Explorative Study of Perceived Social Loafing in VR Group Discussion: A Comparison between The Poster Presentation Environment and The Typical Conference Environment

Koutaro Kamada[1][0000-0002-7585-150X], Ryuya Watarai [2], Tzu-Yang Wang[1], Kentaro Takashima[1], Yasuyuki Sumi[2], and Takaya Yuizono[1]

[1] Japan Advanced Institute of Science and Technology, Nomi, Japan
`kamada@jaist.ac.jp`
[2] Future University Hakodate, Hakodate, Japan

Abstract. Social loafing is a phenomenon in which members of a group reduce individual motivation and effort. We explored the difference between social loafing perceived by the loafer himself/herself (Self Perceived Social Loafing; SPSL) and social loafing perceived by other group members in VR group discussion (Others Perceived Social Loafing; OPSL). We also investigated how this difference changes in two types of group discussion: the poster presentation environment and the typical conference environment. An experiment with a between-participant design was conducted, and participants conducted a desert survival task through VR group discussion. The results showed that, overall, there was only a weak positive correlation and not much agreement between SPSL and OPSL. The results also suggested that there were significant positive correlations between the indicators relating to conversation behavior and OPSL in the typical conference environment but not in the poster presentation environment. In addition, an analysis by Lasso was conducted to examine the relationship between OPSL and these indicators and found that three indicators relating to participants' conversation behavior were selected in the typical conference environment, but none were selected in the poster presentation environment. Our study suggested that, in the typical conference environment, people judged the other people's social loafing through their conversation behavior; on the other hand, people's conversation behavior may not be used as significant indicators for social loafing in the poster presentation environment.

Keywords: Social Loafing · Perceived Social Loafing · VR Group Discussion · Social VR · Metaverse

1 Introduction

VR group discussion has become more important in the last few years because of the development of VR technology and the COVID-19 pandemic. People use avatars to attend in a virtual environment, and they perform group discussion in that environment.

However, just like other group works, social loafing that members in a group reduce individual motivation and effort are also expected to occur in VR group discussion [21]. Indeed, it has previously been shown that social loafing occurs in remote collaborative work [23, 33]. Social loafing has a negative impact on group members and has traditionally been desired to be controlled [21]. If the task involves interaction with other members, as in VR group discussion of this study, the bad effects on the members are likely to occur when they perceive social loafing [1, 25]. Therefore, we focused specifically on perceived social loafing. That is, the criterion is not whether social loafing is actually occurring, but whether or not the person is aware that social loafing was occurring. There are two types of perceived entities: self (Self Perceived Social Loafing; SPSL) and others (Others Perceived Social Loafing: OPSL).

Additionally, we focused on the VR space. This is because VR group discussions take place in a variety of VR environments and perceived social loafing can be affected by them. In fact, group cohesion, which has been reported to be associated with social loafing, has also been found to be affected by VR environments [13, 16].

Considering the above, to understand the elementary mechanism of perceived social loafing in VR group discussions, we set the following Research Questions (RQs).

First of all, the question arose as to how the estimation ability of other people's evaluations is. In other words, even if we think we are taking ourselves seriously, others may not think so, and vice versa. Therefore, we examined how well other people's evaluations match their personal evaluations.

— RQ1: How different is Others Perceived Social Loafing from Self Perceived Social Loafing in VR group discussions?

Next, we considered the possibility that the VR environment could affect others' perceptions of social loafing. Therefore, we examined whether changes in the VR environment affect the degree of agreement between self and others' evaluations.

— RQ2: Does the type of the VR environment influence the degree of agreement between Self Perceived Social Loafing and Others Perceived Social Loafing?

Finally, we focused specifically on the evaluation of others. This was because we consider that, unlike self-evaluations, judgments are made based on objective event cues. Differences in VR environments should affect this cue. That is, we decided to investigate how they perceive social loafing of others and what indicators they use to rate it.

— RQ3: How does the type of the VR environment affect the rating of Others Perceived Social Loafing?

Since there has been no research investigating social loafing in the VR space, these three RQs are a first step toward understanding perceived social loafing in VR.

2 Related Work

2.1 Concepts of Social Loafing

In the real world, performing tasks in groups is essential and one in which the individual members combine their strengths. Groups must be more efficient and effective than when performing tasks individually. This is because groups utilize more resources than individuals. However, in the real world, performing tasks in groups often reduce individual productivity. This loss of productivity was referred to by Steiner as process loss [32]. Social loafing, a form of process loss, has been studied in the field of social psychology. According to them, the traditional definition of social loafing is "the reducing effort or motivation by an individual in a group task". Laboratory experiments and field studies were conducted on a variety of tasks. The results revealed the occurrence of social loafing. Karau and Williams categorized tasks in which social loafing occurs as physical tasks (e.g., rope pulling [21]), cognitive tasks (e.g., decision making [15]), evaluative tasks (e.g., evaluating poetry [26]), and perceptual tasks (e.g., maze performance [14]) [17].

Research on social loafing has also been conducted in the field of HCI and CSCW. It has been observed to occur in cognitive tasks in remote collaboration [23, 33]. However, the mechanisms of social loafing in the VR space have not yet been investigated. With the current emphasis on social VR, we consider it important to understand social loafing in virtual space.

2.2 Evaluation Methodology of Social Loafing in Group Discussion

How to measure social loafing is an important issue and at the same time can be very difficult for some tasks. Unfortunately, group discussion is one of the most difficult tasks to measure social loafing.

In the research on group discussion, measures of effort or motivation can be broadly divided into objective and subjective evaluations. Objective evaluation is a method that measures the quantity of output, while subjective evaluation is a questionnaire-based method.

