IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 31, 2021, accepted April 8, 2021, date of publication April 12, 2021, date of current version April 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3072790

Encouragement of Turn-Taking by Real-Time Feedback Impacts Creative Idea Generation in Dyads

SARINASADAT HOSSEINI[1], XIAOQI DENG[1], YOSHIHIRO MIYAKE[1], (Member, IEEE), AND TAKAYUKI NOZAWA[2]

[1]Department of Computer Science, Tokyo Institute of Technology, Yokohama 226-8502, Japan
[2]Research Institute of Earth Inclusive Sensing, Tokyo Institute of Technology, Tokyo 152-8550, Japan

Corresponding author: Takayuki Nozawa (nozawa.t.ac@m.titech.ac.jp)

This work was supported in part by the Center of Innovation Program from Japan Science and Technology Agency (JST) under Grant JPMJCE1309, and in part by Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP20H03553.

ABSTRACT Building creativity in a remote group is challenging, and various methods to improve it have been investigated. However, the effect of a real-time feedback system on computer-mediated group creativity outcome remains undetermined. In this study, we studied the creativity performance of 20 two-person groups (dyads) while they engaged in a group alternative uses task via video-mediated communication and received real-time feedback cues based on turn-taking. The study included three conditions: turn-taking encouragement feedback (TTF), random time feedback (RF), and no feedback (NF). To probe the underlying mechanism, we assessed creativity outcomes, participants' mood, and the temporal characteristics of the feedback effect type (i.e., slow gradual or instantaneous). The performance results revealed that TTF can enhance creativity outcome in the three dimensions of fluency, originality, and index of convergence. Comparing the self-reported moods indicated that the TTF cues heightened negative valence with high emotional arousal. While the results did not show any slow gradual changes in idea generation and idea quality differences between conditions, the results showed that the feedback cue impact was instantaneous. The results also indicated a significant instantaneous effect of immediate turn-taking on idea generation. Our findings suggest that the TTF condition enhanced cognitive persistence rather than cognitive flexibility during the remote group creativity session. The results of this study can be useful not only in remote human-to-human communications, but also in designing social robots and agents.

INDEX TERMS Group creativity, real-time feedback, remote group, turn-taking, video-mediated communication.

I. INTRODUCTION

Group creativity is an essential tool in educational and industrial settings [1], and communication is a critical point in group creativity [2], [3]. Computer-mediated communication has become a new trend because of rapid technological advancements. This raises the question of how to enhance the creativity session outcome of a remote group, which suffers from improper communication. In this study, inspired by the increase in remote communication during the COVID-19 pandemic, we proposed a feedback system based on turn-taking to improve remote group creativity. We studied the effect of this system on behaviors among group members and on different aspects of remote creative idea generation.

Feedback consists of the actions taken by an agent to provide information regarding some aspect(s) of one's task [4], [5]. Feedback can affect performance, motivation, and mood [6]. Moreover, awareness may heighten attentiveness to our social behaviors, leading to fruitful communication, which is an inevitable part of group creativity [7]. Previous studies on the impact of feedback on communication have reported a positive impact of real-time feedback on collaboration [8]. In addition, several previous studies have addressed the effects of visual feedback on remote communication. Lederman *et al.* reported that measuring conversation time and turn-taking behavior may be useful in

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague.

remote communication [9], and Zolyomi e*t al.* indicated that showing the active speaker helps autistic people during video calls [10].

Here, we chose turn-taking as the feedback stimulus because turn-taking in group creativity is related to the originality of ideas and perspective-taking behaviors [11]. On the other hand, traditionally in turn-taking groups, members can mention one idea per turn, which results in a production block, and removing the blocking effect of turn-taking results in exposure to the ideas of other group members and in increase in idea quantity [12]. Thus, we hypothesized that group creativity can be enhanced by encouraging spontaneous turn-taking through feedback rather than by constraining the group members to take turns as a rule.

To some extent, aspects of turn-taking and visual feedback in communication and group creativity can be inferred. However, to date, no study has examined the impact of encouraging turn-taking audio-based feedback on remote group creativity. Therefore, its potential impact remains unclear. In addition, the effect of feedback on the remote creative process is unknown.

Our study comprised three trials with different feedback conditions: turn-taking encouragement feedback (TTF), random time feedback (RF), and no feedback (NF). Creativity measures of the generated ideas and affective measures of the participants were compared between the TTF condition and the other two control conditions. Furthermore, temporal characteristics of the generated ideas and their quality in relation to the feedback cues were investigated.

To establish the effect of feedback and to understand its working mechanism, we first addressed the situationally induced moods during feedback conditions. Based on the dual-pathway to creativity model [13], creativity is achievable through cognitive flexibility, cognitive persistence, or a combination of both. Induced positive mood enhances individual-level creativity [14], [15]. Moreover, high in arousal dimension moods (regardless of their valence) evoke creative outcomes more than low in arousal dimension moods. Whereas high in arousal dimension with positive valence moods (HAPM) stimulate cognitive flexibility, high in arousal dimension with negative tone moods (HANM) improve persistence [16]. Mood alteration during remote group creativity is rather a complex system; therefore, it was difficult to formulate any precise hypothesis. To assess mood, we measured self-reported mood.

We assumed three possible scenarios. First, TTF affects creativity in both pathways. If so, we would observe enhanced levels of creativity outcome in all dimensions and no significant difference in valence. Second, TTF improves cognitive persistence. If so, results would show creativity enhancement in all categories except for flexibility, and primarily HANM moods. Third, TTF improves cognitive flexibility. The results would indicate creativity enhancement in all categories, and HAPM experienced by participants in this condition.

In this study, we addressed these questions: can turn-taking feedback stimulate turn-taking behavior? What are the effects on the overall outcomes of remote group creativity? Would feedback improve creativity outcomes based on the persistence cognition or flexibility cognition? On what timescale does the feedback affect remote group creativity—does it induce a slow, gradual change, or does it have an instantaneous and fleeting effect? Answers to these questions can help in designing applications to aid in idea sharing and fostering group remote communication. To answer the above questions, we examined the impact of turn-taking-based real-time feedback on group creativity, specifically in the form of idea generation. Here, we used an adapted version of the Alternative Uses Task (AUT), which is originally a measure of individual divergent creativity level [17], to investigate group creativity [18].

II. METHODS

A. ETHICS STATEMENT

The Human Subjects Research Ethics Review Committee of the Tokyo Institute of Technology approved this study procedure. All participants were briefed on the experimental procedure and provided oral and written informed consent before participating in the experiment. Participants were paid ¥2000 (approximately 19 USD) after the experiment as a reward for their time and effort.

B. PARTICIPANTS

Forty international students were recruited (20 males, 20 females; age = 20-35) via flyers and online surveys to participate in this experiment. They were of various nationalities. Participants were randomly grouped into dyads (7 female-female; 6 female-male; 7 male-male).

C. FEEDBACK PROGRAM

To implement feedback based on turn-taking events, we first needed to detect the active speaker per second. For this purpose, we used the live caption functionality of Google Meet. The live caption capability indicates the active speaker's account profile picture along with their speech content per second. Before the experiment, we sent either a solid green or red picture to participants based on their groups with the instruction to use them as profile pictures. With this, detecting the active speaker per second using the OpenCV color filter was sufficiently accurate. After we acquired the active speaker per second, we stored them in an array and assessed the turn-taking events. Here, turn-taking is defined as the switching of a speaker from one participant to the other in a group. It should be noted that we rejected turns that lasted less than two seconds and accepted turn-takings within five seconds of silent time in between. We called the remaining turn-taking events successful turn-taking events.

This study included three feedback modes: TTF, RF, and NF. We used a clap sound as the feedback cue and played the cue during the TTF and RF trials. The feedback cue as a clapping sound was given to participants after three successful turn-taking events during the TTF trials. For the RF trial, we used a set of the same clapping sounds, which was given at random timings; the number of feedback cues was equal, on average, to that in the TTF trials. The participants did not hear clap sounds in the NF mode. We ran the application in all trials to assess the number of successful turn-taking events and feedback cue timings in the TTF trial.

D. EXPERIMENTAL PROCEDURES

In this study, all experimental sessions were held online using Google Meet, and questionnaires were completed via Google survey links distributed through Google Meet chat. Before the experiment, we sent consent forms as well as a video clip (Supplementary video 1) to explain the experimental procedure to the participants. After both participants logged in to Google Meet, we asked them to turn on their video and audio. The experiment began after we briefly explained the experimental procedure once more and read the consent form. We ensured that both participants could hear the audio perfectly and agreed to all terms in the consent form before each experiment.

The experimental procedure consisted of a pre-questionnaire and three trials. The pre-questionnaire was the group preference scale, which consists of 10 items on a 7-point Likert scale and determines one's general preference for working in groups [19]. Upon pre-questionnaire completion, participants had a 5-minute break, and the main experiment started.

