APPENDIX A
PROMPT ENGINEERING

The blue words in the prompts will be replaced by input information or historical data.

A.1   Association Generator

A.1.1   Generate new associations

System prompt
You are an AI assistant. Your task is to recommend associations with a given found object based on the shape similarity. Each association will be constructed as a tuple (subject, component), which means the found object serves as a functional component within the subject.

Order prompt
Step 1: Here I give you an image of the found object. You should recommend {generated_number} associations, which are simple and familiar associations that people can easily recognize and understand. Try to think out of the obvious semantics and imagine what else this shape could represent. For each association, please identify: 1) The main subject: A complete, standalone object (not a part of something else). 2) The found object as a component: A functional part of the subject (e.g., elephant's trunk, bird's beak).

Step 2: You should construct each association as a tuple (subject, component), the subject and component should be simple. For example, if the found object is round such as an orange slice, possible associations include: (car, wheel), (owl, eye), (turtle, shell), etc.

Step 3: All your associations must align with the theme: {theme}. For example, if the theme is 'animal' and the found object is round such as an orange slice, possible associations include: (owl, eye), (turtle, shell) – all subjects are related to the animal theme; But NOT (car, wheel) or (ship, porthole) – as these subjects are not related to the animal theme.

Step 4: Please check the historical record to avoid any duplicate recommendations.
The historical record is: {recommend_history}.

Step 5: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "theme": "<theme_name>",
    "associations":
    [
      {
        "id": "<id of the association>",
        "subject": "<name of the subject>",
        "component": "<name of the component>",
        "explanation": "<explain why make this association in this format: 'The found object is <component> of <subject>,
        because <reason>', maximum 50 words>"
      },
      ...
    ]
}
```

Let's think step by step.

A.1.2   Idea filter

System prompt
You are an AI assistant. Your task is to select proper associations of the found object according to the its placement in the image and the drawing area.

Order prompt
Step 1: Here I give you an original and a {image_type} image of the found object. The area that can be drawn is: {drawable_area}. The area that cannot be drawn is: {undrawable_area}. Some associations with the found object are given: {associations}. You need to select the associations that match the placement of the found object in the image and the drawing area.

Step 2: Here are some dimensions to evaluate whether each association is proper:
1. Check whether the current rotation angle of the found object would match the component in each association, sometimes the shape similarity is high, but the rotation angle is not suitable.
2. Based on the position and proportion of the found object in the image, analyze whether it realistically fits the expected placement and size of that component within its subject. For example, if the found object is proposed as an 'eye', consider whether its size is appropriate (not too large to dominate the whole image) and whether its position (not too low in the image) would allow for a natural completion of the subject. Check whether the current placement would force an awkward or unnatural composition of the final subject.
3. Check every association, understand the relationship between the component and subject, imagine how to integrate the found object into drawings according to the explanation and the clarification of the drawable area. Analyze whether the found object would influence the recognition of the subject in each association, if so, the association may not be suitable.
4. Check every association, the subject should be a standalone object, and the component should be a functional part of the subject.

Step 3: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "selected_associations": [list of association_id]
}
```

Let's think step by step.

A.1.3 Generate similar associations (switch subject)

System prompt
You are an AI assistant. Your task is to recommend associations with a given found object based on shape similarity. Each association will be constructed as a tuple (subject, component), which means the found object serves as a functional component in the subject. You need to recommend associations that have the same functional component as the given association.

Order prompt
Step 1: Here I give you a mask image of the found object and an example of the association: {selected_association}. There are three elements: the main subject, the found object's functional component within the subject, and the explanation to explain the association.

Step 2: You should recommend {generated_number} similar associations according to the found object's shape and the given example, each association should be constructed as a tuple (subject, component). You should keep the found object's functional component same as the example, but change the main subject. Here I give you some examples, if the example is (car, wheel), then the results of similar associations may like: (bicycle, wheel), (motorcycle, wheel), etc.; if the example is (elephant, trunk), then the results of similar associations may like: (anteater, trunk), (tapir, trunk), etc. Please note that the subject should not be a part of some other object, but a standalone object, and the subject should be simple.