Objective evaluation is an important method for measuring social loafing. However, there are few objective evaluation methods for social loafing in group discussions [15]. This is because objective evaluation is thus used when individual and group output can be treated quantitatively such as the strength of a rope pull. In the experiment, the strength of an individual's rope pull was regarded as effort or motivation, and the pull strength per person is compared when the rope is pulled by one person and when the rope is pulled by a group [21]. On the other hand, the few existing studies on group discussion measure social loafing from the amount of information recalled prior to decision making [15]. In the experiment, the task was to assume decision-making regarding a civil trial and to recall as much as possible of the 80 items of information contained in a complex civil litigation scenario. That is, the number of information recalled by individuals and groups is used as an objective measure.

Subjective evaluation method measures an individual's perception of social loafing, rather than actual social loafing, and is measured using a questionnaire. Subjective evaluation can be based on the concept of perceived loafing advocated by Comer [6]. Perceived loafing is "the perceived low contribution of one or more other members to the group" [6]. However, this definition is not sufficient. This is because it only covers methods that measure the perception of others' social loafing and not the perception of self-social loafing [12]. A method to measure one's own social loafing is also an important concept [27]. For tasks that involve interaction among group members, such as group discussions, it is more important whether the social loafing of others is recognized than whether social loafing actually occurs. This is because individual effort or motivation is supposed to be influenced by the perceived effort or motivation of others [1, 25]. For example, when people perceive that group members other than themselves are putting enough effort or motivation into a task, they are known to reduce their effort and motivation, thinking that they can rely on them to do the task [17]. Also, the results were shown for virtual work [23]. Therefore, we decided that it was appropriate to investigate perceived social loafing in a subjective evaluation. Therefore, in this study, we extend perceived loafing and propose two concepts: Self Perceived Social Loafing (SPSL) and Others Perceived Social Loafing (OPSL). We defined SPSL as social loafing perceived by the loafer himself/herself and OPSL as social loafing perceived by other group members in VR group discussion. In this study, these are collectively referred to as perceived social loafing. This paper investigates perceived social loafing (SPSL and OPSL) in VR group discussions using a subjective rating scale.

2.3 VR Environment and Perceived Social Loafing

We consider that we should focus on the VR space. As previously stated, VR group discussions can take place in a variety of VR environments. It is possible that VR space will have an impact on perceived social loafing. Related studies have reported that VR spaces influence social behavior [13]. For example, group cohesion, which is considered an indicator closely related to social loafing, is influenced by VR space (closed and open space) [13, 16].

In this study, we chose two representative types of group discussion: the poster presentation environment and the typical conference environment. The poster presentation environment is an environment that promotes the movement of avatars; the typical conference environment is an environment does not promote participants' movement. Because VR allows users to move freely in the environment, various group discussions can be applied. Although we could not cover all types, we considered the movement affects users' judgment of social loafing and selected the two ends of the group discussion types in this work: the conference type with no movement required and the poster type with movement required. That is, we hypothesized that people's way of perceiving social loafing differs between the two environments.

3　　Preliminary Investigation

In this paper, for answering the three RQs mentioned above, we conducted an experiment to simulate VR group discussion and to investigate perceived social loafing. Before the experiment, we conducted a preliminary investigation to find the following: How long could the discussion last before social loafing occurs and what parameters could be used to judge social loafing based on in remote conferences. The findings were used to set the time of the group discussion and what data would be collected in the experiment. This is because, in order to answer the three RQs, social loafing needs to occur in the experiment, and we would like to find out what potential parameters people used to judge others' social loafing.

Amazon Mechanical Turk was used for the investigation. 19 participants (8 females and 11 males) were recruited randomly. The average age was 40.8 years (SD = 11.6). Each participant received compensation of 1.5 dollars. After explaining the definition of social loafing, the participants recalled one of the most significant group discussion experiences when they intentionally reduced their effort. They then answered an open-ended questionnaire, all self-reported, about the features of that.

As a result, 6 people responded that they had experienced perceived social loafing at one of the remote conferences. Additionally, the average time for participants to start reducing their effort was about 22 min (SD = 17.4). In addition, the result showed that the social loafing happened due to the following reasons: 1) the participants have other works to do, 2) the participants felt being ignored, 3) the participants have low interested in the topic, 4) the participants felt that their ideas were not be accepted and their opinion were different, 5) the participants perceived strong dominance from other members, and 6) the participants considered that they did not have to participate in it. Based on the results, we considered that social presence, cohesion, conflict styles, a quantity of conversation, new idea expression, agreement/disagreement, dominance, reaction and interrupting a turn may be associated with social loafing.

4　　Experiment

4.1　Experimental Design

An experiment with a between-participant design was conducted, and 39 participants conducted VR group discussion in 8 groups of 4-5 people. Each group of participants only participated in one of the two conditions. The topic of discussion was the desert survival task, which participants negotiate about important items for surviving in the desert after an airplane accident (see details in 4.2). The discussion lasted 40 minutes based on the result of the preliminary investigation that the mean time of occurrence of social loafing is about 22 min. To explore the effect of VR environment on perceived social loafing, in this experiment, we considered the movement an important potential factor affecting users' judgment of social loafing and compared the two ends of the group discussion types: the poster presentation environment which promotes participants' movement and the typical conference environment which does not promote

participants' movement (see Fig. 1). Both were configured with Mozilla Hubs using Spoke [24].

The two VR environments involved a poster that explains the desert survival task in the front of the room, a table, 3D models of the items, and posters of the expert opinion of items. For the poster presentation environment, the items were placed in a distributed manner. For the typical conference environment, they were placed on desks. The way these were arranged implicitly promote or demote the participants to move their avatars during the discussion. Additionally, in both environments, to induce social loafing, two objects which are irrelevant to the desert survival task (an aquarium and a whiteboard with the novel written on it) were placed to induce social loafing.

All participants used the avatar in Fig. 1. The simple white avatar without facial expressions was designed to minimize the possibility that the impression to the participants was affected by the impression of the avatar's design, such as the Proteus effect [37], the color of the avatar [8].

Fig. 1. VR Environment.