During each trial, the dyads solved three AUT tasks. We independently randomized the order of the trials and AUT tasks. We also explained to the dyads that in the two trials with the feedback cues, the clapping sound informed them that they scored a point, and we instructed them to obtain the highest possible score. Such an approach allowed us to test under game-like conditions and examine the impact of turn-taking feedback without direct instructions. Before each trial, a link including three object names was given to participants through the chat. Since the experiment had three trials, we prepared three sets of three objects as tasks for the AUTs. The three sets of objects were key, towel, and shoe; tire, cork, and button; and newspaper, spoon, and clothes hanger (Supplementary part 1). For each trial, after participants read the questions and became aware of the objects, they could start the AUT discussion. Each trial consisted of fifteen minutes of discussion time, ten minutes for completing the post-questionnaire, and five minutes for a break.

After each discussion, participants answered the post-questionnaire in terms of their emotions. We used the Positive and Negative Affect Schedule (PANAS), which is a self-report questionnaire, to measure both positive and negative aspects of mood, and each item is rated on a 5-point scale of 1 (very slightly or not at all) to 5 (extremely) [20]. To evaluate creative task performance, the experimenter recorded the discussions using Google Meet recording.

E. AUT OUTCOME ASSESSMENT

Dyad's AUT performance was assessed using four dimensions: fluency, originality, flexibility [17], and index of convergence (IOC) [19]. We scored each dimension as follows: the fluency score was the total number of ideas per trial. To score originality, we first calculated the raw originality as the average of the infrequency counts of each idea for each object in a set [21]. Denoting frequency count of idea ID for and object OB over all groups as $n(OB, ID)$, the raw originality score per object is defined as:

$$OriginalityRaw(OB, ID) = max(n(OB, \cdot)) - n(OB, ID). \quad (1)$$

As the set of three objects assigned to each condition differed by dyads, we calculated the standardized originality score of each idea [23] to eliminate the effect of possible differences in the originality score distribution. Having $fucntion(n(OB, .))$ as the function $maximum, mean, SD$ of ideas' frequency, the *standardized originality* is formulated as below:

$$OriginalityZ\,(OB, ID) = \frac{OriginalityRaw\,(OB, ID) - mean\,(OriginalityRaw\,(OB, .))}{SD\,(OriginalityRaw\,(OB, .))}. \quad (2)$$

Finally, the originality was the average of the calculated standardized originality scores of the mentioned ideas for all three objects per trial.

To assess flexibility and IOC, we first categorized each idea [21]. Then, flexibility was defined as the total number of categories per trial. Similar to originality (to ignore possible differences in flexibility score distribution), we calculated the standardized flexibility. For IOC, we scored ideas based on their convergence quality, which represents the extent to which the dyad combined one's own responses with that of the partner's (perspective-taking behavior) [11]. To score this quality, when an idea mentioned by one participant was in the same category as the previous idea mentioned by the other participant, it was scored as convergent. Finally, the IOC score for each dyad was obtained based on the sum of the convergent ideas over the fluency score per trial. To observe the possible effects of the feedback modes on creativity performance, we used R statistical computing software, version 4.0.2.

F. MOOD ASSESSMENT

To assess mood, we used the self-reported scores from the post-questionnaire, PANAS. Each item in the questionnaire reflected mood in two dimensions: valence and arousal. First, we averaged the scores reported by the dyads per group and per trial. To indicate the hedonic tone, we reversed the items of negative valence and then averaged the scores, resulting in the valence score. Next, to assess the arousal level, we reversed items with low arousal and then averaged all scores based on the arousal dimension. Classifications of valance/arousal are provided in Supplementary Table 1.

Finally, to investigate whether participants' moods were more toward the HAPM or HANM, we formulated them as below:

$$HAPM = Arousal\ Score + Valence\ score \tag{3}$$
$$HANM = Arousal\ score - Valence\ score. \tag{4}$$

G. IDEA GENERATION TRENDS

Here, we aimed to determine whether any feedback mode would result in gradual changes in the shared ideas per trial. To investigate this gradual change in the quantity of generated ideas along with their quality (originality and convergence), we determined the difference between the number of generated ideas, convergence, and originality of the ideas between the first and the second halves of each trial.

H. TEMPORAL COORDINATION BETWEEN FEEDBACK CUE AND IDEA GENERATION

In this part, we aimed to clarify the relationship between creativity outcomes and the feedback cue in the TTF condition. Specifically, we tested whether the feedback cue led to a higher chance of idea generation in the successive moment. We also aimed to test the contributions of the feedback cue in leading to any specific quality of ideas.

To test such temporal coordination, we followed seven steps as previously described in [25].

In step 1, given ft as the time in seconds when a feedback cue was given, we defined the feedback cue sequence (F_t) as follows:

$$F_t = \begin{cases} 1 & t = ft, \\ 0 & otherwise. \end{cases} \tag{5}$$

Similarly, given t_i as the time in seconds when an idea was generated, we defined the idea sequence (I_t), originality sequence (O_t), and convergence sequence (C_t) as

$$I_t = \begin{cases} 1 & t = t_i, \\ 0 & otherwise \end{cases} \tag{6}$$

$$O_t = \begin{cases} Raw\ originality\ score & \text{if } t = t_i, \\ 0 & otherwise \end{cases} \tag{7}$$

$$C_t = \begin{cases} 1 & if\ t = t_i\ and\ the\ idea\ is\ convergent, \\ 0 & otherwise. \end{cases} \tag{8}$$

In step 2, we generated surrogate data (SI_t) of the idea sequence (I_t), defined in (6), by randomizing the order of intervals between two sequential ideas in the original idea sequence 1000 times. This provides a set of 1000 pseudo idea sequences that maintain the density and the interval distribution of the original idea sequence but are severed from the possible temporal coordination with the feedback cues.

We aimed to determine whether the feedback cue would result in a higher probability of idea generation, and if yes, to identify how long it would take a feedback cue to increase the chance of idea generation. Therefore, in step 3, we tested ten types of time-window sequence of F_t sequence with window widths of [0, w], where w = 1 − 10 s, extending the F_t sequence to FW_t sequences. Fig. 1 illustrates examples of generated FW_t sequences with the ten window widths.

In step 4, to evaluate the temporal coordination of F_t and I_t response, we calculated their dot product as $FtoI_w = (FW_t) \cdot (I_t)$. This index captures how likely an idea is generated following a feedback cue with a delay within [0, w] s. Similarly, we calculated the dot product using each instance of the surrogate-based idea sequence SI_t, as $SFtoI_w = (FW_t) \cdot (SI_t)$. The distribution of this index provides a null distribution of temporal coordination "by chance" between feedback and idea generation. Using these scores, we assessed the z-score per group as below:

$$z(FtoI) = \frac{FtoI_w - mean(\{SFtoI_w\})}{SD(\{SFtoI_w\})}. \tag{9}$$

The $z(FtoI)$ was meaningful in two ways. First, it enabled us to test whether we could reject the null hypothesis and support that a feedback cue leads to idea generation. Second, it helped us identify the most relevant time lag between the feedback cue and idea generation response.

In step 5, we addressed the quality of the ideas following a feedback cue in terms of the originality and convergence of those ideas. To assess the quality, along the series of ideas indexed by $j \in 1, 2, \ldots, N$, where N represents the number of ideas generated in a trial, we first defined a "feedback-coordinated idea" sequence ($FtoI_{wj}$) that classifies ideas into those that occurred within w seconds from a feedback cue and those that did not.

$$FtoI_{wj} = \begin{cases} 1 & if\ idea\ j\ was\ generated\ within\ w\ seconds \\ & after\ a\ feedback\ cue, \\ 0 & otherwise. \end{cases} \tag{10}$$

Similarly, a "convergence of idea" sequence (C_j) was defined as

$$C_j = \begin{cases} 1 & if\ idea\ j\ was\ convergent, \\ 0 & otherwise. \end{cases} \tag{11}$$

An "originality of idea" sequence (O_j) was defined as

$$O_j = \text{Raw originality score of idea } j.$$

In step 6, we calculated the Pearson correlation coefficients $r(FtoI_{wj}, C_j)$ between the "feedback-coordinated idea" sequence ($FtoI_{wj}$) and the "convergence of idea" sequence (C_j) as well as $r(FtoI_{wj}, O_j)$ between ($FtoI_{wj}$) and the "originality of idea" sequence (O_j). Owing to the non-Gaussian nature of the sequences, the null distributions of the correlations might be different over groups and trials. Therefore, we generated a set of surrogate data of the sequences (C_j) and (O_j) by permuting the sequence orders in random 1000 times and denoted the generated sequences