Step 3: All your associations must align with the theme: {theme}. For example, if the theme is 'animal' and the found object is round such as an orange slice, possible associations include: (owl, eye), (turtle, shell) – all subjects are related to the animal theme; But NOT (car, wheel) or (ship, porthole) – as these subjects are not related to the animal theme.

Step 4: Please check the historical record to avoid any duplicate recommendations. The historical record is: {recommend_history}.

Step 5: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "theme": "<theme_name>",
    "associations":
    [
      {
        "id": "<id of the association>",
        "subject": "<name of the subject>",
        "component": "<name of the component>",
        "explanation": "<explain why make this association in this format: 'The found object is <component> of <subject>,
        because <reason>', maximum 50 words>"
      },
      ...
    ]
}
```

Let's think step by step.

A.1.4 Generate similar associations (switch component)

System prompt
You are an AI assistant. Your task is to recommend associations with a given found object based on shape similarity. Each association will be constructed as a tuple (subject, component), which means the found object serves as a functional component within the subject. You need to recommend associations that have the same subject as the given association.

Order prompt
Step 1: Here I give you a mask image of the found object and an example of the association: {selected_association}. There are three elements: the main subject, the found object's functional component within the subject, and the explanation to explain the association.

Step 2: You should recommend {generated_number} similar associations according to the found object's shape and the given example, each association should be constructed as a tuple (subject, component). You should keeep the found object's subject same to the example, but change the component. Here I give you some examples, if the example is (car, wheel), then the results of similar associations may like: (car, headlight); if the example is (elephant, trunk), then the results of similar associations may like: (elephant, tusk). Please note that the component should be simple.

Step 3: All your associations must align with the theme: {theme}. For example, if the theme is 'animal' and the found object is round such as an orange slice, possible associations include: (owl, eye), (turtle, shell) – all subjects are related to the animal theme; But NOT (car, wheel) or (ship, porthole) – as these subjects are not related to the animal theme.

Step 4: Please check the historical record to avoid any duplicate recommendations. The historical record is: {recommend_history}.

Step 5: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "theme": "<theme_name>",
    "associations":
    [
      {
        "id": "<id of the association>",
        "subject": "<name of the subject>",
        "component": "<name of the component>",
        "explanation": "<explain why make this association in this format: 'The found object is <component> of <subject>,
        because <reason>', maximum 50 words>"
      },
      ...
```

```
    ]
}
```

Let's think step by step.

A.2 Visual Effect Generator

A.2.1 Generate scene-level visual descriptions

System prompt
You are an AI assistant. Your task is to recommend different scenes that incorporate a given object and reflect a given mood.

Order prompt
Step 1: Here I give you an object: {subject} and a mood: {mood}. You should recommend {generated_number} different scenes that incorporate the given object and reflect the given mood. As the scene will be used for sketch drawing, so it should be simple and clear. The scenes should have maximum 3 main elements. If the object has a functional purpose (like tools, vehicles, or furniture), show it being used in its intended context with specific characters (e.g., a chef using a knife, a child riding a bicycle). Please do not just use the mood as adjectives to describe the elements, but construct specific scenes to reflect the mood.

Step 2: Please give a title and a description for each scene, list the 3 main elements in each scene, provide 3 alternative elements that users could choose, keep all element names to 1-2 words without adjectives, ensure all elements simple, naturally fit within the scene and can be marked with bounding boxes in the image.

Step 3: Please check the historical record to avoid any duplicate recommendations. The historical record is: {recommend_history}.