4.2 Task

A desert survival task was used as our experiment task to simulate a VR group discussion situation. The desert survival task was developed by Lafferty et al and widely used in group discussion research [19, 34]. It is a task that requires the selection of goods necessary for survival in the desert. In this study, we modified and used it. In the existing desert survival task, each item is prioritized. However, in this experiment, we determined that it would be difficult to memorize the priorities of all the items, so we asked the participants to choose five items from a list of ten items that were necessary for survival.

The participants were told the task was to select the items for survival because the plane they were on had crashed in the desert. The plane was off the flight plan and it was very far from the crash site to a nearby residence. The desert was too hot, flat plain with nothing but cacti. They were also informed that they wear light clothing (shirt with short sleeves, pants, socks, and sneakers).

The ten items adopted are as follows: 1L water per person, compass, dressing mirror, one coat per person, pistol, 2L vodka, flashlight, table salt, books on desert animals suitable for food, and aerial picture of the surrounding area.

4.3 Participants

40 participants (8 females and 32 males) were recruited from Japan Advanced Institute of Science and Technology (25 people) and Future University Hakodate (15 people). They were native speakers of Japanese. However, one of them withdrew from the experiment before participating, and resulted in 39 people. The average age was 22.6 years old, and the standard deviation was 1.55. At the end of the experiment, each participant received a compensation of 3000 yen.

4.4 Procedure

All participants used their laptop computer with a web camera and participated in the experiment at locations where they could naturally participate in the remote conference (e.g., their own room). The experimenter asked participants to use WebEx and share their screens and activate their webcams. They were recorded.

The experiment was described as follows. After informing them that this was not an experiment to measure their ability to discuss, they were instructed to discuss the task freely and build a consensus as a group. Taking notes during the task was prohibited. They did not reveal any personally identifiable information, such as self-introductions. Therefore, anonymous names were used for the discussion. Each of the 4-5 participants was asked to identify themselves as A, B, C, D (, E). In addition, since we wanted to observe the usual remote discussions, we emphasized that there were no penalties and that the rewards did not fluctuate based on results and asked the participants to keep in mind that they were to discuss as usual. In the VR environment, because Mozilla hubs have various functions, we restricted their use except for those we instructed them to use.

Before the experiment, each participant was individually briefed about the experiment, agreed with the experiment to make an ethical review, set up their computers, and practiced operating their avatars. After that, all experimental conditions proceeded as presented in Fig. 2. In the beginning, the participants were asked to connect to WebEx [36] and Mozilla Hubs [24]. Then, using WebEx, the experiment and the tasks were explained again. Next, the participants were asked to submit their own opinions of selected items for the task using Google Forms. In a submission of individual opinions, they were asked to describe the five selected items and why they chose or did not choose them for all items. After confirming everyone's submissions, they started a 40-minute discussion with Mozilla Hubs. After the discussion began with an opening signal, the participants were informed about the time left (20 minutes, 10 minutes, 5 minutes, and 1 minute) before the end of the session. Apart from that, we did not interrupt the participants' group discussion. After the discussion, each participant submitted the results of the group consensus and completed a questionnaire. In submission of the group's opinions, they were asked to report on the five items on which the group had reached consensus and to describe, along with reasons, whether they were satisfied with their opinions on all items. Finally, a focus-group interview was conducted. The interviews were semi-structured. The entire session was approximately two hours.
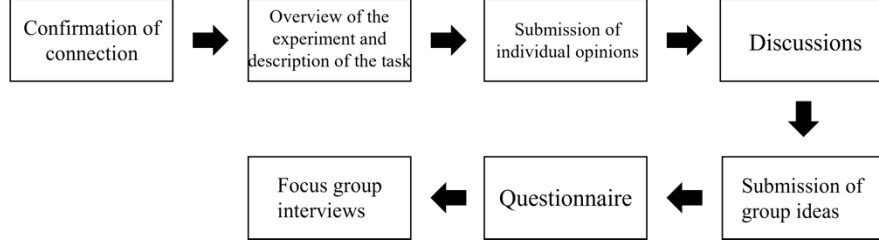
Fig. 2. Experiment Flow.

4.5 Measure

We used a variety of indicators, and it included three broad categories: indicators for measure of perceived social loafing, indicators for measure of discussion quality, and indicators that Amazon Mechanical Turk and previous studies have suggested related to social loafing.

Perceived social loafing was measured using questionnaires. Discussion quality was measured from a variety of indicators: perceived consensus and objective consensus. Indicators suggested to be associated with social loafing include the following: social presence, perceived cohesion, conflict management styles, conversation analysis and morpheme. Those whose units are percentages are divided by the total of all group members and expressed as a percentage.

Perceived Social Loafing. We used the two questionnaires to assess both Self Perceived Social Loafing and Others Perceived Social Loafing. Table 1 is the English version of our questionnaires. This questionnaire was developed based on George [12], and Petty and Williams [27], with modifications to fit the discussion task.

- Self Perceived Social Loafing (SPSL): measures how much effort the participant themselves perceived they put into the discussion. Initially, there were 23 items, but 10 were eliminated due to ceiling and floor effects. Consequently, there were 13 items. The average score of the items were considered as the SPSL of the participants.
- Others Perceived Social Loafing (OPSL): measures other participants perceived a participant whether he/she put effort into the discussion. Each participant judged and selected other participants whether they put effort into the discussion or whether they were involved in the discussion. For each participant, the average of number of being selected was considered as the OPSL of the participant. (0 = Strong social loafing to 4 = Week social loafing). Note that for the group with four participants, we adjusted the score by multiplying it by 4/3.

Table 1. Questionnaire of Perceived Social Loafing.

Self Perceived Social Loafing (SPSL)
I tried as hard as I could.
I did not contribute much to the discussion (R).
I did the best I could with the abilities I had.