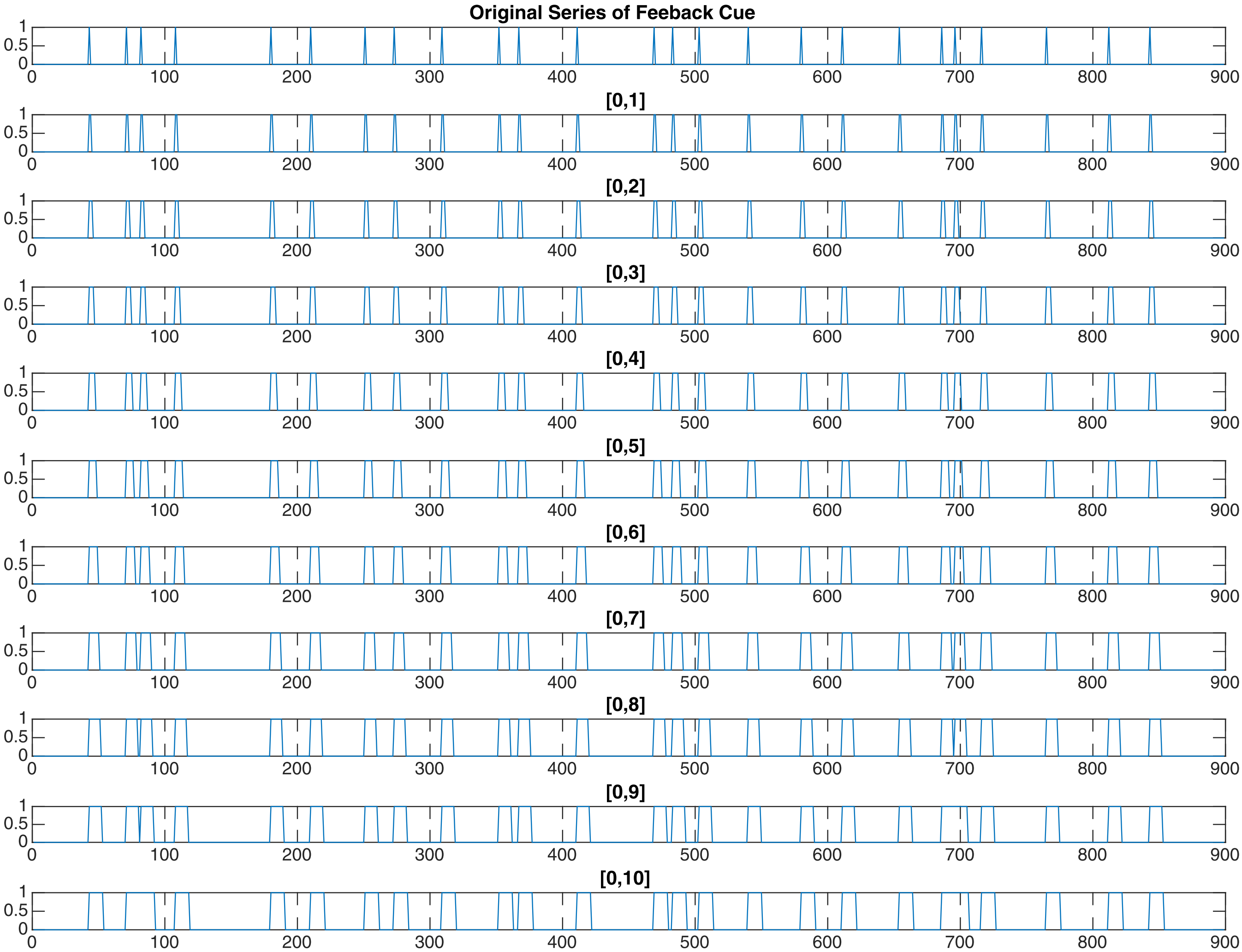

FIGURE 1. Example of original feedback cue sequence (top) and its window-adopted versions, with 10 time-window settings [0, w] (w = 1 – 10 s).

as (PC_j) and (PO_j), respectively. Then, we calculated Pearson correlation coefficients $r(FtoI_{wj}, PC_j)$ and $r(FtoI_{wj}, PO_j)$ for each of the surrogate sequences. The sets of correlation coefficients, $\{r\left(FtoI_{wj}, PC_j\right)\}$ and $\{r\left(FtoI_{wj}, PO_j\right)\}$, provide the null distributions for $r(FtoI_{wj}, C_j)$ and $r(FtoI_{wj}, O_j)$, respectively.

Finally, in step 7, using the null distributions, we z-standardized the correlation values. This results in the feedback-coordinated convergence z-score $z(r\left(FtoI_{wj}, C_j\right))$ and the feedback-coordinated originality z-score $z(r\left(FtoI_{wj}, O_j\right))$, defined as follows:

$$z\left(r\left(FtoI_{wj}, C_j\right)\right) = \frac{r\left(FtoI_{wj}, C_j\right) - mean(\{r(FtoI_{wj}, PC_j)\})}{SD(\{r\left(FtoI_{wj}, PC_j\right)\})} \quad (12)$$

$$z(r\left(FtoI_{wj}, O_j\right)) = \frac{r\left(FtoI_{wj}, O_j\right) - mean(\{r(FtoI_{wj}, PO_j)\})}{SD(\{r(FtoI_{wj}, PO_j)\})}. \quad (13)$$

I. TEMPORAL COORDINATION BETWEEN IMMEDIATE TURN-TAKING AND IDEA GENERATION

It has been suggested that a long gap after a question indicates that the recipient may have a problem with the question [22]. Here, we considered every turn-taking that occurred within a second as immediate turn-taking. Finally, we hypothesized that idea sharing occurs more frequently in response to an immediate turn-taking event. If this is the case, immediate turn-taking has a temporally spot-like (i.e., instantaneous and fleeting) impact on idea sharing.

To explore this, given tt as the time in seconds when an immediate change of speaker happened, we defined the immediate turn-taking sequence T_t as follows:

$$T_t = \begin{cases} 1 & if\ t = tt \\ 0 & otherwise. \end{cases} \quad (14)$$

Next, to compute $(ITtoI_w)$, which captures how an idea is likely to be generated following immediate turn-taking

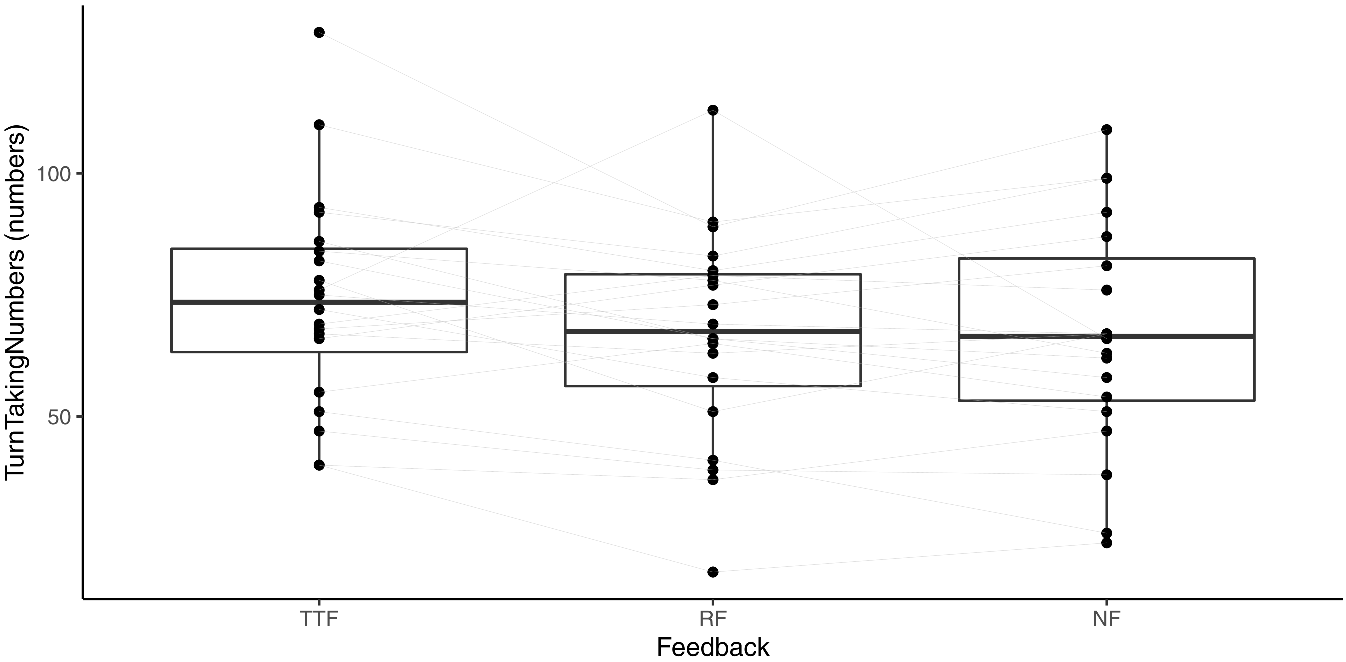

FIGURE 2. Effects of feedback mode on turn-taking.

with delay within [0, w] s, we repeated steps 2 to 4 from Section II.H, by substituting the F_t to T_t sequence. As we had three trials per group, in step 4, after computing the $z(ITtoI)$, we also calculated the mean value as follows:

$$mean(z(ITtol)) = \frac{\sum_{i=1}^{3}(z(ITtol))_i}{3}. \quad (15)$$

Using the $mean(z(ITtol))$ for each time window, we could reject or accept our hypothesis over the temporal coordination of idea generation and immediate turn-taking. Comparing the results from different time windows, we could also suggest the most optimized time lag between these two sequences and determine whether any feedback mode had enhanced this temporal coordination. Lastly, we assessed the immediate turn-taking-coordinated convergence z-score and the immediate turn-taking-coordinated originality z-score. These other two scores could explain whether the generated ideas immediately after taking turns were different in quality compared with the other ideas.

