I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18058

The Effect of Visual Openness in Meeting Rooms on Team Productivity and Communication Quality: Evidence from a Randomized Controlled Trial

Shinsuke Asakawa
Akinori Shimono
Mikiko Takahashi
Shoko Yamane
Masaru Sasaki

AUGUST 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18058

The Effect of Visual Openness in Meeting Rooms on Team Productivity and Communication Quality: Evidence from a Randomized Controlled Trial

Shinsuke Asakawa
Saga University

Akinori Shimono
COMANY INC

Mikiko Takahashi
COMANY INC

Shoko Yamane
The Papalaka Research Institute, Ltd. and The University of Osaka

Masaru Sasaki
The University of Osaka and IZA

AUGUST 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Effect of Visual Openness in Meeting Rooms on Team Productivity and Communication Quality: Evidence from a Randomized Controlled Trial*

This study investigates whether a meeting environment's visual openness influences team productivity and communication quality. We conducted a randomized controlled trial with participants assigned to discussions held in a transparent glass meeting room (treatment) or a fully curtained room (control). Team productivity was evaluated based on the quality of participants' policy proposals. Communication quality was assessed using transcript- based indicators such as laugh frequency and topic diversity. We found that groups in visually open meeting rooms received significantly higher proposal ratings and exhibited greater emotional positivity and topic diversity, highlighting that within-session dynamics expose how environmental design affects group interaction.

JEL Classification: C93, J24, M54

Keywords: team productivity, communication quality, visual openness, randomized controlled trial, workplace design, natural language processing

Corresponding author:
Shinsuke Asakawa
Saga University
1 Honjōmachi
Saga, 840-8502
Japan
E-mail: asakawas@cc.saga-u.ac.jp

* This study was approved by the Ethics Committee of the Graduate School of Economics at the University of Osaka and was conducted as a joint research project between COMANY INC., the University of Osaka, and Saga University. We are grateful to all participants involved in the experimental sessions, and especially to Saki Ueda and Kohei Ohura of COMANY INC. for their dedicated support in organizing and implementing the study. We also thank Yuta Kuroda, Kentaro Nakajima, Hideo Owan, Kohei Takahashi, and participants at the JEA 2023 Autumn Meeting, as well as seminar attendees at Kwansei Gakuin University, Waseda University (Personnel Economics Workshop), and Osaka Metropolitan University, for their valuable feedback and suggestions. Financial support from the Japan Society for the Promotion of Science (JSPS KAKENHI: 20K20757 and 23K01557 for Sasaki; 25K16706 for Asakawa) and from Saga University is gratefully acknowledged.

1. Introduction

High team productivity is a cornerstone of organizational performance and a central concern in labor economics. Traditional production functions typically model labor as a homogeneous and separable input, but modern workplaces increasingly rely on team-based production, where output depends not only on individual skills and effort but also on team composition and within-team interactions. Extensive research has documented peer effects, demonstrating that coworkers influence individual performance through mechanisms such as knowledge spillovers, peer pressure, and coordination (Cornelissen, Dustman, and Schönberg 2017; Falk and Ichino 2006; Herbst and Mas 2015; Mas and Moretti 2009). In parallel, the growing literature on team production emphasizes the importance of skill complementarities, task allocation, and collective problem-solving (Boudreau, Lacetera, and Lakhani 2011; Hamilton, Nickerson, and Owen 2003; Lazear and Shaw 2007; Weidmann and Deming 2021).

However, identifying the specific drivers of team productivity remains challenging, as it is shaped by a range of long-term factors. Meetings, capturing critical collective decision-making and information exchange, provide a controlled setting for isolating short-term, causal effects on team outcomes. Much existing meeting research has focused on interpersonal dynamics and process-related factors—such as leadership, participation, relationship building, time management, and proactive engagement (Allen, Lehmann-Willenbrock, and Landowski 2014; Allen and Lehmann-Willenbrock 2023; Berg and Kauffeld 2024; Lehmann-Willenbrock, Allen, and Kauffeld 2013; Rogelberg et al. 2006). However, how the physical environment, including spatial openness, affects team interactions and performance is overlooked (Sicotte et al. 2019; Tanaka et al. 2022). This

represents a critical gap, as meetings' physical environments in can directly shape team dynamics, affecting collaborative processes.

To address this gap, this paper examines whether a manipulable feature of the meeting environment—its visual openness—can causally influence team productivity and communication quality. From an economic perspective, workspace design constitutes a form of capital input in the production function. While the effects of human capital and organizational practices on productivity have been extensively studied, the role of spatial design in shaping intra-team collaboration remains underexplored. Particularly, visual openness, may influence perceptions of psychological safety, social accountability, and visibility, thereby shaping how individuals coordinate.

To investigate this relationship, we conducted a randomized controlled trial (RCT) between 2022 and 2024, involving 220 university students. Participants were randomly assigned to 30-minute small-group discussions in either a visually open meeting room with transparent glass walls (treatment; 104 participants) or a visually closed room with curtains (control; 116 participants). Team productivity was assessed by 200 independent raters evaluating each group's policy proposal. Communication quality was measured through natural language processing applied to meeting transcripts, generating six objective indicators: laugh frequency, silence frequency, silence duration, word count, emotional polarity, and topic diversity.

We found that visual openness improves productivity by approximately 4.4 percentage points. Although average 30-minute communication measures show no significant differences, fixed-effects models at five-minute intervals reveal that openness enhances communication during the middle discussion phase—by increasing emotional

positivity and topic diversity. Other indicators such as laughter, word count, or silences remained unaffected, suggesting temporally specific effects.

This study contributes to the literature on team productivity, communication quality, and workplace design in three ways. First, it provides causal evidence for the impact of a simple, scalable spatial feature on collaborative performance and interaction, overcoming selection bias concerns in observational studies. Second, it employs objective, behavior-based indicators—rather than self-assessments—to assess team output and communication dynamics. Third, it highlights the importance of within-meeting communication dynamics by showing that treatment effects are time-specific and concentrated during the most productive phase of team discussion. Together, these contributions offer novel insights into how workplace environments shape the micro-foundations of productivity.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature. Section 3 describes the experimental design. Section 4 outlines the measurement strategy. Section 5 explains the estimation methods. Section 6 presents the empirical results. Section 7 concludes.

2. Literature Review and Hypotheses

2.1. Team Productivity: Peer Dynamics, Meeting Quality, and Physical Environment

Team productivity has long been recognized as a cornerstone of organizational performance and a central topic in labor economics. Traditional models often treated labor as a homogeneous and separable input, but mounting evidence reveals that productivity in modern workplaces depends not only on individual skills and effort but also on team composition, peer dynamics, interpersonal process quality, and physical

environment design. Thus, we review four strands of research informing our study: (i) individual peer effects, (ii) team-level dynamics, (iii) meeting quality, and (iv) meeting space design.

Research outlines that coworkers' behavior and characteristics shape individuals' productivity. Peer effects operate through mechanisms such as knowledge spillovers, social pressure, and coordination incentives. Mas and Moretti (2009) found that grocery store employees exert more effort when surrounded by productive peers within visual proximity. Falk and Ichino (2006) show that pairing workers increases output, particularly by inducing lower-productivity individuals to adjust upward toward their peers. Herbst and Mas (2015) corroborate these findings, showing that peer effects estimated in laboratory settings generalize to field environments. These studies collectively suggest that peer environments meaningfully alter individual behavior and performance, independent of formal incentive structures.

Beyond individual-level influences, team productivity emerges from member interactions, skill complementarities, and collective problem-solving processes. Hamilton, Nickerson, and Owen (2003) show that introducing team-based production in a garment factory increased output by approximately 14%, particularly among heterogeneous teams that could exploit skill complementarities. Lazear and Shaw (2007) theorize that teams are most productive when workers specialize based on comparative advantage rather than absolute ability. Boudreau, Lacetera, and Lakhani (2011) find that although innovation contests may reduce individual effort, the likelihood of achieving breakthrough solutions for uncertain problems is increased, illustrating how collective dynamics can offset individual disincentives. More recently, Weidmann and Deming (2021) identify "team players"—individuals with high social intelligence who consistently boost their teams'

performance beyond what is predicted by cognitive ability alone. These studies show that team production cannot be understood simply by summing individual contributions; it depends critically on within-team interactions and skill complementarities.

While team composition and peer dynamics are crucial, the quality of interpersonal processes—especially meetings—is vital in shaping team outcomes. Meetings are the primary forums through which teams coordinate, deliberate, and make collective decisions. Rogelberg et al. (2006) find that shorter, more focused meetings are associated with higher job satisfaction and well-being, particularly for workers whose tasks are less interdependent. Allen and Lehmann-Willenbrock (2023) identify five key dimensions of effective meetings—leadership, dialogue, time management, participation, and relationship building—that significantly enhance team productivity and satisfaction. Lehmann-Willenbrock et al. (2013) further show that procedural facilitation, such as agenda setting and summarization, can reduce dysfunctional communication patterns and improve meeting outcomes. Micro-level behaviors such as humor (Lehmann-Willenbrock, Allen, and Kauffeld 2013), proactive engagement (Berg and Kauffeld 2024), and pre-meeting small talk (Allen, Lehmann-Willenbrock, and Landowski 2014) also foster a positive communicative climate. As such, well-structured and relationally supportive meetings can amplify the benefits of effective team composition.

Finally, research suggests that a meeting's physical environment may influence team collaboration and productivity. While traditional research emphasized team composition and behavioral dynamics, recent studies explore how manipulable environmental features affect interaction quality. Sicotte et al. (2019) finds that satisfaction with large, well-equipped meeting rooms is positively associated with both team creativity and effectiveness. Furthermore, Tanaka et al. (2022) show that individuals perform better on

creative tasks in visually open or semi-open spaces, as compared to closed rooms, highlighting the potential importance of spatial openness for idea generation.

However, most existing evidence relies on correlational or self-reported data, limiting causal inference. As a result, it remains unclear whether specific environmental features causally enhance team outcomes in cognitively demanding tasks. To address this gap, our study tests the following hypothesis:

Hypothesis 1: Meeting rooms with greater visual openness yield higher productivity in idea-generation tasks.

2.2. Workplace Environment and Communication Dynamics

In addition to influencing productivity, environmental cues can shape communication tone and content within meetings. For instance, Li et al. (2021) show that warmer lighting conditions reduce negative emotions, while Smolders, de Kort, and Cluitmans (2012) find that higher light intensity improves alertness. Thus, spatial cues such as lighting and openness may affect verbal behaviors and affective tone during collaborative interactions. Building on these, we examine whether brighter and more open meeting environments encourage more positive forms of communication.

Hypothesis 2: Bright, open meeting rooms increase the frequency of positive statements.

2.3. Communication as a Channel Linking Environment and Productivity

Research links communication patterns within teams to productivity outcomes. Field studies show that frequent, effective communication promotes better problem-

solving and team productivity (García-Morales, Matías-Reche, and Verdú-Jover 2011; Giri and Pavan Kumar 2010; Nakajima, Uehara, and Tsuru 2018; Salis and Williams 2008). Further, workspace structure and proximity affect communication; Battiston, Blanes i Vidal, and Kirchmaier (2021) and Battiston et al. (2023) find that physical proximity to supervisors or colleagues enhances responsiveness and coordination. Emanuel, Harrington, and Pallais (2023) report that junior engineers receive more mentoring when seated near senior engineers.

Flexible office layouts such as free-address seating and activity-based working also increase communication frequency (Boutellier et al. 2008; Wohlers and Hertel 2018), although they may reduce depth or continuity of interactions (Haapakangas et al. 2018; Seddigh et al. 2014). While much of this research emphasizes long-term relationships or static teams, less is known about communication dynamics in ad hoc or short-lived teams, especially in relation to physical features like visual openness.

Additionally, many existing studies rely on observational data from subjective survey, limiting causal interpretation. Moreover, spatial design features are often conflated with behavioral variables such as seat choice, making it difficult to isolate their effects. To address these limitations, we use a randomized controlled trial and time-resolved communication metrics to assess how visual openness influences meeting discussion flow. Particularly, we test whether the effect of openness is most salient during the central, idea-generating phase of meetings.

Hypothesis 3: Open meeting rooms facilitate a higher volume of ideas and dynamic discussions, especially during the middle phase of sessions.

3. Experimental Design

3.1. Experimental Setting

We evaluate the impact of meeting layout openness on the team productivity and quality of communication. To achieve this objective, we compared a group of four people communicating in a glass room (semi-closed condition, hereafter referred to as the treatment group) with a group of four people communicating in a closed room fitted with curtains (closed condition, hereafter referred to as the control group). As shown in Figure 1, the primary difference between these two conditions is the degree of visibility of the outside environment. Both rooms were soundproofed to eliminate any potential influence of external noise and ensure that the only salient difference was the visual openness of each layout. Furthermore, to control for extraneous factors, all experiments were conducted in the same room by changing the physical setup (i.e., removing or adding curtains) between the control and treatment conditions.

(Figure 1)

Participant compensation was standardized to avoid any confounding effects related to performance- or outcome-based incentives. Each participant in the main experiment received a flat fee of JPY5,000, including a transportation allowance. Reserve participants who were called upon to join a session were also paid JPY5,000. Those who were not ultimately needed for participation received a fixed transportation allowance of JPY2,000.

At the time of recruitment, reserve participants were informed of the minimum guaranteed amount of JPY2,000 for transportation, as well as an additional waiting fee of JPY1,500 per hour. Since each experimental session lasted approximately two hours,

reserve participants were compensated based on the number of sessions they were asked to remain available for. For example, if they waited for one session, they received JPY5,000 (JPY3,000 for waiting fee and JPY2,000 for transportation); if for two sessions, JPY8,000 (JPY6,000 for waiting fee and JPY2,000 for transportation). This fixed remuneration structure was intended to minimize any incentive for participants to alter their behavior in hopes of receiving higher compensation.

3.2. Experiment Schedule

The experiment was conducted over three periods between 2022 and 2024 at the offices of COMANY INC., a partition manufacturer in Komatsu City, Ishikawa Prefecture, Japan. The first period covered September 5–9, 2022, he second period November 28–29 and December 18–21, 2023, and the third period March 14–15 and 20–21, 2024. During each period, the control and treatment conditions were alternated to distribute any external influences—such as slight seasonal variations or other environmental factors—across both experimental conditions as evenly as possible.

In the first period, nine control groups completed the experiment first, after which the curtain was removed to create a glass-walled environment for the 10 treatment groups. In the second period, 11 control groups ran sessions before transitioning to nine treatment groups. In the third period, nine control groups were followed by seven treatment groups after the curtain was removed. Table 1 presents overall experimental schedule, including the number of groups for each condition in each period, outlining how the experiment was conducted over time.

(Table 1)

3.3. Experiment Participants

A total of 270 university students, recruited from 11 universities in Ishikawa Prefecture, participated in the experiment on a per-session basis. These participants were assigned randomly to one of three groups: a treatment group consisting of 104 participants (26 groups of four), a control group consisting of 108 participants (27 groups of four), and a reserve pool of 50 participants[1]. The reserve participants were scheduled at a rate of two per morning session and two or three per afternoon session to ensure that any absent or late arrivals among the main participants could be replaced promptly.

Measures were taken to minimize preexisting relationships among group members. Specifically, students identified as acquaintances were deliberately placed in separate groups to avoid interpersonal bias. After addressing these known relationships, the remaining participants were assigned randomly to groups.

3.4. Experiment outline

On the day of the experiment, participants typically arrived at the study site by either car or train. For those traveling by train, COMANY INC. provided transportation from the station to the experimental venue using either a company vehicle or taxi. Because some university students required more than one hour to travel each way, this transportation service helped ensure that the sessions began as planned.

Once on site, the main and reserve participants were assigned to separate waiting rooms. They were instructed to remain silent until the official start of the experiment and were explicitly discouraged from engaging in private conversations. However,

[1] The number of sessions conducted under the closed condition was 29; however, two groups presented more than one proposal (groups were asked to submit one), making it impossible to obtain third-party evaluations for those plans. Therefore, these two groups were excluded from the analysis.

participants were permitted to leave the waiting room for restroom breaks and could use electronic devices if needed. The goal of this arrangement was to mitigate any advance discussion or social bonding among participants before data collection.

Following this waiting period, the experimenter spent approximately 10 min explaining the first half of the experimental document to the main participants, providing the study details and answering any questions. At this point, each participant was asked to provide informed consent to proceed. Once all participants provided consent, they were guided to the designated experiment room, where the second half of the experimental document was explained in detail. As soon as this explanation was concluded, a stopwatch was activated to mark the start of the 30-min experimental period.

During these 30 min, two experimenters walked the perimeter of the experiment room at 5-min intervals, conversing on a one-way route. This "walking stimulus" was repeated five times to simulate a modest, controlled level of external movement, akin to a real-world office where people occasionally pass by meeting rooms. At the conclusion of the 30-min session, participants were directed back to their original waiting rooms, where they were asked to complete a post-experiment questionnaire. Once the questionnaires were collected from all participants, they were compensated according to the remuneration structure described above, thus concluding the session.

3.5. Discussion Theme

The central task assigned to each group during the 30-min discussion was to develop ideas for "creating a new national holiday." This topic was chosen to reduce the likelihood of prior knowledge or personal expertise unduly influencing participants' contributions, thereby encouraging relatively spontaneous and creative group discussions. Each group

received a worksheet to record their proposed holiday and the supporting rationale by the end of the session. Participants were also given an experiment manual containing the definition and current list of national holidays, along with a brief outline of the study's objectives and protocols. They were asked to determine what kind of new holiday might benefit Japanese people and why such a holiday should be recognized officially, without any restrictions on the date itself. The worksheet instructed participants to write down one proposed holiday along with the reason for its selection (see Figure A1 for the worksheet format).

To maintain a focus on face-to-face interaction and ensure that conversations were driven by participants' own ideas, use of electronic devices such as smartphones was prohibited during the 30-min discussion. Only writing materials were provided, and participants were encouraged to discuss and record their ideas collaboratively. Interactions were audio and video recorded. The recorder was placed strategically at the center of the table, as shown in Figure 1, while two cameras, one positioned at the angle depicted in Figure 1 and another placed behind the group, were concealed near the ceiling to observe and record the participants' discussions without drawing undue attention. This multimedia recording enabled a comprehensive analysis of both the verbal and nonverbal aspects of communication, thereby contributing to a more detailed understanding of how room layout and openness affect group interactions.