I tried as hard as I could to offer my opinion.
I actively tried to speak up.
I was not concentrating on the task (R).
I tried to be actively involved in the discussion.
I tried to maximize the ability I had.
I tried to understand the thoughts of the other members.
I participated in the discussion, organizing the opinions of the entire group.
Relying on other members, I did not express my opinion much (R).
I didn't make an effort harder than the other members (R).
I was more determined to do well than the other members.

Others Perceived Social Loafing (OPSL)

Who did you think put the effort into the discussion? Please select 0-4 people, excluding yourself.
Who did you think was involved in the discussion? Please select 0-4 people, excluding yourself.

Note: item with (R) is the reversed item.

Questionnaire. Questionnaires were conducted to investigate regarding the items in Table 2.

To examine the subjective quality of consensus, two types of consensus were measured: Perceived consensus and objective consensus. In perceived consensus, we adopted the scale developed by DeStephen and Hirokawa [7]. There are five component constructs and each of the items was rated on a 7-point Likert scale. A higher value means a higher perceived consensus. The objective consensus was divided into two variables. Participants submitted their individual opinions before the discussion. This was a selection of 5 from 10 items. After the discussion, they individually submitted their group consensus results (selection of 5 from 10 items) and answered whether or not they were satisfied with each selected item. From these data, the percentage of changing opinions and the percentage of agreement were calculated. The percentage of changing opinions is the total number of items whose opinion changed before and after the discussion divided by 10 and expressed as a percentage. Before discussion opinions refer to the five items selected in the individual opinion submission. After discussion opinions are the items selected and agreed upon in the group opinion, and the items not selected and disagreed upon in the group opinion. The percentage of agreement is the percentage of the number of items not agreed upon in the group opinion submission. The total number of group opinions agreed upon is divided by 10 and expressed as a percentage.

Social presence was measured using a questionnaire developed by Biocca and Harms [3]. All items were not used, only items related to co-presence (perception of self and perception of the other) were used. It defines as the "degree of salience of the other person in the interaction and the consequent salience of the interpersonal relationships" [31]. These measure the degree to which the users feel as if they are together in the same space. Each of the items was rated on a 7-point Likert scale. A higher value means a higher social presence.

In this experiment, we measured cohesion, especially task cohesion. To measure

perceived task cohesion, we adopted the scale developed by Carless and De Paola [4]. Each of the items was rated on a 7-point Likert scale. A higher value means a higher perceived cohesion. Previous studies have indicated that it is closely related to social loafing [20].

Conflict management styles was a questionnaire developed by Rahim and adapted to suit this task [28]. The items were tailored to examine what conflict management styles were used by the participants in the discussions. Each of the items was rated on a 5-point Likert scale. However, the compromise score was removed because of $\alpha = .125$.

- Avoidance score: is measured with three items. "I tried to avoid stating my opinion in order not to create disagreements.", "I kept my opinions to myself if they disagree with others' opinions." and "I tried to avoid disagreements with others." ($\alpha = .793$).
- Accommodation score: is measured with three items. "When there was a disagreement, I tried to satisfy the needs of the other.", "I went along with the desires of others in a conflict situation." and "I went along with the recommendations of others in a conflict." ($\alpha = .656$).
- Confrontation score: is measured with three items. "I used my influence to get my position accepted by others.", "I was firm about advocating my side of an issue." and "I stuck to my position during a conflict." ($\alpha = .582$).
- Collaboration score: is measured with three items. "I tried to discuss an issue with others to find a solution acceptable to all of us.", "I like to work with others to find solutions to a problem that satisfy everyone." and "To resolve a conflict, I tried to blend the ideas of all of the people involved. "($\alpha = .776$).

Table 2. Questionnaire of Several Indicators.

Questionnaire	Variables
Perceived Consensus	Feelings regarding the group decision
	Feelings regarding the decision process
	Feelings regarding group member relationships
	Feelings regarding individual effectiveness
	Feelings regarding individual opportunity to participate
Objective Consensus	The percentage of changing opinions
	The percentage of agreement
Social Presence	Co-presence (Perception of self)
	Co-presence (Perception of the other)
Perceived Cohesion	Task cohesion
Conflict Management Styles	Avoidance score
	Accommodation score
	Confrontation score
	Collaboration score

Conversation Analysis. All conversations that took place in the discussion were coded by two people. The conversation scripts were created by Whisper [2], which mechanically converted the utterances into text and then manually adjusted them. The speech was reproduced as verbatim as possible, and the speaker was also identified. Even the errors of speech are scripted. We conducted the following conversation analysis in conjunction with the task of this experiment. At first, we coded TCU (turn constructional units) for all conversational scripts. In those coding, we followed the principles advocated by Sacks et al [30] that a turn is all speech from the time one speaker begins to speak until another person takes over the chance to speak. However, we did not count as turns any utterances that did not have a smooth response or turn changeover during the utterance. The analysis is at the individual level, not at the group level. Then, for these turns, we further coded the following: turn, turns involving new idea expression, interrupting the other's turn, failed turn-taking, times selecting the member, selected times by the member, agreement, disagreement, reactive token, and floor time. Each operation is described below easily.

Turns involving new idea expressions were counted as turns that stated an idea that no one else in the discussion has yet stated. In other words, it represents the number of unique ideas. Interrupting the other's turn was counted as the number of times the participant interrupted and started talking while someone else was speaking. Times selecting the member counted the number of times the speaker was named by a specific person or an unspecified number of persons when making a speaker change. Conversely, Selected times by the member counted the number of times the member was named. The code of agreement/disagreement was counted as agreement or disagreement with the immediately before opinion. The definition of the reactive token is "short utterance by an interlocutor acting as a listener while another interlocutor is speaking" [5]. For example, in Japanese, "Ah", "Um" and so on. The definition of floor is "Time and space in which the speaker perceives that he has the right to speak" [9]. In this study, floor time was defined as the time for someone in particular to lead the conversation. In many cases, this is the person who created a particular topic or raised an issue. We measured the time from the start of floor's speech until the next floor replaced it. We then calculated the total floor time for each member. The floor (%) was then used as a measure of the percentage of the conversation that was dominated. Failed turn-taking counted the number of times the participant was unable to perform turn-taking smoothly. We have performed morphological analysis on all utterances in the conversation script by using MeCab [22]. The number of morphemes was used as an indicator of message quantity. This is because, in Japanese, spacing is not based on words.

