R4 - Design Evaluation and Validation

Team Members

Avery Ao

Holly Sun

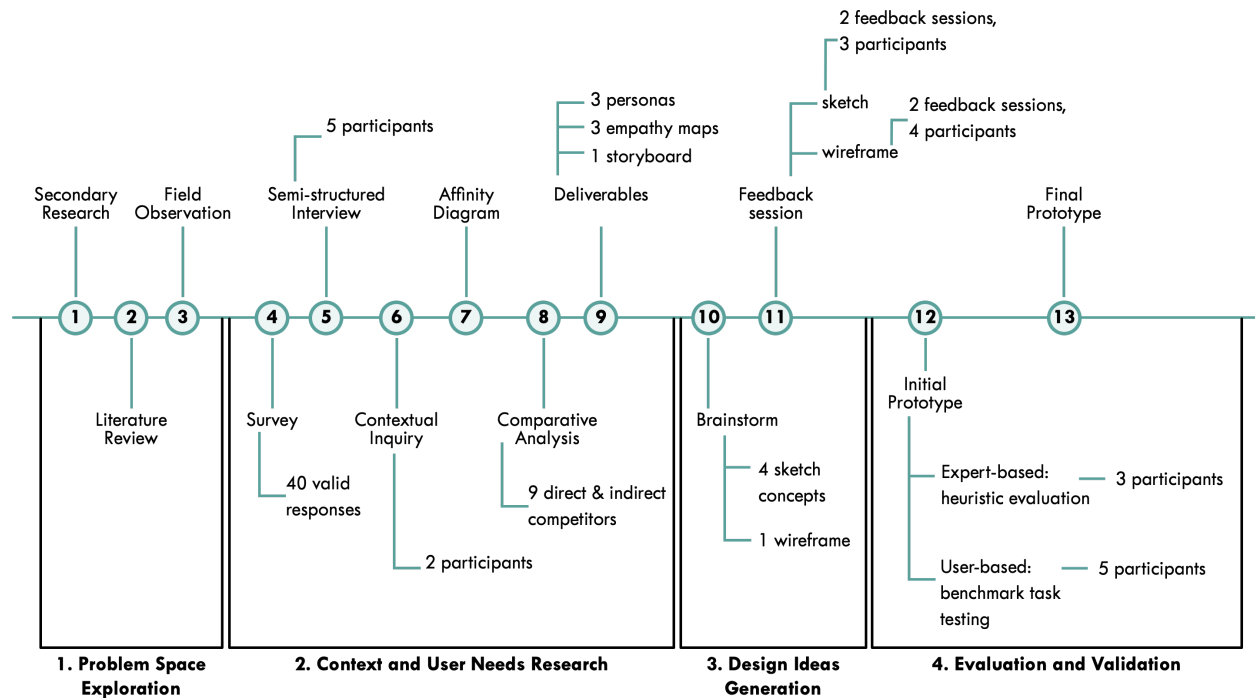Abhinav Thukral

Catherine Yang

Problem Statement

> 🖼️ **Help individual visitors (aged between 18-34 years) without relevant art background to be more engaged in their visiting experience at the High Museum of Art during the COVID-19 pandemic**

Introduction

Timeline diagram with the following phases and steps:

1. Problem Space Exploration
- 1. Secondary Research
- 2. Literature Review
- 3. Field Observation

2. Context and User Needs Research
- 4. Survey — 40 valid responses
- 5. Semi-structured Interview — 5 participants
- 6. Contextual Inquiry — 2 participants
- 7. Affinity Diagram
- 8. Comparative Analysis — 9 direct & indirect competitors
- 9. Deliverables — 3 personas, 3 empathy maps, 1 storyboard

3. Design Ideas Generation
- 10. Brainstorm — 4 sketch concepts, 1 wireframe
- 11. Feedback session — sketch (2 feedback sessions, 3 participants), wireframe (2 feedback sessions, 4 participants)

4. Evaluation and Validation
- 12. Initial Prototype — Expert-based: heuristic evaluation (3 participants), User-based: benchmark task testing (5 participants)
- 13. Final Prototype

Phase 1: Problem space exploration

Based on our research, the role of cultural exhibition spaces such as museums and art galleries has largely shifted from the traditional preservation of artifacts towards a more leisure-based experience, and the demand for attracting visitors and offering engaging experiences has been growing due to the pandemic crisis. Thus it's important to provide solutions to improve visitors' visiting experience while considering the concerns for social distancing. We found that visitors come to the art museums for different purposes including viewing the artwork, relaxation, education, and socialization. The engagement during visits to the art museum thus could be introduced by helping visitors to get their desired information and offering them opportunities to connect with a larger visitor group that offers them opportunities to socialize.

Considering current problems and needs, the objective of our project is to help individual visitors (aged between 18-34 years) without relevant art backgrounds to be more engaged in their visiting experience at the High Museum of Art during the COVID-19 pandemic.

Phase 2: Context and user needs research

We identified some common needs and key issues across our researches. Users express a large interest in the information provided by the museum, while the currently provided resources of information in the museum are inadequate to satisfy the users. In addition, users reported needs for interactions while the High Museum is lacking improvements on it. We also found that users need a way to collect visit-related information and material for future references and socialization.

In addition to those findings, we synthesized 4 key user goals and related needs below.

1. **Knowledge**

    a. Prepare for the visit

    b. Find out museum-related information

    c. Want more organized, concise, and understandable information about artworks

    d. Want to see connections between artwork and exhibitions

2. **Socializing**

    a. Share or receive information with/from others

    b. Share their experience

3. **Navigation**

    a. Find how to navigate around in the High Museum of Art

4. **Documentation**

    a. Keep a record of certain artwork

Phase 3: Design ideas generation

We generated 4 sketch concepts based on research findings, and conducted feedback sessions to identify issues. Based on the feedback, we integrated all concepts into a system with 3 key features, addressing key user goals mentioned above:

1. **AR scanning for artwork-related information**

    a. allows users to easily access artwork information that is understandable and informative to users of all levels of background knowledge in the art

    b. offers storytellings that present backstage connections of artworks, artists and exhibitions

    c. AR offers a channel to access information through a more interactive method, and the AR elements are more understandable for visual learners

2. **Comments on artworks**

    a. provide a platform for users to share their thoughts and read ideas from others contactless

3. **Creating postcards and diaries embedded with artwork information**

    a. as sharable content can be directly sent to other people or platforms

b. offer methods for users to collect artwork-related content for personal use

Navigation need is also addressed by embedding the museum's map into our system as a secondary feature.