4. **Data**

4.1. Team Productivity: Third-Party Evaluation

To objectively evaluate the productivity of team discussions, this study employed a third-party assessment conducted via an online survey. The respondents consisted of 200

individuals (100 men and 100 women) registered as monitors with a web-based research company, Cross Marketing Inc. Participants were recruited based on their registered prefecture of residence, and recruitment was terminated once the predetermined number of respondents for each prefecture was reached.

The evaluation was conducted over three days, from May 7 to May 9, 2024. The assessment targeted 53 out of the 55 experimental groups that successfully completed the task of formulating one holiday proposal. Of these, 27 belonged to the closed condition and 26 to the semi-closed condition. Each group's proposal was presented to the respondents, who were then asked to rate it based on how convincing they found it.

The evaluation employed a four-point scale; responses were coded numerically from 4 ("agree") to 1 ("disagree") and treated as an ordinal outcome variable reflecting the perceived quality of each group's final proposal. In addition to their evaluations, respondents provided demographic information including gender, age, residential prefecture, occupation (classified into nine categories), marital status, and whether they had children. Thus, we could examine how individual characteristics may relate to assessment patterns, though the primary aim of this evaluation was to offer an impartial, external perspective on the effectiveness of each team's outcome.

4.2. Communication Quality: Text Data

Although a third-party evaluation of discussion outcomes is useful, it has two primary limitations. First, it is subjective, as it inevitably reflects the social circumstances at the time of assessment and the specific cultural context of COMANY Inc., including regional and industrial characteristics. Second, this evaluation focuses solely on the final outcome, offering no information on how the communication quality change over time

during the discussion. To address these constraints, we conducted a detailed analysis of the text data obtained from participants' statements. These text data provided more objective insights into the communication process itself, rather than just its final product.

All participants' utterances were recorded using a voice recorder, and the resulting audio data were converted into text using a four-step procedure: Step 1) audio files sent to a professional transcription service for initial draft scripts; Step 2) two student research assistants reviewed and cross-checked transcribed content accuracy, correcting potential errors in spelling or segmentation; Step 3) the validated transcripts were split into 5-min intervals for a more granular analysis; Step 4) variables were created from these textual data through natural language processing techniques, specifically employing *RMeCab* for morphological analysis and the Google Cloud Natural Language application programming interface ($Cloud\ NL\ API$) for sentiment analysis.

Before this, the transcription company, Tokyo Hanyaku Co.,Ltd., was instructed to (1) identify each speaker using a unique code, (2) note the timing of laughter, (3) record the duration of silences, and (4) mark unclear or inaudible audio segments in the transcripts. Once the text was fully verified, it was aggregated by group, individual participants, or time (in 5-min increments) for further examination. Based on these transcripts, four primary outcome variables were created to capture the different dimensions of communication: "number of laughs," "number and duration of silences," and "number of words in statements," "emotional polarity value" and "number of communication topics."

In many groups, participants took a few minutes at the start of the experiment to introduce themselves or, conversely, engaged in casual chatting once they finalized their holiday proposals. To distinguish between these potentially off-topic conversational

segments and the core discussion, two versions of the outcome variables were created: one that included all 30 min of each session (“with self-introductions and chats”) and one that excluded text identified as self-introductions or chatting (“without self-introductions and chats”). The following sections detail the creation of each outcome variable.

4.2.1. Number of Laughs and Silences

To test the hypothesis that a more open environment leads to livelier discussions, this study examined three variables at the group level for every 5-min interval: (1) number of laughs, (2) number of silences, and (3) total duration of silence. While laughs are recorded in the transcript, it is difficult to precisely identify who laughed, making it appropriate to aggregate this measure by group. Similarly, silences are a collective phenomenon occurring when all participants stop speaking simultaneously; therefore, both the frequency and total length of silences were analyzed at the group level.

4.2.2. Character Data

We evaluated whether room openness influences communication volume, proxied by the total number of words spoken. Specifically, we hypothesized that a more open room induces participants to produce more words. Because of occasional instances of inaudible speech, these segments were removed or edited to maintain intelligible text. Nevertheless, the main analysis retained most individual statements, counted both in 5-min increments and across the entire 30-min session, thus capturing each participant’s verbal output.

4.2.3. Emotional Polarity Data

To further explore whether a more open room fosters more energetic or positive discussions, the study uses "emotional polarity value" as a dependent variable. This reflects the degree of positive sentiment expressed in the participants' words, as determined by the sentiment analysis functionality of the Google Cloud NL API.[2] This approach draws on recent research (Islam, 2018; Mondal et al., 2022; Pandey et al., 2022) and relies on an advanced natural language processing model known as *BERT* (Devlin et al. 2019).

Compared to conventional sentiment dictionaries—for instance, the "Japanese Evaluative Polarity Dictionary"[3] by Inui and Suzuki at Tohoku University and the "Word Polarity Table"[4] by Takamura at the Tokyo Institute of Technology—the Cloud NL API can detect context-dependent variations in meaning more accurately, since it calculates sentiment polarity on a sentence-by-sentence basis.[5] This ensures that no utterances are excluded, even if they include words absent from conventional dictionaries. The Cloud NL API assigns a sentiment polarity score ranging from -1 to 1 in increments of 0.1, and a separate sentiment intensity (magnitude) value of 0 or higher. Following Pandey et al. (2022), two additional dummy variables are derived: one indicating "positive word usage" (score ≥ 0.2 and magnitude ≥ 0.25) and another indicating "negative word usage" (score

[2] The reference URL is https://cloud.google.com/natural-language/docs/sentiment-analysis-client-libraries?hl=ja, date of access: April 19, 2023. In this study, we utilized Python code to apply the [2]Cloud NL API to text data for conducting sentiment analysis. Python code can be shared upon request.

[3] The reference URL is https://www.cl.ecei.tohoku.ac.jp/Open_Resources-Japanese_Sentiment_Polarity_Dictionary.html, date of access: April 19, 2023.

[4] The reference URL is http://www.lr.pi.titech.ac.jp/~takamura/pndic_en.html, date of access: April 19, 2023.

[5] The conventional sentiment polarity dictionaries calculate sentiment polarity values on a word-by-word basis (Takamura, Inui, and Okumura 2006).

≤ -0.2 and magnitude ≥ 0.25). These dummies allow us to account for instances where the overall average polarity may mask strongly positive or negative statements.

4.2.4. Topic Modeling Data

To explore an additional dimension of how participants interacted, we performed topic modeling on the nouns used in the statements of each individual. Specifically, we measured the number of topics each participant expressed over defined time intervals (5-min segments and the entire 30-min discussion period). The text-mining procedure began with a morphological analysis; each statement was broken down at word level and filtered to retain only nouns. This step generated a document-term matrix composed of nouns, excluding stop words, and utilized the International Phonetic Alphabet (IPA) Neologd dictionary to accommodate new or emerging words in Japanese.[6] We then treated each speaker's utterances as separate "documents" according to [1] 5-min increments per individual (i.e., each person's nouns in each 5-min window) and [2] the entire 30-min block per individual (i.e., all nouns spoken by one person across the discussion).

Using these noun-based documents, we performed latent Dirichlet allocation—one of the most common approaches to topic modeling—to identify the optimal number of topics. To determine optimality, we considered two aspects: predictive accuracy (favoring

[6] IPA Neologd is a custom dictionary for MeCab that incorporates newly coined words and named entities gathered from extensive web-based linguistic resources. It contains approximately 3.19 million entries (including duplicates) and is updated at least twice a week, ensuring rapid inclusion of emerging terms. By integrating data from sources such as Hatena Keyword dumps, postal code databases, station names, and personal name lists—and by adding entries for adverbs, adjectives, and interjections not covered in the standard IPA dictionary—it supports more accurate segmentation for expressions otherwise difficult to parse. The dictionary also accounts for spelling variations and informal "broken" spellings frequently found on social networking sites and in news articles. Besides these regular updates, continuous patches correct misreadings, resulting in a consistently up-to-date and high-quality environment for morphological analysis.

a low perplexity score) and interpretability (favoring a high coherence value). Practically, we relied on the *ldatuning* package in R, which calculates several metrics to guide topic-number selection (Arun et al. 2010; Cao et al. 2009; Deveaud, SanJuan, and Bellot 2014; Griffiths and Steyvers, 2004). For Cao et al. (2009) and Arun (2010), a lower score indicates a better fit, while for Griffiths and Steyvers (2004) and Deveaud, SanJuan, and Bellot (2014), higher values are preferable. By examining these four metrics in combination, we arrived at an estimated optimal number of topics that served as an upper bound. We then calculated the number of identified topics each participant expressed in each document during the discussion. This procedure allowed us to capture not only the quantitative aspects of discourse—such as the frequency of words or silences—but also the qualitative variety in speech content through the lens of topic diversity.

4.3. Participants' Personal Traits

After the experiment, participants were requested to complete a brief questionnaire. Among the questionnaire items, only gender and age were associated with the participants' personal traits. Consequently, a female dummy was established, where 1 represented women and 0 represented men, and the age responses were applied in the analysis.

4.4. Descriptive Statistics

4.4.1. Team Productivity: Third-Party Evaluation

Table 2 reports individual characteristics of the 200 evaluators who completed the third-party evaluation questionnaire.

(Table 2)

The average age of the evaluators was 40.2 years with a standard deviation of 10.9, and their ages ranged from 22 to 59 years. The sample included an equal number of men and women, as gender was evenly split at 50 percent each. In terms of family status, 48 percent of respondents were married, and approximately 30.5 percent reported having children. Participants were recruited from across Japan and were distributed across six broad regions. The largest proportion resided in the Kanto area, accounting for 43.5 percent of the sample. This was followed by the Kansai area with 21 percent, the Chubu area with 10 percent, the Hokkaido and Tohoku region with 9.5 percent, Kyushu and Okinawa with 9 percent, and the Chugoku and Shikoku region with 7 percent. Regarding occupational background, most respondents (56 percent) identified as company staff. Temporary or contract workers constituted 11 percent of the sample, while company managers accounted for 9 percent. Other professional categories included self-employed individuals (5.5 percent), medical professionals (6.5 percent), public sector or nonprofit workers (4.5 percent), and company executives (2 percent). Smaller proportions were engaged in legal or business professions (1 percent), agriculture or fisheries (1 percent), or classified their occupation as "other" (3.5 percent).

Table 3 presents the descriptive statistics for these variables by room type. Approximately half of the evaluators rated the proposals as "agree/somewhat agree," while the other half rated them "not so agree/disagree," mirroring the earlier observation that respondents were fairly divided in their opinions. In line with the initial findings, a t-test comparing the "strong disagreement dummy" (where 1 indicates "disagree") between the two room types showed that the semi-closed condition received significantly fewer strong disagreements at the 5% significance level ($t = 2.023^{**}$). Although other agreement-related measures (e.g., "strong agreement dummy" and "agreement dummy")

did not exhibit statistically significant differences, the lower incidence of strong disagreement suggested that the semi-closed layout may have elicited a somewhat more favorable reaction from third-party evaluators.

(Table 3)

4.4.2. Communication Quality: Text Data

Table 4 provides the descriptive statistics at 30-min intervals (group and individual levels). In general, no large disparities emerged between the two room types in the 30-min averages. For instance, the group-level mean numbers of laughs, silences, and total seconds of silence did not differ significantly ($t = 0.43, t = 0.287, and\ t = 0.348$, respectively). Likewise, the individual-level mean number of words spoken and sentiment scores did not present statistically significant gaps when aggregated over 30 min (e.g., $t = 1.051$ for the number of words spoken). Nevertheless, although none of these differences rose to strong statistical significance for the 30-min interval, the semi-closed condition exhibited slightly lower means for laughter, silences, and negative words, but slightly higher means for polarity value and positive words.

(Table 4)

A closer look at Table 5, which presents descriptive statistics aggregated at 5-min intervals, reveals several subtle yet noteworthy patterns. When analyzing the discussion over shorter time spans, distinct differences emerged between the semi-closed (glass-walled) and closed (curtained) conditions. First, the number of words spoken per individual was slightly higher in the closed condition ($t = 1.919*, p < 0.1$). Additionally, negative sentences were more prevalent in this setting ($t = 1.884^{*}, p <$

0.1), suggesting that participants in a visually restricted environment may engage in slightly more negative communication. Conversely, the sentiment polarity score, particularly when the magnitude exceeded 0.25, indicated a borderline significant difference favoring the semi-closed condition ($t = -1.788^{*}, p < 0.1$). This suggests that discussions in more visually open rooms tend to exhibit a more positive tone over time.

(Table 5)

Overall, while the 30-min aggregated data showed minimal differences, the 5-min interval analysis highlighted small yet potentially meaningful variations in communication styles. These results imply that room openness may influence conversational dynamics at specific points in the discussion, rather than uniformly throughout the session.

4.4.3. Data of Participants

Table 6 outlines the personal characteristics of participants according to room type. The proportion of female participants ranged between 35% and 39%, and the average age was roughly 21.4 years, with the youngest participant aged 19 years and the oldest aged 25 years. There was no notable difference in the percentage of female participants across room types, and any age differences appeared modest overall. However, additional tests showed a borderline significant difference in age at the 10% level ($t = 1.745^{*}$), suggesting that participants in one condition were marginally younger, on average.

(Table 6)

5. Empirical Specifications

5.1. Analysis of Team Productivity

This study aimed to determine the impact of room type on discussion productivity by randomly assigning participants to the semi-closed or closed condition in a controlled trial framework. We first measured discussion productivity through a third-party evaluation using the OLS method. As the outcome of interest, $Assessment_{ig}$, was measured on an ordinal scale (four levels), we employed two types of estimators for robustness: (1) the conventional OLS framework and (2) ordinal-response models (ordinal logit and ordinal probit). This dual approach allowed us to verify whether the results were consistent across different assumptions regarding the underlying data-generating process.

5.1.1. Baseline OLS Model

We began with a baseline OLS specification to examine whether group g being in the semi-closed condition predicted higher (or lower) third-party assessments. Denoting the outcome variable by $Assessment_{ig}$—the rating that evaluator i assigns to group g—we estimate

$$Assessment_{ig} = \alpha_1 SemiClose_g + \alpha_2 x_i + \varepsilon_{ig}, \quad (1)$$

where $SemiClose_g$ is a dummy variable taking the value of 1 if group g is assigned to the semi-closed condition and 0 if it is assigned to the closed condition. Vector x_i includes control variables for evaluator i's characteristics, such as evaluator i's fixed effects, gender (female dummy), age, residential region (seven categories), occupation (nine

categories), marital status (dummy), and parental status (dummy). The term ε_{ig} is an error component capturing unobserved factors.

Although OLS can provide an intuitive benchmark for the average treatment effect of room type, it does not fully account for the fact that $Assessment_{ig}$ is an ordinal variable with four distinct categories. Consequently, we estimate the same relationship using ordinal response models to verify the robustness of the results.

5.1.2. Ordinal Logit Model

To better reflect the ordinal nature of $Assessment_{ig}$, we used an ordinal logit specification. In this framework, we modeled the probability that evaluator i's assessment of group g falls at or below category j ($where\ j \in \{1,2,3,4\}$) as follows:

$$logit(P\big(Assessment_{ig} \leq j\big|SemiClose_g, x_i\big) = \alpha_0' + \alpha_1' SemiClose_g + \alpha_2' x_i. \quad (2)$$

Here, $logit(\,)$ denotes the logit link function, $SemiClose_g$ and x_i are as defined above, and α_0' represents thresholds or cut-off points that vary with j. This model exploits the ordered nature of the dependent variable by assuming that "agree," "somewhat agree," "less so agree," and "disagree" reflect increasing levels of disagreement while still allowing for a non-linear probability function.

5.1.3. Ordinal Probit Model

As an additional robustness check, we estimated an ordinal probit model using a probit link function rather than a logit function. Specifically,

$$Probit(P(Assessment_{ig} \leq j|SemiClose_g, x_i) = \alpha''_0 + \alpha''_1 SemiClose_g + \alpha''_2 x_i. \quad (3)$$

While both ordinal logit and ordinal probit specifications address the ranked nature of $Assessment_{ig}$, the differences in distributional assumptions (logistic vs. normal) can lead to subtle variations in coefficient estimates. By comparing the results of these three approaches (OLS, ordinal logit, and ordinal probit), we gained a comprehensive view of the sensitivity of our findings to the methodological choices.

5.1.4. Stepwise Inclusion of Control Variables and Fixed Effects

To examine whether the effect of the semi-closed room condition persisted once we accounted for various evaluator characteristics, we introduced personal attributes into x_i in stages. We started with a baseline model that includes only the room-type dummy ($SemiClose_g$) and no other controls. In the second stage, we added certain demographic features, such as a female dummy and age-group dummies (e.g., 20s, 30s, 40s, 50s, and 60s or older, referencing the teen group). This allowed us to observe whether the evaluators' age or gender systematically influenced their scoring behavior and, in turn, the estimated coefficient for $SemiClose_g$.

In the third stage, we included a more comprehensive set of evaluator attributes, such as residential region, occupation, marital status, and parental status. By doing so, we could detect whether the effect of room type remained significant after controlling for potential differences in evaluators' backgrounds that might drive their assessments. Finally, in the models that include fixed effects, we captured any unobserved factors tied to each evaluator that do not vary across the ratings they gave. Collectively, these stepwise additions revealed whether the estimated effect of the semi-closed room condition was robust or diminished once heterogeneity among evaluators was considered.

Following this, the procedures for OLS, ordinal logit, and ordinal probit estimations ensured that our findings were not overly sensitive to one methodological framework or omitted demographic variables. If the coefficient of $SemiClose_g$ remained stable across all levels of control, the observed relationship was unlikely to be an artifact of unobserved evaluator traits or other confounding factors. Conversely, if the semi-closed effect dissipated once certain controls were introduced, pre-existing differences among evaluators—rather than the room condition itself—may explain the initial results.

5.2. Analysis of Communication Quality

Table 5's data demonstrates that there were no significant differences in the outcome variables based on room type in the 30-min averages. However, Figures A2–A10, which display the means of the outcome variables for each group at 5-min intervals, indicate no significant differences in the means of the outcome variables between the time points. This could be due to the nature of the discussion, where idea generation tended to occur in the first half of the session and summary work in the second half.

Thus, this part of the study explored whether the semi-closed room condition influenced communication quality, as observed in the text data. We conducted two main types of analyses: one based on 30-min aggregated data (capturing overall group- or individual-level outcomes) and another using 5-min interval data (capturing the evolution of communication over time). By comparing these two approaches, we could distinguish between broad, session-wide differences and more detailed, time-specific patterns in group discussions. In both cases, we employed an OLS framework with potential fixed and year fixed effects to account for cross-year variations in the experimental conditions.