5 Result

5.1 Correlation between SPSL and OPSL
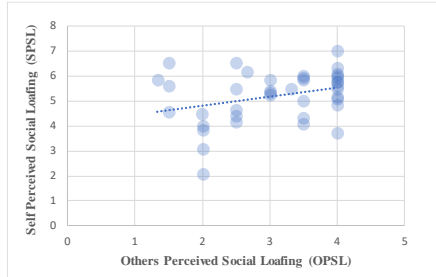
Under two environments, we conducted Pearson's Correlation Analysis between SPSL (Mean = 5.19, SD = 1.02) and OPSL (Mean = 3.06, SD = 0.894) (see Fig. 3). There was a weak positive correlation between the two variables (r = .320, p = < .05, N = 39).

Fig. 3. Scatterplot of the relationship between SPSL and OPSL. Overlapping points are represented darkly.

5.2 Correlation between Perceived Social Loafing and Several Indicators

For each environment, we conducted Pearson's Correlation analysis (two-sided) between perceived social loafing (SPSL and OPSL) and collected data, and then the correlation coefficient difference test (see Table 3).

There was a positive correlation between SPSL and OPSL in the poster presentation environment (r = .246, P = .309) and in the typical conference environment (r = .444, P = .05), and no significant difference in the correlation between the two environments (P = .516).

There tended to be more than moderate correlations between SPSL and subjective indicators. In the poster presentation environment, the variables for which strong correlations were obtained were following: individual effectiveness (r = .725, P < .01), individual opportunity to participate (r = .675, P < .01), confrontation score (r = .709, P < .01). The moderately and significantly correlated variables were as following: the group decision (r = .525, P < .05), avoidance score (r = -.589, P < .01), accommodation score (r = -.513, P < .05). In the typical conference environment, the variables for which strong correlations were obtained were following: co-presence (perception of self) (r = .639, P < .01), group member relationships (r = .790, P < .01), individual effectiveness(r = .687, P < .01), individual opportunity to participate (r = .720, P < .01), confrontation score (r = .619, P < .01), collaboration score (r = .678, P < .01). The moderately and significantly correlated variables were as following: co-presence (perception of the other) (r = .575, P < .05), the decision process (r = .511, P < .05). On the other hand, the just OPSL and subjective indicators that obtained more than moderate correlations were collaboration score (the poster presentation environment r = .484, P < .05), group member relationships (the poster presentation environment r = .497, P < .05; the typical conference environment r = .452, P < .05).

There tended to be more than moderate correlations between OPSL and objective indicators in the typical conference environment. In the typical conference

environment, the variables for which strong correlations were obtained were Selected times by the member (r = -.638, P < .01). The moderately and significantly correlated variables were as following: morpheme (r = .559, P < .05), morpheme (%) (r = .580, P < .05), turn (%) (r = .538, P < .05), turns involving new idea expression (%) (r = .578, P < .05), times selecting the member(%) (r = .451, P < .01), reactive token(%) (r = .456, P < .05). On the other hand, in the poster presentation environment, none of the variables were strongly correlated. The only one variable that showed moderately significant correlations were times selecting the member (r = -.546, P < .05).

Indicators for which the difference in correlation with SPSL was significant in the two environments were as follows: co-presence (perception of self) (P < .05), group member relationships (P < .05), agreement (P < .05). For OPSL, the indicators were as following: the percentage of changing opinions (P < .05), morpheme (P < .01), morpheme (%) (P < .05), turn (P < .05), turn (%) (P < .01), turns involving new idea expression (%) (P < .05), times selecting the member (P < .01), times selecting the member (%) (P < .01) agreement (P < .05). Marginally significant differences were as following: co-presence (Perception of self) (P < .1), turns involving new idea expression (P < .1), interrupting the other's turn (%) (P < .1), times selecting the member (P < .1), floor time (%) (P < .1).

Table 3. Correlation Analysis Result.

Variables	Poster Presentation Environment				Typical Conference Environment				Correlation coefficient difference test between two environments (p-value)	
	SPSL	OPSL	Mean	SD	SPSL	OPSL	Mean	SD	SPSL	OPSL
SPSL	—	.246	5.35	0.900	—	.444*†	5.04	1.13	—	.516
OPSL	.246	—	2.86	0.800	.444*†	—	3.25	0.950	.516	—
Co-presence (Perception of self)	-.110	-.301	5.42	1.16	.639**‡	.346	5.00	1.64	.013	.054
Co-presence (Perception of the other)	.192	-.124	4.87	1.26	.575**†	.145	4.83	1.14	.186	.437
The group decision	.525*†	.236	6.05	0.891	.443†	.101	5.68	1.20	.758	.689
The decision process	.387	.387	5.63	1.08	.511*†	.023	5.48	1.47	.655	.269
Group member relationships	.286	.497*†	5.25	1.41	.790**‡	.452*†	4.72	1.64	.026	.868