Significance

Addressing those goals and needs is important. By increasing the efficiency of obtaining information during the visits, visitors' demand for learning could be satisfied, and the knowledge that they learned will have a lifelong benefit. Offering socializing channels allows visitors to exchange thoughts and connect to others freely even during the pandemic, making the visiting experience more interactive. Providing a guide on navigation is necessary to ensure a smooth visiting experience. And documentation helps visitors to retrieve and share their visits, which may also be triggering factors for repeating visits. These together prove that the museum experience is more than a physical trip to observe and learn about artworks, but also an opportunity to interact with artworks and other visitors through multiple forms. Our design may thus have a long-term positive impact on users, demonstrating the significance of our topic.

Aside from the benefits on an individual level mentioned above, our design also has a large social impact. Inviting viewers to interact with artworks could potentially increase the public's interest in cultural studies, and promote the level of education related to this field. Our approach towards the museum experience is also innovative in terms of how we combine the cultural and social aspects of museum visits and our acknowledgment of a community of visitors.

Phase 4: Evaluation and validation

For this phase, we created our prototypes and conducted evaluation sessions to assess usability issues. The prototype was iterated throughout the entire process. More details will be discussed in the following sections of this report.

Prototype Description

In this section, we are going to describe our final prototype. We also want to explain that, this final prototype refers to the version of our design before we did our evaluation, and we actually made some changes during our evaluation. In order to maintain a clear and chronological order of the presentation of our R4 stage, we will first describe the prototype before any changes and will present the changes in the "Design Implication" section.

Link to the description of the prototype: https://www.figma.com/file/2jv3gsTFvYxOhHjHIPnprK/R4---Prototype-Description?node-id=0%3A1

Link to demos of primary functions: https://faint-fragrance-7a0.notion.site/Project-Demo-16f9e712ef13460dbe551f5ae0dd566f

Evaluation Activities

Expert-Based: Heuristic Evaluation

Evaluation goals

For our expert-based heuristic evaluation, we hope to have experts review the prototype and offer suggestions for further modifications on prototype. More specifically, we have the following evaluation goals to guide our evaluation sessions:

- Evaluate the overall functionality of the prototype

- Identify usability and accessibility issues of the design

- Test the ease of use of the prototype among expert users

- Assess the learnability of each features o the prototype

- Assess the error recovery features of the system

Method justification

Strength

Organized evaluation standards

Standardized principles/rules for finding usability problems are used, which help to organize the evaluation process and data. Applying the correct heuristic can also help suggest the best corrective measures to the design team.

Domain-related knowledge

Experts who are more familiar with target domains could help to identify usability and accessibility problems that might be ignored by design teams and users, and offer more objective feedback from a third-party perspective.

Quick and flexible

Expert evaluation requires fewer people to participants compared to user testing, allowing us to gather feedback relatively more quickly. It can also be used along with other usability testing methods and can be used early in the design process, making it a fairly flexible method to be used.

Limitation

Recruitment of experts takes time and effort

In order to obtain reliable and useful evaluation data, knowledge and experiences to apply the heuristic effectively are necessary, and trained usability experts are hard to find. In our case, we had two of our HCI professors and one HCI first-year student as the experts. While professors have the systematic expertise to successfully apply the heuristic rules, the HCI student might lack enough experience to assess the prototype independently. The student was recruited due to the lack of time and convenience of access, which we shall admit the limitation of data collected and analyze the results accordingly.

Fewer insights from actual users

Heuristic evaluation primarily assesses problems from domain-related experts' perspectives, who may not belong to the target user group, thus not offering insights from actual users' perspectives. Additional user

testings are required to fully assess the usability issues.

Method details

Participants

It's necessary to recruit experts who have knowledge of usability evaluation and are able to use digital devices to interact with our prototypes without barriers. Our screen criteria are listed below:

- Participants should be familiar with Nielsen's 10 Heuristics
- Participants should have adequate knowledge of usability evaluation
- Participants should be comfortable using phones or computers to assess prototype

Based on our criteria, we recruited 3 experts for our evaluation:

- 2 MS-HCI professors
- 1 MS-HCI first-year student

Task/task scenarios

We prepared 4 tasks for each expert during heuristic evaluations. Those tasks aim to help experts understand the full functionalities supported by our system. While given the tasks, experts were expected to explore features and flows freely and use think-aloud techniques that allow us to understand their thoughts when identifying problems.

Tasks	Rationales
Find the information of the art + bookmark the art and add notes	- To understand the learnability of AR scanning the artwork to access information; - To understand the learnability of bookmark and note features; - To identify accessibility issues of the design - To explore the navigation flow among different information sections
Read comments and send reactions	- To explore users' understanding of the comment page's design and layout - To identify usability and accessibility issues of the comment page
Find previous bookmarks and notes	- To explore the navigate flow to the bookmarks and note page from the scanning screen - To identify confusions caused by the design - To identify accessibility issues of the design
Create and share a Postcard made from the artwork	- To explore the flow of creating a postcard - To identify confusions caused by the design - To identify accessibility issues of the design

Justification

For every expert evaluation, we provided partially different task lists throughout the process. For the first expert, we offered 6 tasks but realized the redundancy of giving too many tasks, so we decided to combine the first three tasks into one flow. However, after the second expert session, we realized that combining all three tasks led to ignorance of details in evaluation, so for the third expert session we divided the first set of tasks into two. The difference of task list was only limited to the first set of flows, and the rest of the tasks were the same for all experts, as shown in the graphic below.
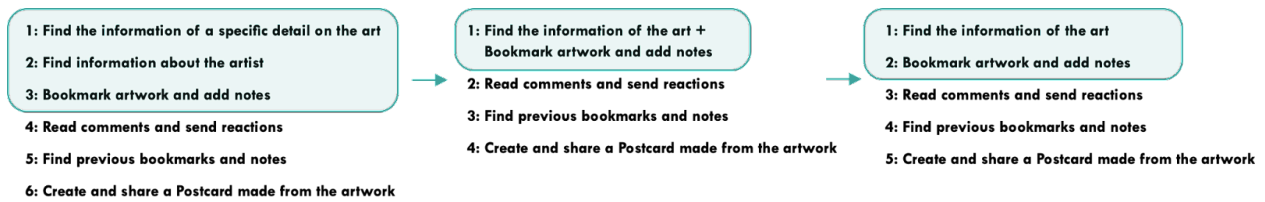
Tasklist variations throughout the process

Such inconsistency is due to our lack of experience in conducting expert evaluations and attempting to adjust the task list after previous sessions' feedback and performance. We shall admit that such inconsistency may cause inconvenience and inaccuracy in data analysis, and we would avoid such situations in the future.

For data analysis purposes, as we lack severity ratings for certain tasks due to the inconsistency of task design across different experts, we chose to use the second task list that combines all separated tasks into one as a standard to calculate quantitative data. We calculated the mean value of task 1+2+3 for list one and task 1+2 for list three to yield a value align with the data set of list two. Therefore, our data analysis is based on **task list two that has 4 tasks**. More details on data analysis are in the following section.

