

Module 4 | Deliverable C: Create a Plan and Script for Usability Testing (Team)

TEAM: Nasa AI Vanguard (Team D1)

COURSE: 05410/05610 User-Centered Research and Evaluation

Research Questions:

Any specific questions to be answered through the Think-aloud, identified from your research to date, or from a quick heuristic evaluation of the app

Overarching Goals:

- Identify specific difficulties in the UI/UX when participants interact with TAIGA/other GenAI platforms
- Understand what users are thinking about, logically and emotionally, when using websites like TAIGA
- Understand user perceptions and experiences of bias
- Understand user behavior in reporting a bias through TAIGA/other GenAI platforms
- Understand some basic demographic information about our participants
- Understand demographic group's preferences on specific topics
- Understand how we can empower users to audit Generative AI

Specific Research Questions:

1. What is the user profile? (basic info, previous knowledge about bias, familiarity with algorithmic systems, age group, and frequency of use.)
2. What are the GenAI platforms that participants visit most frequently, if any? How often do users use these platforms? What do they use them for?
3. How can this website contribute to their awareness of bias?
4. Do users find biases, and if they do, do they report biases they have experienced in these systems?
5. What reasons/rationale is responsible for users choosing their input prompts?
6. What info can people get from these websites/apps?
7. What feature(s) / function(s) of TAIGA are the most important to users?
8. What feature(s) / function(s) of TAIGA are the most attractive (entry point) to users?
9. How do users like the web design of this page? (UI, transition, content, visual, etc.)
10. Are users able to navigate the website without explicit, elaborate instructions? If not, what are their pain points?
11. What do users feel about the website in general?
12. What values do users perceive of the website?
13. Are certain age groups more sensitive to biases, and if so, are they more likely to report it than others?
14. If they choose to report it, do they prefer to report it using the platforms' report options if they exist, or do they choose alternative methods like social media?

The Users

Given the circumstances, most users we will investigate will likely be other CMU students. Due to the nature of our study, we will endeavor to conduct user research on people of diverse backgrounds; to this end, we should record as many participant demographics as possible. These demographics would include fields such as age, race/ethnicity, gender, and sexual orientation, given that one of our primary objectives is to understand better if specific demographics are more sensitive to certain types of bias than others.

We will also be piloting and conducting the study with our parents, relatives, and family to draw inferences from the usability study across age groups. We will ensure we pick a sample of participants with varying levels of familiarity with GenAI platforms.

We will also record how long users take to perform each task and their thought processes as they navigate the website. For the latter, we will ask them to provide their thoughts as they perform the tasks in a Think-Aloud. In this process, our participants can voice their thoughts on the website's usability, such as whether things are easy to navigate, whether certain features respond the way they are expected to, and so on.

The Assigned Tasks

Task 1: Navigation

Participants start by visiting the TAIGA web platform and locating the '**Trending Threads**' section. This step tests the platform's navigational design and ease of access to its features. Initially, participants are asked to browse casually through the intro and trending thread examples. This allows them to get accustomed to the types of images presented and understand what they would perform for the test.

Task 2: Input Prompt

Participants are asked to input prompts for the genAI and hit the "**Generate**" button to get the result images.

Task 3: Identify Patterns or Biases

The objective of this task is for participants to scrutinize the images more closely to identify any recurring themes, subjects, colors, or any biases towards certain representations (e.g., gender, race, settings). This could involve noticing if certain types of images are overrepresented or underrepresented or if there's a noticeable skew in how certain subjects are portrayed.

Task 4: Generate Feedback

Use the '+ New Thread' Button to Record Observations. Upon identifying a pattern or bias, participants must use a 'Feedback' button (presumably located near each image or within the gallery interface) to record their observations. This step is crucial for collecting data on user perceptions and interactions with the AI-generated content.

Task 5: Verbalize Thoughts Using the Think-Aloud Process

Participants are asked to verbalize their thoughts, observations, decisions, and feelings throughout the task. This includes when they notice something interesting when they decide to record feedback and their rationale behind identifying something as a pattern or bias. This process helps us understand the participants' thought processes and how they interact with the platform and its content.