III. RESULTS

A. EFFECTS OF FEEDBACK MODE ON TURN-TAKING

To check our hypothesis regarding the effect of feedback mode on turn-taking behavior, we evaluated the total number of turn-taking events during each trial for each dyad. A one-way repeated measures ANOVA was conducted on three feedback conditions, and a quasi-significant general tendency was found across the conditions (F(2,38) = 3.08, p = 0.057; Fig. 2). Post hoc tests (with Bonferroni correction) showed that turn-taking events in the TTF condition (M = 74, SD = 22.29) were not significantly different from those in the RF condition (M = 66.75, SD = 21.85; p = 0.06) but significantly higher than in the NF condition (M = 66.65, SD = 23.43; p = 0.02). There was no significant different between the RF and NF conditions (p = 0.97).

As a supplementary analysis to check for possible confounding effects, we conducted a repeated measurement ANOVA on task sets (three sets of objects), and no significant effect was found (F(2,38) = 0.7, p = 0.50). In addition, to check for possible confounding effects of gender, a mixed-design ANOVA with the gender combination as a between-group factor and feedback mode as a within-group factor was conducted. No significant difference was found on either of the factors or their interaction (Supplementary part 2).

B. EFFECTS OF FEEDBACK MODE ON CREATIVITY OUTCOMES

One-way repeated measures ANOVAs performed on the three feedback conditions showed significant differences across conditions in fluency (F(2,38) = 25.6, p < 0.0001; Fig. 3), originality (F(2,38) = 8.11, p = 0.001; Fig. 4), and IOC (F(2,38) = 8.06, p = 0.001; Fig. 5) but not in flexibility (F(2,38) = 2.03, p = 0.14). Post hoc tests (with Bonferroni correction) showed that fluency in the TTF condition (M = 31.70, SD = 15.14) was significantly higher than in the RF condition (M = 22.70, SD = 11.59; p < 0.0001) and NF condition (M = 22.95, SD = 11.07; p = 0.0001). There was no significant difference between the RF and NF conditions (P = 0.82). Further, post hoc tests showed that the originality in the TTF condition (M = 0.12, SD = 0.21) was significantly higher than that in the RF condition (M = −0.15, SD = 0.37; p = 0.001) and NF condition (M = −0.20, SD = 0.22;

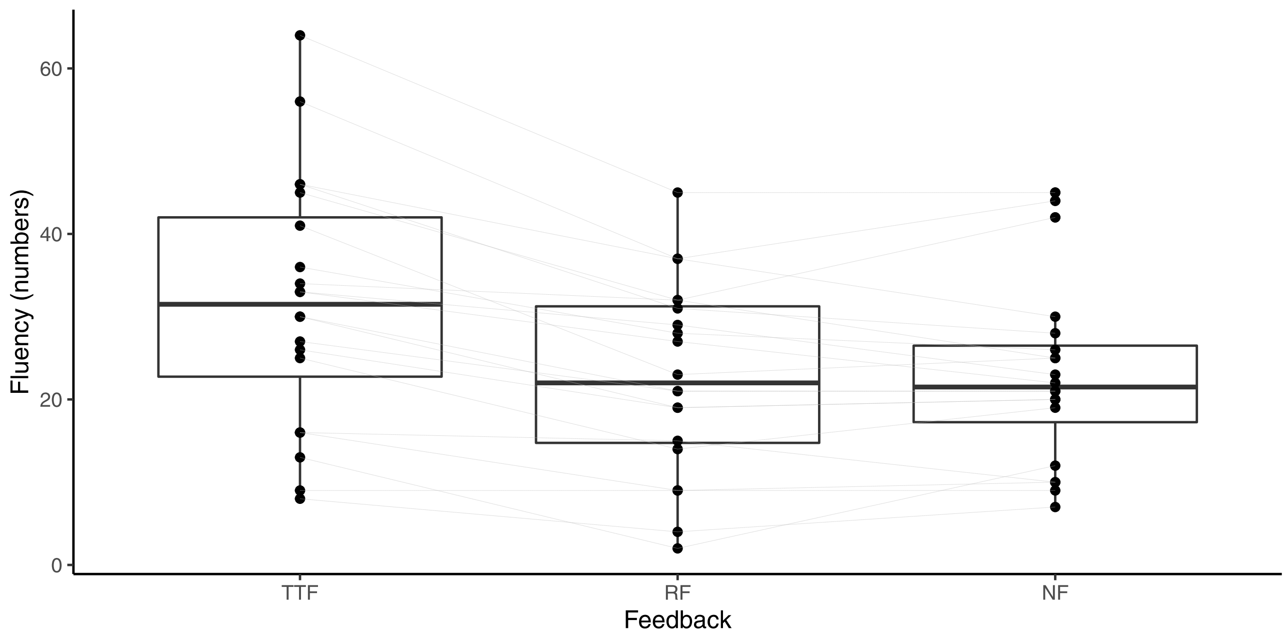

FIGURE 3. Effects of feedback mode on the fluency dimension of group creativity outcomes.

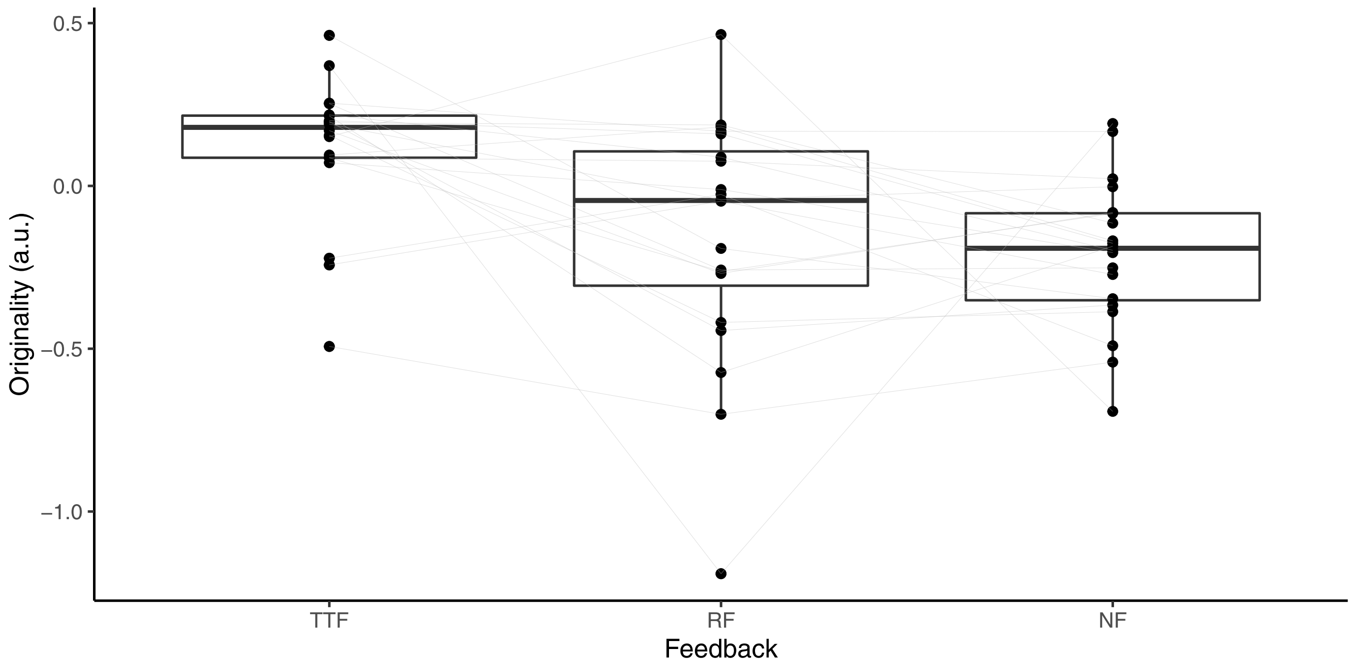

FIGURE 4. Effects of feedback mode on the originality dimension of group creativity outcomes.

$p < 0.0001$). Additionally, there was no significant different between RF and NF conditions ($p = 0.6$). Finally, post hoc test results on IOC showed that the convergence index in the TTF condition ($M = 0.16$, $SD = 0.08$) was significantly higher than in the RF condition ($M = 0.07$, $SD = 0.07$; $p = 0.002$) and significantly higher than in the NF condition ($M = 0.08$, $SD = 0.08$; $p = 0.01$). There was no significant different between the RF and NF conditions ($p = 0.41$).

Furthermore, as supplementary analyses to check for possible confounding effects, we conducted a repeated measurement ANOVA on task sets. No significant effect was found for fluency ($F(2,38) = 1.32$, $p = 0.27$), originality ($F(2,38) = 0.04$, $p = 0.95$), IOC ($F(2,38) = 0.21$, $p = 0.80$), or flexibility ($F(2,38) = 0$, $p = 1$). To check for possible confounding effects of gender, mixed-design ANOVA with the gender combination as a between-group factor and feedback mode as a within-group factor was conducted. No significant difference was found on either of the factors or their interaction for creativity outcome dimensions, except for IOC (Supplementary part 3).