Metrics

We chose to use **Nielsen's 10 Usability Heuristics** as our metrics for the evaluation since it offers a series of quick and standardized principles to assess usability. We decided to remove "Help and Documentation" from the evaluation principle list because the documentation didn't apply to our app design and other forms of "help" could be assessed from the "Help users recognize, diagnose, and recover from errors" principle.

To assess each heuristic, we provided experts with a table listing each principle, and let them rate each heuristic with a **severity rating** (1-5, 1 = I don't agree that this is a usability problem at all, 5 = Usability catastrophe: imperative to fix this before the product can be released).

In addition to the heuristic principles, we also include the **Accessibility Concerns** section to ask for feedback on our design's accessibility. A metric table is offered based on each task.

Heuristic Principles	Rationales
Awareness of System Status	- To explore if we have enough feedback and notifications to inform users about the system status
Match between system and the real world	- To understand if there are any confusions caused by the design of icons, buttons, and other design elements- To explore users' mental models and expectations for each feature
User control and freedom	- To identify any lack of emergency exit and undo/redo features
Consistency and standards	- To assess the consistency of the internal design system- To assess the learnability of features
Error prevention	- To assess the use of notifications, warnings, and constraints to reduce errors
Recognition rather than recall	- To access practices help reduce users' cognitive load
Flexibility and efficiency of use	- To explore improvements for customizations and personalizations used in our design
Minimalist design	- To assess the content and UI design - To identify any information overloading

Heuristic Principles	Rationales
Help users recognize, diagnose, and recover from errors	- To assess the efficiency of the use of error messages
Accessibility concerns	- To identify missing accessibility issues and obtain feedback on possible practices

Procedure

We recruited our experts through Slack messages and emails and scheduled 1.5-2 hrs sessions with each participant. 2 sessions were conducted in person, and 1 was conducted remotely while the expert shared her screen. The prototype was evaluated within the Figma on a computer or a mobile phone device.

All four team members participated in all sessions. For each session, we assigned one **moderator** who briefed the expert on our project's focus, and gave instructions on tasks, and answered questions when the expert encountered difficulties navigating through our system. The moderator also guided the expert to key features that were missed in the inspection at the end of the session to ensure we receive feedback on all target features. Two team members were assigned as **notetakers**, and one was the **observer**.

During each session, after the moderator briefed the problem space to the expert and presented the task list, the expert started to inspect the system freely and applied the think-aloud method to say their thoughts out. At the end of each task, we provided the expert with the heuristic evaluation list and recorded their scores, violations, and recommendations for each principle along with Accessibility concerns. Additional questions were also addressed by the moderator and other team members at the end of each session.

Data Analysis

Procedure

Quantitative

We collected heuristic severity ratings for each of the 4 tasks during each expert session, resulting in 12 severity rating tables in total. Then we calculated the **mean values of severity ratings across all experts for each task**, resulting in 4 tables in total (listed below). The average rating of each principle helps us understand which principles are the most violated ones.

The **average overall ratings of each task** were also calculated to help us understand which task flow is the most problematic.

Qualitative

We collected and listed out all violations, recommendations, and accessibility concerns for each task as the qualitative data. After finding out top violated principles and problematic task flow based on Avg. severity ratings, we looked into corresponding qualitative data to gain insights on what caused the problems.

We also summarized the Accessibility Concerns list to gain a more systematic understanding of the accessibility issues we had in the design.

Data

> ## Task 1: Find the information of the art + bookmark the art and add notes

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Awareness of System Status	1.2	- Would assume tapping on the cross on the top right to go back when I am at the comment page - The "save" pop up is too short, can stay on the screen for longer - It started scanning before I take an action to do it - I don't know how broad I need to scan (not sure how to scan) - Not sure how the photo is scanned	- User should also be able to swipe up to get rid of notifications - Have an entry to access all the notes - Auto-save note indication could be more obvious - Need indication that I have taken notes already
Match between system and the real world	1.4	- Unknown how the notes could be managed based on current fidelity - I'm not sure what "timeline" means	- May give a quick onboarding for the first-time user - Allow swipe to switch between info sections
User control and freedom	1.8	- Want to go back to the original status (seeing fewer comments) after clicking on "see more comments" on the comments page - Users' control option is limited when scanning the painting - Need to manually delete notes within the artworks now, want to manage all notes in one place - Would be nice to have a place jump out to other places without restarting the scanning - Looks like the audio automatically started to play and I have to force it to stop	- Instead of using the "view more comment' button, load more comments by dragging down the page with infinite loading, but be able to come back to the initial comment display by dragging back - Global availability of menu - For additional info, may be hard to build into the system, but might be easier to link to external info
Consistency and standards	1.2	- The take note icon looks like hinting me to draw on the image	- Use a notebook icon instead of a pen icon
Error prevention	1.4	- May erroneously click on the cross button on the top right - A lot of interaction options after scanning the painting, overwhelming	- Cross button is currently parallel with the switching tab, which makes it seem functionally parallel. Maybe put the switching tab a bit downwards - Have a last viewed floating window on the camera page - More clear indication if the tag is clickable or not
Recognition rather than recall	1		- Show the time for the previous note for repeating users
Flexibility and efficiency of use	1.2		- Global availability of menu
Minimalist design	1	- Text is too small, may not be readable to some	- Change the size of the text — maybe try zoomable text section
Help users recognize, diagnose, and recover from errors	1.2		

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Avg. Overall Rating	1.27		

Top violation principles:

- User control and freedom (1.8)
- Match between system and the real (1.4)
- Error prevention (1.4)

Task 2: Read comments and send reactions

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Awareness of System Status	1	- Couldn't see how many comments there are - Cannot see which user sent which emoji	- Show how many comments, the order of the comment
Match between system and the real world	1.3	- In the real world there might be a discussion(back and forth) about things, but the current page looks quite linear	
User control and freedom	1.3	- Couldn't comment on comments	- May need a limitation on comment word count
Consistency and standards	1		
Error prevention	1	- Think X lead to the close of the comment section and I would be able to see the image again -- counter-intuitive	
Recognition rather than recall	1		
Flexibility and efficiency of use	1.3	- Couldn't comment on comments	
Minimalist design	1		- Zoom in comments, pinch to view

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Help users recognize, diagnose, and recover from errors	2	- Not sure if I can send multiple emojis - Not sure the difference between emoji on the screen and emoji from the message entering part - Think reaction is the filter of comments -- no clear indication of the emoji function	- Consider how Slack send reactions
Avg. Overall Rating	1.21		

Top violation principles:

- Help users recognize, diagnose, and recover from errors (2)

- Match between system and the real world (1.3)

- User control and freedom (1.3)

- Flexibility and efficiency of use (1.3)