Step 4: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "mood": "<mood_name>",
    "ideas":
    [
      {
        "title": "<title of the scene>",
        "description": "<scene description, maximum 30 words>",
        "elements": ["<element1>", "<element2>", "<element3>"],
        "alternative_elements": ["<alternative_element1>", "<alternative_element2>", "<alternative_element3>"]
      },
      ...
    ]
}
```

Let's think step by step.

A.2.2 Update scene description

System prompt
You are an AI assistant. Your task is to generate a cohesive scene description based on the given elements and mood.

Order prompt
Step 1: Here I give you some elements: {elements}, and a mood: {mood}. Please generate a cohesive scene description that only incorporates all these elements naturally and reflect the given mood. As the scene will be used for sketch drawing, it should be simple and clear. Please do not just use the mood as adjectives to describe the elements, but construct a scene to reflect the mood.

Step 2: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "mood": "<mood_name>",
    "result":
    {
      "title": "<title of the scene>",
      "description": "<scene description, maximum 30 words>",
    }
}
```

Let's think step by step.

A.2.3 Load more elements

System prompt
You are an AI assistant. Your task is to recommend elements with a given image description that reflects a given mood.

Order prompt
Step 1: Here I give you an image description: {description}, and a mood that the image reflects: {mood}. The current elements in the image are: {elements}. You should recommend another 3 simple elements that can be included in the image and maintain the mood, keep all element names to 1-2 words without adjectives. The recommended elements should be different from these ones: {alternative_elements}.

Step 2: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "mood": "<mood_name>",
    "elements": ["<element1>", "<element2>", "<element3>"]
}
```

Let's think step by step.

A.2.4 Generate object-level visual descriptions

System prompt
You are an AI assistant. Your task is to recommend different variations of a given object that reflect a given mood.

Order prompt
Step 1: Here I give you an object: {subject} and a mood: {mood}. You should recommend {generated_number} different
variations of the object that reflect the given mood. For example, if the object is a car and the mood is 'whimsical',
you could recommend variations like a flying car with wings. If the mood is 'energetic', you could recommend variations
like a sports car. As the object will be used for sketch drawing, so it should be simple and clear. Please focus on
describing the object's appearance and ignore any environmental elements or descriptions. Please do not just use the
mood as adjectives to describe the object.

Step 2: Please give a title and a description for each variation, list the object as the only element, keep the
object's name to 1-2 words without adjectives.

Step 3: Please check the historical record to avoid any duplicate recommendations. The historical record is:
{recommend_history}.

Step 4: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "mood": "<mood_name>",
    "ideas":
    [
      {
        "title": "<title of the object>",
        "description": "<object description, maximum 30 words>",
        "elements": ["<element1>"]
      },
      ...
    ]
}
```

Let's think step by step.

A.3 Layout Generator

A.3.1 Generate new layouts

System prompt
You are an AI assistant. Your task is to generate bounding box layout variations of some elements based on an image
description, elements in the image, and one element's component's mask image and its bounding box. You need to analyze
the image description and elements to determine the bounding box of each element, and also consider the component's
relationship with its parent element.

Order prompt
Step 1: Here I give you an image description: {selected_visual_description}, elements in the image: {elements},
the bounding box of the component: {bounding_box}, and the component's relationship with its parent element: {component}
of {subject}. You should generate {generated_number} layout variations of the bounding boxes of each element and the
component in the image. The bounding boxes should be in the format of "[x1,y1,x2,y2]", where x1,y1 is the top-left
corner of the bounding box, and x2,y2 is the bottom-right corner of the bounding box. The image size is 512*512.

Step 2: When creating these layouts, please carefully consider:
1. The number and name of the elements in layout should have an exact match with the given elements.
2. The component must be positioned within or properly connected to its parent element, maintaining a natural spatial
relationship. For example, if the component is 'trunk', place it at the front of an elephant; if the component is
'wheel', place it at the bottom of a car. The bounding box of the component should be unchanged.
3. The relative sizes of elements should be realistic and proportional (e.g., a hat should be smaller than the
person wearing it).
4. Elements should not overlap excessively unless it makes sense in the context (e.g., a person holding an object).
5. Ensure proper visual hierarchy with the main subject being prominent.
6. Distribute elements across the canvas in a balanced way.
7. Consider typical spatial relationships (e.g., a hat above a head, glasses on a face).
Each layout variation should feel natural while maintaining the integrity of the original description.