Task 6 (optional): Demographic information

Incorporating an optional task to collect demographic information (Age, Gender, Race, Sexuality, Personal Experience) from participants in a usability study, like the one for the TAIGA platform, is crucial for understanding the diversity of the participant pool and how different groups might perceive patterns and biases in AI-generated images.

Expected Usability Performance

Task Completion

Knowing whether participants can complete the tasks, such as navigating the intro page, inputting prompts for GenAI, and using the “thread” to record their observations. The better the participants completed the task, the better it was usable.

Error Rate

The number of errors participants make while attempting to complete the task, such as difficulty disabling comparisons or using the 'New Threads' button. A low error rate suggests that the platform is intuitive and user-friendly.

Time on Task:

The time participants take to complete each task. This includes navigating the platform, using the prompt to generate images, and providing feedback. Shorter times in the navigating and generating sessions indicate that the platform is easy to use, while longer times suggest areas where the design or task instructions may need improvement.

Participants Engagement:

The extent to which participants effectively verbalize their thoughts, feelings, and actions while completing the task. High engagement in the Think-Aloud process suggests that the instructions were clear and that participants were comfortable with the methodology.

Ability to Report Bias:

Can users use the website to identify and report biases effectively? We will measure this by monitoring how they navigate the website and identify the process to report the bias. We will note down how many participants report the bias out of those who identify them in the first place.

Script

Hello! My name is [YOUR NAME]. Thank you so much for agreeing to participate in our study today. During this study, I will observe you using a [TAIGA/other GenAI platform] website. Your feedback is invaluable for improving user experience and ensuring the platform is inclusive and accessible. Remember, there are no right or wrong answers here; we're interested in your honest thoughts and experiences.

ASK FOR CONSENT

Before we begin, we would like to ask you for your permission to record the video and audio of the interview and a video of your screen when applicable. This is outlined in the consent form we have sent you. We would also like to have you verbally confirm your consent.

Please note that all data and recorded information will be maintained in strict confidence and anonymized. The information gathered will solely be utilized for the User-Centered Research and Evaluation course project.

Link to consent form: [Module 4 Deliverable Consent Form](#)

REASSURANCE

During this interview, we will not assess you or your expertise in GenAI systems, algorithmic biases, or any related topics. Should anything not proceed as planned, please know it is not a reflection on you. Remember, you are free to discontinue your participation at any time.

BACKGROUND QUESTIONS (optional)

Before we dive into today's activities, I'd like to ask you a few questions about your previous experiences with artificial intelligence, specifically generative AI technologies.

Your answers will help us better understand your background and how it might influence your interaction with the platform we're testing.

Have you had any previous experience using generative AI technologies? This includes interactions with platforms or tools that generate text, images, code, or other content based on user inputs. Some examples are ChatGPT, DALL·E for image creation, [Sora](#) for text-to-video generation, AI coding assistants like GitHub co-pilot, or smart-text editing assistive tools like Grammarly.

If Yes,

That's great to hear! We appreciate your knowledge of these concepts.

If No,

No worries at all! Generative AI technologies refer to systems that can create new content based on learning from large datasets. Examples include creating text responding to user prompts, generating images from descriptions, and more. Today's session might include interactions with similar AI technologies.

Explain Why You're Asking

We're asking these questions to understand how your previous experiences might influence your approach and reactions during today's session. Whether you're a frequent user or if this is your first encounter with such technology, it is perfectly fine.

TRANSITION TO MAIN ACTIVITY

Thank you for sharing that information. Understanding your background helps us ensure that our study accommodates various experiences and perspectives.

HOW TO THINK-ALOUD

(Make sure the participants understand the process of ThinkAlouds.)

During today's session, we'll be using a technique called '**Think-Aloud.**' Before we proceed, I'd like to check in with you.