Task 3: Find previous bookmarks and notes

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Awareness of System Status	2	- Bookmark interaction confusing	
Match between system and the real world	1.7		
User control and freedom	1.7	- Bookmark and notes management are confusing - Don't have access to the yellow info tags	- Global availability of the menu
Consistency and standards	1.5	- There is both a note and bookmark icon on the menu page, but only a bookmark icon when clicking on the artwork - Art info is presented in different formats in scanning page VS. bookmarked page - Pencil and bookmark icons are used differently in card VS on page	
Error prevention	1.7	- Not sure if un-bookmark artwork will delete the notes taken with the artwork - Pencil and bookmark icons are used differently in card VS on page	
Recognition rather than recall	1		

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Flexibility and efficiency of use	1.7	- Not sure what will happen if un-bookmarked artwork has notes under the note section - Don't have access to the yellow info tags	
Minimalist design	1	- Text on the bookmarked card is small and lack contrast	
Help users recognize, diagnose, and recover from errors	1		
Avg. Overall Rating	1.47		

Top violation principles:

- Awareness of System Status (2)
- Match between system and the real world (1.7)
- User control and freedom (1.7)
- Error prevention (1.7)
- Flexibility and efficiency of use (1.7)

Task 4: Create and share a Postcard made from the artwork

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Awareness of System Status	1.7	- Steps are confusing after editing	- Other images should be greyed out once I choose one artwork to make a postcard (since we only allow one image at a time in postcard making)
Match between system and the real world	1.3	- Small interactions are not really clear	
User control and freedom	1.3	- Couldn't share postcards from the collection page - Not sure how much content can be put into the card	- I might want my avatar image on the postcard

Heuristic Principles	Avg. Severity Rating (1-5)	Violation	Recommendation
Consistency and standards	1	- Share and other buttons transpose, not sure in terms of the whole design	
Error prevention	1.3	- After I create the postcard, the back button takes me to the last page, not the landing page. I want a button that can take me all the way back - Use of "done" not "next" after editing makes me imagine I would be brought back to the menu page	- Do not need a done button at the end of the process
Recognition rather than recall	1.7	- When I enter from the diary/postcard page, I do not realize that all the images I can select are those that I have bookmarked	
Flexibility and efficiency of use	1.7	- Extra steps reduce efficiency	- Need share option on postcard collection page
Minimalist design	1	- Template seems to be too small and I would probably not be able to understand the template	
Help users recognize, diagnose, and recover from errors	1.3		- Need share option on postcard collection page
Avg. Overall Rating	1.37		

Top violation principles:

- Awareness of System Status (1.7)
- Recognition rather than recall (1.7)
- Flexibility and efficiency of use (1.7)

Task severity rating orders (descending):
- Task 3: Find previous bookmarks and notes (1.47)
- Task 4: Create and share a Postcard made from the artwork (1.37)
- Task 1: Find the information of the art + bookmark the art and add notes (1.27)
- Task 2: Read comments and send reactions (1.21)

Accessibility concerns summary:

- Text size is too small for some users, maybe offer to zoom of the text on art info, comments, CTA buttons

- Need descriptions of visual elements within artworks in addition to the audio guide

- Consider info dots' contrast over the artwork

- Consider users speaking foreign languages

    - Translate comments in foreign languages

- Consider speech to text for note-taking feature

User-based evaluation: think-aloud + semi-structured interview

Evaluation Goals

For user-based evaluation, we seek feedback from users regarding how well our system satisfies their needs. While several evaluation goals might be similar to the expert-based evaluation, this time we try to approach such goals from a user's perspective. Specifically, the goals include:

- Understand if there is any confusion with using our system

- Understand whether our system is easy to use

- Check user's likeliness of using our system and the reasons behind

- Check user's expectations of our system and whether the current prototype satisfies these expectations

Method Justification

We use a combination of think-aloud and semi-structured interviews for user-based evaluations. We will discuss the strength and limitations of such a combination in this section. Details will be introduced in the next session.

Strength

Comprehensive

While think-aloud gives users much freedom in sharing their thoughts, a combination with the semi-structured interview ensures that we cover all the evaluation goals in each feedback session. Some users may be comfortable with sharing their thoughts and may talk a lot, while others may not speak up as much. As a result, simply conducting a think-aloud may not give us enough feedback since the effectiveness of this method is very dependent on the user. The addition of a semi-structured interview will encourage some users to share more thoughts and ensure that all goals are satisfied.

More standardized

All participants are asked to complete the same task during the think-aloud, and the moderator will ensure that we got the answer to all the questions we prepared. In this way, the result of each feedback session is more standardized since a combination of the two methods ensures that we get all the information needed across different participants.

Flexible

Both think aloud and semi-structured interview allows participants and the moderator to conduct the session in ways that are suitable to a specific participant. Participants do not face much restrictions during the think-

aloud, which gives them freedom in navigating our app and sharing their thoughts. As for a semi-structured interview, the moderator may determine when and how to ask the questions based on the circumstance. For example, if the participant is very active in sharing their thoughts and has already covered our evaluation goals, then the moderator would not ask certain questions in order to avoid being repetitive.

Limitation

Potentially guiding/leading

The combination of a semi-structured interview and a think-aloud means that the moderator may ask questions during the think aloud. There is thus a possibility that these questions might appear leading and bias the result of the think aloud. Therefore, the effectiveness of the method is relatively dependent on the moderator.

Need to educate users

Users might not be comfortable with the think-aloud. Some users do not speak up a lot and the moderator has to pop in and ask questions to guide users to share. This increases the likeliness of a biased result since the moderator might need to interrupt the think-aloud process several times in order to encourage users to talk more.

Method Details

Participants

We screen our participants based on whether they fit into our target user group. They satisfied four criteria:

- All participants have visited the High Museum of Art in the recent year

- All participants are categorized as individual visitors, meaning that they appreciated the artworks on their own

- All participants fall into the age range of 18 to 34

- Participants may have some knowledge in art, but they are surely not art experts

Since our system is designed for this group of people, users who fit such criteria can better help us identify if our design satisfies their needs. This is because these criteria would affect the user's needs. Visiting the High Museum of Art in the recent year means that all participants visited the museum during the COVID-19 pandemic and have experienced an alternate museum experience as a result. They would thus be more concerned with changes brought by the pandemic, such as matters of social distancing. Being an individual visitor suggests that they do not have familiar people to chat with or gain information from when they are appreciating artworks. Therefore, they may have more needs in getting information in other ways as well as needs of socialization. The age group of 18-34 not only suggests that they are the major visiting group of the High Museum of Art, but also indicates that they are very likely to be tech-savvy. This would affect our accessibility concerns as certain features may be easily understandable to young users but not elderlies (because people aged 18-34 are likely to be familiar with systems such as TikTok and Instagram, which introduces them to certain types of interactions that seniors may not be familiar with). Participants are not experts in art, which means that they need the information to be concise and easily understandable. While an art expert might be looking for in-depth knowledge and might be willing to spend plenty of time before each artwork, people who do not have much background in art prefer information that is more introductory and

succinct. Such preference would surely affect our design when it comes to what type and amount of information to include. With all these being discussed, it is important for us to have participants from our target user group since people who fit into these four criteria will have their corresponding user needs that our design attempts to address.