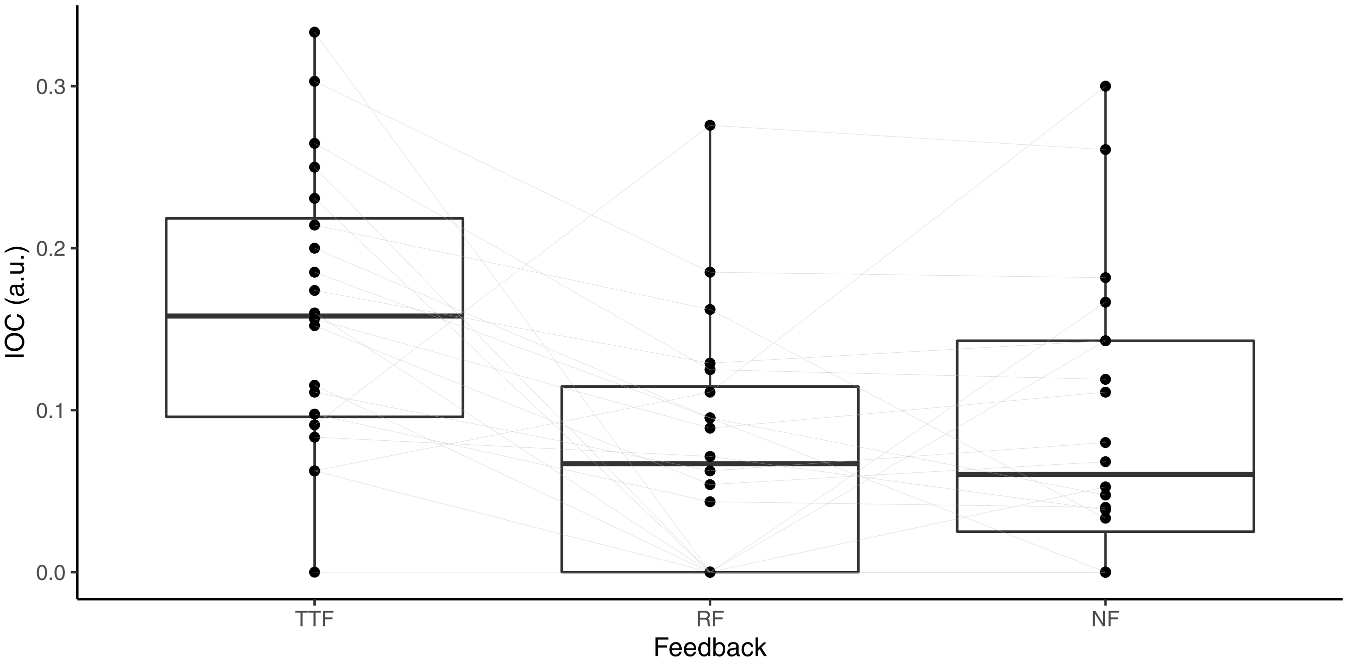

FIGURE 5. Effects of feedback mode on the index of convergence (IOC) dimension of group creativity outcomes.

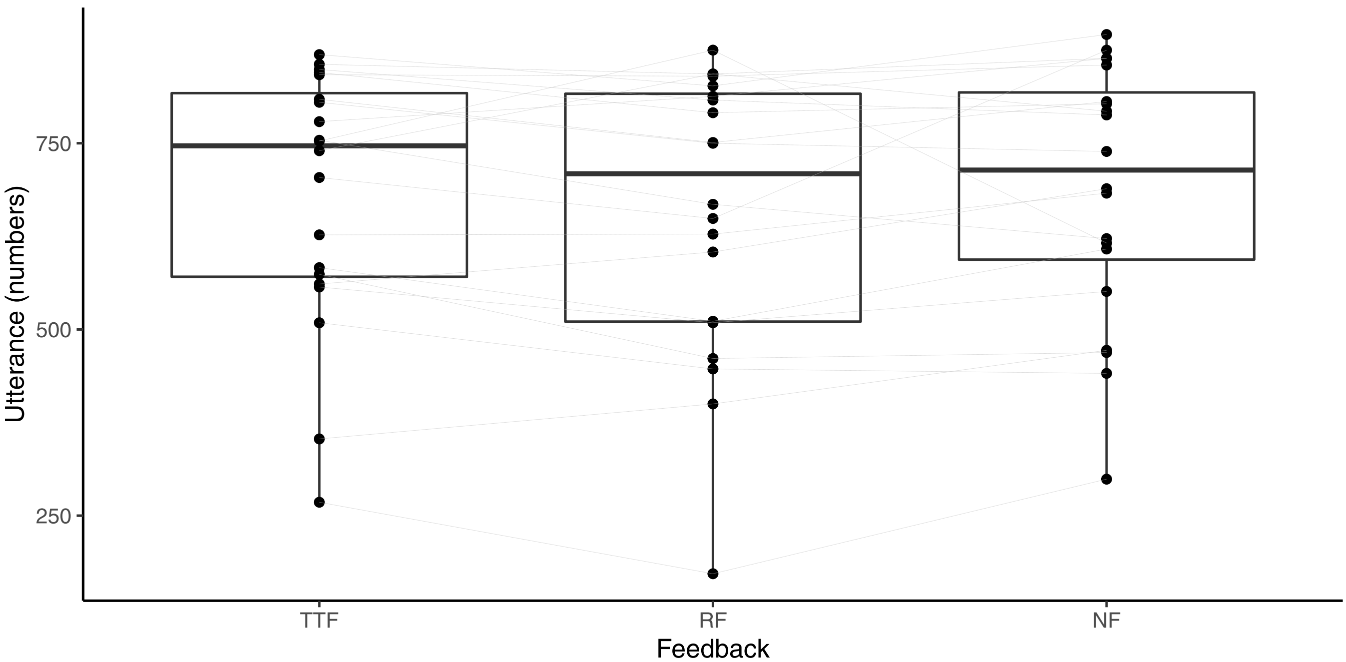

FIGURE 6. Effects of feedback mode on utterance time.

Finally, to clarify whether feedback specifically affected idea generation or generally the whole communication, we measured the effect of feedback conditions on the total utterance time. The utterance time includes not only idea generation, but also irrelevant speech. The one-way repeated measures ANOVA performed on the three feedback conditions showed no significant differences across conditions ($F(2,38) = 1.27$, $p = 0.29$; Fig. 6).

C. EFFECTS OF FEEDBACK MODE ON MOOD

To test the effect of feedback mode on mood, we evaluated the self-reported scores during each trial for each dyad. A one-way repeated measures ANOVA was conducted on valence ($F(2,38) = 2.36$, $p = 0.108$), arousal ($F(2,38) = 0.23$, $p = 0.79$), HAPM ($F(2,38) = 0.91$, $p = 0.40$), and HANM ($F(2,38) = 4.51$, $p = 0.01$, Fig. 7). We found a significant difference in HANM between the conditions. Post hoc tests

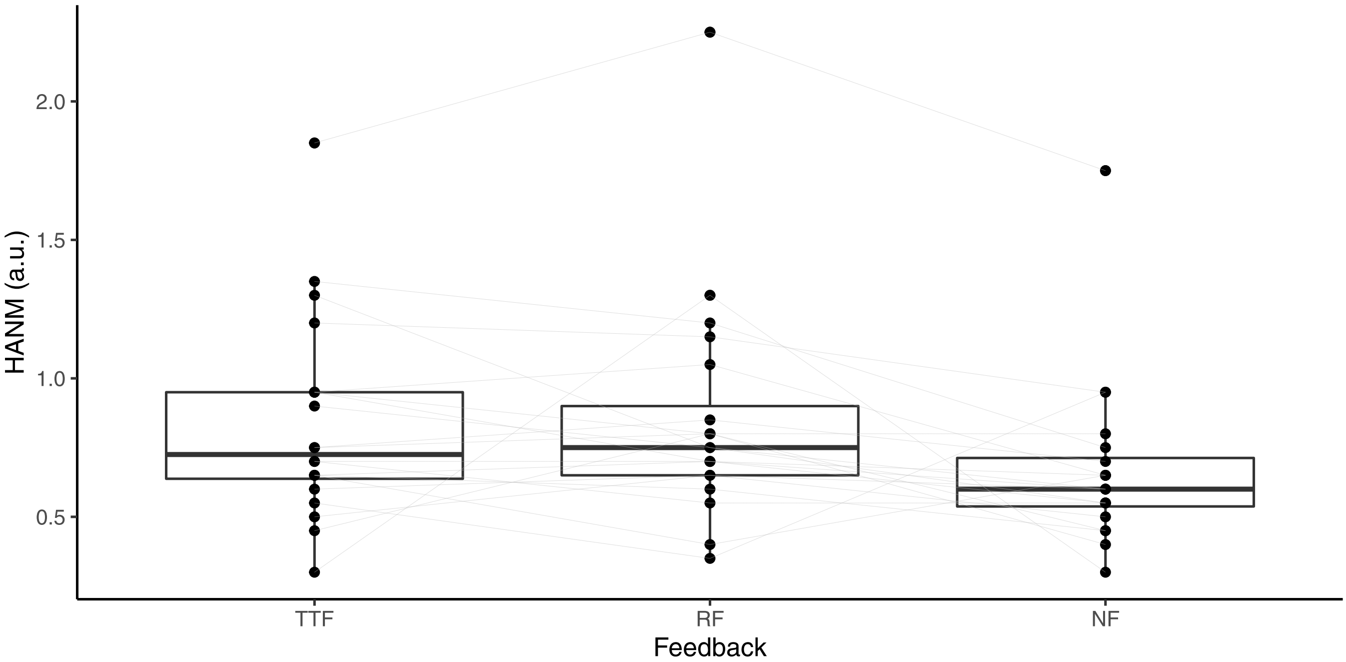

FIGURE 7. Effects of feedback mode on HANM from the self-reported mood.