Have you ever participated in a task where you were asked to verbalize your thoughts while completing it? This could have been in a previous study, a workshop, or any other setting where you were asked to speak aloud about your thoughts, feelings, and actions in real-time.

If Yes,

That's amazing! Could you please briefly describe your experience with the Think-Aloud technique? What did you do, and how did it feel to verbalize your thoughts while engaging in a task?

If No,

Thank you for sharing that. It's totally common. The **Think-Aloud** process involves verbalizing your thoughts as you interact with [product/service]. We ask that you say aloud what you're thinking and trying to do and how you feel about what you're experiencing as you go along. This could include your impressions, confusion, and any assumptions you're making.

(Providing Examples)

For example, if you're trying to locate a feature on our app, you might say, 'I'm looking for the settings menu so I can change my privacy options. I think it might be under this icon, but I'm unsure. Oh, here it is. That was a bit hard to find.' This gives us insight into your thought process and helps us understand how people interact with what we're testing.

(Emphasizing the Importance of Honesty and Detail)

It's essential to be as open and detailed as possible. There are no right or wrong answers here. We're interested in your genuine experience and thoughts, even if it's 'I'm not sure what to do here.' This helps us identify what's working well and what might need improvement.

Before we dive into the main activities, would you like to do a quick practice run using the Think-Aloud technique? It can be an excellent way to get comfortable verbalizing your thoughts before we start

Does that make sense? Do you have any questions about the Think-Aloud process before we begin?

USABILITY TEST OF TAIGA

I will share a link to a website our research team is interested in learning about. Once you arrive at the homepage, take a moment to study it.

As you perform each task, we ask that you use the Think-Aloud process. This means we would like you to verbalize your thoughts, observations, and any reasoning behind your actions as you navigate the platform. For example, if you notice a pattern in the images, describe what you see and why you think it's a pattern. If something strikes you, explain your thought process.

Feel free to ask questions anytime during your think-aloud process. There are no right or wrong answers to this task. We're interested in your genuine observations and thought processes. Your insights will be invaluable in helping us understand how users perceive the content generated by TAIGA and how we can improve the platform.

Individual Task Scripts

(subject to change after pilot testing iterations)

Each question will be prompted based on the completion of the prior question or if the user cannot proceed without one of the evaluators stepping in. We keep the first questions broad without revealing too much information and then add some terminology to guide the user if necessary.

Task 1: Navigation

1. Please start by visiting the TAIGA web platform link.
2. Now, browse the website and its content.
3. What's your initial impression of the website?
4. What do you think is the primary focus of this website?
5. How would you navigate this website to achieve its intended purpose?
6. Did you find anything trending now (Ability to find the 'Trending Threads sections')?

Task 2: Input Prompt

1. Next, we'd like you to use the platform's GenAI capabilities.
2. Please describe the images you would like to produce.
3. Describe your prompt and why you chose it.
4. Observe the generated images. How do they align with your expectations based on your prompt?

Task 3: Identify Patterns or Biases

1. Take a closer look at the images generated from your prompt and others. Are there any patterns, themes, or issues you notice?
2. This might relate to subjects, colors, or representations of gender, race, or settings.
3. If something stands out, please explain what you've noticed and why it caught your attention.

Task 4: Generate Feedback

1. Please record your observations regarding any patterns or biases you identified (Can they use the “New Thread” feature?). Think of this as providing feedback that could help improve the platform.
2. As you record your feedback, continue to verbalize your thoughts. What are you noting down, and why do you think it's important?

Task 5: Verbalize Thoughts Using the Think-Aloud Process

Throughout all these tasks, please keep verbalizing your thoughts, observations, decisions, and feelings about the interface and the content. Your insights into your interaction process are incredibly valuable to us.

Task 6 (optional): Demographic Information

Lastly, there's an optional task: we collect some demographic information from you to help us better understand your work in this study. This helps us understand the diverse perspectives of our user base. If you're comfortable, please provide details like age, gender, race, sexuality, and any personal experiences that might influence how you perceive AI-generated content through this form.

LINK to form: [User Demographic Optional Form](#)