Step 3: Please give the results in JSON format, as shown below.
```
{
    "chain-of-thought": "<cot_content>",
    "layouts":
    [
      {
        "component": [x1,y1,x2,y2],
        "element_name1": [x1,y1,x2,y2],
        "element_name2": [x1,y1,x2,y2],
        ...
      },
      {
        "component": [x1,y1,x2,y2],
        "element_name1": [x1,y1,x2,y2],
        "element_name2": [x1,y1,x2,y2],
        ...
      },
      ...
    ]
}
```

Let's think step by step.

A.4 Prompt Generator

A.4.1 Generate scene-level T2I prompts

System prompt
You are an AI assistant. Your task is to generate high-quality Stable Diffusion 1.5 text-to-image prompts with a given scene description, elements, and style.

Order prompt
Step 1: Here I give you a scene description: {selected_visual_description}, elements in the scene: {elements}, and a style: {style}. You should generate a concise, descriptive prompt based on the scene description. Ensure the prompt accurately reflects the provided scene description and contains all elements in the scene. Each element name must have an exact match in the prompt. Format the prompt in this structure: '[scene information], [style tags], [quality tags]'. Do not use adjectives to describe the elements. The prompt should be simple and maximum 20 words.

Step 2: Please give the results in JSON format, as shown below.
{
   "chain-of-thought": "<cot_content>",
   "style": "<style_name>",
   "prompt": "<prompt_content, maximum 20 words>"
}

Let's think step by step.

A.4.2 Generate object-level T2I prompts

System prompt
You are an AI assistant. Your task is to generate high-quality Stable Diffusion 1.5 text-to-image prompts with a given object description, elements, and style.

Order prompt
Step 1: Here I give you an object description: {selected_visual_description}, object: {elements}, and a style: {style}. You should generate a concise, descriptive prompt based on the object description. Ensure the prompt accurately reflects the provided object description. The object name must have an exact match in the prompt. Format the prompt in this structure: '[object information], [style tags], [quality tags]'. Do not use adjectives to describe the elements and ensure a blank background. The prompt should be simple and maximum 20 words.

Step 2: Please give the results in JSON format, as shown below.
{
   "chain-of-thought": "<cot_content>",
   "style": "<style_name>",
   "prompt": "<prompt_content, maximum 20 words>"
}

Let's think step by step.

APPENDIX B
USER STUDY DETAILS

B.1 Demographic Information

TABLE B1
Demographic information of the user study participants, including gender, age, background, and drawing skills.

Participant	Gender	Age	Background	Drawing Skills
P1	Male	31	Design	Experienced
P2	Male	26	New Media Art	Limited
P3	Male	27	Human-Computer Interaction	Intermediate
P4	Female	27	Art	Experienced
P5	Female	29	Animation	Experienced
P6	Male	25	Human-Computer Interaction	Intermediate
P7	Male	28	Human-Computer Interaction	Limited
P8	Female	28	Human-Computer Interaction	Limited
P9	Male	28	Human-Computer Interaction	Intermediate
P10	Male	27	Human-Computer Interaction	Limited
P11	Male	30	Computer Science	Limited
P12	Male	27	Computer Science	Limited

B.2 Questionnaires

Part 1: NASA-TLX

Q1: (Mental Demand) The task required a lot of mental effort.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q2: (Physical Demand) The task required a lot of physical effort.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q3: (Temporal Demand) I felt rushed or under time pressure while performing the task.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q4: (Effort) I had to work hard to achieve a good result in this task.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q5: (Performance) I am not satisfied with how I performed in the task.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q6: (Frustration) I felt frustrated, annoyed, or stressed while doing the task.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Part 2: SUS