Individual effectiveness	.725***‡	.251	4.75	1.80	.687***‡	.311	4.48	1.46	.828	.852
Individual opportunity to participate	.675***‡	.207	5.17	1.02	.720***‡	.264	4.86	1.40	.801	.862
Perceived Cohesion	.378	.215	6.21	0.850	.422†	-.186	5.74	1.13	.880	.243
Avoidance score	-.589***†	-.204	2.40	0.821	-.358	.173	2.23	1.22	.387	.273
Accommodation score	-.513*†	-.363	2.89	0.910	-.349	-.084	2.68	0.820	.561	.395
Confrontation score	.709***‡	.241	3.04	0.744	.619***‡	.270	2.77	0.750	.642	.929
Collaboration score	.321	.484*†	4.05	0.536	.678***‡	.181	3.52	1.15	.157	.322
The percentage of changing opinions	-.213	.218	22.11	18.73	.093	-.430†	25.50	16.38	.374	.050
The percentage of agreement	.283	.521*†	91.58	13.85	.349	.304	93.50	11.37	.833	.449
Morpheme	.000	-.319	363.00	220.57	.110	.559*†	388.80	236.56	.751	.006
Morpheme(%)	.104	-.178	—	0.120	.251	.580**†	—	0.090	.662	.016
Turn	-.146	-.295	58.26	38.79	.426†	.364	52.40	28.76	.084	.049
Turn(%)	.079	-.346	—	0.110	.428†	.538*†	—	0.100	.277	.006
Turns involving new idea expression	-.036	-.192	9.68	6.06	.326	.441†	7.75	4.04	.282	.055
Turns involving new idea expression(%)	.144	-.087	—	0.130	.270	.578**†	—	0.094	.705	.032
Interrupting the other's turn	-.092	-.068	2.21	2.49	.272	.217	1.90	2.29	.286	.407
Interrupting the other's turn(%)	.083	-.297	—	0.190	.276	.268	—	0.200	.566	.095
Failed turn-taking	-.299	-.023	4.84	3.79	-.050	.081	2.75	2.57	.458	.765

Failed turn-taking(%)	-.085	-.189	—	0.120	.138	.260	—	0.120	.520	.189
Times selecting the member	.081	-.546**†	7.16	6.32	.331	.275	7.80	8.64	.451	.010
Times selecting the member(%)	-.098	-.442†	—	0.200	.393	.451**†	—	0.180	.140	.006
Selected times by the member	-.307	-.446†	2.32	1.86	-.214	-.638**‡	2.65	3.75	.774	.430
Selected times by the member(%)	-.119	-.248	—	0.170	-.102	-.267	—	0.250	.961	.953
Agreement	-.362	.112	11.58	10.34	.363	.007	8.15	6.34	.029	.762
Agreement(%)	-.133	-.307	—	0.100	.265	.038	—	0.130	.245	.308
Disagreement	-.142	-.053	1.68	1.73	-.035	.151	0.950	1.19	.757	.556
Disagreement(%)	.013	.161	—	0.240	.039	.178	—	0.260	.940	.960
Reactive Token	-.168	-.007	20.47	14.79	.361	.348	19.70	16.65	.116	.288
Reactive Token(%)	.047	-.049	—	0.120	.287	.456**†	—	0.140	.476	.120
Floor Time (%)	-.014	-.119	—	0.160	.303	.438†	—	0.210	.348	.091

Note: *P<.05; **P<.01. Correlation coefficient：.00-.19 = very weak; .20 - .39 = weak; †.40 - .59 = moderate (light red); ‡.60 - .79 = strong (deep red).

5.3 Explanatory Variables Selected by Lasso

Lasso (Least Absolute Shrinkage and Selection Operator) was a method for estimation in linear models and was developed by Robert [29]. It minimizes the residual sum of squares while limiting the sum of the absolute values of the coefficients. That is, it can used to automatically select more relevant coefficients. In this study, Lasso was used to find explanatory variables intrinsically related to OPSL.

To conduct the Lasso, we used the R package glmnet [10]. The objective variable was OPSL, and 11 objective indicators were entered as explanatory variables: Morpheme (%), turns involving new idea expression (%), interrupting the other's turn (%), failed turn-taking, failed turn-taking (%), times selecting the member (%), selected times by the member (%), agreement (%), disagreement (%), reactive token (%), and floor time (%). The reason these indicators are percentage unit is that we determined that this social loafing was evaluated relative to the group members.

As a result, no variable had explanatory power in the Poster condition. For the Conference condition, three indicators were selected, namely: morpheme (%), turns involving new idea expression (%), and reactive token (%).

6 Discussion

6.1 Relationship between SPSL and OPSL

Throughout, there was a significant weak positive correlation between SPSL and OPSL. This means that SPSL and OPSL are rarely in coincidence (RQ1). In addition, when the correlation coefficients between SPSL and OPSL were compared separately between the two environments, the coefficient of the typical conference environment was higher than the coefficient of the poster conference environment. Although the difference was not significant, the result suggested that the difference between SPSL and OPSL was affected by the form of VR environment, and the typical conference environment had a smaller difference (RQ2).

The existence and degree of difference between SPSL and OPSL may be attributed to the extent to which participants know each other. It has been reported that close acquaintances are less likely to have disagreements between self-assessments and others' assessments of them in the related work about personality judgment [11]. In this experiment, the group discussions were conducted anonymously, and group members did not know each other. Thus, the degree of agreement between the SPSL and OPSL might be affected by the VR environments which provide participants with different levels of opportunity to know other participants. In the poster presentation environment, the avatars moved around in the VR space, so the time spent looking at other avatars was likely to be short and difficult to fully observe. Therefore, participants may not get to know other participants better than in the typical conference environment, and it resulted in a larger difference between SPSL and OPSL.

6.2 Relationship between Perceived Social Loafing and Several Indicators

Pearson's Correlation analysis shows that interestingly, in both environments, SPSL tends to correlate significantly with subjective indicators. However, there were more exceptions in the poster presentation environment than in the typical conference environment. For example, co-presence or perceived cohesion and SPSL were only weakly correlated in the poster presentation environment. These indicators were found to be moderately or highly correlated in the preliminary study and related work [23].

On the other hand, our results indicated that OPSL was more likely to be significantly correlated with objective indicators in the typical conference environment while OPSL shows little correlation with indicators in the poster presentation environment. This shows that the participants judged others' social loafing differently in the two environments. The finding is also supported by Lasso's results that three variables are selected in the typical conference environment and no variables are selected in the poster presentation environment. Furthermore, the selected three variables (morpheme,

turns involving new idea expression, and reactive token) showed that the participants might have used conversational cues to determine OPSL in the typical conference environment. Conversely, the participants might not have used conversational cues to determine OPSL in the poster presentation environment (RQ3).