(with Bonferroni correction) showed that HANM in the TTF condition (M = 0.83, SD = 0.36) was not significantly different from that in the RF condition (M = 0.85, SD = 0.41; p = 0.83), but it was significantly higher than that in the NF condition (M = 0.67, SD = 0.30; p = 0.01). HANM in the RF condition was also significantly higher than that in the NF condition (p = 0.01).

We conducted a supplementary repeated measurement ANOVA on task sets, and no significant effect was found ($F(2,38) = 0.29$, $p = 0.74$). In addition, mixed-design ANOVA with the gender combination as a between-group factor and feedback mode as a within-group factor showed no significant effects (Supplementary part 4).

D. EFFECTS OF FEEDBACK MODE ON IDEA GENERATION TREND

We hypothesized that the feedback modes may have had a gradual impact on the rate of idea sharing and the quality of ideas. However, by comparing the changes from the first to the second halves of each trial, we found no significant effect of the feedback mode on the trend of the number of shared ideas ($F(2,38) = 0.42$, $p = 0.65$), convergence quality of shared ideas ($F(2,38) = 1.27$, $p = 0.29$), or the originality of the ideas ($F(2,38) = 0.106$ $p = 0.89$) over time. These results indicate that the feedback mode did not affect the gradual trend in the remote creativity outcomes.

E. TEMPORAL COORDINATION BETWEEN FEEDBACK CUE AND IDEA GENERATION

Our alternative hypothesis on the temporal scale of the feedback effect was that it may have worked in a spot-like (i.e., instantaneous and fleeting) manner to enhance idea generation. In other words, ideas may have been generated more frequently in response to a feedback cue during the TTF mode. To test this, we used the surrogated-based z-scores for each time window [0, w] (w = 1 to 10 s), and we conducted one-sample t-tests on 10 surrogate-based z-scores. The results suggested that idea generation is most likely to occur within a second after a feedback cue ($t(19) = 5.03$, $p < 0.0001$). Additionally, there was a significant effect of the time window on this temporal coordination ($F(9,190) = 4.21$, $p < 0.0001$; Fig. 8).

Finally, we checked whether the quality of the ideas generated immediately after the feedback cue was different from the quality of the ideas that were not temporally adjacent to the feedback cue. To do so, we assessed the z-score of the originality and convergence of ideas that were mentioned within a second from the feedback cue (see Section II.H. for more details). Excluding data of sessions with no feedback-coordinated idea generation, separate t-tests were conducted. Separated t-tests over the quality showed no significant difference for either the convergence ($t(17) = 0.79$, $p = 0.43$) or the originality ($t(18) = -0.79$, $p = 0.43$) of the ideas that were mentioned immediately after feedback.

F. TEMPORAL COORDINATION BETWEEN IMMEDIATE TURN-TAKING AND IDEA GENERATION

To check our hypothesis that ideas were generated more frequently in response to an immediate turn-taking, we used the surrogated-based mean-z-scores for each time window [0, w] (w = 1 to 10 s). Next, we conducted one-sample t-tests on 10 surrogate-based mean z-scores. The results

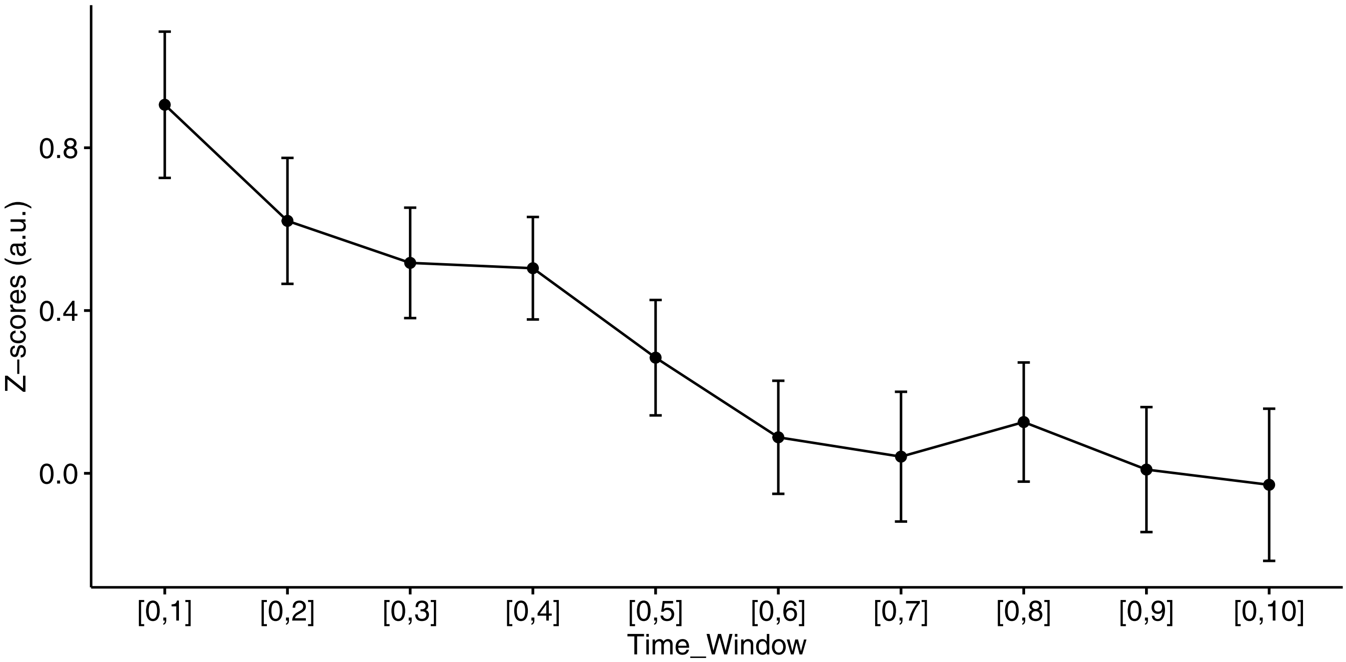

FIGURE 8. Surrogate-based z-scores of feedback cue-to-idea generation, with 10 time-window settings [0, w] (w = 1 – 10 s). Error bars represent standard error of the mean.

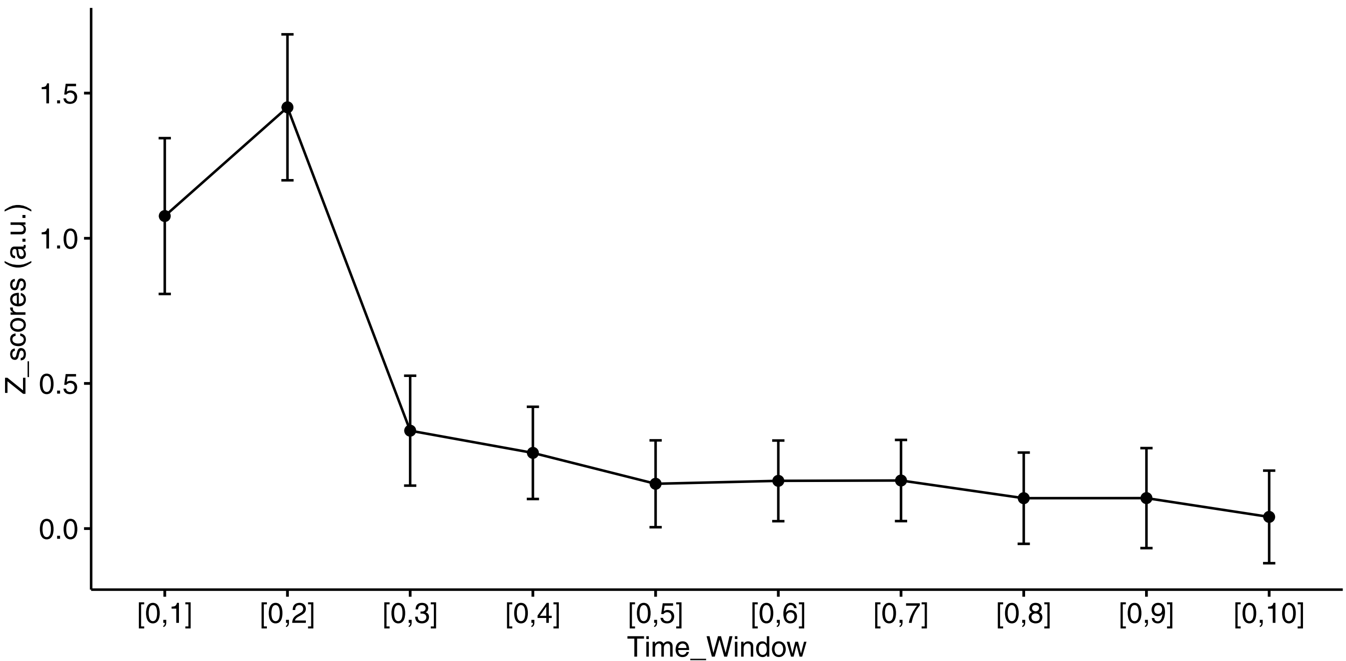

FIGURE 9. Surrogate-based z-scores of immediate turn-taking-to-idea generation, with 10 time-window settings [0, w] (w = 1 – 10 s). Error bars represent standard error of the mean.