Taken together, these two analyses provided a multifaceted view of whether, how, and when the semi-closed condition mattered. The consistent finding of higher (or lower) communication outcomes in the semi-closed setup, across both the overall and time-segmented models, points to a robust effect of room openness. Conversely, if the semi-closed effect appeared only at specific intervals or disappeared once we considered the time dimension, we could gain insights into the temporal nuances of how the physical environment shapes group interactions.

5.2.1. Thirty-Minute Aggregated Analysis

To identify whether a more open (semi-closed) layout yielded systematically different outcomes over the entire 30-min discussion, we began by aggregating each outcome variable at the group or individual level. For group-level outcomes (e.g., number of laughs, number of silences, and total duration of silence), we estimated

$$y_g = \boldsymbol{\beta_1^G} SemiClose_g + \boldsymbol{\beta_2^G} MeanAge_g + v_{year} + \epsilon_g, \quad (4)$$

where y_g denotes the outcome of interest (e.g., average laughs for group g over the entire session), $SemiClose_g$ indicates whether group g was assigned to the semi-closed (1) or closed (0) condition, and $MeanAge_g$ is the average age of group members. The term v_{year} controls for year fixed effects, capturing any unobserved differences across the multiple years in which the experiment was conducted, and ϵ_g is an error term.

In cases where the outcome variable is aggregated at the individual level (e.g., total number of words spoken by individual *iii* across 30 min), we analogously write

$$y_i = \boldsymbol{\beta_1^I} SemiClose_i + \boldsymbol{\beta_2^I} MeanAge_G + v_{year} + \epsilon_i, \quad (5)$$

where y_i refers to an individual-level measure, $SemiClose_i$indicates whether individual i is in a group assigned to the semi-closed room, and ϵ_i is an individual-level error term.

5.2.2. Five-Minute Interval Analysis

While examining 30-min totals highlighted the overall differences, it may have obscured time-varying effects—for instance, if the semi-closed condition primarily affected early brainstorming or late-stage summarizing. To capture these dynamics, we further broke the discussion into six intervals of 5 min each, introducing time dummies $Time_t$ for $t = 1, \dots, 6$. This approach enabled us to observe whether communication differed by room type at specific phases of the discussion.

For group-level outcomes measured every 5 min (e.g., laughs or silences aggregated at the group level within each interval), the model can be written as

$$y_{g,t} = \sum_{t=1}^{6} \boldsymbol{\beta}_{\mathbf{1},\boldsymbol{t}}^{\boldsymbol{G}} SemiClose_g * Time_t + \boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{G}} MeanAge_g + \mu_t + v_{year} + \epsilon_{g,t}, \quad (6)$$

where $y_{g,t}$ denotes the outcome at time interval t for group g, and $Time_t$ are dummy variables capturing each 5-min segment (e.g., 0–5, 5–10, …). Here, $\boldsymbol{\beta}_{\mathbf{1},\boldsymbol{t}}^{\boldsymbol{G}}$ measures the effect of being in a semi-closed room specifically at interval t, while $\mu_t$is a time fixed effect capturing any general trends across all groups for interval t. We again included year fixed effects (v_{year}) to adjust for differences across experimental years.

If the variable of interest was recorded at the individual level for every 5-min segment (e.g., number of words individual *iii* speaks in interval *t*), we can specify

$$y_{i,t} = \sum_{t=1}^{6} \boldsymbol{\beta}_{\mathbf{1},\boldsymbol{t}}^{\boldsymbol{I}} SemiClose_g * Time_t + \boldsymbol{\beta}_{\mathbf{2}}^{\boldsymbol{I}} MeanAge_g + \mu_t + \mu_i + v_{year} + \epsilon_{i,t}, \quad (7)$$

where $y_{i,t}$ is the outcome for individual i in group g at time t. Notice that $\mu_i$can capture individual fixed effects, accounting for unobserved characteristics of each participant (e.g., personality and communication style), while μ_t again captures time-specific effects shared by all participants. As before, v_{year} is a year fixed effect, and $\epsilon_{i,t}$ is the idiosyncratic error.

6. **Estimated Results**

6.1. Effects on Team Productivity

We began by examining whether the semi-closed room condition influenced external reviewers' evaluations. Table 7 reports estimated results using the OLS, ordered logit, and ordered probit models for a four-level dependent variable (e.g., "agree," "somewhat agree," "less so agree," and "disagree"). The coefficient of interest is a semi-closed dummy, and the four columns in Table 7 present the results for varying sets of controls. These include no controls (Column 1), individual-level dummies (Column 2), a female dummy plus mean age (Column 3), and additional demographic covariates (Column 4).

(Table 7)

Across all specifications, the semi-closed indicator remains positive and statistically significant. In the OLS-based fixed effects models, the semi-closed condition raised the evaluation score by approximately 0.038 points (significant at either the 1% or 5% level, depending on the controls). The ordered logit and ordered probit models also yielded positive estimates, implying that the probability of attaining a higher evaluation category increased when discussions occurred in a semi-closed room. These results suggest that a semi-closed layout significantly improved third-party perceptions of

discussion outcomes. Considering that the outcome mean of the control group was approximately 2.47, this effect size was nontrivial and remained robust when additional demographic covariates were included.

6.2. Effects on Communication Quality

The 30-min aggregated results broadly disclose how the semi-closed condition influenced communication across the full discussion, as shown in Tables 8–11. Table 8 summarizes the impact of the semi-closed condition on the number of laughs, number of silences, silence duration, and number of words spoken during the 30-min session. Table 9 focuses on the sentiment polarity score, proportion of positive and negative sentences, sentiment polarity (magnitude > 0.25), and number of topics discussed in the same 30-min window. For the 5-min interval analyses, Table 10 reports the estimates for laughs, silences, silence duration, and words spoken in each time segment, whereas Table 11 examines sentiment polarity, positive/negative sentence rates, sentiment polarity (magnitude > 0.25), and number of topics from a time-segmented perspective.

(Tables 8–11)

These outcomes were estimated under two specifications in the above tables. Column 1 in each table controls only for time-fixed effects [$FE(T)$], suitable for identifying causal effects in this RCT setting. Column 2 introduces additional controls to check robustness: time and year fixed effects plus mean age for group-level outcomes [$FE(T,Y) + Mean\,Age$], whereas for individual-level variables, it further adds individual fixed effects: [$FE(T,Y,I) + Mean\,Age$]. The first specification provides a direct estimate of the causal effect of room openness, while the second ones serve as a

robustness check considering potential year-to-year variations and differences in group-level demographics.

Figures 2–10 illustrate these results visually, with the left side of each figure displaying the 30-min aggregated analysis and the right side showing the 5-min interval analysis. This distinction is maintained throughout our discussion. When referring to the figures, the appropriate side of the panel is specified to clarify whether the effect was observed over the full session or within specific time segments.

(Figures 2–10)

6.2.1. Thirty-Minute Aggregated Findings

The 30-min aggregated analysis provides an overall perspective on whether the semi-closed condition led to sustained differences in communication patterns throughout the discussion. Table 8 and the left panels of Figures 2–5 indicate that the semi-closed condition did not significantly impact the number of laughs, number of silences, duration of silence, or number of words spoken over the full discussion period. The coefficients of the semi-closed dummy remain statistically insignificant across both the baseline [$FE(T)$] and robustness check [$FE(T,Y) + Mean\ Age$ for group-level data; $FE(T,Y,I) + Mean\ Age$ for individual-level data] specifications. Similarly, Table 9 and the left panels of Figures 6–10 reveal no systematic effect on sentiment polarity, the proportion of positive or negative sentences, sentiment polarity among high-intensity statements (magnitude > 0.25), or the number of topics covered in the discussion.

The absence of significant findings in these aggregated models contrasts with the positive effect of the semi-closed condition observed in the third-party evaluations. This suggests that the external reviewers may be responding to qualitative aspects of the

discussion not captured by these communication measures, or that the effects of room openness are not distributed uniformly over the session but rather emerge in specific phases of the discussion. To explore this possibility further, the 5-min interval analysis examined whether communication patterns differed between the two conditions at specific time points.

6.2.2. Five-Minute Interval Findings

The 5-min interval analysis uncovered more nuanced patterns in how semi-closed layouts influenced discussion dynamics. Table 10 and the right panels of Figures 2–5 show the laughs, silences, silence durations, and words spoken in discrete time segments. While most intervals did not exhibit significant differences between the semi-closed and closed rooms, the final 5-min window (25–30 min) showed a marginally significant decrease in silence duration [coefficient $\approx -32.4, p < 0.10$ in the $FE(T,Y) + Mean\ Age$ specification]. This result suggests that semi-closed groups may be more engaged at the end of the discussion, although this did not appear consistently across all model variations.

Table 11 and the right panels of Figures 6–10 shed light on the sentiment and topic diversity over the discussion. In the initial interval (5–10 min), the semi-closed condition produced a slightly lower sentiment polarity score, indicating a more negative tone, although these estimates lacked robust statistical significance. Moving to the middle phase (10–20 min), polarity significantly increased in semi-closed rooms by 0.034–0.038 ($p < 0.10$ or $p < 0.05$, depending on the exact time segment and specification), suggesting heightened positivity mid-discussion. Moreover, the proportion of positive sentences increased by 0.042–0.047 ($p < 0.05$) in the 15–20 min window, whereas the proportion of

negative sentences decreased by 0.046–0.048 ($p < 0.05$) in the 10–15 min segment. These shifts imply that although semi-closed rooms may start off slightly more negatively, they transitioned into a more positive emotional atmosphere as the discussion progressed.

Regarding topic diversity, Table 11 shows that the number of topics briefly climbed in the 10–15 min interval—with an increase of 2.07–2.32 ($p < 0.1$) in one specification—only to drop sharply in the final 25–30 min segment by -8.12 to -7.59 ($p < 0.1$). This stark decline suggests that semi-closed groups converged on fewer, more focused ideas as they wrapped up, potentially reflecting more decisive or streamlined decision-making.

Overall, the 5-min interval analysis demonstrates that the semi-closed condition did not exert uniform effects throughout the session but rather influenced particular stages of the conversation. Although the aggregated data over 30 min failed to show clear differences, the segment-level findings revealed an initial phase (5–10 min) with mildly negative sentiments, a mid-phase (10–20 min) with increasing positivity and engagement, and a final phase (25–30 min) with heightened topic convergence. These temporal patterns show the dynamic nature of the influence of room openness on communication.

6.2.3. Synthesis of Findings

Combining these results leads to three key conclusions. First, the 30-min aggregated models detected no statistically significant impact of the semi-closed condition on laughter, silence, word count, sentiment, or topic measures, in contrast with the consistently positive third-party evaluations. This suggests that reviewers may be reacting to qualities not fully captured by these text-based metrics, or that room openness exerted time-specific effects that became less visible when averaged over the entire session.

Second, a 5-min interval analysis revealed that semi-closed discussions evolved differently at various points. There is a minor negativity in the opening (5–10 min), a significant upswing in positive sentiment by the midpoint (10–20 min) of roughly 0.034–0.038 in polarity ($p < 0.1$ or $p < 0.05$), and a notable contraction in the number of topics (-8.12, $p < 0.01$) in the final segment (25–30 min). Additionally, the proportions of positive and negative sentences moved in opposite directions mid-discussion, underlining the dynamic emotional shift that occurs in semi-closed rooms.

Third, the robustness checks [i.e., $FE(T, Y) + Mean\ Age$ for group-level data and $FE(T, Y, I) + Mean\ Age$ for individual-level data] confirmed that the observed patterns were not artifacts of year-to-year variations or demographic differences. Although the baseline FE(T) specifications were sufficient, in principle, for an RCT, these additional controls strengthen confidence in the findings by demonstrating that semi-closed effects persisted once potential confounders were accounted for.

Overall, room openness influenced the discussion in a nonuniform, phase-dependent manner, consistently enhancing sentiment or focusing on certain intervals without uniformly reshaping the entire conversation. This temporal insight aligns with the notion that spatial layout can promote or dampen engagement at critical moments, rather than having a static effect across the full 30-min session.

7. **Concluding Remarks**

This study investigated whether a manipulable spatial feature of the meeting environment—its visual openness—can causally influence team productivity and communication quality. Leveraging a RCT, we compared small-group discussions conducted in visually open meeting rooms (transparent glass walls) with those in visually

closed rooms (covered by curtains). Unlike prior studies relying primarily on self-reported or observational assessments, our analysis combined third-party evaluations of team output with objective, transcript-based measures of communication quality, including laugh frequency, silence frequency and duration, word count, emotional polarity, and topic diversity.

Our findings yield four main insights. First, teams in visually open rooms produced policy proposals that received significantly more favorable evaluations from third-party raters, suggesting enhanced team productivity. Second, communication was more dynamic during the middle phase (10–20 minutes) of discussions in the visually open condition, as evidenced by more laughter and less silence. Third, participants in the visually open condition spoke more words overall during the first two-thirds of the session, indicating greater verbal engagement. Fourth, although negative sentiment was initially higher in visually open rooms, emotional tone shifted positively midway through the session, implying a transition toward a more constructive atmosphere. Together, these results suggest that visual openness enhances the quality and quantity of group communication, particularly during the most productive portion of the meeting, thereby improving the collaborative output.

Despite these promising results, four limitations merit discussion. First, external validity remains a concern. Our study focused on university students engaged in a creative discussion task within a controlled setting. Whether similar effects would hold in real workplaces with ongoing task interdependence and professional stakes is unclear. In contrast, research by Nakajima, Uehara, and Tsuru (2018) measured labor productivity in a real workplace using wearable sensors to capture communication volume. Consequently, it remains to be seen whether the positive effects of room openness we uncovered would

scale or translate directly to genuine workplace productivity. Future research should assess the causal impact of visual openness on productivity in real organizational contexts.

Second, although our text-based measures of communication provide granular, time-sensitive indicators, their interpretive validity requires further scrutiny. Measures such as laugh frequency and emotional polarity serve as useful behavioral proxies but may not fully capture subjective engagement or interpersonal dynamics. Future studies could validate and enrich these proxies through triangulation with participant surveys or detailed behavioral coding of video recordings.

Third, our analysis does not capture potential motivational pathways within team interactions. For example, a positive reaction from one participant to another's idea might boost the original speaker's motivation to contribute further, potentially creating a feedback loop that reinforces collaborative behavior. While our findings may partly reflect the influence of meeting room layout on the magnitude of such peer effects, this possibility remains untested in the present study. Future research should explore whether room openness influences not only direct productivity effects but also these more complex motivational pathways.

Finally, the absence of pre-treatment communication data introduces potential endogeneity in our within-group analyses. Although robustness checks with group fixed effects support the direction of our results, unobserved team-level characteristics may still confound estimated magnitudes. Collecting baseline conversational data before treatment assignment would allow stronger identification in future experiments.

In summary, this study advances the literature on team production, workplace design, and communication by offering causal evidence that visual openness can improve team discussions and their perceived output. Our results underscore the importance of

meeting-room design as a potential lever for enhancing collaboration and productivity, while also pointing to the need for future studies in naturalistic settings with richer outcome measurement and baseline data collection.

References

Allen, J. A., & Lehmann-Willenbrock, N. (2023). The key features of workplace meetings: Conceptualizing the why, how, and what of meetings at work. *Organizational Psychology Review*, *13*(4), 355–378. https://doi.org/10.1177/20413866221129231

Allen, J. A., Lehmann-Willenbrock, N., & Landowski, N. (2014). Linking pre-meeting communication to meeting effectiveness. *Journal of Managerial Psychology*, *29*(8), 1064–1081. https://doi.org/10.1108/JMP-09-2012-0265

Arun, R., Suresh, V., Madhavan, C. E. V., & Murty, M. N. (2010). On finding the natural number of topics with Latent Dirichlet Allocation: Some observations. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6118 LNAI*(PART 1), 391–402. https://doi.org/10.1007/978-3-642-13657-3_43

Battiston, D., Blanes I Vidal, J., & Kirchmaier, T. (2021). Face-to-Face Communication in Organizations. *Review of Economic Studies*, *88*(2), 574–609. https://doi.org/10.1093/restud/rdaa060

Battiston, D., Blanes I Vidal, J., Kirchmaier, T., & Szemeredi, K. (2023). Peer Pressure and Manager Pressure in Organisations. *Centre for Economic Perfomance*, *1924*, 70.

Berg, A. K., & Kauffeld, S. (2024). Proactive verbal behavior in team meetings: Effects of supportive and critical responses on satisfaction and performance. *Current Psychology*, *43*(23), 20640–20654. https://doi.org/10.1007/s12144-024-05806-y

Boudreau, K. J., Lacetera, N., & Lakhani, K. R. (2011). Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. *Management Science*, *57*(5), 843–863. https://doi.org/10.1287/mnsc.1110.1322

Boutellier, R., Ullman, F., Schreiber, J., & Naef, R. (2008). Impact of office layout on communication in a science-driven business. *R and D Management*, *38*(4), 372–391. https://doi.org/10.1111/j.1467-9310.2008.00524.x

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, *72*(7–9), 1775–1781. https://doi.org/10.1016/j.neucom.2008.06.011

Cornelissen, T., Dustmann, C., & Schönberg, U. (2017). Peer effects in the workplace. *American Economic Review*, *107*(2), 425–456. https://doi.org/10.1257/aer.20141300

Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective Latent Concept Modeling for ad hoc information retrieval. *Document Numerique*, *17*(1), 61–84. https://doi.org/10.3166/dn.17.1.61-84

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference*, *1*(Mlm), 4171–4186.