Session arrangement

Participant	Roles Assigned
User 1	Moderator: Abhinav Notetaker: Catherine, Holly Observer: Holly
User 2	Moderator: Holly Notetaker: Catherine, Abhinav Observer: Abhinav
User 3	Moderator: Catherine Notetaker: Holly, Abhinav Observer: Abhinav
User 4	Moderator: Avery Notetaker: Catherine, Abhinav Observer: Abhinav
User 5	Moderator: Avery Notetaker: Catherine, Abhinav Observer: Abhinav

Procedure

For each feedback session, the moderator will give an opening script. They will first briefly introduce our system, and then introduce the user to the procedures of this feedback session. The moderator will explain what a think-aloud is and ask the user to share their thoughts while they are navigating through our system. The moderator will then ask the user to conduct certain tasks and tell them that we will ask questions during and after each task. In this process, the notetaker will note down what the user shared, and the observer will observe the user's behavior to check if there's any confusion with using the system or any subconscious behavior that could potentially suggest the user's thoughts.

Tasks and Rationale

The tasks are designed to cover all essential functions of our system. Having participants going through these tasks will ensure that the feedback session is comprehensive. Assigning the same tasks to all participants also helps standardize the evaluation.

The scenario under each task refers to the scenario we ask users to imagine when they are testing out these features.

Task and Scenario	Rationale
Task 1: Scan artwork and find information Scenario: The user is visiting the High museum and is physically inside the museum, standing before an artwork.	This task covers the central feature of our system, which is to make the visiting experience more engaging by providing users with easier ways to gain art-related information. By asking users to complete this task, we explore how users feel about the current form of information presentation, and whether they are satisfied with the variety of information that our system provides.
Task 2: Bookmark artwork and add notes Scenario: The user is visiting the High museum and is physically inside the museum, standing before an artwork.	This task covers features that are very related to our information presentation feature. When users scan artwork and see the information page, they can also bookmark this artwork and add notes for personal records. We want to know if users find these features to be helpful, and whether there are problems with our current way of implementing them.

Task and Scenario	Rationale
Task 3: Read comments and send reactions Scenario: The user is visiting the High museum and is physically inside the museum, standing before an artwork.	This task runs users through a feature that is presented side by side with the information page. When users swipe left on the information page, they will enter a comments page where they can see other visitors' comments and reactions to the artwork. We want to know if users find this feature to be helpful and intuitive to use (whether our current design is easily understandable).
Task 4: Find previous bookmarks and notes Scenario: 1) User is in the museum in front of another artwork, but wants to refer back to a previous artwork that they saw on another floor. 2) User is now back at home, but wants to refer back to their museum visit.	This task asks users to go through the documentation features within our system. By asking them to find the bookmarks and notes they added previously, we explore whether users find the documentation to be helpful and effective during the visit and post-visit.
Task 5: Create and share a Postcard made from the artwork Scenario: 1) User is viewing an artwork in the museum, and wants to immediately share it with their friend. 2) User is now back at home, and wants to share their visiting experience with others.	This task asks users to explore the sharing features of our system. Our system allows users to create and share a postcard directly from the information page, and from the main menu. We guide users to go through both processes by providing them with different scenarios. When users are figuring out this task, we observe whether the two ways of making a postcard are explicit to the users. We also observe if users have a preference and why. Last but not least, we explore if users are satisfied with the current form of shareable content, and whether they find the making and sharing process to be easy.

Questions and Rationale

We ask questions on two occasions: 1) when users ask us whether they are doing something correctly and 2) when users finish each task. We try to not provide any guiding information even when users get stuck. Therefore, we would ask them what they expect to happen before giving out any information. We also want to explore users' overall feelings about this function. While users are completing the tasks, they would share their instant feelings, but some may not share an overall comment. Therefore, we prepared some questions to make sure that the latter aspect is covered in the feedback sessions.

Question	Rationale
What were you expecting to happen for this specific action?	When users get stuck and ask us for help, or when users question if they are using the system properly, we will ask them this question before telling them an answer. We want to know what confuses users and what are they expecting when they figure out that our system does not suffice their needs.
How would you interpret this function?	This is a follow-up question after the last one. After we explain their confusion, we would ask users how they would interpret this function based on the information given. We want to know how they would use this function (if their expectation is different from ours) and how they think about it in general.
Do you have any recommendations for this?	When users raise their confusions or problems with our system, we would ask this follow-up question to check whether they have any inspiration in how to resolve the problem. We acknowledge users may not have a recommendation and we tell them that this is a completely optional question and ask them not to feel stressed about the question in advance.

Question	Rationale
What would you think should be included in this page?	We ask them this question after they finish tasks in a specific interface. This is to check whether there is anything that our system does not have right now, but that users would anticipate us to include.
Do you think you will use this function?	We ask this question after each task. Users will always go through the whole system as we asked them to, but this does not mean they would prefer to use all the features if the system is actually implemented. Therefore, we ask this question to check what features they would use in real-world scenarios and what features they are less likely to use.

Data

We went through all the notes, observe problems and issues that appear as clusters, take notes of prominently severe issues, and picked notes that inspires us to make changes to the prototype. For issues, we would consider any design that causes repetitive confusion or errors on participants and their level of severity would be analyzed based on the rating for expert's evaluation and the amount of confusion and errors for the user testing. The notes we picked are in the following tables, the findings will be discussed in the design implications section.

Task 1: Scan artwork and find information

User 1	"how do I go back to the first screen?" "I thought the info tags are just telling me that it is a scanned page, and it will go away after the screen is done" "it could be more clear why artworks are related"
User 2	"for related artworks, I want to know if this artwork is in the High museum and where can i find it" "I want an explanation of why this artwork is related" "I want the menu icon to be global and consistent"
User 3	"In my experiences, I used to read the labels and I remember sometimes they had a QR code — confusing when you do not have a clear instruction in the museum about scanning the art piece, QR code is intuitive." "I might need to know that I can scan the art piece" "Audio should go more obvious instead of hiding in the section"
User 4	"am I supposed to scan a barcode or the artwork?" "I like that you don't scan the barcode but the actual artwork, but it might be confusing in the beginning cause I will expect to scan a barcode" "for related artworks, how are they related?" "I thought the info tags are clickable!" "I expect to have zoomed details not from my photo, but from an actual scanned image" "I don't have the actual text right now so I am confused what the info tags are about"
User 5	"I expect the bookmark would allow me to share my comments in social media, and to share my museum experience"