revealed that idea generation is most likely to occur within two seconds after taking a turn ($t(19) = 5.77$, $p < 0.0001$). There was also a significant difference across time windows ($F(9,190) = 6.78$, $p < 0.0001$; Fig. 9). Next, we compared z-scores with $w = 2$ s using a one-way repeated measures ANOVA on three conditions to check whether any feedback mode enhanced this temporal coordination. A significant difference was found across the conditions ($F(2,37) = 4.95$, $p = 0.01$; Fig. 10). Post hoc tests (with Bonferroni correction) showed that the chance of idea generation after immediate turn-taking in the TTF condition ($M = 2.14$, $SD = 1.75$) was significantly higher than that in the RF condition ($M = 1.18$,

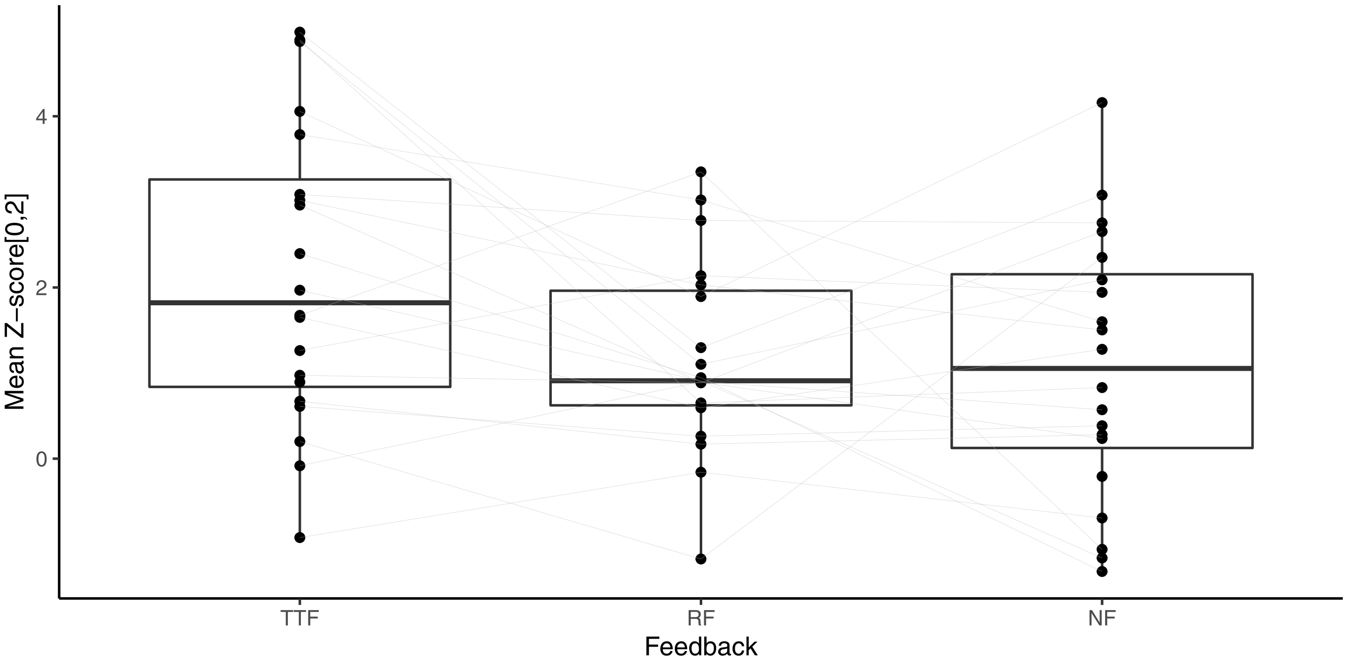

FIGURE 10. Effects of feedback mode on the temporal coordination of immediate turn-taking-to-idea generation.

SD = 1.13; p = 0.01) and NF condition (M = 1.06, SD = 1.53; p = 0.01). There was no significant different between the RF and NF conditions (p = 0.89).

We conducted a repeated measurement ANOVA on task sets, and no significant effect was found ($F(2,37) = 1.75$, $p = 0.18$). A supplementary mixed-design ANOVA with the gender combination as a between-group factor and feedback mode as a within-group factor was conducted. Difference by gender combination was found across the conditions (Supplementary part 5).

Furthermore, we checked whether the quality of the ideas generated immediately after turn-taking shows any difference in terms of originality and convergence. The results were computed in the same manner as in Section III.E. We computed the average scores per group. A significant difference over the quality of those ideas in terms of convergence ($t(18) = 4.17$, $p = 0.0005$) and a quasi-significant difference over originality ($t(19) = -2.07$, $p = 0.05$) was found.

IV. DISCUSSION

In this study, we aimed to determine the impact of turn-taking-based feedback on remote group creativity. We compared the remote group creativity outcomes for three conditions: TTF, RF, and NF. The results showed that the TTF condition improved the number of shared ideas and their quality in terms of the fluency, convergence, and originality but did not significantly increase flexibility and irrelevant speech. This indicates that the feedback cue specifically affected the communication goal through the persistence path.

In the following sections, we discuss our results in two parts: the general impact of feedback on creativity outcomes and mood, and the temporal characteristics of the effect type (i.e., slow gradual change vs. instantaneous), which can provide information about the underlying mechanism of the impact.

A. FEEDBACK IMPACT ON CREATIVITY OUTCOME AND MOOD

Results showed that there is only a quasi-significant difference between all conditions over the total number of turn-taking events, while TTF significantly increased turn-taking events compared with NF. This quasi-significant difference may be because some participants preferred not to receive feedback cues while they were speaking during the TTF condition. Another possibility is that RF could also encourage turn-taking but not as significantly as the TTF condition.

We found that TTF improved creativity outcomes through an encouragement of sharing ideas, in three dimensions, namely, fluency, originality, and IOC but not in flexibility. This indicates that the TTF cues specifically encouraged participants to delve more into one category rather than consider new ones. One possible reason is that during TTF, participants were encouraged to take turn more actively, which resulted in higher IOC scores compared with those in RF. This can be further validated by [11] where turn-taking and cooperation in groups can lead to higher group creativity outcomes in terms of the IOC and originality of the ideas. Originality improvement by delving deeper in a few categories is supported by previous research [13]. Nijstad *et al.* reported that while HAPM enhances flexibility, HANM improves creativity through persistence. Here, the competitive nature induced by the feedback could be related to cognitive persistence. Our results also indicate that the TTF cues increased the participants' HANM. Thus, it can be argued that the TTF cues enhanced HANM and participants' cognitive persistence rather than their cognitive flexibility.

B. UNDERLYING MECHANISMS OF THE EFFECT

We also investigated gradual and instantaneous changes in the quantity and quality of generated ideas to investigate how the feedback worked in the temporal viewpoint. Whereas we found no difference between feedback conditions in terms of slow gradual changes from the first to the second halves of each session, the instantaneous and fleeting effect of the TTF cues was significantly supported.

Our findings show that immediately after a turn-based feedback cue, the dyad would exhibit a higher tendency of sharing an idea, regardless of its quality. Therefore, it is arguable that regardless of the duration of the creativity session of the group (if groups have sufficient time to generate ideas), feedback cues can be beneficial, and the feedback did not take time to show its effect. Our results imply that immediate taking of turns would promote idea sharing. Those ideas would likely be convergent and slightly unoriginal. Previous studies reported the speech pause to be useful for speech only in the presence of cognitive complexity [23]. Pause time is positively correlated to the level of the required processing time [24] and reflects the planning of the next lines of speech [25]. Therefore, it is consistent with the literature that ideas immediately following turn-taking are convergent and unoriginal, regardless of the feedback mode. Moreover, our results show that the TTF condition would specifically increase idea sharing after immediate turn-taking. While both feedback conditions (TTF and RF) enhanced HANM, only the TTF condition induced a higher level of fluency, convergence, and originality on the session level. Combining these results, in addition to the immediate and fleeting effect, we can infer a more contextual or diffusive effect of TTF. Following is a tentative interpretation for such diffusive effect. First, TTF encouraged sharing of ideas after immediate turn-taking. Although those ideas could be relatively convergent and unoriginal, along with HANM, they promoted persistence and stimulated idea generation, leading to the overall enhancement of fluency, convergence, and originality. The actual mechanism of such an effect would be an interesting topic for future studies.

Finally, because of the specific setting of our results, there is a concern about the sustained effect of the developed TTF. For example, in longer interactions, people might get used to or become bored of the feedback and stop responding further. This possibility should be addressed in future research. On the other hand, it should also be noted that in this study, each trial lasted for fifteen minutes, which is a relatively long time for AUT tasks [26]; an AUT session duration typically lasts for five to six minutes [27]. Additionally, it might be expected that people might have had ideas and were waiting for their turn to utter the ideas. Despite this possibility, it should be noted that this study focused on group creativity improvement through cooperation rather than individual creativity. Therefore, even if TTF enhanced the willingness to share ideas rather than the quality of an individual's ideas, our aim was still achieved because more ideas were produced with higher originality as groups.