Emanuel, N., Harrington, E., & Pallais, A. (2023). The Power of Proximity to Coworkers: Training for Tomorrow or Productivity Today? *National Bureau of Economic Research Working Paper Series, 31880*. https://doi.org/10.2139/ssrn.4638214

Falk, A., & Ichino, A. (2006). Clean Evidence on Peer Effects. *Journal of Labor Economics*, *24*(1), 39–57. https://doi.org/10.1086/497818

García-Morales, V. J., Matías-Reche, F., & Verdú-Jover, A. J. (2011). Influence of Internal Communication on Technological Proactivity, Organizational Learning, and Organizational Innovation in the Pharmaceutical Sector. *Journal of Communication*, *61*(1), 150–177. https://doi.org/10.1111/j.1460-2466.2010.01530.x

Giri, V. N., & Pavan Kumar, B. (2010). Assessing the impact of organizational communication on job satisfaction and job performance. *Psychological Studies*, *55*(2), 137–143. https://doi.org/10.1007/s12646-010-0013-6

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(SUPPL. 1), 5228–5235. https://doi.org/10.1073/pnas.0307752101

Haapakangas, A., Hallman, D. M., Mathiassen, S. E., & Jahncke, H. (2018). Self-rated productivity and employee well-being in activity-based offices: The role of environmental perceptions and workspace use. *Building and Environment*, *145*(August), 115–124. https://doi.org/10.1016/j.buildenv.2018.09.017

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team Incentives and Worker Heterogeneity: An Empirical Analysis of the Impact of Teams on Productivity and Participation. *Journal of Political Economy*, *111*(3), 465–497. https://doi.org/10.1086/374182

Herbst, D., & Mas, A. (2015). Peer effects on worker output in the laboratory generalize to the field. *Science*, *350*(6260), 545–549. https://doi.org/10.1126/science.aac9555

Islam, N. (2018). A Novel Framework Using Machine Learning to Effectively Analyze the Faculty Evaluations. *Journal of Education & Social Sciences*, *6*(2), 40–52. https://doi.org/10.20547/jess0621806204

Lazear, E. P., & Shaw, K. L. (2007). Personnel Economics: The Economist's View of Human Resources. *Journal of Economic Perspectives*, *Volume 21*(Number 4), 91–114.

Lehmann-Willenbrock, N., Allen, J. A., & Kauffeld, S. (2013). A Sequential Analysis of Procedural Meeting Communication: How Teams Facilitate Their Meetings. *Journal of Applied Communication Research*, *41*(4), 365–388. https://doi.org/10.1080/00909882.2013.844847

Li, Y., Ru, T., Chen, Q., Qian, L., Luo, X., & Zhou, G. (2021). Effects of illuminance and correlated color temperature of indoor light on emotion perception. *Scientific Reports*, *11*(1). https://doi.org/10.1038/s41598-021-93523-y

Mas, A., & Moretti, E. (2009). Peers at Work. *American Economic Review*, *99*(1), 112–145. https://doi.org/10.1257/aer.99.1.112

Mondal, A. S., Zhu, Y., Bhagat, K. K., & Giacaman, N. (2022). Analysing user reviews of interactive educational apps: A sentiment analysis approach. *Interactive*

Learning Environments, *0*(0), 1–18. https://doi.org/10.1080/10494820.2022.2086578

Nakajima, K., Uehara, K., & Tsuru, T. (2018). The impact of intra-company communication networks on productivity: A quantitative evaluation using wearable sensors. *Keizai Kenkyu (Hitotsubashi University)*, *69*, 18–34.

Pandey, D., Pradhan, B., Wangmo, & Wairya, S. (2022). Understanding COVID-19 response by twitter users: A text analysis approach. *Heliyon*, *8*(8), e09994. https://doi.org/10.1016/j.heliyon.2022.e09994

Rogelberg, S. G., Leach, D. J., Warr, P. B., & Burnfield, J. L. (2006). "Not another meeting!" Are meeting time demands related to employee well-being? *The Journal of Applied Psychology*, *91*(1), 83–96. https://doi.org/10.1037/0021-9010.91.1.83

Salis, S., & Williams, A. M. (2008). Knowledge sharing through face-to-face communication and labour productivity: Evidence from british workplaces. *Proceedings of the European Conference on Knowledge Management, ECKM*, 753–762. https://doi.org/10.2139/ssrn.1162296

Seddigh, A., Berntson, E., Bodin Danielson, C., & Westerlund, H. (2014). Concentration requirements modify the effect of office type on indicators of health and performance. *Journal of Environmental Psychology*, *38*, 167–174. https://doi.org/10.1016/j.jenvp.2014.01.009

Sicotte, H., De Serres, A., Delerue, H., & Ménard, V. (2019). Open creative workspaces impacts for new product development team creativity and effectiveness. *Journal of Corporate Real Estate*, *21*(4), 290–306. https://doi.org/10.1108/JCRE-10-2017-0039

Smolders, K. C. H. J., de Kort, Y. A. W., & Cluitmans, P. J. M. (2012). A higher illuminance induces alertness even during office hours: Findings on subjective measures, task performance and heart rate measures. *Physiology and Behavior*, *107*(1), 7–16. https://doi.org/10.1016/j.physbeh.2012.04.028

Takamura, H., Inui, T., & Okumura, M. (2006). Latent variable models for semantic orientations of phrases. *EACL 2006 - 11th Conference of the European Chapter of*

the Association for Computational Linguistics, Proceedings of the Conference, 201–208.

Tanaka, T., Chikamoto, T., Sakata, H., & Kobayashi, Y. (2022). A study on the effects of differences in workspaces in ABW-oriented offices on intellectual productivity and psychological/physiological measures (Report 1): Overview of the mid-term and summer subject experiments and results of simulated work performance. *Proceedings of the Annual Conference of the Society of Heating, Air-Conditioning and Sanitary Engineers of Japan, 2022 (Kobe)*, *Vol. 8*(Performance Verification and Actual Condition Survey), 73–76.

Weidmann, B., & Deming, D. J. (2021). Team Players: How Social Skills Improve Team Performance. *Econometrica*, *89*(6), 2637–2657. https://doi.org/10.3982/ECTA18461

Wohlers, C., & Hertel, G. (2018). Longitudinal effects of activity-based flexible office design on teamwork. *Frontiers in Psychology*, *9*(OCT). https://doi.org/10.3389/fpsyg.2018.02016

List of Tables

Table 1: Experiment Schedule by Room Condition (Closed vs. Semi-Closed)

Exp. No.	Closed (control)		Semi-closed (treatment))	
	Date	Time Range	Date	Time Range
1	2022-09-05	12:45–14:00	2022-09-07	14:45–16:00
2	2022-09-05	14:45–16:00	2022-09-07	16:45–18:00
3	2022-09-05	16:45–18:00	2022-09-08	10:45–12:00
4	2022-09-06	08:45–10:00	2022-09-08	12:45–14:00
5	2022-09-06	10:45–12:00	2022-09-08	14:45–16:00
6	2022-09-06	12:45–14:00	2022-09-08	16:45–18:00
7	2022-09-07	08:45–10:00	2022-09-09	08:45–10:00
8	2022-09-07	10:45–12:00	2022-09-09	10:45–12:00
9	2022-09-07	12:45–14:00	2022-09-09	12:45–14:00
10	2023-11-28	11:30–12:45	2022-09-09	14:45–16:00
11	2023-11-28	13:30–14:45	2023-12-19	15:30–16:45
12	2023-11-28	15:30–16:45	2023-12-20	09:30–10:45
13	2023-11-29	09:30–10:45	2023-12-20	11:30–12:45
14	2023-11-29	11:30–12:45	2023-12-20	13:30–14:45
15	2023-12-18	09:30–10:45	2023-12-20	15:30–16:45
16	2023-12-18	11:30–12:45	2023-12-21	09:30–10:45
17	2023-12-18	13:30–14:45	2023-12-21	11:30–12:45
18	2023-12-19	09:30–10:45	2023-12-21	13:30–14:45
19	2023-12-19	11:30–12:45	2023-12-21	15:30–16:45
20	2023-12-19	13:30–14:45	2024-03-15	16:45–18:00
21	2024-03-14	08:45–10:00	2024-03-20	11:30–12:45
22	2024-03-14	10:45–12:00	2024-03-20	13:30–14:45
23	2024-03-14	12:45–14:00	2024-03-20	15:30–16:45
24	2024-03-14	14:45–16:00	2024-03-21	11:30–12:45
25	2024-03-14	16:45–18:00	2024-03-21	13:30–14:45
26	2024-03-15	08:45–10:00	2024-03-21	15:30–16:45
27	2024-03-15	10:45–12:00		
28	2024-03-15	12:45–14:00		
29	2024-03-15	14:45–16:00		

Table 2: Descriptive Statistics of Third-Party Evaluators

Variable	Obs	NA	Mean	S.D.	Min.	Max.
Age	200	0	40.2	10.9	22	59
Married (Dummy)	200	0	0.48	0.501	0	1
Female (Dummy)	200	0	0.5	0.501	0	1
Has Children (Dummy)	200	0	0.305	0.462	0	1
Variable	**Obs**	**Percentage**				
Region	200					
... (1) Hokkaido and Tohoku Area	19	9.5%				
... (2) Kanto Area	87	43.5%				
... (3) Chubu Area	20	10%				
... (4) Kansai Area	42	21%				
... (5) Chugoku and Shikoku Area	14	7%				
... (6) Kyushu and Okinawa Area	18	9%				
Occupation	200					
... Company Staff	112	56%				
... Company Manager	18	9%				
... Company Executive	4	2%				
... Public Sector/Nonprofit	9	4.5%				
... Self-Employed	11	5.5%				
... Medical Professional	13	6.5%				
... Legal/Business Professional	2	1%				
... Agriculture/Fisheries	2	1%				
... Temp/Contract Employee	22	11%				
... Other	7	3.5%				

Table 3: Third-Party Evaluation Results by Room Condition

Room type (treatment status)	Closed (control)						Semi-closed (treatment)						
Variable	Obs	NA	Mean	S.D.	Min	Max	Obs	NA	Mean	S.D.	Min	Max	T-test
Dummy: Agreement (Response = 4 or 3)	5400	0	0.500	0.500	0	1	5200	0	0.516	0.500	0	1	t=-1.604
Dummy: Strong Agreement (Response = 4)	5400	0	0.156	0.363	0	1	5200	0	0.163	0.370	0	1	t=-0.979
Dummy: Somewhat Agreement (Response = 3)	5400	0	0.344	0.475	0	1	5200	0	0.352	0.478	0	1	t=-0.930
Dummy: Somewhat Disagreement (Response = 2)	5400	0	0.293	0.455	0	1	5200	0	0.293	0.455	0	1	t=-0.013
Dummy: Strong Disagreement (Response = 1)	5400	0	0.207	0.405	0	1	5200	0	0.191	0.393	0	1	t=2.023**

Note: The mean difference test (Welch's t-test) was conducted, and the test statistics and p-values are reported, with statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01.

Table 4: Descriptive Statistics of Communication Outcomes (Aggregated over 30 Minutes)

Room type (treatment status)	Closed (control)						Semi-closed (treatment)						
Variable	Obs	NA	Mean	S.D.	Min	Max	Obs	NA	Mean	S.D.	Min	Max	T-test
Group-level outcomes													
Number of laughters	27	0	63.4	41.7	9	194	26	0	58.6	38.8	19	152	t=0.43
Number of silences	27	0	49.9	19.3	14	83	26	0	48.3	22.5	11	90	t=0.287
Durations of silence	27	0	566	289	46	1256	26	0	539	280	87	979	t=0.348
Individual-level outcomes													
Number of words spoken	108	0	1099	860	44	4856	104	0	986	712	18	4430	t=1.051
Positive sentences	108	0	0.143	0.0713	0	0.4	104	0	0.145	0.0735	0	0.391	t=-0.226
Negative sentences	108	0	0.121	0.0561	0	0.318	104	0	0.11	0.0545	0	0.279	t=1.428
Sentiment Polarity Score	108	0	0.0186	0.0571	-0.114	0.21	104	0	0.0289	0.0618	-0.133	0.267	t=-1.255
Sentiment Polarity Score (magnitude > 0.25)	107	1	0.0753	0.195	-0.525	0.543	103	1	0.107	0.206	-0.6	0.7	t=-1.148
Number of topics	108	0	6.75	3.78	1	14	104	0	7.29	4.09	1	14	t=-0.994

Note: "T-test" is used to assess differences between means, with the t-value and p-value serving as statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01

Table 5: Descriptive Statistics of Communication Outcomes (5-Minute Intervals)

Room type (treatment status)	Closed (control)						Semi-closed (treatment)						
Variable	Obs	NA	Mean	S.D.	Min	Max	Obs	NA	Mean	S.D.	Min	Max	T-test
Group-level outcomes													
Number of laughters	162	0	10.6	10.1	0	43	156	0	9.77	9.38	0	43	t=0.724
Number of silences	162	0	8.32	5.87	0	28	156	0	8.04	6.17	0	25	t=0.409
Duration of silence	162	0	94.3	72	0	311	156	0	89.8	74.2	0	335	t=0.551
Individual-level outcomes													
Number of words spoken	648	0	183	184	0	995	624	0	164	168	0	882	t=1.919*
Positive sentences	536	112	0.146	0.14	0	1	494	130	0.151	0.146	0	1	t=-0.516
Negative sentences	536	112	0.124	0.12	0	1	494	130	0.11	0.116	0	1	t=1.884*
Sentiment Polarity Score	536	112	0.0194	0.118	-0.7	0.9	494	130	0.0317	0.116	-0.375	0.7	t=-1.691*
Sentiment Polarity Score (magnitude > 0.25)	488	160	0.0721	0.324	-0.8	0.9	450	174	0.113	0.364	-0.8	0.9	t=-1.788*
Number of topics	471	177	11.4	6.53	1	22	433	191	11.5	6.35	1	22	t=-0.419

Note: "T-test" is used to assess differences between means, with the t-value and p-value serving as statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01

Table 6: Descriptive Statistics of Experimental Participants by Room Condition

Room type (treatment status)	Closed (control)						Semi-closed (treatment)						
Variable	Obs	NA	Mean	S.D.	Min	Max	Obs	NA	Mean	S.D.	Min	Max	Test
Female dummy	104	4	0.413	0.495	0	1	104	0	0.442	0.499	0	1	t=-0.419
Age	108	0	21.1	1.35	19	25	104	0	20.8	1.42	19	25	t=1.745*
Experiment fiscal year	108						104						X2=2.025
... FY2022	32	29.6%					40	38.5%					
... FY2023	40	37%					36	34.6%					
... FY2024	36	33.3%					28	26.9%					

Note: In "Test" column, the mean difference test (Welch's t-test) was conducted for numerical data, while the chi-square (χ^2) test was performed for categorical data. The test statistics and p-values are reported, with statistical significance markers: * p<0.1; ** p<0.05; *** p<0.01.

Table 7: Effect of Semi-Closed Room Condition on Third-Party Evaluation Scores (OLS and Ordinal Models)

Dependent Variable	Third-Party Evaluation Score (4 Levels)			
	(1)	(2)	(3)	(4)
OLS/FE	0.038**	0.038***	0.038***	0.038***
	(0.019)	(0.01)	(0.01)	(0.01)
Ordered logit	0.069**	0.103**	0.069*	0.073**
	(0.035)	(0.04)	(0.035)	(0.035)
Ordered probit	0.041**	0.069***	0.042**	0.043**
	(0.021)	(0.023)	(0.021)	(0.021)
Control Variables				
Evaluator FE		✓		
Female D + Age			✓	✓
Other variables				✓
Number of Observations	10600	10600	10600	10600
Evaluators' mean	2.468	2.468	2.468	2.468

Note: Only coefficients for the Semi-Close Dummy are shown here. Other variables include "Marriage D", "Has children D", "Regional FE (6 categories)", "Occupation FE (9 categories)". Values in parentheses indicate standard errors. Statistical significance markers are as follows: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 8: Effect of Semi-Closed Condition on Laughs, Silences, and Word Count (30-Minute Aggregates)

Outcome Variables	Number of Laughs		Number of Silences		Silence Duration		Number of Words Spoken	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Semi-Close	-4.75	4.40	-1.66	1.76	-27.2	-76.9	-113.7	-88.2
	(11.0)	(11.9)	(5.77)	(5.48)	(78.1)	(83.8)	(121.1)	(154.5)
Control Variables								
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE		✓		✓		✓		✓
Mean Age		✓		✓		✓		✓
Individual FE								✓
Observations	53	53	53	53	53	53	212	212
Outcome Mean of Control	61.0	61.0	49.1	49.1	552.8	552.8	1,043.6	1,043.6

Note: Parentheses indicate cluster-robust standard errors (SEs) at the group level. "Mean Age" represents the average age of the group. Statistical significance markers are as follows: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 9: Effect of Semi-Closed Condition on Sentiment and Topic Diversity (30-Minute Aggregates)

Outcome Variables	Sentiment Polarity Score		Sentiment Polarity Score (Magnitude > 0.25)		Positive Sentence Rate		Negative Sentence Rate		Number of Topics (MAX=14)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Semi-Close	0.010	0.009	0.032	0.029	0.002	0.002	-0.011	-0.010	0.538	0.828
	(0.009)	(0.011)	(0.032)	(0.040)	(0.011)	(0.012)	(0.009)	(0.012)	(0.871)	(0.977)
Control Variables										
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE		✓		✓		✓		✓		✓
Mean Age		✓		✓		✓		✓		✓
Individual FE		✓		✓		✓		✓		✓
Observations	212	212	210	210	212	212	212	212	212	212
Outcome Mean of Control	0.024	0.024	0.091	0.091	0.144	0.144	0.115	0.115	7.01	7.01

Note: Parentheses indicate cluster-robust standard errors (SEs) at the group level. "Mean Age" represents the average age of the group. Statistical significance markers are as follows: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table 10: Effect of Semi-Closed Condition on Laughs, Silences, and Word Count (5-Minute Intervals)

Outcome Variables	Number of Laughs		Number of Silences		Silence Duration		Number of Words Spoken	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Semi-Close × 0–5 mins.	-0.792	0.734	-0.043	0.527	-4.40	-12.7	-9.08	-4.83
	(2.71)	(2.81)	(1.75)	(1.69)	(19.5)	(21.2)	(25.6)	(29.6)
Semi-Close × 5–10 mins.	0.242	1.77	-0.719	-0.150	10.3	1.98	-19.7	-15.5
	(3.20)	(3.18)	(1.59)	(1.55)	(18.7)	(18.3)	(24.0)	(27.7)
Semi-Close × 10–15 mins.	-1.45	0.079	0.336	0.906	-8.52	-16.8	-5.00	-0.757
	(2.54)	(2.58)	(1.40)	(1.28)	(15.8)	(18.7)	(24.1)	(28.6)
Semi-Close × 15–20 mins.	-0.313	1.21	0.490	1.06	-10.9	-19.2	-22.0	-17.7
	(2.47)	(2.48)	(1.51)	(1.41)	(17.7)	(18.4)	(26.5)	(31.1)
Semi-Close × 20–25 mins.	-0.821	0.705	-0.489	0.081	10.5	2.22	-26.8	-22.6
	(2.49)	(2.62)	(1.53)	(1.56)	(23.1)	(22.9)	(29.6)	(33.2)
Semi-Close × 25–30 mins.	-1.62	-0.098	-1.23	-0.663	-24.1	-32.4*	-31.1	-26.9
	(1.82)	(2.11)	(1.24)	(1.29)	(19.3)	(17.4)	(29.3)	(34.3)
Control Variables								
Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE		✓		✓		✓		✓
Mean Age		✓		✓		✓		✓
Individual FE								✓
Observations	318	318	318	318	318	318	1272	1272
Outcome Mean of Control	10.2	10.2	8.19	8.19	92.1	92.1	173.9	173.9

Note: Parentheses indicate cluster-robust standard errors (SEs) at the group level. "Mean Age" represents the average age of the group. Statistical significance markers are as follows: *** p < 0.01, ** p < 0.05, * p < 0.10.

