation increase as group size increases [27]. The social loafer differs from the free rider in that the former lacks the motivation to add to the group performance, whereas the latter attempts to profit from others while minimizing essential contributions [27]. Therefore, 5-person groups are not feasible for discussion on SNSs.

4.3 Optimal Group Size

Based on these considerations, we attempted to identify optimal group size for collaborative argumentation using SNS.

First, a 5-person group is not feasible, because of the possibility of a social loafer and because this number of individuals does not facilitate conversations or discussions. We did not observe many differences between 3-person and 4-person groups, especially regarding the transition probabilities between two talkers. The number of utterances in 4-person groups was less than the number of utterances in 3-person groups during the first practice. No significant difference was observed during the second practice. Regarding the ease of reaching consensus, statistically significant results were obtained for the 4-person groups. Therefore, 4-person groups were observed to be particularly apt for collaborative argumentation using SNS.

5 Conclusion

This study aimed to determine the optimal group size for collaborative argumentation using SNS for educational purposes. Based on the transition probabilities between two talkers, the number of utterances, and the task performance, the results showed that a 4-person group was the optimal group size for collaborative argumentation using SNS for educational purposes. Identifying the optimal group size for collaborative argumentation using SNS for educational purposes is critical to improve group performance and enhance the learning.

This study also demonstrated that the number of utterances per person in the SNS condition was significantly less than the number of utterances per person in the FTF condition. Therefore, a limitation of this study was that we could not obtain a sufficient number of utterances to calculate the transition probabilities. Another limitation was that we did not consider gender differences while assigning participants to groups.

We recommend that further research should consider the differences between the groups after providing a sufficient amount of time to finish the discussion in the SNS condition. Additionally,

researchers could ask students for their impressions and evaluations regarding collaborative argumentation using a questionnaire and examine the correlation between the answers and group size.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 16H00224.

References

- [1] J. Andriessen, M. Baker, “Arguing to Learn,” in “The Cambridge handbook of the learning sciences, second edition,” Cambridge University Press, 2014, pp. 439-460.
- [2] Ministry of Education, Culture, Sports, Science and Technology, “the new Course of Study (Public notice in March, 2017),” from http://www.mext.go.jp/a_menu/shotou/new-cs/1383986.htm.
- [3] C.A. Chinn and D.B. Clark, “Learning through collaborative argumentation,” in “International Handbook of Collaborative Learning,” New York: Taylor & Francis, 2013, pp. 314-332.
- [4] T. Sumner and S.B. Shum, “Open peer review & argumentation: Loosening the paper chains on journals,” *Ariadne*, 5, 1996.
- [5] A.C. Jeong and J. Lee, “The effects of active versus reflective learning style on the processes of critical discourse in computer-supported collaborative argumentation,” *British Journal of Educational Technology*, Vol. 39, 2008, pp. 651–665.
- [6] Ministry of Internal Affairs and Communications, “White Paper on Information and Communication 2018” from <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/pdf/30honpen.pdf>.
- [7] Y. Nakahashi, “A Study on 'Reading Ability' in Social Media,” *Information science and technology*, Vol. 68, No. 8, 2019, pp. 406-411.
- [8] D. Tsovaltzi, T. Puhl, R. Judele and A. Weinberger, “Group awareness support and argumentation scripts for individual preparation of arguments in Facebook,” *Computers & Education*, Vol. 76, 2014, pp. 108-118.
- [9] C. Holland and L. Muilenburg, “Supporting student collaboration: Edmodo in the classroom,” In *Society for Information Technology & Teacher Education International Conference*, Vol. 2011, No. 1, 2011, pp. 3232-3236.
- [10] M. Sugai, T. Horita and Y. Wada, “Characteristics of High School Students' Collaborative Problem Solving using SNS for Educational Purposes,” *Information Processing Society of Japan Transactions, Education and Computer (TCE)*, Vol. 4, No. 1, 2018, pp. 43-56.
- [11] P. Dillenbourg, “What do you mean by collaborative learning?,” *Collaborative-learning: Cognitive and Computational Approaches*, Oxford: Elsevier, 1999, pp.1-19.

