SOFTENG 702 Group 7 Pilot Study Protocol

Project T2
Tutor: Tamil

Research Goals

With this study, our group aims to investigate the use of AI-mediated feedback in augmenting group performance in remote meetings. The impact of AI-mediated feedback has been investigated on improving individual cognitive performance and is well-established, but this approach has not been studied in a group environment, where several individuals interact with each other to perform a task or achieve a common goal. Hence, our research aims to fill this gap by observing the impact of feedback provided by AI on a group of individuals performing a brainstorming activity.

Our research question:

Does AI-mediated feedback impact group members' engagement and idea generation during online remote meetings?

The goals of the research are thus as follows:

- Evaluate the presence of AI feedback mechanisms
- Fill gaps in the literature
- Investigating the Relationship Between Human Cognitive Performance and AI Feedback During Brainstorming Sessions

Study Procedure

Study-design:

8 recruited participants will be divided into 2 groups, consisting of 4 people each. Each group will undergo two sessions, in which one will be used as the control and the other will be used to give AI-mediated feedback. To counterbalance the effects, the two sessions will be conducted in reverse order: half of the groups will be going through the controlled session first, then the one with AI feedback, while the rest of the groups will go in the other way around.

Users will not be directly interacting with the artefact developed; instead the artefact will only be used by the facilitator to get the feedback on the participants engagement level and then use ai to provide feedback. The artefact developed is a deep-learning-based emotion detection tool that captures human facial expressions and evaluates emotional states based on six fundamental emotions according to Ekman's theory: happiness, sadness, disgust, anger, surprise, and fear. The procedure will be as follows: once a session starts, the artefact will be initialised and start to detect peoples' emotions. Once there is an indication that some of the group members are having difficulties in the process of generating ideas (detected by the artefact), the facilitators will come in and provide some AI-mediated feedback to those particular participants.

Gathering Informed Consent

We will provide a consent form to every participant at the initial introduction of our project, which occurs a few days before the actual study. Participation in this study is completely voluntary, and participants can withdraw at any point. All data collected will remain confidential and anonymized. The participants will sign the consent form whenever they are ready, before contributing to the study.

Script

Step 1. Add all participants in the zoom meeting

Introduction (all participants)
Hello, everyone! Thank you for agreeing to participate in this study. Today, we are exploring how feedback during group brainstorming sessions can impact the ideas generated by teams working together remotely. Your participation will help us better understand how different approaches to brainstorming can influence group performance.

Step 2. Divide participants into breakout rooms, and share the link to the Miro board

In today's session, you'll be participating in two brainstorming tasks. The goal is to generate as many ideas as you can on the topic provided.

Each session will last around 15 minutes, and we ask that you actively collaborate with your group to generate ideas. You'll write your ideas down on the Miro board. Please ensure that if you come up with an idea, you write that down on the Miro board yourself. After each session, we'll give you a brief questionnaire to gather feedback on your experience.

Conducting task 1

Step 3. Add sem-wizard of oz into the meeting if AI-session

For the first task, we'd like you to brainstorm ideas around the following topic:

"As we gather to discuss the unique challenges faced by postgraduate students, let's remember that mental health awareness is a vital part of creating a supportive academic environment.

Postgraduate students often face unique challenges, including high academic expectations, research pressures, and the transition to independent study. Promoting mental health awareness can help create a supportive environment and foster resilience. Within this session, please try to come up with some specific suggestions for improving the mental health awareness for each individual based on the context."

You'll have 3.5 minutes to read the task brief and then 15 minutes to brainstorm as a group. Please make sure to note down all your ideas on the Miro board.

Thank you for completing the first brainstorming task. Please fill out the questionnaire based on your experience during the session.

Step 4. Send them google form link to answer questionnaire in group chat

Conducting task 2

Step 5. Add sem-wizard of oz into the meeting if AI-session

Now, let's move on to the second brainstorming task. Just like before, you'll have 3.5 minutes to read the brief, and then 15 minutes to brainstorm ideas around a different topic. During this session, you might receive some feedback to help guide your group in generating ideas.

Here’s the topic for the second session:

“As you know, the Canvas page is a crucial platform for both staff and students at the University of Auckland, playing an essential role in how we communicate academic information within our large educational community.