Table 11: Effect of Semi-Closed Condition on Sentiment and Topic Diversity (5-Minute Intervals)

Outcome Variables	Sentiment Polarity Score		Sentiment Polarity Score (Magnitude > 0.25)		Positive Sentence Rate		Negative Sentence Rate		Number of Topics (MAX=22)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Semi-Close × 0–5 mins.	-0.002	-0.006	-0.020	-0.029	-0.022	-0.026	-0.003	-0.002	1.10	1.33
	(0.015)	(0.015)	(0.051)	(0.049)	(0.019)	(0.018)	(0.017)	(0.018)	(1.25)	(1.34)
Semi-Close × 5–10 mins.	-0.010	-0.013	-0.058	-0.064	-0.004	-0.009	0.008	0.009	-0.551	-0.290
	(0.019)	(0.020)	(0.059)	(0.064)	(0.023)	(0.025)	(0.017)	(0.018)	(1.12)	(1.13)
Semi-Close × 10–15 mins.	0.036*	0.032	0.102	0.096	0.006	0.002	-0.048**	-0.046**	2.07*	2.32*
	(0.020)	(0.020)	(0.068)	(0.069)	(0.019)	(0.019)	(0.021)	(0.022)	(1.08)	(1.19)
Semi-Close × 15–20 mins.	0.038**	0.034*	0.143**	0.135**	0.047**	0.042*	-0.017	-0.015	-0.576	-0.266
	(0.018)	(0.019)	(0.056)	(0.060)	(0.019)	(0.021)	(0.022)	(0.024)	(1.52)	(1.63)
Semi-Close × 20–25 mins.	0.001	-0.003	0.023	0.012	0.003	-0.002	-0.013	-0.012	1.03	1.30
	(0.021)	(0.021)	(0.065)	(0.066)	(0.028)	(0.027)	(0.015)	(0.016)	(1.40)	(1.51)
Semi-Close × 25–30 mins.	0.026	0.020	0.100	0.087	0.019	0.011	0.0004	0.003	-8.12***	-7.59***
	(0.060)	(0.061)	(0.138)	(0.147)	(0.070)	(0.069)	(0.033)	(0.037)	(1.98)	(2.34)
Control Variables										
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Year FE		✓		✓		✓		✓		✓
Mean Age		✓		✓		✓		✓		✓
Individual FE		✓		✓		✓		✓		✓
Observations	1,030	1,030	938	938	1,030	1,030	1,030	1,030	1,272	1,272
Outcome Mean of Control	0.025	0.025	0.091	0.091	0.148	0.148	0.117	0.117	173.9	173.9

Note: Parentheses indicate cluster-robust standard errors (SEs) at the group level. "Mean Age" represents the average age of the group. Statistical significance markers are as follows: *** p < 0.01, ** p < 0.05, * p < 0.10.

List of Figures

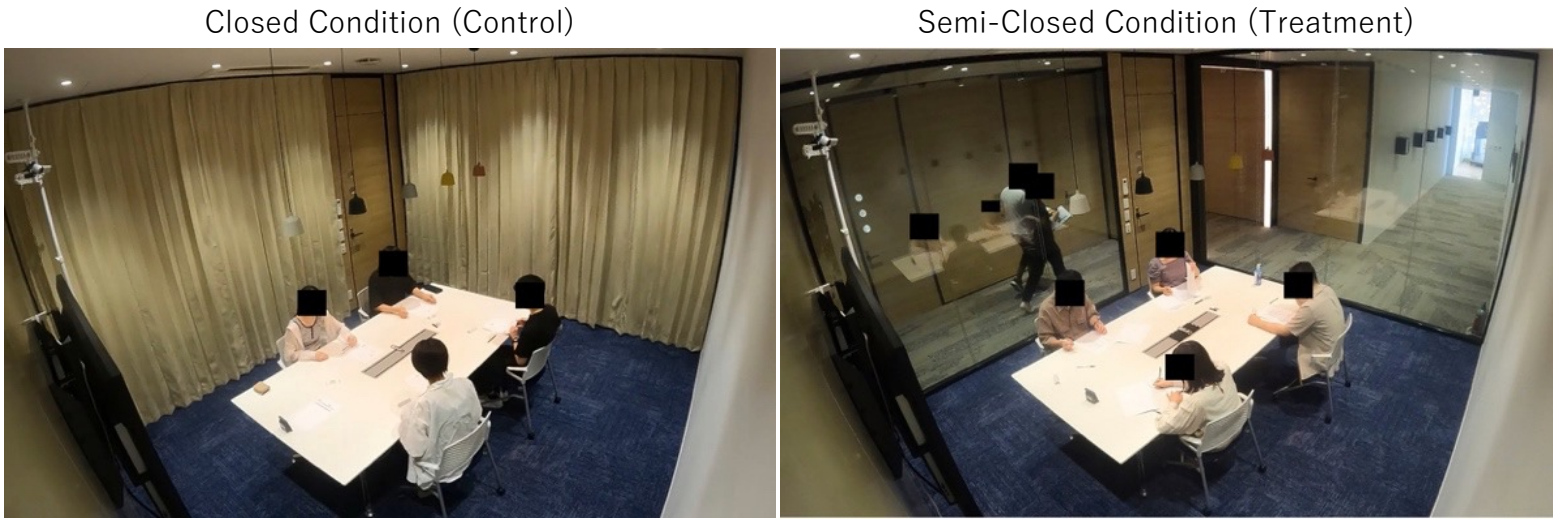

Figure 1: Photograph of the Experimental Room Setup

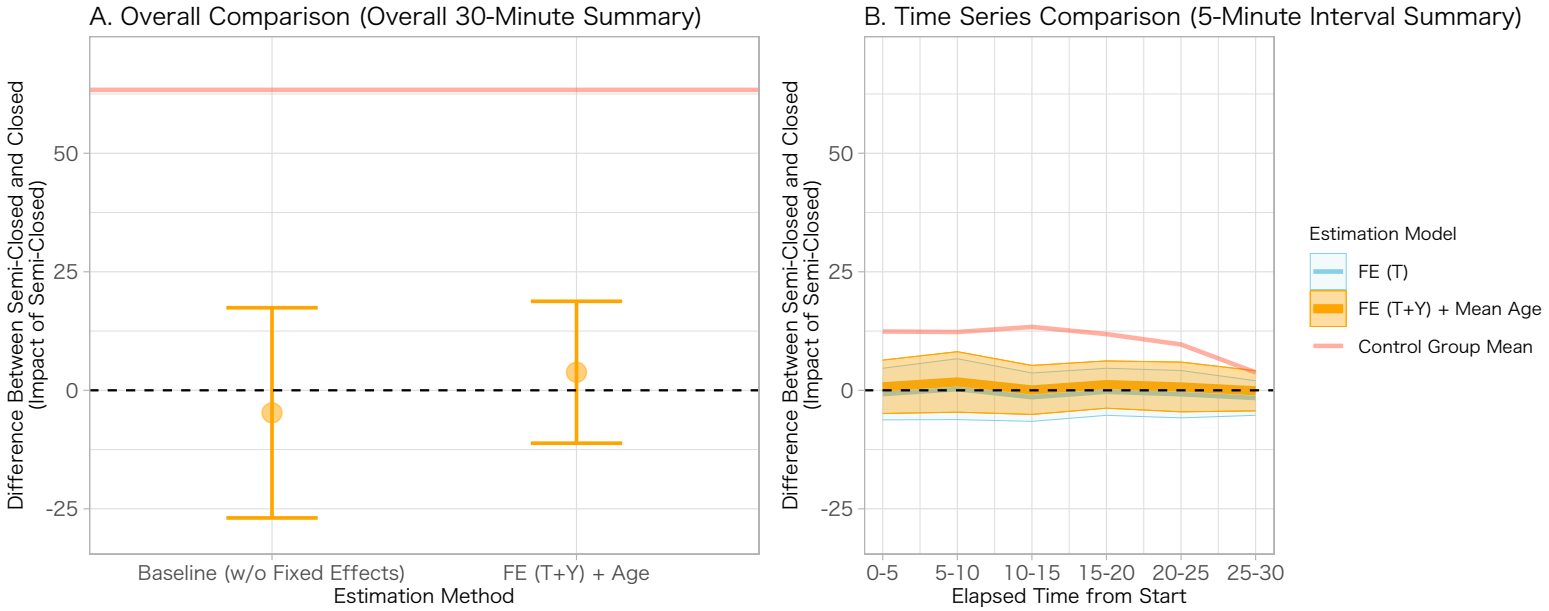

Figure 2: Effect of Semi-Closed Condition on Number of Laughs (30-Minute and 5-Minute Analyses)

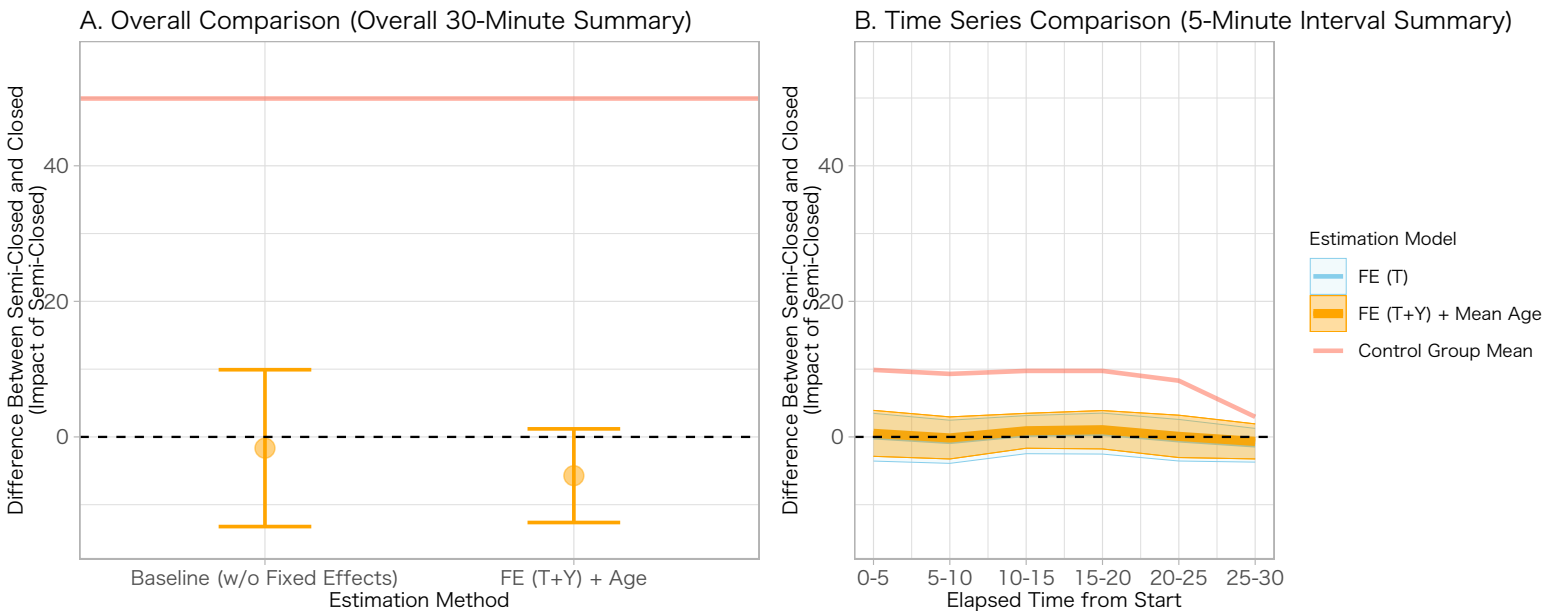

Figure 3: Effect of Semi-Closed Condition on Number of Silences (30-Minute and 5-Minute Analyses)

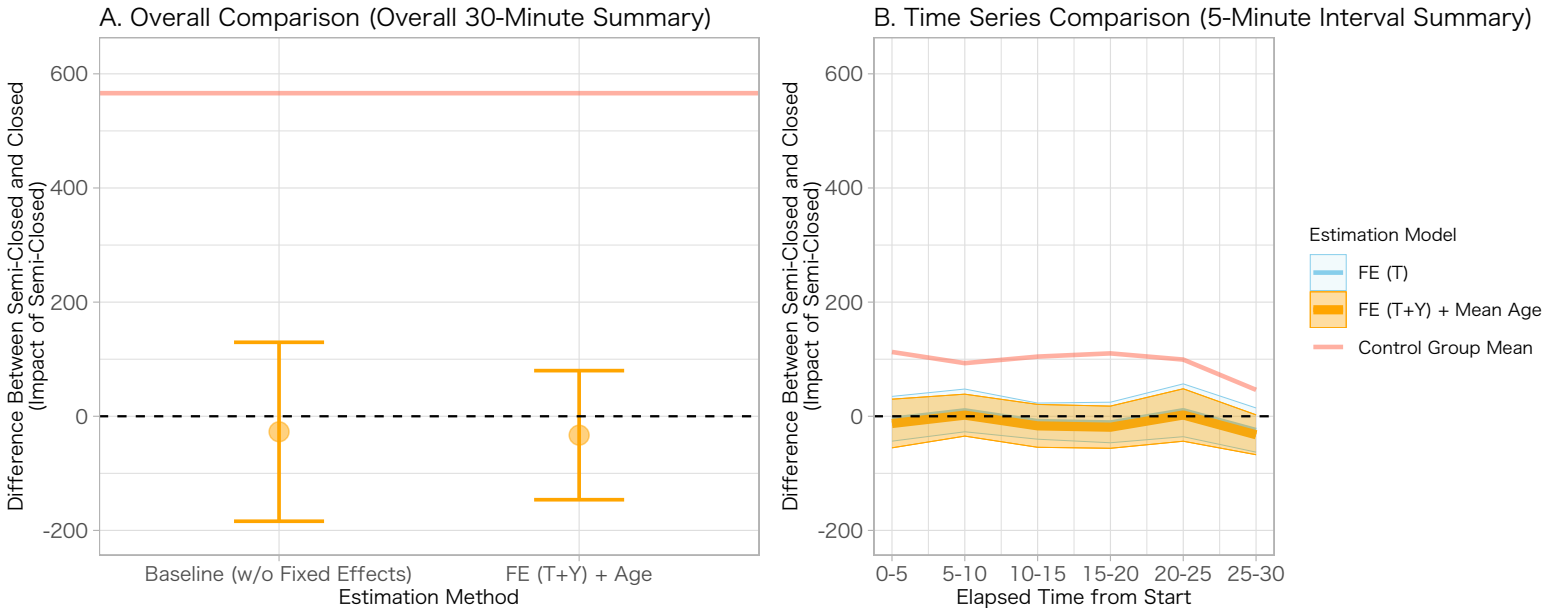

Figure 4: Effect of Semi-Closed Condition on Duration of Silences (30-Minute and 5-Minute Analyses)

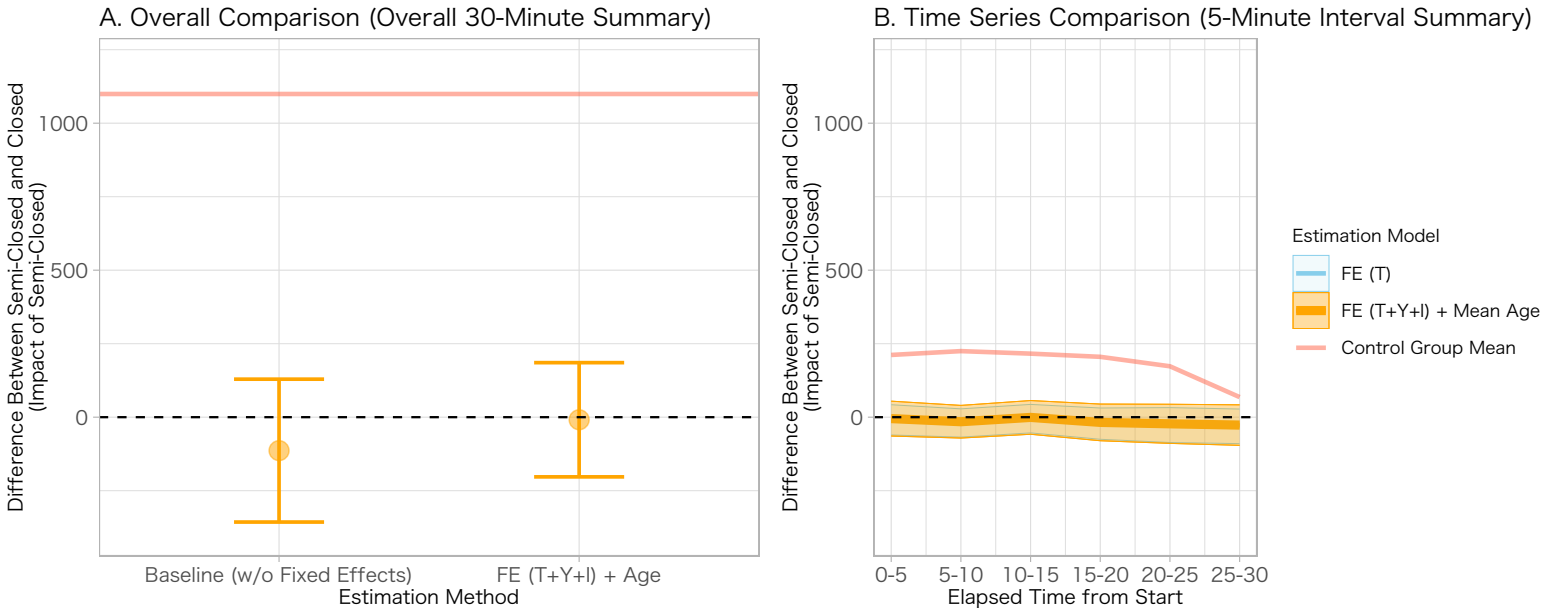

Figure 5: Effect of Semi-Closed Condition on Number of Words Spoken (30-Minute and 5-Minute Analyses)

