**Part 1 – Ethics**

I confirm that I have read and understood the *Massey Code of Ethical Conduct for Research*, *Teaching and Evaluations involving Human Participants*. Both of my participants are aged 16 years or older. All completed ethics documents (ethical compliance statement, low-risk checklist, information sheet, and consent forms) are filed in the “Ethics documents” folder.

**Part 2 –** **Usability evaluation plan**

**Goal**

Assess how easy it is for new users to create sketches using the Sketchpad website. Will they be able to successfully use the Crop & Resize, Text, Pen, Fill Vector, Star, Clipart tools. Also, will new users be able to save a sketch or export an existing one.

**Questions**

* How effectively can you use drawing tools?
* Can users find the eraser in the current site layout?
* Will they be able to save and export the sketch without prompting?
* Will they be able to use stickers and drawing tools using layers?

**Equipment/Materials/Environment**

* **Equipment**

1. Laptop
2. Microphone
3. OBS – Screen and audio recording
4. Phone – Using a stopwatch and digital touch sheet

* **Materials**

1. Consent form (oner per participant)
2. Task sheet (oner per participant)
3. Paper for notes – I will leave notes on how successfully or unsuccessfully users were able to use the website.
4. Pens, highlighter

* **Environment**

A quiet, well-lit room with two chairs and a table where I'll place my laptop. Stable and good Wi-Fi.

**Participants**

**Ideal Participants**

* Age 16-45, mixed gender
* Basic computer literacy; no previous experience with Sketchpad
* Students and professionals who need to use such a tool in their work or studies

**Tasks**

Participants must complete six tasks. Each task has success criteria and a time limit for independent completion (until help is offered).

* **Task 1 – Freehand drawing (Pencil)**
* **Instruction:** “Select the pencil, change the stroke color to red and draw a start shape”
* **Success criteria**: Pencil selected, red star visible on canvas
* **Time limit before help**: 4 minutes
* **Task 2 – Select and Resize tools**
* **Instruction:** “Select a star and increase or decrease its size depending on the original size of the star”
* **Success criteria**: Start selected and changed star size
* **Time limit before help**: 4 minutes
* **Task 3 – Eraser usage**
* **Instruction:** “Switch to the eraser and erase one edge of the star”
* **Success criteria**: The edge of the star has been successfully erased
* **Time limit before help**: 6 minutes
* **Task 4 – Increase the size of the star line**
* **Instruction:** “Select a star and increase the star line size to your choice”
* **Success criteria**: The star line size has been increased
* **Time limit before help**: 6 minutes
* **Task 5 – Layers, drawing tools, stickers**
* **Instruction:** “Select any other drawing tool. Draw any two objects. Can use a sticker and place one object behind the star and the other in front. That is, use different layers.”
* **Success criteria**: One object is behind the star; the other is in front
* **Time limit before help**: 10 minutes
* **Task 6 – Export**
* **Instruction:** “Save the image to your computer in png format”
* **Success criteria**: The file has been successfully saved in png format
* **Time limit before help**: 5 minutes

**Test procedure**

Each session lasts approximately 45-55 minutes. The following checklist will ensure consistency between sessions.

**Before participant arrives**

* Make sure recording equipment is ready (video recorder, microphone).
* Prepare printed consent forms and task sheets.

**Session flow**

* **Welcome & consent (3-5 min)**
* Greeting and consent
* Greeting the participant; confirm age ≥16 years.
* Provide information sheet; questions.
* Participant signs consent form; assign participant ID (P01, P02).
* Practice thinking out loud
* **Setup & recording (2 min)**
* Launch video recorder screen (OBS) + audio recording; check the sound (30 sec test)
* **Pre-task instructions (1-2 min)**
* Explain the think-aloud protocol and encourage verbal expression.
* Provide the task sheet and explain the time limit and that help will be provided after the time limit.
* **Performing the task (30-35 min)**
* Introduce the tasks in order.
* For each task: start the timer, watch silently, take notes (start time, end time, success (yes/no), help (yes/no), mistakes).
* If a participant is silent for more than 20 seconds, ask him to think out loud.
* **Questionnaire (5 min)**
* The participant completes the SUS test and three short Likert scale tests.
* **Post-task interview (3-5 min)**
* Ask open-ended questions: “What was the most difficult?” “What did you like?” “Any suggestions for improvement?”
* **Wrap-up (1 min)**
* Stop recording; confirm saving the exported file; thank the participant and award him.

**Data to be collected**

* **Quantitative**
* Time spent on task (seconds)
* Task success (passed/partially/unsuccessfully).
* Number of errors (misclicks, wrong tool selections)
* SUS score (0–100)
* Number of Undo/Redo uses.
* **Qualitative**
* Audio recordings of thinking out loud (no face)
* Screen recording (video)
* Observational notes
* Ongoing comments and responses
* **Interview questions**
* *What did you like about Sketchpad?*
* *What did you find annoying or inconvenient?*
* *What 1-2 improvements would you suggest?*

**Data analysis**

* **Quantitative analysis**
* Calculate for each task: average time, median, standard deviation
* Calculate the percentage of successful task completion (task success rate) If success < 80% - mark as a critical issue
* SUS: calculate the final score using the standard procedure (0-100), average and spread
* **Qualitative analysis**
* Record critical incidents in recordings (long pauses, repeated clicks)
* Transcribe short excerpts (main comments and difficulties)
* For key issues, assign a severity rating on a three-level scale: high/medium/low

**Part 3 – Usability pilot study**

I conducted a pilot study of the Sketchpad web app evaluation procedure. A total of 2 participants (P01, P02) were recruited from friends/students; the participants ranged in age from 20 to 45 years. Both participants had no previous experience with similar drawing tools, which allowed them to test the comprehensibility of the interface for beginners.

Each session lasted from 11 to 17 minutes. Before the start, each participant was given an information sheet, questions were answered, and a consent form was signed; participants' faces were not recorded, but screen and voice (audio) were recorded. No ethical issues arose: all participants confirmed their understanding of the study objectives and gave written consent for the recording.

Upon subsequent review of the recordings, two technical and two procedural issues were noted that affected the quality of the data and the interaction of the participants. Technical issue: the built-in microphone of the laptop recorded the participant's voice quite well, while my voice was of poor quality and sometimes interfered with the analysis of dialogues. It was found in the procedure that newcomers often ask for help too early and instinctively look at me - they are probably afraid of "looking stupid", so they think out loud quietly or wait for a hint, which reduces the completeness of verbal recording.

Based on the pilot, the following changes were made to the usability evaluation plan:

1. Increase the allotted time for each task so that participants do not rush and can think out loud before the time runs out.
2. Strictly do not provide help before the time for the task runs out; in case of severe blocking, can remind about think-aloud, but do not direct actions.
3. Add a short practice of think-aloud before the start of tasks, so that the participant gets used to saying thoughts out loud.
4. Improve the audio check: before the start of each session, test the levels and, if possible, use an external/headset microphone for the moderator or record the moderator separately to keep the audio track clear.

**Part 4 – Heuristic evaluation**