Task 2: Bookmark artwork and add notes

User 1	"I want the bookmark notification to stay longer" "How is the note saved?" "I consider sharing bookmarks to be more educational because it is informative"
User 2	"I am confused about the bookmark and note buttons" "there should be an onboarding telling me why I would want to save this artwork into bookmark" "I want to directly get to the all bookmarks and notes page without having to unbookmark and bookmark again" "would be helpful if I can copy/paste from info pages and add those to notes" "I fear to tap on the notes part because I fear that I would delete it" "how do you know your notes are saved when you close it?"
User 3	"Adding notes is confusing" "I don't know what to say for notes" "Telling the bookmarked is redundant because I'm already in there" "Adding photos/videos does not make sense" "I will never use notes in any of my visits"
User 4	"is the notes private or public?" "I expect the photo to also be saved in my camera folder on my phone" "Where does the image in the bookmark come from? it is confusing" "I expect to see the scanned image(not with the frame,

	scanned by the museum), or a picture taken from the frontal aspect" "for photos, I add to the notes, I want to see it when I swipe left right in the bookmarks and notes page" "I would like to highlight information that I find interesting, and I want them to be a part of my notes, or maybe highlight them in the bookmark"
User 5	"Can I edit my notes afterward?"

Task 3: Read comments and send reactions

User 1	"how do I react to a particular comment?" "Viewing emojis and comments side by side is confusing to me" "I want to know the top comments based on popularity"
User 2	"I am confused about the meaning of some emojis" "Is the comment presented by popularity or time?"
User 3	"comments are not adding values" "emojis are personalized and misleading" "sad face could be prime people that they would be reluctant to see the art piece" "My purpose for using this app is to get as much information as possible in such a short time frame instead of spending time on the comments" "Are notes and comments the same thing?" "does not really add value to the visit"
User 4	"I would like to see when they visited, not just the time of commenting" "I want to know if there are live comments, and how many people are in the museum while I am there" "the artboard emoji is confusing" "I'm expecting to see other people's picture in the comments section"
User 5	"I thought notes and comments are the same things, but they are not" "I assume that the comments and emojis are rolling automatically so I don't need to scroll"

Task 4: Find previous bookmarks and notes

User 1	"I want to share my bookmarks"
User 2	"The bookmark and notes page give me more info than I expected cause it also contains the notes page. I expect to see my bookmarks only" "I want to swipe down the info session to expand the image of artwork" "I prefer to have the info in the info tags are also shown in this page"
User 3	NA
User 4	"for photos I add to the notes, I want to see it when I swipe left right in the bookmarks and notes page"
User 5	"I expect all notes and videos will be compressed into one section, and I expect that I can swipe the photo on the top to see photos and videos that I took"

Task 5: Create and share a Postcard made from the artwork

|	
User 1	"I would've never shared a postcard because I use this app for educational purpose and would prefer to share bookmarks"
User 2	"I want to make a postcard from an image of the artwork only, not including the frame and the wall" "I want to zoom in and out to adjust the image" "When did the system save my postcard?"
User 3	"thought it was something at the gift shop"
User 4	"I expect to see artworks that I have already scanned" "I would not press the share button to download the image"
User 5	"I expect to add my notes to my postcard and share it with friends" "I expect that the postcard will be saved in the app, and I can choose to save it to my phone as well"

Design Recommendations

In the R4 stage, we have collected a lot of useful feedback about our design and we got some chances and time to work on improving the design, so in this section, we are going to discuss those improvements that we've already worked on and then those ones that we did not have time to touch on but had some ideas in our plan.

Changes Made During Testing

As we were conducting our expert testing and user testing, we were also improving and updating our design based on the feedback we got. We did not change everything that got negative feedback or recommendation because participants' feedback may also be biased or less reliable. Instead, we identified problems and issues that happened the most and caused the biggest confusion and ranked them by the level of priority of changing. In this way, we improved our design to avoid wasting testing time on discussing repetitive problems or clarifying the same confusion.

Change 1: Remove Timeline

Before   **After**
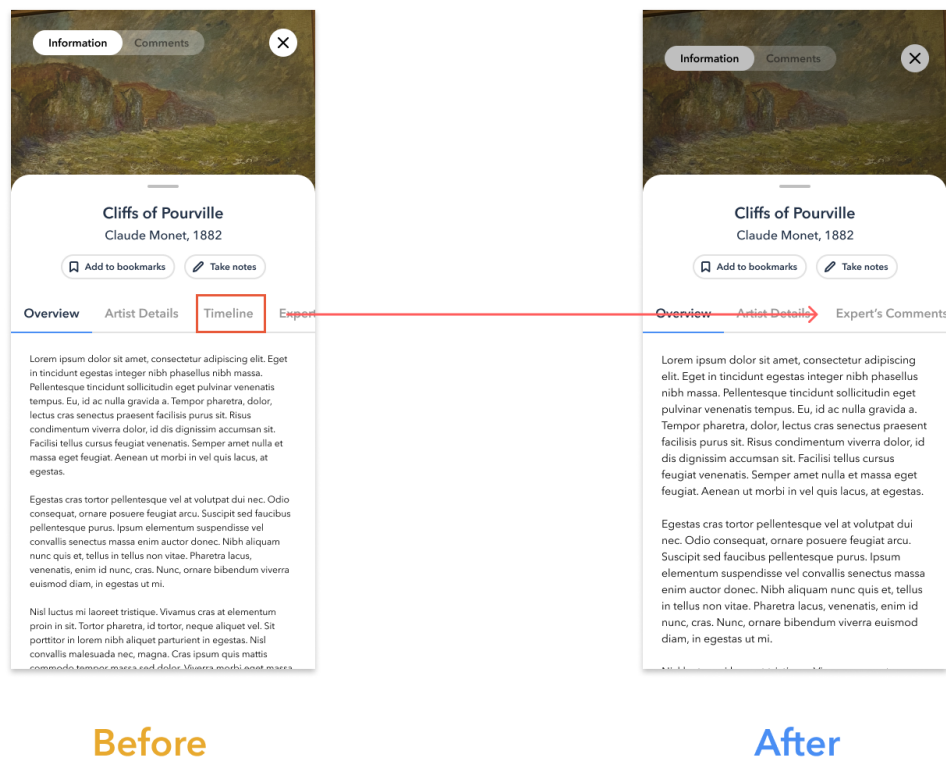
In our previous idea, we wanted to include the "Timeline" section which would provide the related background time information of the artwork, for example, when the artwork was created, what happened during that time, what period/art style was dominant during that time. However, as we were doing our first expert-based testing,

our participant **found the word "timeline" to be confusing** and in our later discussion, we found that the information provided by the "timeline" section could fit well into other sections. Therefore, we decided to take off the "timeline" section for less confusion and more clear information structure.