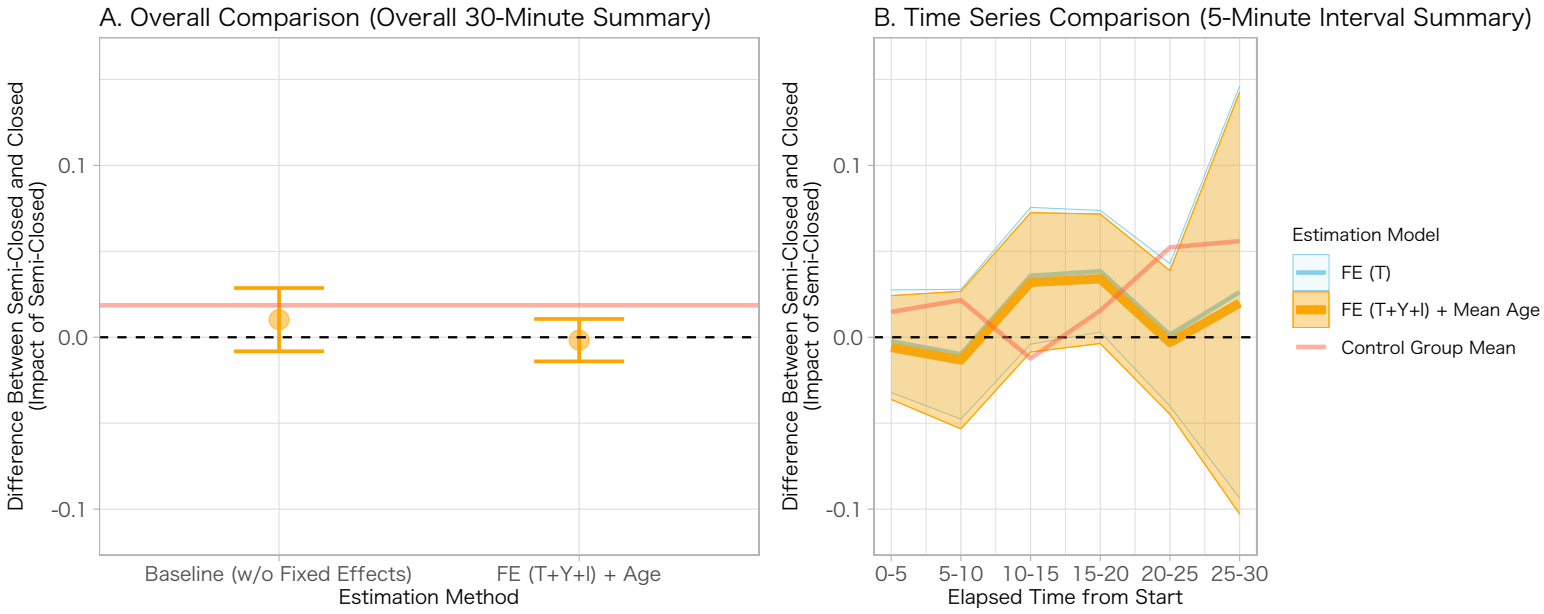

Figure 6: Effect of Semi-Closed Condition on Sentiment Polarity Score (30-Minute and 5-Minute Analyses)

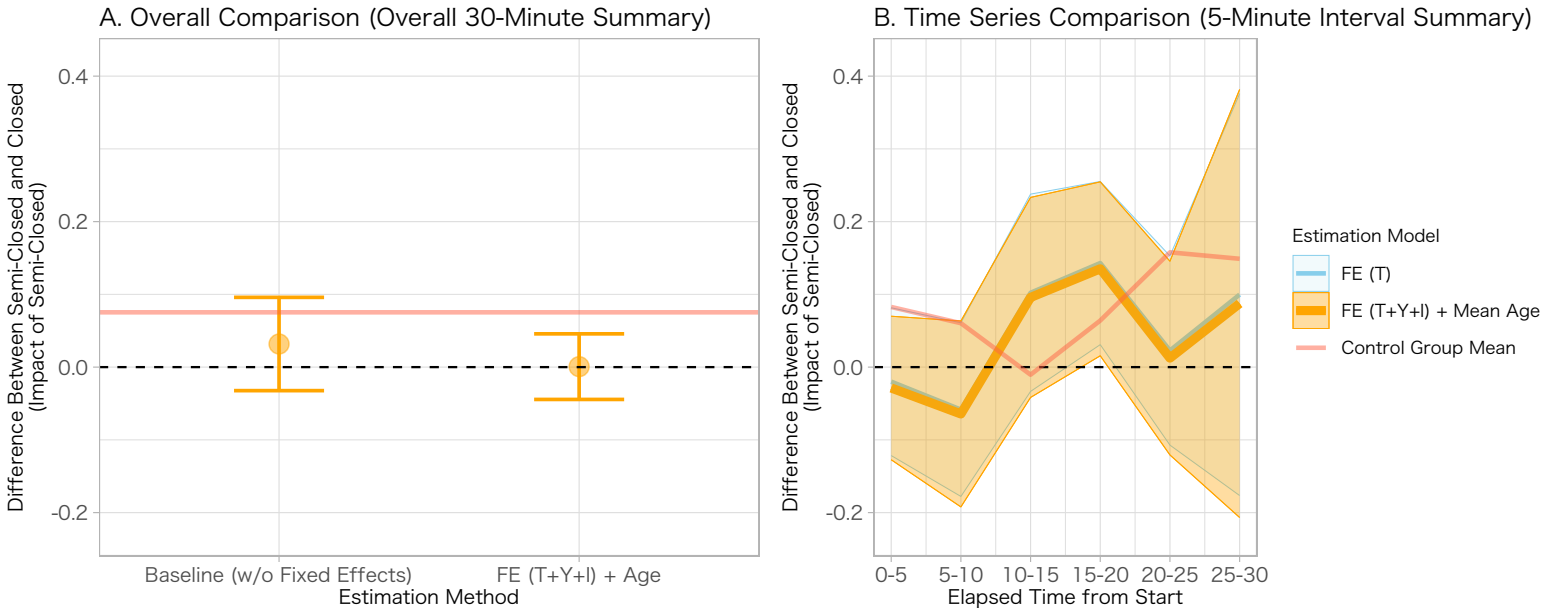

Figure 7: Effect of Semi-Closed Condition on Proportion of Positive Sentences (30-Minute and 5-Minute Analyses)

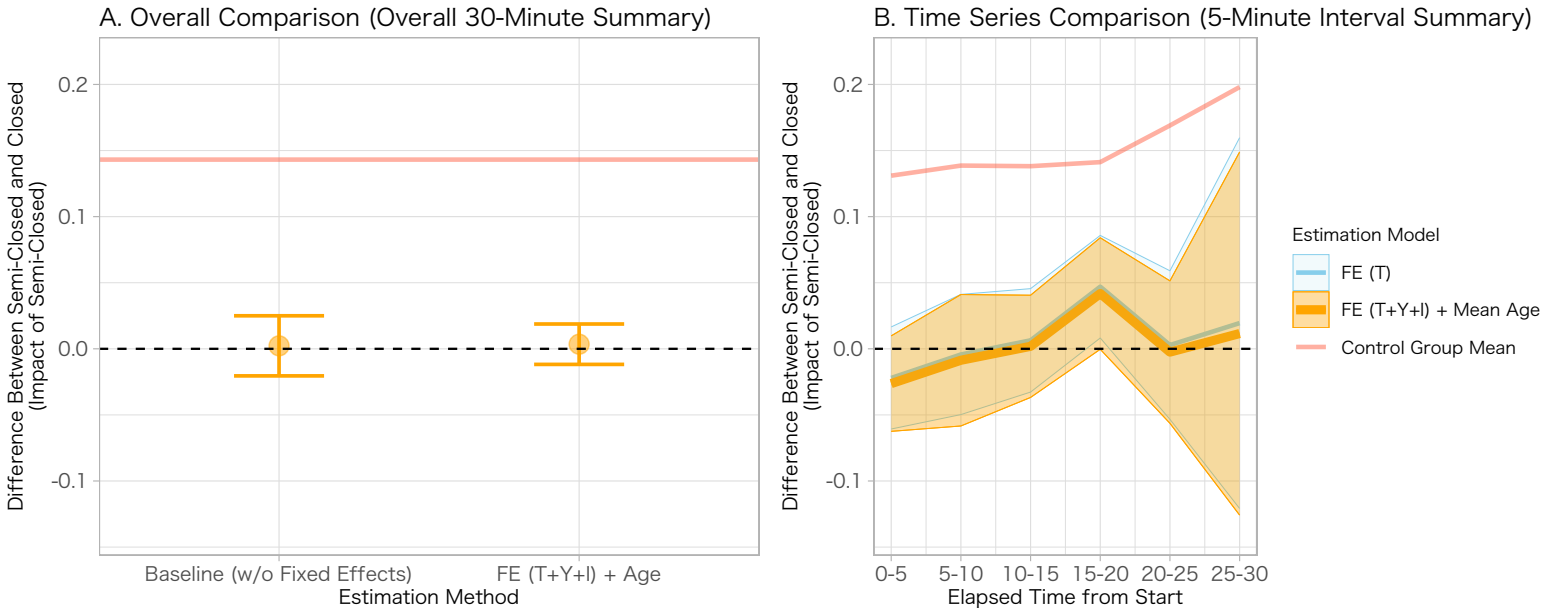

Figure 8: Effect of Semi-Closed Condition on Proportion of Negative Sentences (30-Minute and 5-Minute Analyses)

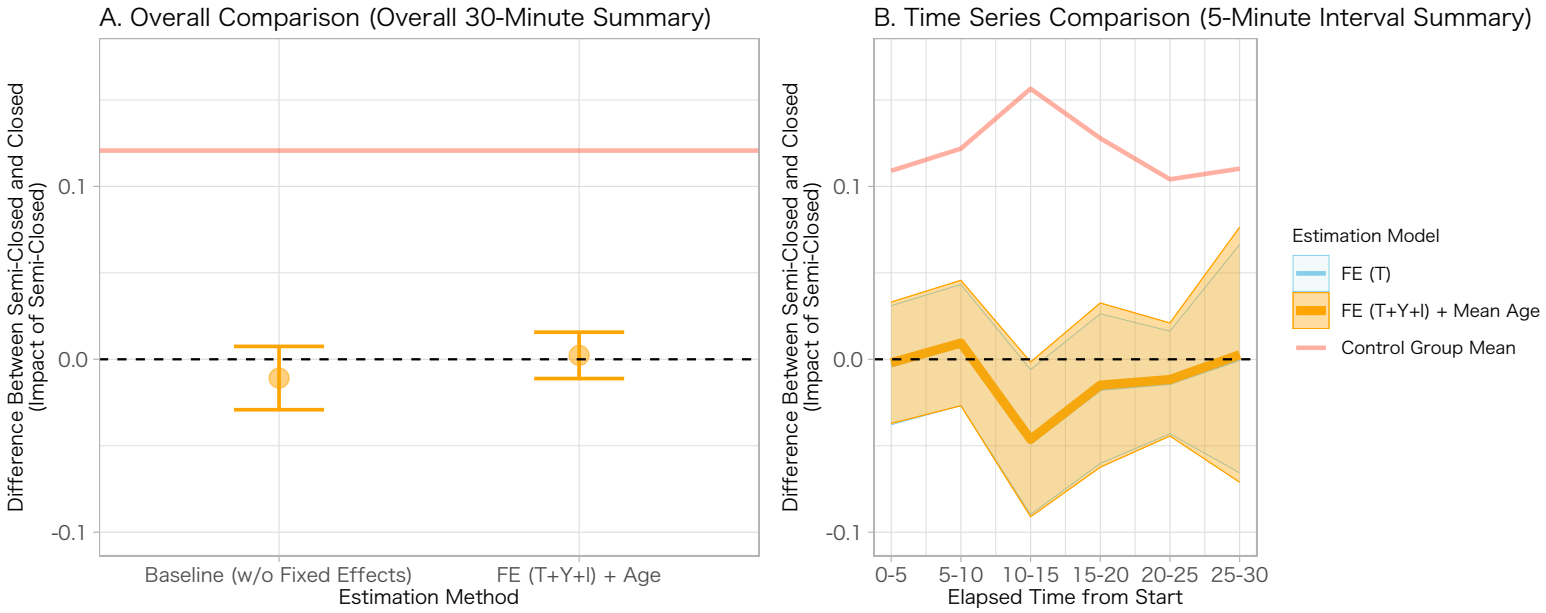

Figure 9: Effect of Semi-Closed Condition on High-Magnitude Sentiment Polarity (30-Minute and 5-Minute Analyses)

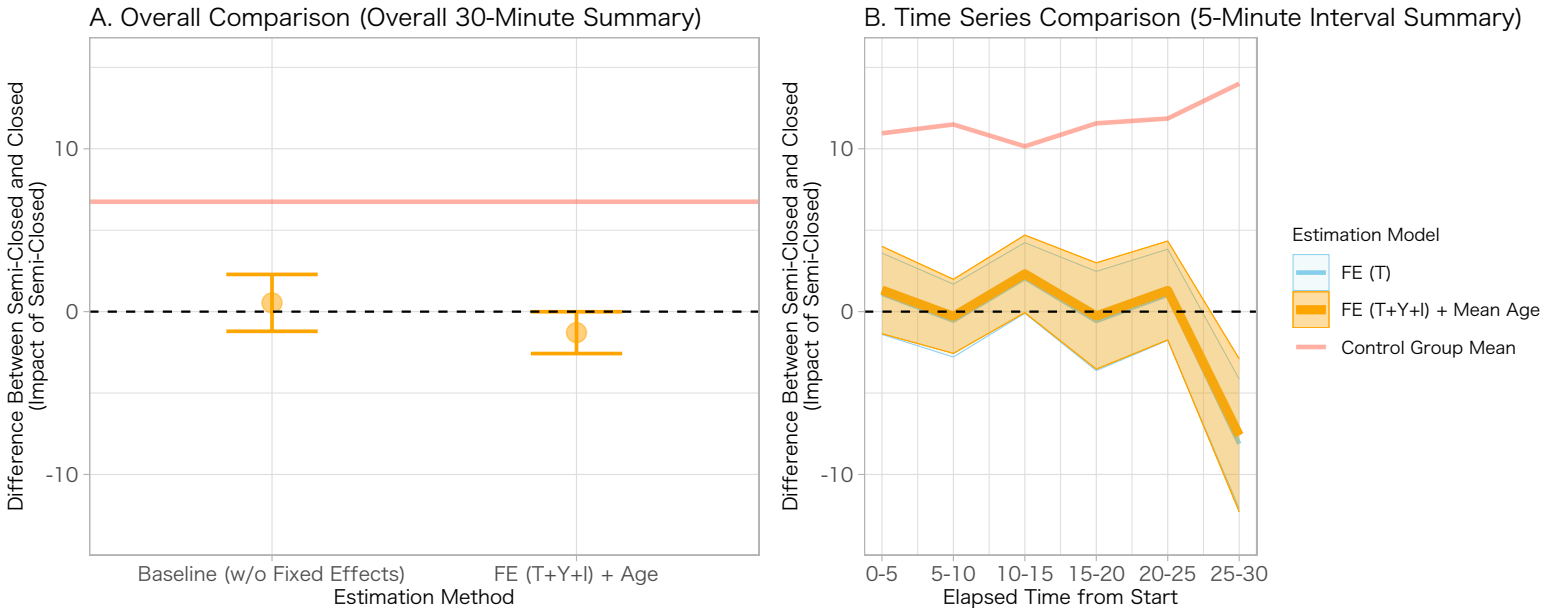

Figure 10: Effect of Semi-Closed Condition on Number of Topics Discussed (30-Minute and 5-Minute Analyses)