Q1: I think that I would like to use the tool frequently.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q2: I find the tool to be simple.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q3: I think the tool was easy to use.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q4: I think that I could use this tool without the support of a technical person.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q5: I find that the various functions in this tool were well integrated.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q6: I think that there was a lot of consistency in this tool.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q7: I would imagine that most people would learn to use this tool very quickly.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Q8: I find the tool very intuitive.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree
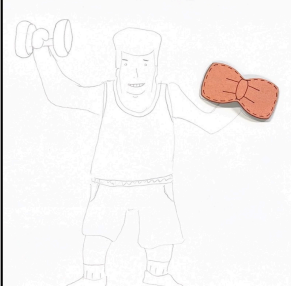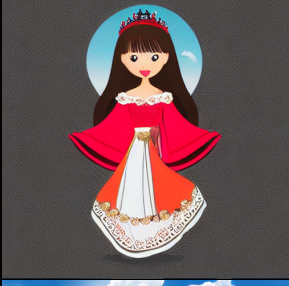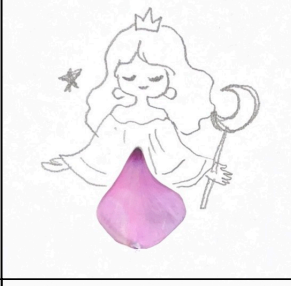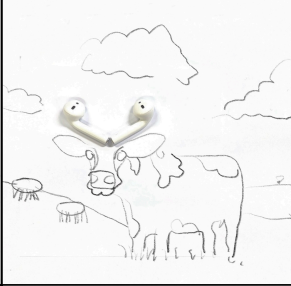
Q9: I feel very confident using the tool.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree
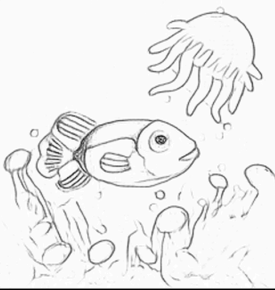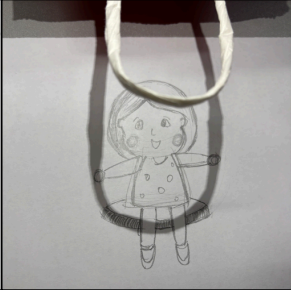
Q10: I could use this tool without having to learn anything new.
Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

B.3 Open-ended Study Results
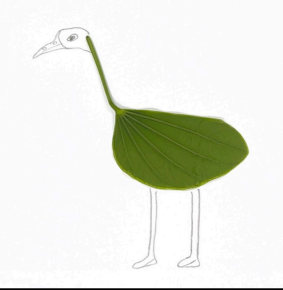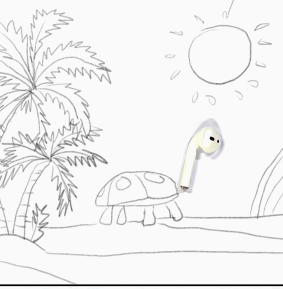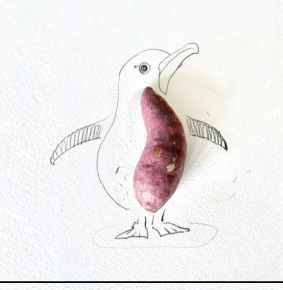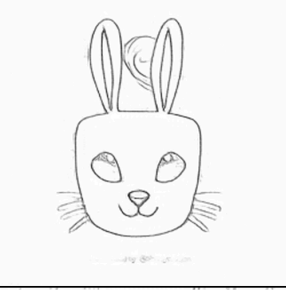

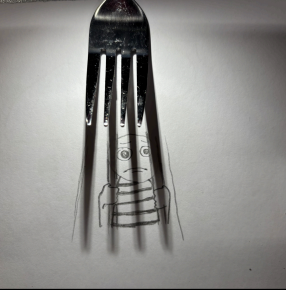| Found Object | Generated Reference Image | Pencil Sketch | Participants' Drawings |

	Found Object	Generated Reference Image	Pencil Sketch	Participants' Drawings
P7				
P8				
P9				
P10				
P11				
P12				

Found Object	Generated Reference Image	Pencil Sketch	Drawings