Change 2: Display all comments and make the reactions static

Before   **After**
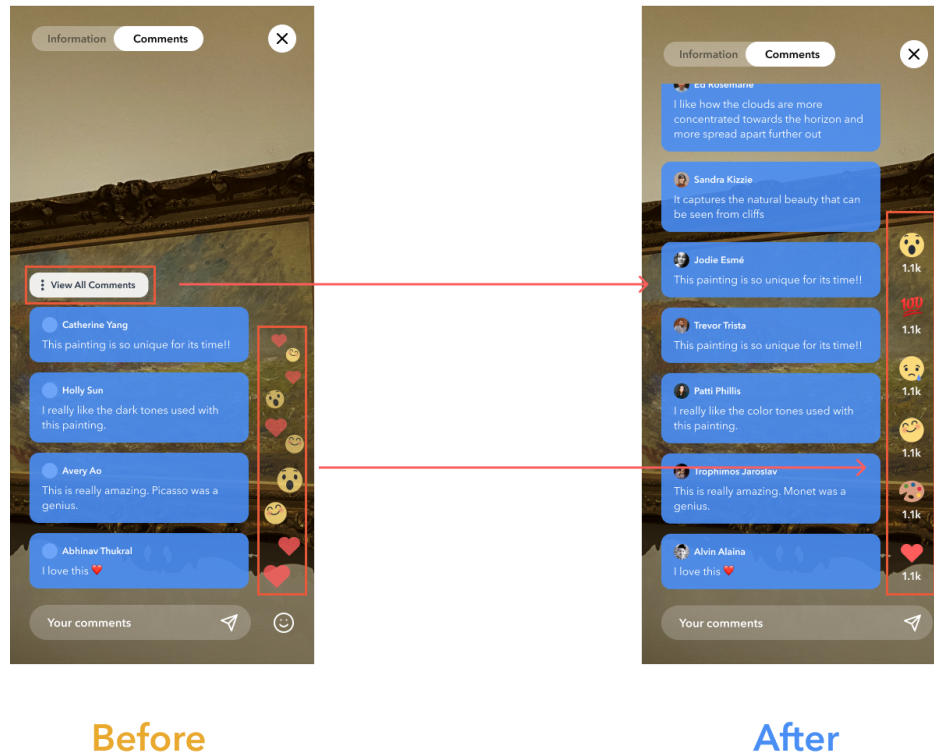
During our first expert-based testing, our participant pointed out that **the system is lacking user control** in the way that users could not return to the original folded comments page after they tap on view all comments. In this case, we removed the extra page with the folded comments and instead present users with all the comments as a feed that allows them to explore in a higher level of freedom. Furthermore, the reactions' animation was also confusing to our participants, so we decided to change it to a static status.

Change 3: Provide more visibility feedback

Before   **After**
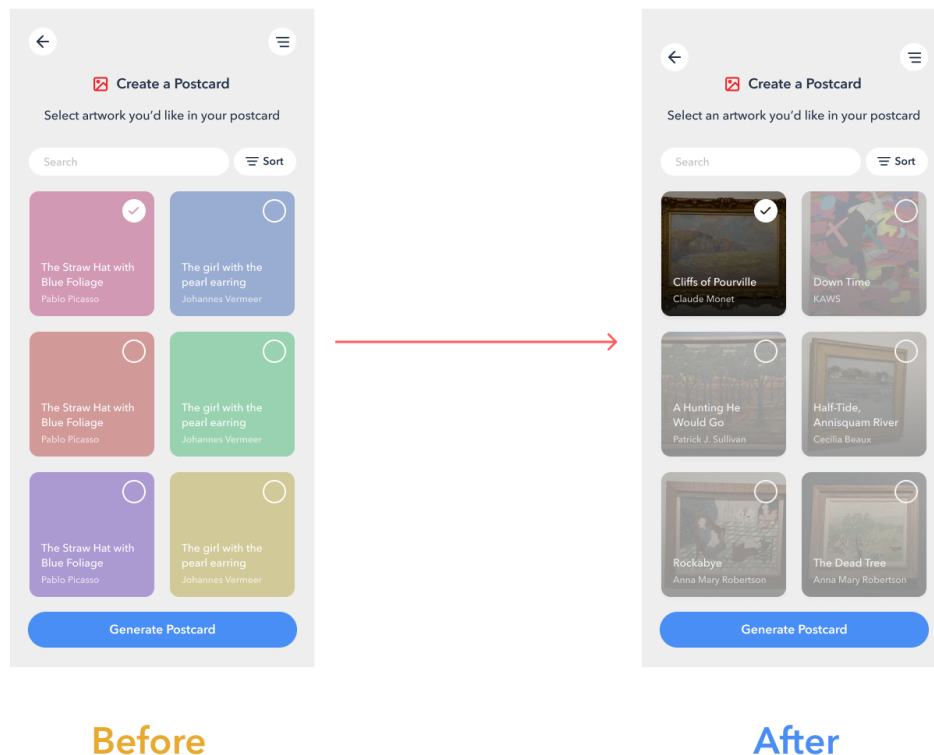
Our first expert also helped us identify **the lack of visibility of the status** within the system where when users choose the postcard, the system did not "tell" them that they could not choose others anymore. To provide more feedback to increase the usability, we updated the prototype to grey out all the unselected postcards when users already choose one.

Change 4: Remove the whole diary function

Before **After**
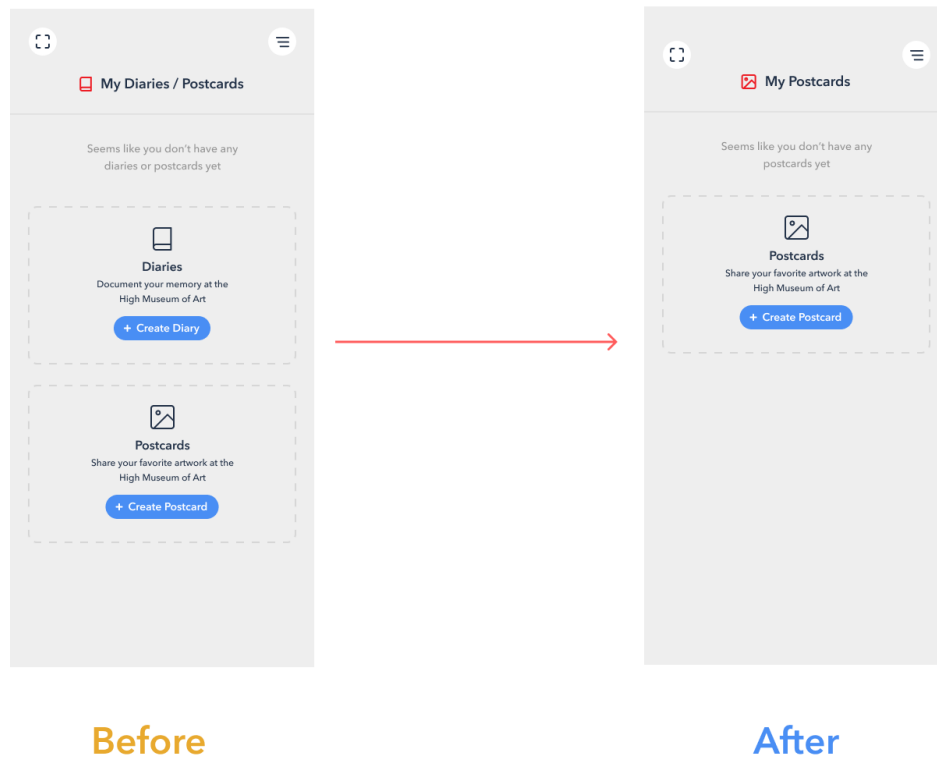
The "diary" function, which was not preferred by lots of participants since our feedback sessions, has already been a problem. We did not take it off in the R3 stage because we thought our participants consisted of only the "social" and "education" personas but no "documentation" persona for which the "diary" function was designed. This time, our first expert gave their feedback from the perspective of the "documentation" persona. However, the feedback did not appear to be very positive and consider the "diary" function as indispensable. Furthermore, the "diary" function is causing confusion. Therefore, after our discussion, we decided to take off the "diary" function based on these primary considerations: 1) It is not needed by the majority of participants. 2) Our design was for increasing engagement in the museum while the "diary" function helps more in the post-visit scenarios. 3) We still could not reach an agreement on what content the diary should contain and to which degree of freedom the users could have on editing the content.