List of Appendix Figures

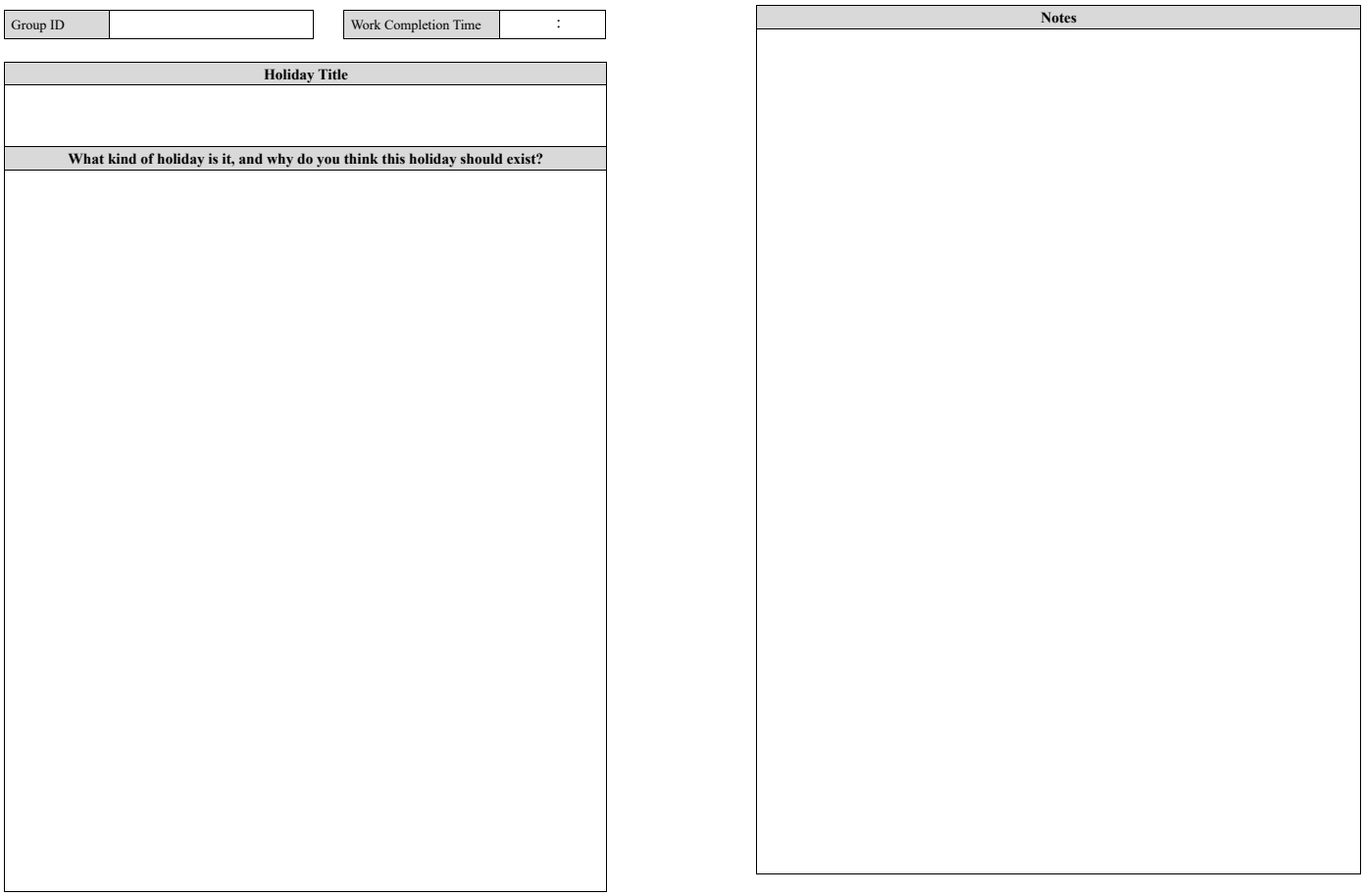

Figure A1: Image of Worksheet

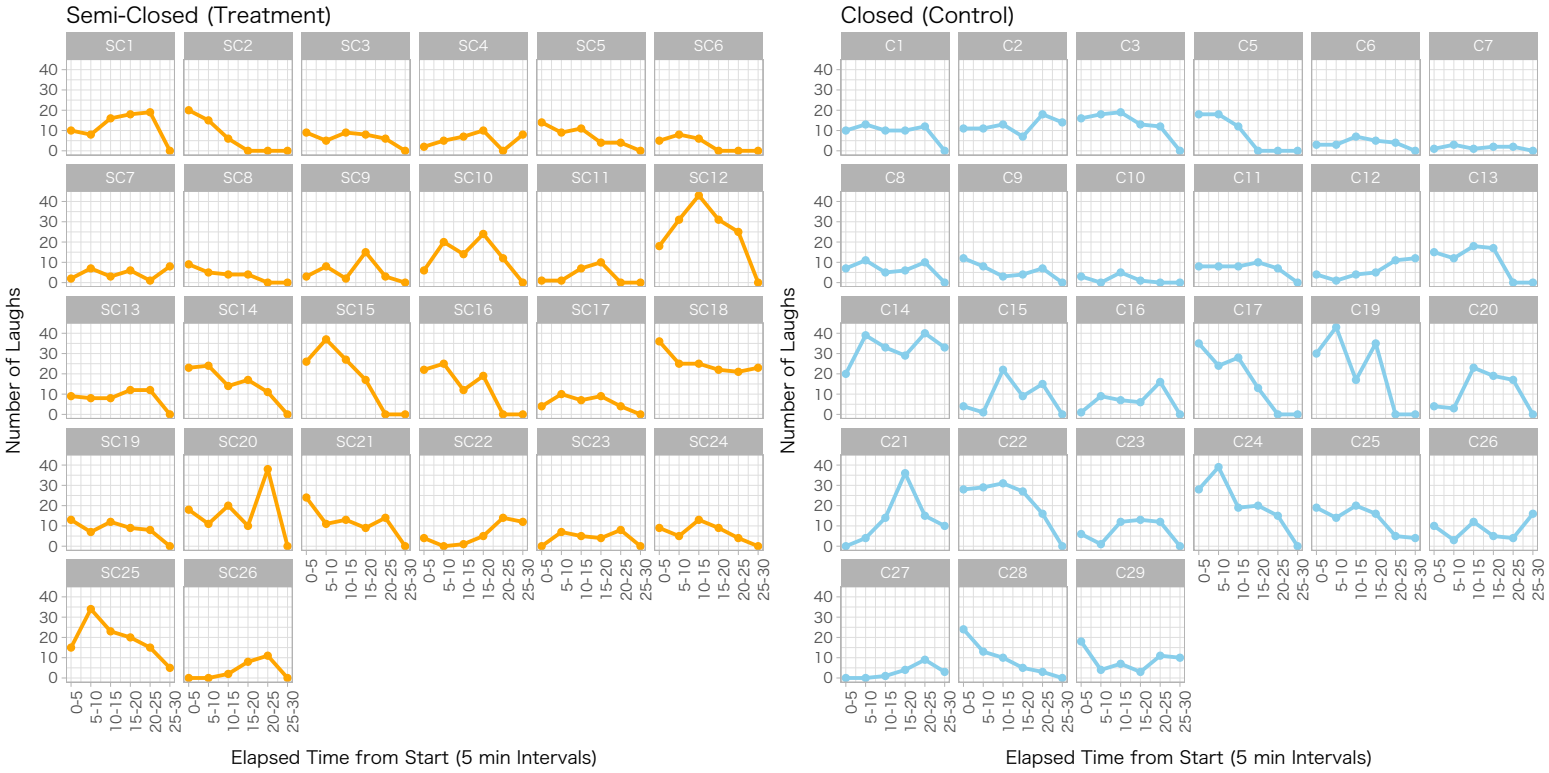

Figure A2: Change in the Number of Laughs by Group (5-Minute Intervals)

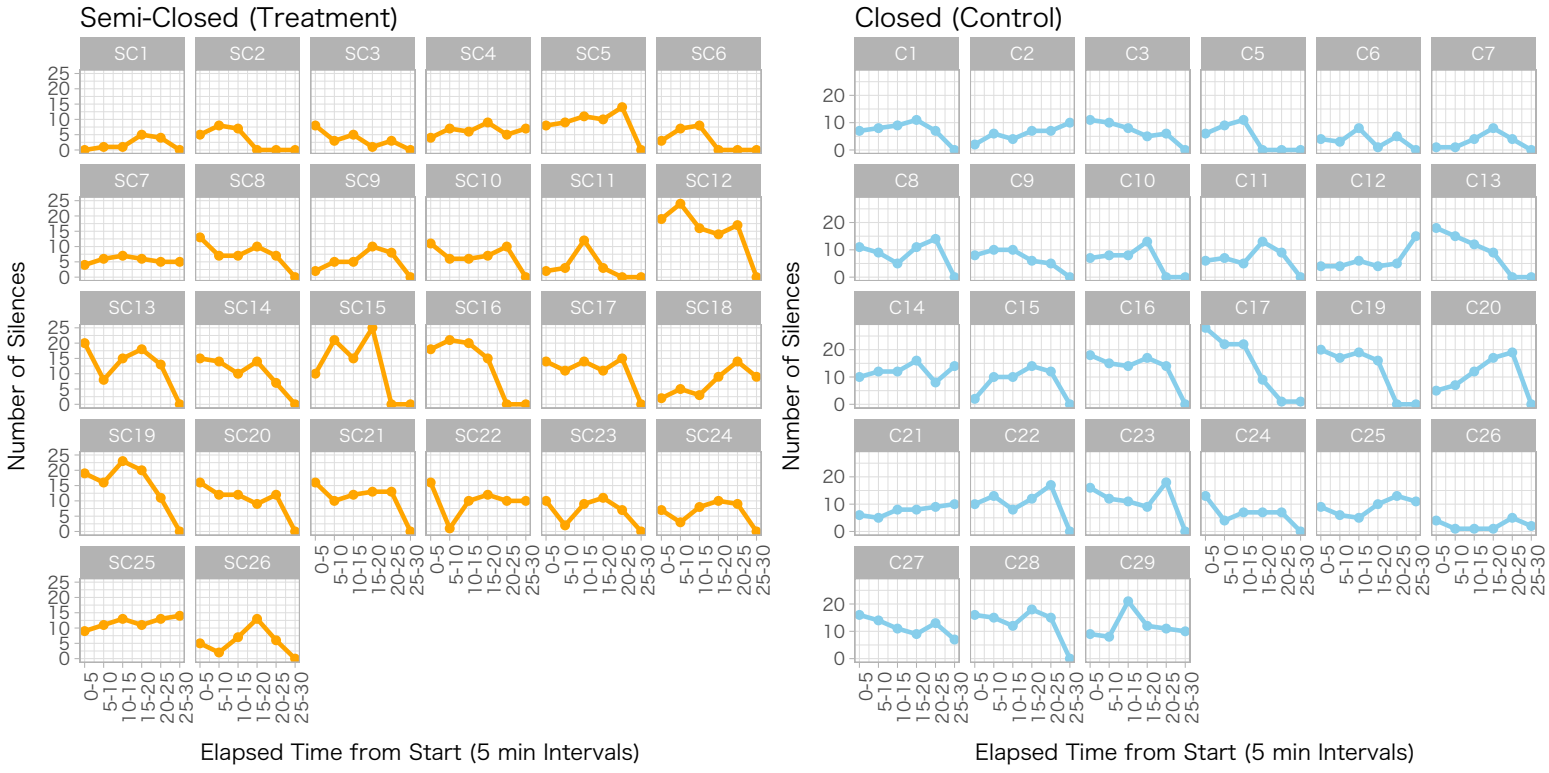

Figure A3: Change in the Number of Silences by Group (5-Minute Intervals)

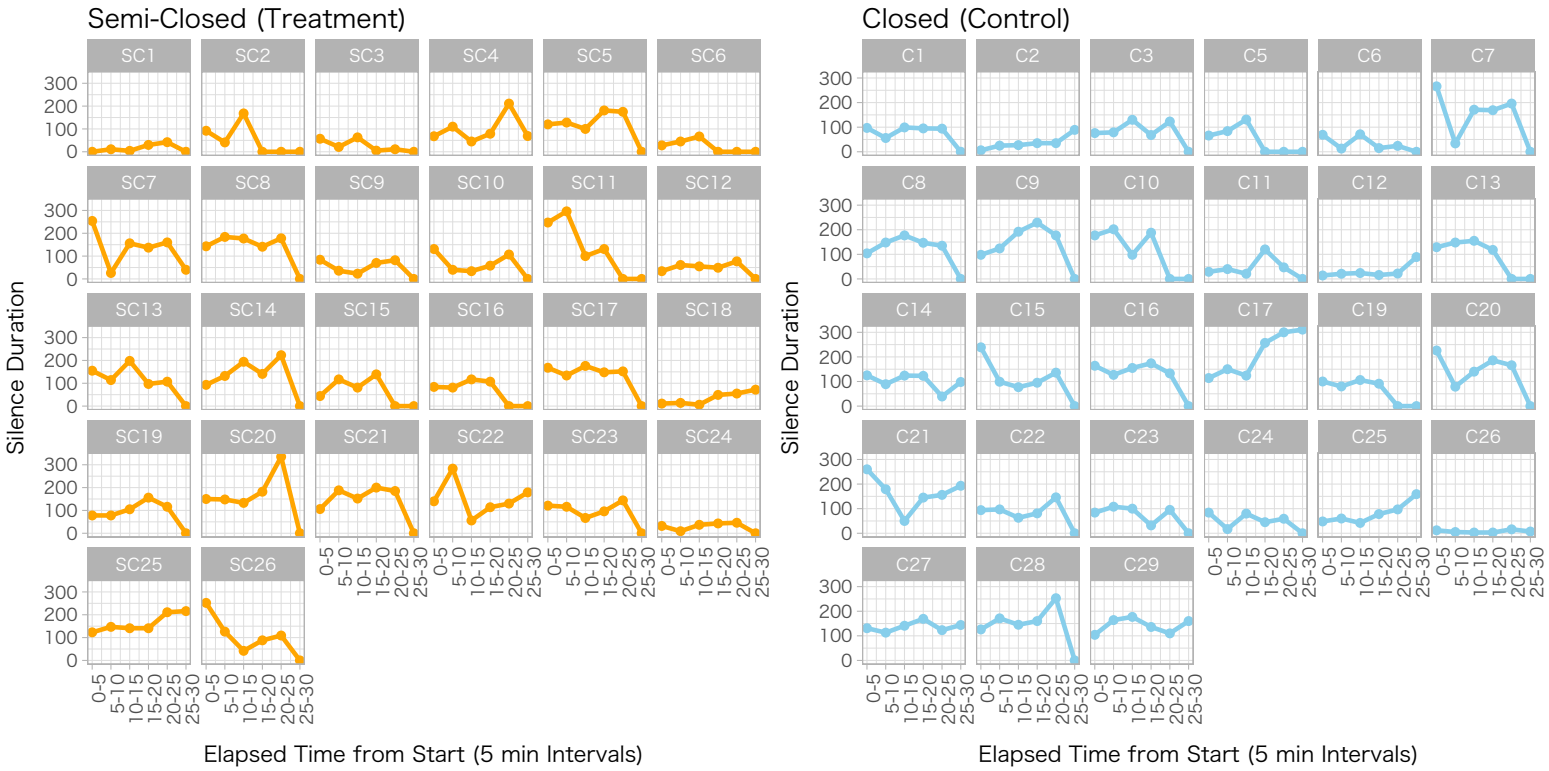

Figure A4: Change in the Duration of Silences by Group (5-Minute Intervals)

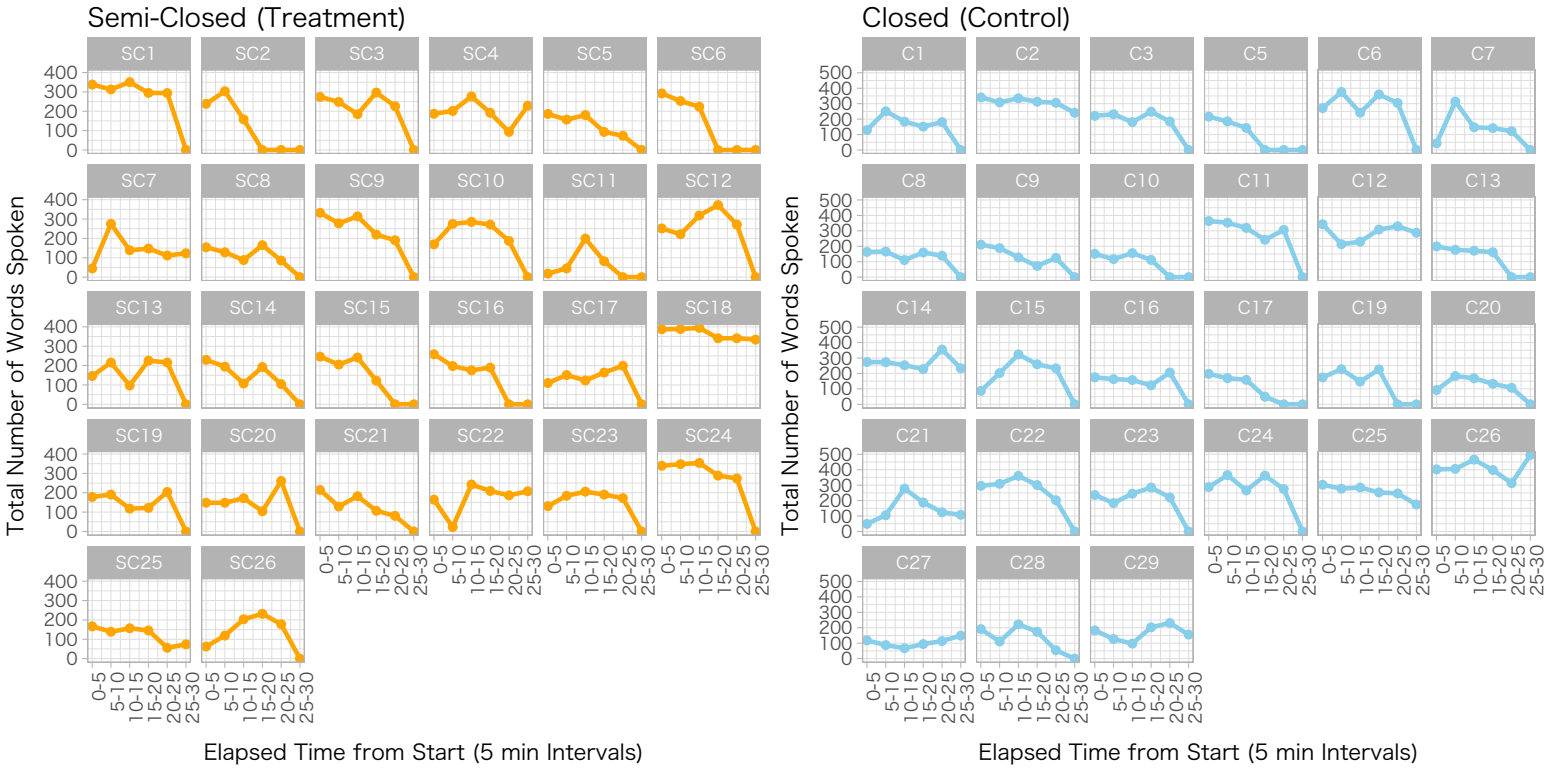

Figure A5: Change in the Number of Words Spoken by Group (5-Minute Intervals)