Today, we’ll be diving into the pros and cons of the Canvas website. We’ll also brainstorm suggestions for improvements to enhance our experience. Your insights and experiences are invaluable, so let’s create an open environment where everyone feels comfortable sharing their thoughts.”

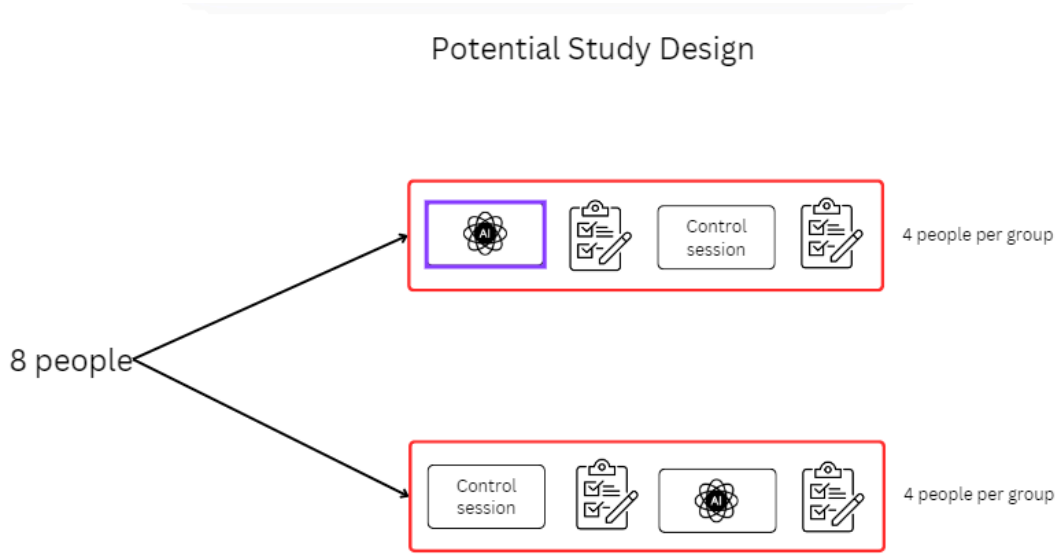
Step 6. Send out a Google form link to answer the questionnaire in the Zoom group chat

Great job! Please fill out this final questionnaire about your experience in this session.

Wrap-up

Step 7. Get participants out of breakout rooms in one zoom meeting

Thank you so much for your participation today! Your insights will be valuable for helping us understand how different approaches to brainstorming can impact group performance. If you have any questions or would like more information about the study, feel free to reach out.



Questionnaire Design and Administration

When to administer

The self-evaluation questionnaire will be administered to the participants after they complete each brainstorming session. This will record their perceived experience of the brainstorming activity that they just completed.

Draft of survey

Post-session questionnaires (will have two questionnaires)

Without AI session (all on a scale of 1-6):

1. How mentally demanding was the task?
2. How physically demanding was the task?
3. How successful do you believe you were in completing the task?
4. How hard did you work to accomplish your level of performance?
5. How would you rate your ideas quality during the session?(average)
6. How would you rate your engagement level during the session?
7. How frustrated, insecure, and discouraged were you during the session?

With AI session:

1. How often do you use AI tools in your daily life?
2. How useful do you find the AI tool to be during the session?
3. How distracting did you find the AI intervention to be?
4. How mentally demanding was the task?
5. How physically demanding was the task?
6. How successful do you believe you were in completing the task?
7. How hard did you work to accomplish your level of performance?
8. How would you rate your ideas quality during the session?(average)
9. How would you rate your engagement level during the session?
10. How frustrated, insecure, and discouraged were you during the session?

Rationale

In the post-study questionnaire, we included two distinct sets of questions for the two different sessions: one for the AI-mediated task and one for the task without AI. For the AI session, we measured users' responses using Likert scales, where 1 represents the lowest score and 7 represents the highest.

This survey is designed by combining well-established scales: TAM, NASA-TLX, and Godspeed, each of which has been widely validated. TAM focuses on the perceived usefulness and ease of AI, NASA-TLX addresses task load and performance, and Godspeed examines the emotional and trust-based relationship with AI.