We considered that the strategy for default usage of the environment influenced whether or not conversational information was used to determine OPSL. The poster presentation environment is to provide information to others or to explore information, while the typical conference environment is to exchange information or interact with others. In the poster presentation environment, the participants paid little attention to others' conversational behaviors; thus, those behaviors may be less important as cues for OPSL. In contrast, the participants in the typical conference environment expected others and themselves to engage in the discussion; thus, the conversational behaviors were being focused more; therefore, these behaviors may be critical cues for OPSL.

6.3 Design Principles for Designers to Facilitate Communication in VR

From the findings of the result, we discuss some design principles for designers to facilitate communication in VR.

Many facilitating functions focused on how their functions make participants engage more in communication, and they evaluated the objective outcome and the participants' engagement with self-report. For example, focusing on the amount of speech as an outcome, and a subjective self-evaluation using a questionnaire [18]. However, our research findings suggested the importance of considering OPSL while designing facilitation functions. OPSL is an indicator that affects group members and should not be taken lightly [25]. In addition, the results of this study show that OPSL has a complex relationship with SPSL or simple output. Even if SPSL or simple output is improved, it does not necessarily mean that OPSL can be improved.

Furthermore, based on our finding that the type of the VR environment affects the OPSL, we also argue that it is important to consider the interaction between the type of the VR environment and the facilitating functions during design. For example, a facilitating function inducing users to speak more in a typical conference environment may improve his/her OPSL; however, it might not work well in a poster presentation environment. This is because, according to the results of this experiment, in a poster presentation environment, people's interest is not in conversational behavior, and a person who speaks more may still not be perceived to perform less social loafing.

6.4 Limitation and Future Work

In this research, we are aware that there are some limitations and future work.

First, the results of this study were based on a single type of discussion, and the generalizability is one of the limitations. We adopted the desert survival task in the experiment. This kind of negotiation task has a clear procedure and goal, so it may be easier for participants to distinguish others' social loafing. However, real-world group discussion type varies, and they often involve more open-ended topics and involve more physical movement, such as people drawing plans and writing down ideas on

whiteboards in a topic involving planning [35]. Moreover, people may also play different roles. People might difficultly judge others' social loafing, and their strategy might change. Thus, it is necessary to conduct experiments with different types of discussion to investigate how SPSL and OPSL differ in different situations and determine robust indicators for estimating perceived social loafing.

Second, we are aware that the individual differences might be biases of our experiment. However, we did not collect much demographic information, such as the expertise level of VR, the knowledge about the tasks, and the participants' personalities. As a future work, it is necessary to investigate whether these individual differences affect the result or not.

Third, in the poster presentation environment, OPSL was not determined based on conversational information. It is possible that OPSL was determined based on other factors. For example, the avatar's gaze and gait information. It is necessary to clarify these factors in future work.

7 Conclusions

The purpose of this paper was to understand the mechanism of perceived social loafing in VR group discussions. In particular, we explored the difference between SPSL and OPSL. We also investigated how this difference changes in the poster presentation environment and the typical conference environment. Therefore, we formulated the three RQs.

For RQ1, the overall SPSL and OPSL had only a significantly weak positive correlation. This means that the SPSL and OPSL are not so consistent in VR conferences.

For RQ2, the correlation coefficients between SPSL and OPSL were higher in the typical conference environment than in the poster presentation environment. However, in this experiment, the results of the correlation coefficient difference test were not significantly different, so at the level of suggestion.

For RQ3, we looked at the relationship between OPSL and various discussion related indicators in the two VR environments. The results showed that in the VR space of a typical conference, there were significant correlations with objective indicators that seemed to be related to Social Loafing. In addition, regression analysis by Lasso calculated three indicators with explanatory power. Thus, those who participated in discussions in the typical conference environment were found to be cued by several objective indicators. On the other hand, in the poster presentation environment, there was little correlation with indicators that might be associated with social loafing. Regression analysis by Lasso also showed no objective indicators with explanatory power. It can be said that in the poster presentation environment, those who participated in the discussion were at least making judgments without using conversational information as a cue. Therefore, we suggest that the criteria for OPSL are different in the poster presentation environment and the typical conference environment.