- [12] D.W. Johnson, R.T. Johnson and E.J. Holubec, "Circles of learning: Cooperation in the classroom, 5th edition," Edina, MN: Interaction Book Company, 2002.
- [13] S. Sugie, "Introduction to collaborative learning: understanding the basics and 51 ideas," Kyoto: Nakanishiya Publishing, 2011.
- [14] J.W. Strijbos, R.L. Martens and W.M.G. Jochems, "Designing for interaction: Six steps to designing computer-supported group-based learning," *Computers and Education*, Vol. 42, No. 4, 2004, pp. 403–424.
- [15] A.Valero Haro, O. Noroozi, H. Biemans, and M. Mulder, "First-and second-order scaffolding of argumentation competence and domain-specific knowledge acquisition: a systematic review," *Technology, Pedagogy and Education*, Vol.28, No. 3, 2019, pp. 1-17.
- [16] O. Noroozi, A. Weinberger, H.J. Biemans, M. Mulder and M. Chizari, "Argumentation-based computer supported collaborative learning (ABCSCCL): A synthesis of 15 years of research," *Educational Research Review*, Vol. 7, No.2, 2012, pp. 79-106.
- [17] L.S. Fuchs, D. Fuchs, S. Kazdan, K. Karns, M.B. Calhoon and C.L. Hamlett, et al., "Effects of workgroup structure and size on student productivity during collaborative work on complex tasks," *Elementary School Journal*, Vol. 100, 2000, pp. 183–212.
- [18] S. Yasunaga, K. Ejima, and M. Fujikawa, "Development of Discussion Skill Scale," *Bulletin of the Faculty of Literature, Kurume University*, Vol. 12, 1998, pp. 43-58.
- [19] J.C. McCroskey, M.J. Beatty, P. Kearney and T.G. Plax, "The Content Validity of the PRCA-24 as a Measure of Communication Apprehension across Communication Contexts," *Communication Quarterly*, Vol. 33, No 3, 1985, pp. 165-173.
- [20] J. Hall and W.H. Watson, W.H., The effects of a normative intervention on group decision-making performance. *Human relations*, Vol. 23, No. 4, 1970, pp. 299-317.
- [21] R. Sakaita, M. Suwa and N. Noguchi, "Exploration of body intelligence to activate communication space: Mr. Kenkoba's communication skill in variety show," *The 11th Annual Conference on Artificial Intelligence (SIG-SKL)*, 2011.
- [22] P. Bordia, (1997). "Face-to-face versus computer-mediated communication: A synthesis of the experimental literature," *The Journal of Business Communication*, Vol. 34, No.1, 1997, pp. 99-118.
- [23] K.Y.M. Lin, "Chatting online to learn: repair sequences in text-based L1-L2 interaction in CMC," PhD thesis, School of Education, Communication and Language Sciences, Newcastle University, 2015.
- [24] J.M. Sykes, "Synchronous CMC and pragmatic development: Effects of oral and written chat," *CALICO journal*, 2005, pp. 399-431.
- [25] Y. Chiba and K. Nishimoto, "Investigation of activation of discussion and possibility of multi-viewpoint by iteration of different forms of communication," *Information Processing Society of Japan Research Report*, Vol. 2011-HCI-143, No. 5, 2011, pp. 1-7.

- [26] I. Daibo, S. Matsuyama and K. Fujiwara, "The Communication behavior and interpersonal perception in task oriented small group -Social skills and interpersonal relationships-," IEICE Technical Report Human Communication Science (HCS), Vol. 111, No 393, 2012, pp. 21-26.
- [27] K. Kreijns, P.A. Kirschner and W. Jochems, "Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research," *Computers in human behavior*, vol. 19, No.3, 2003, pp. 335-353.