C. GENERAL DISCUSSION AND CONCLUSION

In the first part of this paper, we discussed that the TTF condition encouraged participants to share their ideas and enhanced idea creativity outcomes in all dimensions except flexibility. Reported emotions revealed that participants experienced higher HANM during the TTF condition compared with the NF condition. Therefore, we assumed that TTF stimulated cognitive persistence by activating negative tone moods. The fact that the TTF type improved sharing of ideas without the need for a long pause for turn-taking is yet another evidence of our assumption that turn-taking-based feedback enhanced cognitive persistence in dyads. It should be noted that audio feedback in this study could not enhance the dimension of flexibility as audio-related TTF enhanced creativity through persistence. However, there is a future possibility of overall creativity outcome enhancement (in all dimensions including flexibility) through the activation of positive mood in participants.

The turn-taking-based feedback we devised in this study is a very simple, easy-to-implement method in the context of remote group creativity. We hope that it can help people who are trying to connect with others and to create innovations through remote interactions, particularly under current COVID-19 pandemic situations. We also hope that results of this study can be helpful in designing human–robot as well as human–computer interactions.

REFERENCES

[1] E. W. A. Elias, P. Chamakiotis, T. J. Howard, E. Dekoninck, and S. Culley, "Can a virtual design environment enhance group creativity and the use of stimuli?" in *Proc. 3rd Int. Conf. Res. Design Eng. (ICORD)*, Bengaluru, India, 2011, pp. 1–8.

[2] S. Harvey, "A different perspective: The multiple effects of deep level diversity on group creativity," *J. Exp. Social Psychol.*, vol. 49, no. 5, pp. 822–832, Sep. 2013.

[3] H. Xue, K. Lu, and N. Hao, "Cooperation makes two less-creative individuals turn into a highly-creative pair," *Neuroimage*, vol. 172, pp. 527–537, May 2018.

[4] J. Annett, *Feedback and Human Behaviour: The Effects of Knowledge of Results, Incentives and Reinforcement on Learning and Performance*. Harmondsworth, U.K.: Penguin, 1969.

[5] A. N. Kluger and A. DeNisi, "The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory," *Psychol. Bull.*, vol. 119, no. 2, p. 254, 1996.

[6] C. Baadte and W. Schnotz, "Feedback effects on performance, motivation and mood: Are they moderated by the learner's self-concept?" *Scand. J. Educ. Res.*, vol. 58, no. 5, pp. 570–591, Sep. 2014.

[7] J. M. DiMicco, K. J. Hollenbach, A. Pandolfo, and W. Bender, "The impact of increased awareness while face-to-face," *Hum.-Comput. Interact.*, vol. 22, nos. 1–2, pp. 47–96, 2007.

[8] J. M. DiMicco, K. J. Hollenbach, and W. Bender, "Using visualizations to review a group's interaction dynamics," in *Proc. CHI Extended Abstr. Hum. Factors Comput. Syst.*, 2006, pp. 706–711.

[9] O. Lederman, A. Mohan, D. Calacci, and A. S. Pentland, "Rhythm: A unified measurement platform for human organizations," *IEEE Multimedia Mag.*, vol. 25, no. 1, pp. 26–38, Jan. 2018.

[10] A. Zolyomi, A. Begel, J. F. Waldern, J. Tang, M. Barnett, E. Cutrell, D. McDuff, S. Andrist, and M. R. Morris, "Managing stress: The needs of autistic adults in video calling," *Proc. ACM Hum.-Comput. Interact.*, vol. 3, pp. 1–29, Nov. 2019.

[11] K. Lu, T. Yu, and N. Hao, "Creating while taking turns, the choice to unlocking group creative potential," *Neuroimage*, vol. 219, Oct. 2020, Art. no. 117025.

[12] W. Stroebe, B. A. Nijstad, and E. F. Rietzschel, "Beyond productivity loss in brainstorming groups: The evolution of a question," in *Advances in Experimental Social Psychology*, vol. 43. Amsterdam, The Netherlands: Elsevier, 2010, pp. 157–203.

[13] B. A. Nijstad, C. K. W. De Dreu, E. F. Rietzschel, and M. Baas, "The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence," *Eur. Rev. Social Psychol.*, vol. 21, no. 1, pp. 34–77, Mar. 2010.

[14] F. G. Ashby and A. M. Isen, "A neuropsychological theory of positive affect and its influence on cognition," *Psychol. Rev.*, vol. 106, no. 3, p. 529, 1999.

[15] M. Baas, C. K. W. De Dreu, and B. A. Nijstad, "A meta-analysis of 25 years of mood-creativity research: Hedonic tone, activation, or regulatory focus?" *Psychol. Bull.*, vol. 134, no. 6, p. 779, 2008.

[16] C. K. W. De Dreu, M. Baas, and B. A. Nijstad, "Hedonic tone and activation level in the mood-creativity link: Toward a dual pathway to creativity model," *J. Pers. Social Psychol.*, vol. 94, no. 5, p. 739, 2008.

[17] J. P. Guilford, "Creativity: Yesterday, today and tomorrow," *J. Creative Behav.*, vol. 1, no. 1, pp. 3–14, Jan. 1967.

[18] H. Sarinasadat, Y. Hattori, Y. Miyake, and T. Nozawa, "Music valence and genre influence group creativity," in *Proc. Int. Conf. Hum.-Comput. Interact.*, 2019, pp. 410–422.

[19] T. S. Larey and P. B. Paulus, "Group preference and convergent tendencies in small groups: A content analysis of group brainstorming performance," *Creativity Res. J.*, vol. 12, no. 3, pp. 175–184, Jul. 1999.

[20] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: The PANAS scales," *J. Pers. Social Psychol.*, vol. 54, no. 6, p. 1063, 1988.

[21] C. Dippo and B. Kudrowitz, "Evaluating the alternative uses test of creativity," in *Proc. Nat. Conf. Undergraduate Res. (NCUR)*, 2013, pp. 427–434.

[22] S. C. Levinson, "Turn-taking in human communication–origins and implications for language processing," *Trends Cogn. Sci.*, vol. 20, no. 1, pp. 6–14, Jan. 2016.

[23] D. Aaronson, "Temporal course of perception in an immediate recall task," *J. Exp. Psychol.*, vol. 76, p. 129, Jan. 1968.

[24] S. S. Reich, "Significance of pauses for speech perception," *J. Psycholinguistic Res.*, vol. 9, no. 4, pp. 379–389, Jul. 1980.

[25] T. T. J. Kircher, M. J. Brammer, W. Levelt, M. Bartels, and P. K. McGuire, "Pausing for thought: Engagement of left temporal cortex during pauses in speech," *Neuroimage*, vol. 21, no. 1, pp. 84–90, Jan. 2004.

[26] N. Hao, M. Wu, M. A. Runco, and J. Pina, "More mind wandering, fewer original ideas: Be not distracted during creative idea generation," *Acta Psychologica*, vol. 161, pp. 110–116, Oct. 2015.

[27] K. P. Madore, H. G. Jing, and D. L. Schacter, "Divergent creative thinking in young and older adults: Extending the effects of an episodic specificity induction," *Memory Cogn.*, vol. 44, no. 6, pp. 974–988, Aug. 2016.

SARINASADAT HOSSEINI was born in Rasht, Iran, in 1994. She received the B.S. degree in electrical engineering from Shahid Beheshti University, Tehran, Iran, in 2016, and the M.S. degree in computer science from the Tokyo Institute of Technology, Japan, in 2019. Her research interests include human–computer interaction, multimodal affective computing, learning analytics, and educational technologies.

XIAOQI DENG received the B.S. degree in biology from Nanjing University, China, in 2009, and the M.S. degree in finance from the EDHEC Business School, France, in 2014. She is currently pursuing the Ph.D. degree in computer science with the Tokyo Institute of Technology, Japan. Her research interests include human–computer interaction, nonverbal communication behaviors, and cognitive science.

YOSHIHIRO MIYAKE (Member, IEEE) received the Ph.D. degree from The University of Tokyo, Japan, in 1989. He has been studied the mechanism of human communication from the viewpoint of co-creation systems. He is currently working as a Professor with the Department of Computer Science, Tokyo Institute of Technology. His current research interests include the visualization of implicit communication based on nonlinear dynamics, the communication support based on human–computer interaction, and gait assist robot based on rhythmic interaction.

TAKAYUKI NOZAWA received the B.S. degree in applied physics from the Tokyo Institute of Technology, Japan, in 1997, and the M.S. and Ph.D. degrees in science from the Interdisciplinary Graduate School of Science and Technology, Tokyo Institute of Technology, in 1999 and 2002, respectively. He was affiliated with Kyoto Sangyo University, the National Institution for Academic Degrees and University Evaluation, Japan Women's University, the Tokyo University of Agriculture and Technology, and Tohoku University. He is currently a specially appointed Associate Professor with the Research Institute of Earth Inclusive Sensing, Tokyo Institute of Technology. His research interests include cognitive neuroscience, communication science, complex systems, human–computer interaction, multimodal affective computing, learning analytics, and educational technologies.