References

1. Albanese, R., Van Fleet, D.D.: Rational behavior in groups: The free-riding tendency. Academy of Management Review, 10, 244–255 (1985). https://doi.org/10.2307/257966
2. Alec, R., Jong, W.K., Tao, X., Greg, B., Christine, M., Ilya, S.: Robust speech recognition via large-scale weak supervision (2022). https://doi.org/10.48550/arXiv.2212.04356
3. Biocca, F., Harms, C.: Networked minds social presence inventory (Scales only version 1.2) (2003) http://cogprints.org/6742/
4. Carless, S., De Paola, C.: The measurement of cohesion in work teams. Small group research, 31(1), 71–88 (2000). https://doi.org/10.1177/104649640003100104
5. Clancy, P.M., Thompson, S.A., Suzuki, R., Tao, H.: The conversational use of reactive tokens in English, Japanese, and Mandarin. Journal of Pragmatics, 26(3), 355–387 (1996). https://doi.org/10.1016/0378-2166(95)00036-4
6. Comer, D. R.: A model of social loafing in real work groups. Human Relations, 48(6), 647–667 (1995). https://doi.org/10.1177/001872679504800603
7. DeStephen, R., Hirokawa, R.Y.: Small group consensus: Stability of group support of the decision, task process, and group relationships. Small Group Behavior 19(2), 227–239 (1988). https://psycnet.apa.org/doi/10.1177/104649648801900204
8. Domínguez, I., Roberts, D.: Asymmetric virtual environments: Exploring the effects of avatar colors on performance. Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, 10(3), 8–14 (2021). https://doi.org/10.1609/ai-ide.v10i3.12746
9. Edelsky, C.: Who's got the floor? Language in Society, 10(3), 383–421 (1981). https://doi.org/10.1017/S004740450000885X
10. Friedman, J., Hastie, T., Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent. J Stat Softw, 33(1), 1–22 (2010). https://doi.org/10.18637/jss.v033.i01
11. Funder, D. C., Colvin, C.R.: Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. Journal of personality and social psychology, 55(1), 149–158 (1988). https://doi.org/10.1037/0022-3514.55.1.149
12. George, J.M.: Extrinsic and intrinsic origins of perceived social loafing in organizations. Academy of Management Journal, 35,191–202 (1992). https://doi.org/10.2307/256478
13. Han, E., Miller, M.R., DeVeaux, C., Jun, H., Nowak, K.L., Hancock, J.T., Ram, N., Bailenson, J.N.: People, places, and time: A large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. Journal of Computer-Mediated Communication (2022). https://doi.org/10.1093/jcmc/zmac031
14. Harkins, S.G.: Social loafing and social facilitation. Journal of Experimental Social Psychology, 23(1), 1–18 (1987). https://doi.org/10.1016/0022-1031(87)90022-9
15. Henningsen, D.D., Cruz, M.G., Miller, M.L.: Role of social loafing in predeliberation decision making. Group Dynamics: Theory, Research, and Practice, 4(2), 168–175 (2000). https://doi.org/10.1037/1089-2699.4.2.168
16. Høigaard, R., Säfvenbom, R., Tønnessen, F. E.: The relationship between group cohesion, group norms, and perceived social loafing in soccer teams. Small Group Research, 37(3), 217–232 (2006). https://doi.org/10.1177/1046496406287311
17. Karau, S.J., Williams, K.D.: Social loafing: A meta-analytic review and theoretical integration. Journal of personality and social psychology, 65(4), 681–706 (1993). http://dx.doi.org/10.1037/0022-3514.65.4.681

18. Kim, S., Eun, J., Seering, J., Lee, J.: Moderator chatbot for deliberative discussion: Effects of discussion structure and discussant facilitation. Proceedings of the ACM on Human-Computer Interaction. 5, CSCW1, 87, 1–26 (2021). https://doi.org/10.1145/3449161
19. Lafferty, J.C., Eady, P., Elmers, J.: The desert survival problem. Plymouth, Michigan: Experimental Learning Methods (1974).
20. Lam, C.: The role of communication and cohesion in reducing social loafing in group projects. business and professional communication quarterly, 78(4), 454–475 (2015). https://doi.org/10.1177/2329490615596417
21. Latané, B., Williams, K.D., Harkins, S, G.: Many hands make light the work: The causes and consequences of social loafing. Journal of Personality and Social Psychology, 37(6), 822–832 (1979). https://psycnet.apa.org/doi/10.1037/0022-3514.37.6.822
22. MeCab: Yet another part-of-speech and morphological analyzer, https://taku910.github.io/mecab/, Accessed 18 Jan. 2023.
23. Monzani, L., Ripoll, P., Peiró, J.M., Dick, R.V.: Loafing in the digital age: The role of computer mediated communication in the relation between perceived loafing and group affective outcomes. Computers in Human Behavior, 33, 279–285 (2014). https://doi.org/10.1016/j.chb.2014.01.013
24. Mozilla Hubs, https://hubs.mozilla.com/, Accessed 18 Jan. 2023.
25. Mulvey, P.M., Klein, H.J.: The impact of perceived loafing and collective efficacy on group goal processes and group performance. Organizational Behavior and Human Decision Processes,74(1), 62–87 (1998). https://doi.org/10.1006/obhd.1998.2753
26. Petty, R.E., Harkins, S.G., Williams, K.D., Latane, B.: The effects of group size on cognitive effort and evaluation. Personality and Social Psychology Bulletin, 3(4), 579–582 (1977). https://doi.org/10.1177/014616727700300406
27. Petty, R.E., Harkins, S.G., Williams, K.D.: The effects of group diffusion of cognitive effort on attitudes: An information-processing view. Journal of Personality and Social Psychology, 38(1), 81–92 (1980). https://doi.org/10.1037/0022-3514.38.1.81
28. Rahim, M.: A measure of styles of handling interpersonal conflict. Academy of Management Journal, 26(2), 368–376 (1983). https://doi.org/10.2307/255985
29. Robert, T.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) 267–288 (1996). https://doi.org/10.1111/j.2517-6161.1996.tb02080.x
30. Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language, 50(4), 696–735 (1974). https://doi.org/10.2307/412243
31. Short, J., Williams, E., Christie, B.: The social psychology of telecommunications. London: John Wiley and Sons (1976).
32. Steiner, I.D.: Group process and productivity. Academia Press (1972).
33. Suleiman, J., Warson, R.T.: Social loafing in technology-supported teams. Computer Supported Cooperative Work (CSCW), 17(4), 291–309 (2008). https://doi.org/10.1007/s10606-008-9075-6
34. Wang, T., Kawaguchi, I., Kuzuoka, H., Otsuki, M.: Effect of manipulated amplitude and frequency of human voice on dominance and persuasiveness in audio conferences. Proceedings of the ACM on Human-Computer Interaction.2, CSCW, 177, 1–18 (2018). https://doi.org/10.1145/3274446
35. Wang, T., Noaki, Y., Kuzuoka, H.: Exploring how to display referential action to support remote group discussion. In Asian CHI Symposium 2021. Association for Computing Machinery, 89–96 (2021). https://doi.org/10.1145/3429360.3468188
36. WebEx, https://www.webex.com/ja/index.html, Accessed 18 Jan. 2023.

37. Yee, N., Bailenson, J.: The effect of transformed self-representation on behavior. Human Communication research, 33, 271–290 (2007). https://doi.org/10.1111/j.1468-2958.2007.00299.x