Within the survey, we tried to assess the cognitive load required to complete the task, and to determine the effort participants invested in the task. These help us understand how challenging it was from a mental perspective. It provides insight into participants' experiences and the complexity of the task. We also tried to understand the participants' engagement level in the task, which can affect outcomes and overall satisfaction. During the sessions, measuring negative emotions such as frustration and insecurity provides insight into participants' emotional experiences. This information is crucial for identifying factors that may hinder performance and for developing strategies to improve the task experience.

System Usability Scale (SUS) is a commonly used scale to evaluate the usability of products and systems, however since this is a semi-wizard of oz study and users do not directly interact with the artefact developed, we considered to not use this scale, moreover we found NASA-TLX to be more suitable for task-related effort. We also explored creating very custom questions but adapting existing instruments towards our task allowed us to ensure reliability with existing research.

Data Collection and Metrics

Data to Collect:

Ideas generated per group.

Questionnaire responses (Likert scale ratings before and after both sessions).

Metrics:

Average number of ideas per individual (derived from the total number of ideas generated from the group).

Level of engagement

Cognitive / mental load

Quality of ideas generated for each session

Helpers

This study is a semi-wizard of Oz type of study design; during each session, one facilitator (helper) will focus on monitoring participants' emotional states as reflected by the AI tool, as well as the duration of idle time for each participant. Meanwhile, the other two facilitators will engage with GPT and relay feedback to participants whenever low engagement is detected.

Data Analysis

In this study, we will analyse the data collected on four key metrics to assess the impact of AI-mediated feedback during online group brainstorming sessions: the number of ideas generated, engagement level, quality of ideas, and cognitive load, using within-study design.

1. **Number of Ideas Generated**: We will begin by calculating the total number of ideas generated by each participant in both the AI-mediated and control sessions. Descriptive statistics, including the mean and standard deviation, will be computed to summarise the data. We will then conduct paired samples t-tests to compare the number of ideas generated between the two condfitions, allowing us to determine if AI-mediated feedback significantly enhances idea generation.
2. **Engagement Level**: Engagement levels will be assessed using survey responses collected after each session. We will calculate descriptive statistics to capture the overall engagement levels of participants in both conditions. Similar to the idea generation metric, we will use paired samples t-tests to analyse any differences in engagement levels between the AI and control sessions, providing insight into how feedback influences participant involvement.
3. **Quality of Ideas Generated**: The quality of ideas will be evaluated using a predefined rubric (provided below) to assign scores based on creativity, relevance, and feasibility. We will calculate the mean quality scores for ideas generated in both the AI-mediated and control sessions. Paired samples t-tests will again be utilised to compare the quality of ideas, helping us to understand whether AI feedback leads to higher-quality contributions.
4. **Cognitive Load**: To assess cognitive load, we will use a self-reported scale adapted from previous scales, such as NASA-TLX, to be completed by participants after each session. We will analyse the responses to calculate the average cognitive load scores for both conditions. By employing paired samples t-tests, we will compare the cognitive load experienced during AI-mediated sessions versus control sessions, allowing us to examine the impact of feedback on participants' perceived mental effort.

By analysing these four metrics, we aim to provide a comprehensive understanding of how AI-mediated feedback influences idea generation, engagement, idea quality, and cognitive load in online brainstorming sessions.

Rubic to evaluate the quality of ideas generated:

	Exceeds expectation (2)	Meets expectation (1)	Below expectation (0)
Q1.	The idea is highly innovative, offering unique and groundbreaking solutions.	The idea is moderately innovative, offering new insights with a mix of conventional thinking.	The idea lacks creativity, repeating common or overused suggestions.
Q2.	The idea is highly realistic, and could easily be implemented in real life scenarios.	The idea is moderately realistic, which could be implemented in real life scenarios, with some challenges.	The idea is not realistic at all, which could be very difficult to implement in real life scenarios.
Q3.	The idea highly greatly with the topic and addresses the topic well.	The idea moderately aligns with the topic and somewhat addresses the topic.	The idea does not align with the topic at all and could not address the context in which the brainstorming takes place.

Q1. How new is the idea (creativity)?

Q2. To what extent can this idea be applied to solve a real world problem?

Q3. How relevant is this idea to the topic in the context?