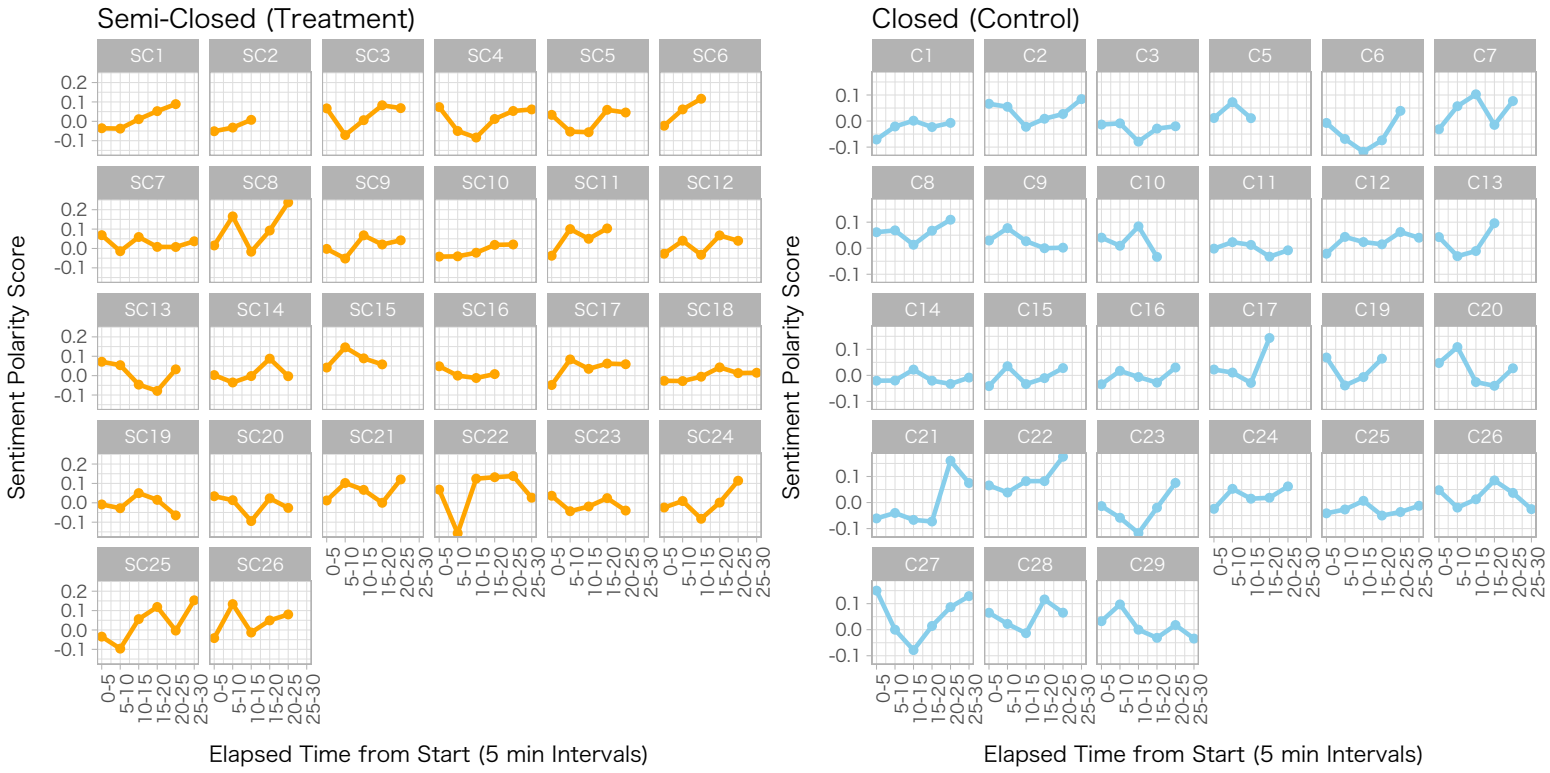

Figure A6: Change in Emotional Polarity Score by Group (5-Minute Intervals)

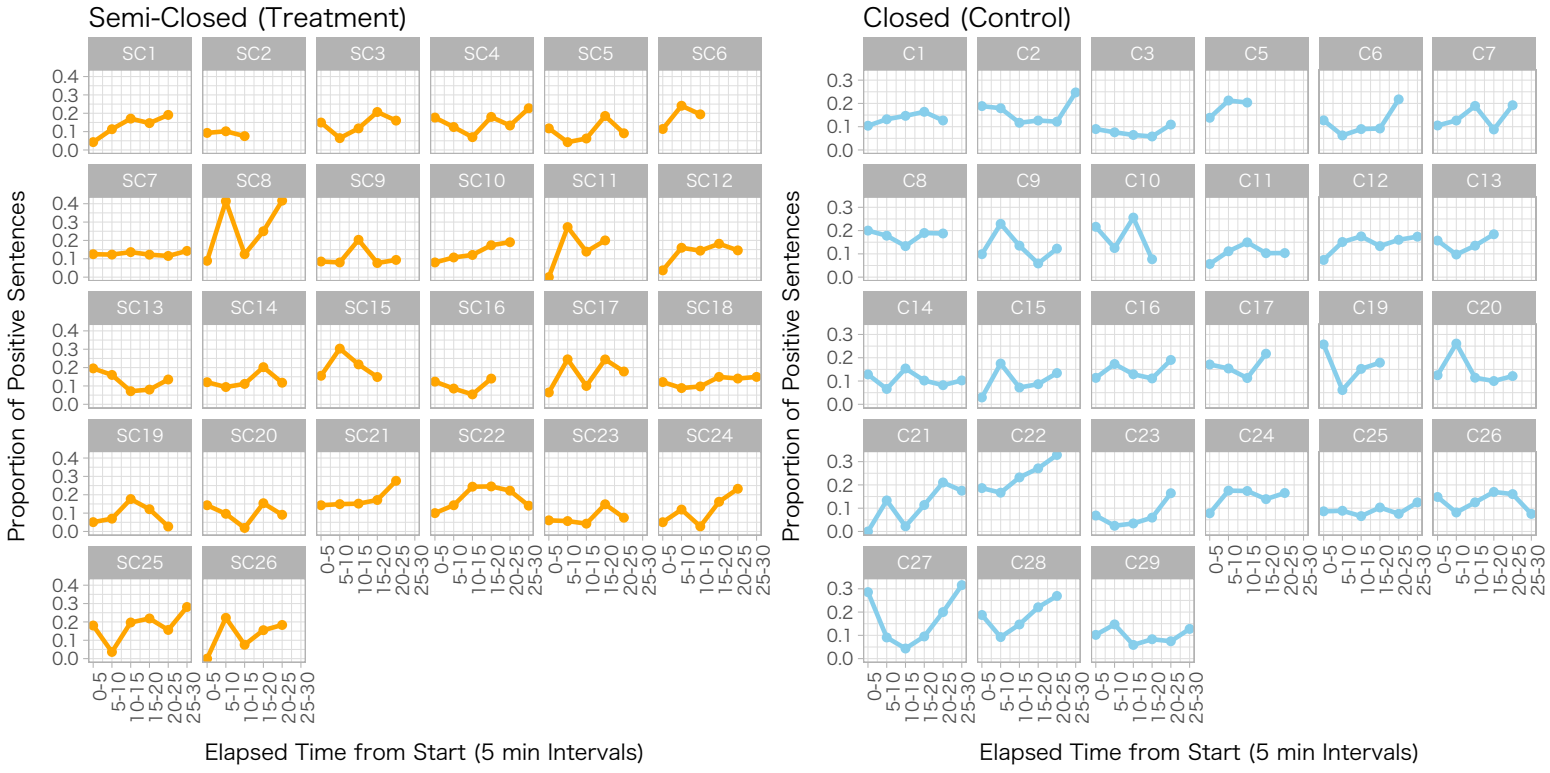

Figure A7: Change in the Proportion of Positive Sentences by Group (5-Minute Intervals)

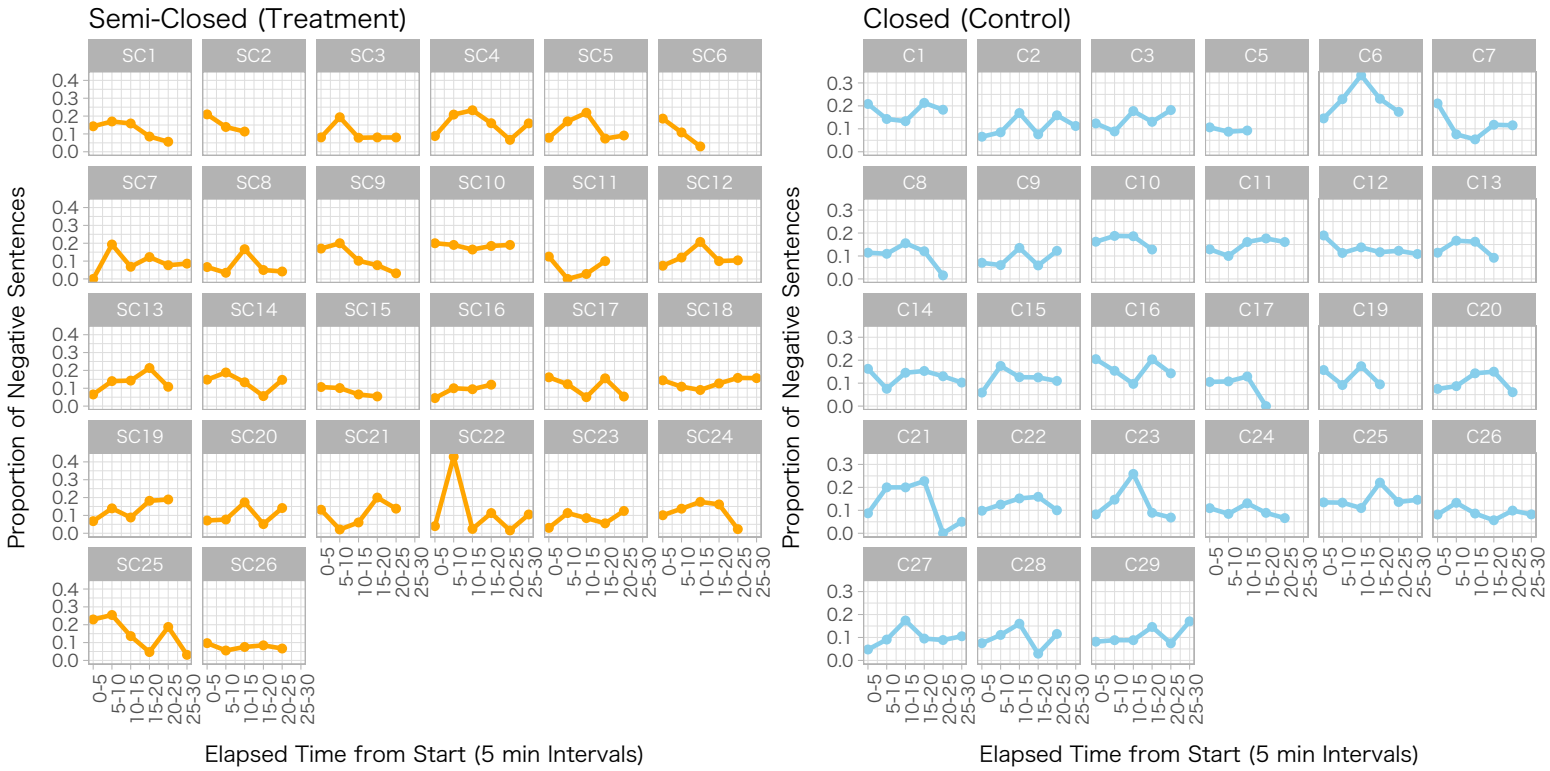

Figure A8: Change in the Proportion of Negative Sentences by Group (5-Minute Intervals)

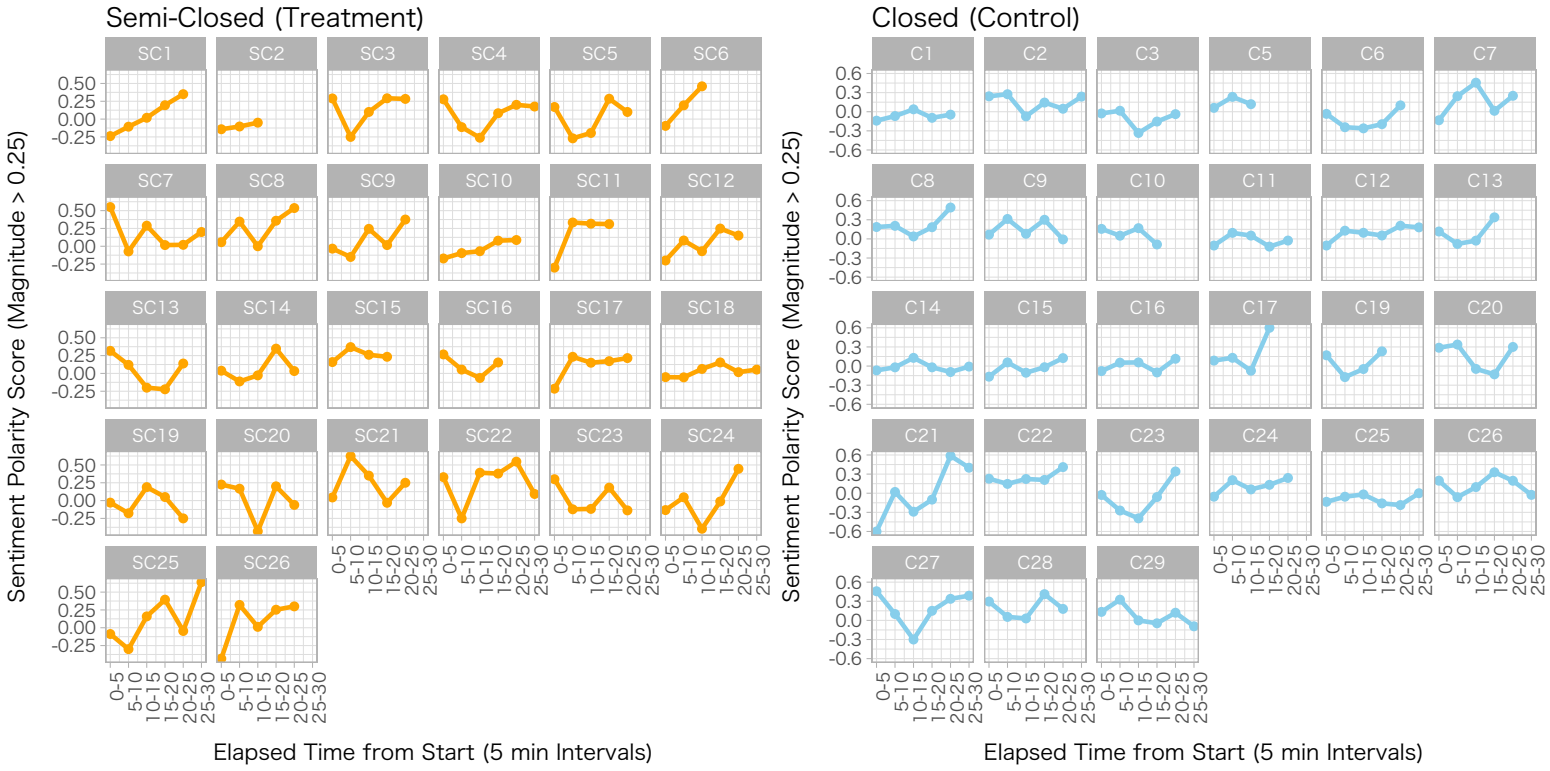

Figure A9: Change in Sentiment Polarity Score (Magnitude > 0.25) by Group (5-Minute Intervals)

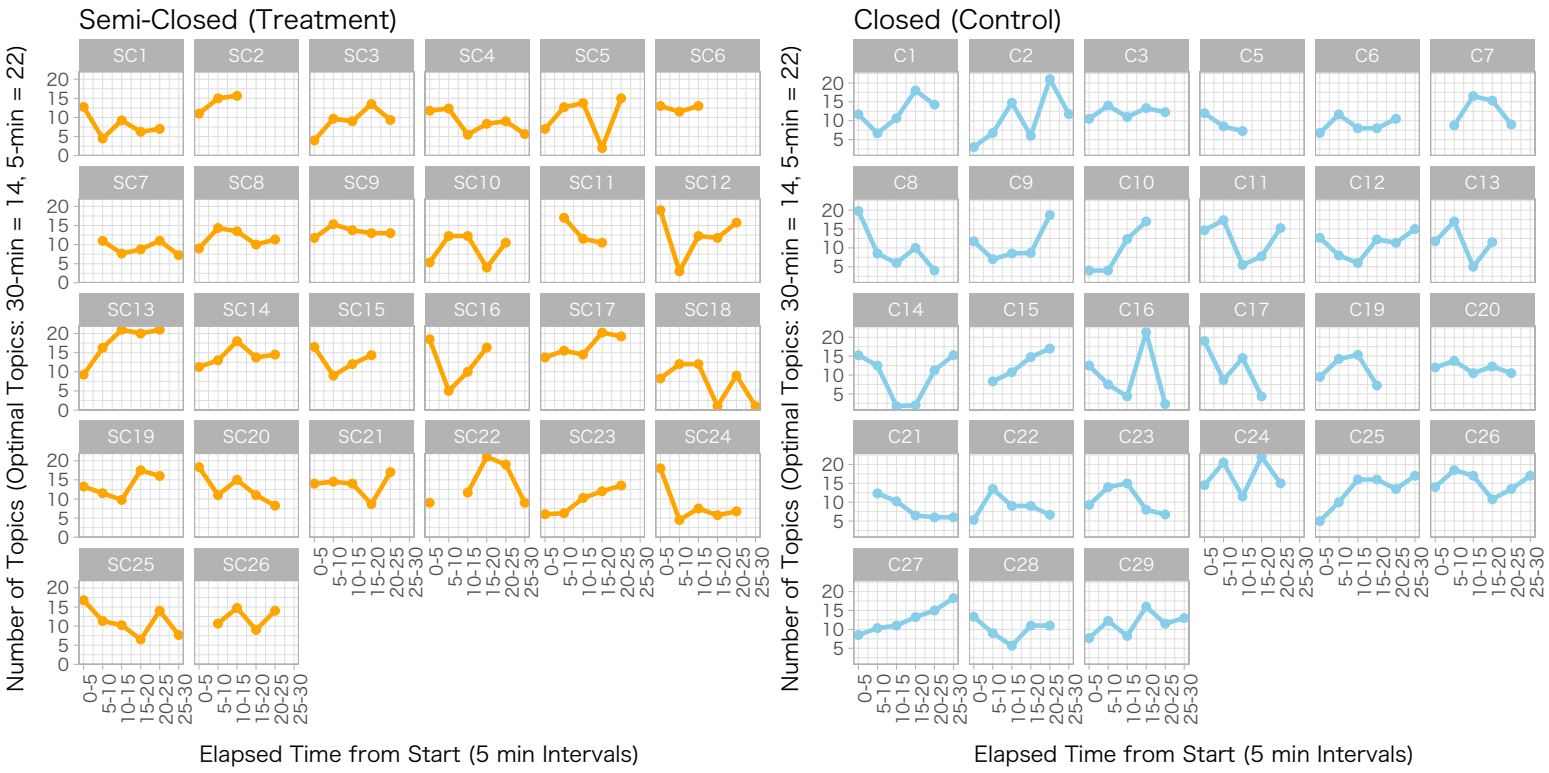

Figure A10: Change in the Number of Topics by Group (5-Minute Intervals)