Change 5: Add the "Recently Viewed" section

**Before**   **After**
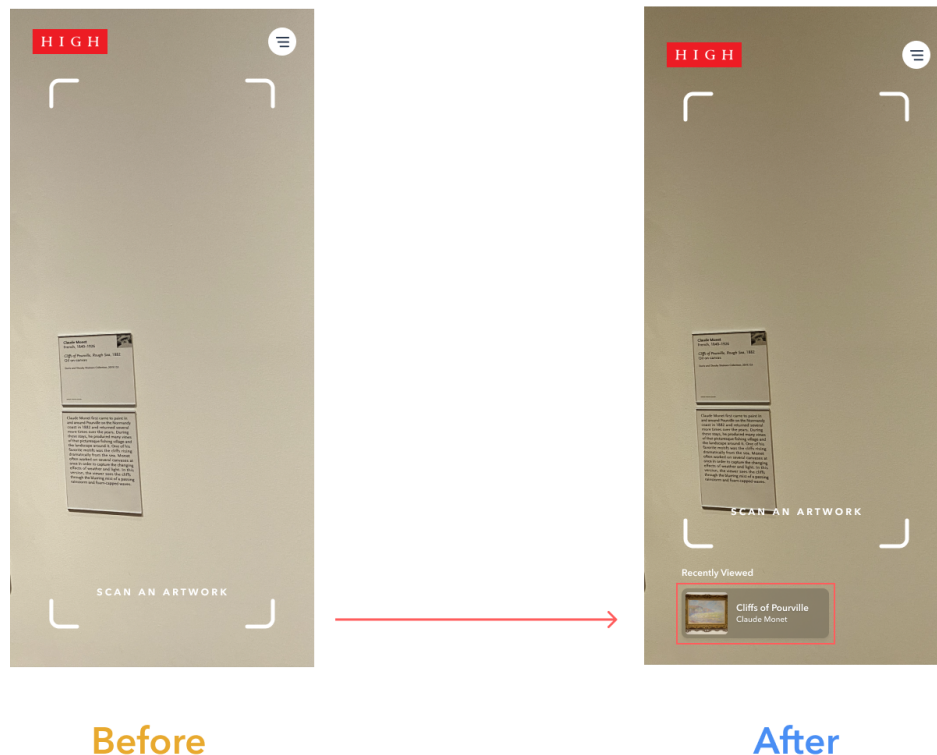
In all of our expert's testing, we realized that we **did not do enough job on the heuristic that helps users recover from errors**, which in this case refers to bringing users back to their previously viewed page when they accidentally close it (which happened during our testing). Therefore, we added an additional section that contains the card to bring users back to their last viewed artwork.

Change 6: Adding onboarding for clarification

Before   **After**
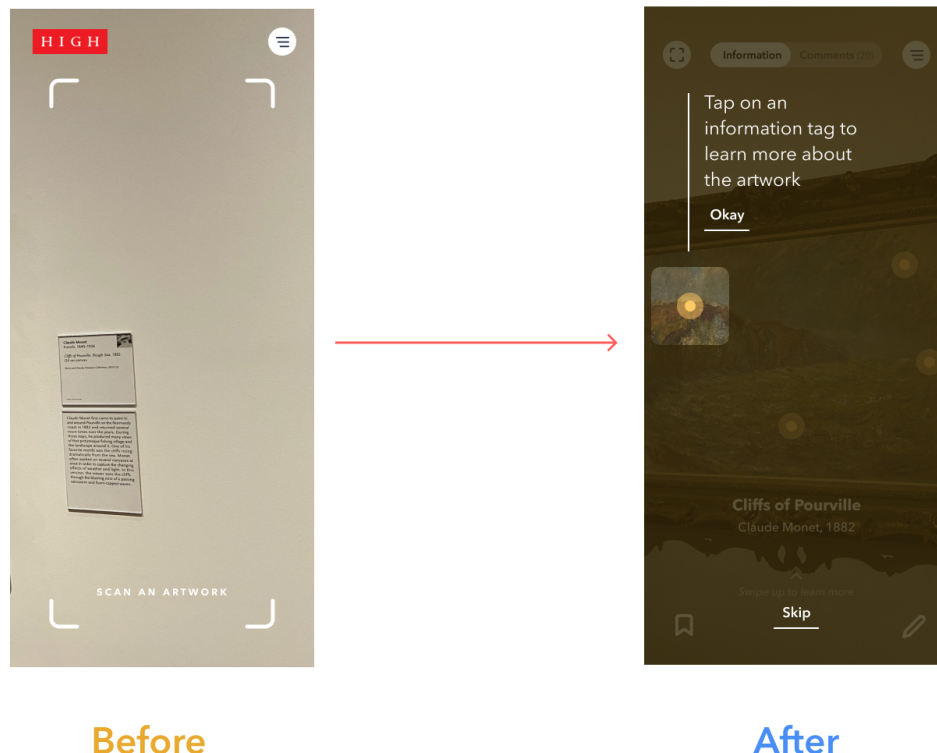
One of the biggest issues we discovered during both the expert's testing and user's testing is that the participants failed to notice that the information tags are tappable and have information folded in them, so the system **still needs more visibility of the status**. There was also other minor confusion disturbing the participants' exploration of our design. Therefore, we accepted the recommendation from our participants to add an onboarding before the users enter the app to help the users get to know the system and some features that they may not be familiar with or aware of.

Change 7: Highlight Info-tags

Before   **After**
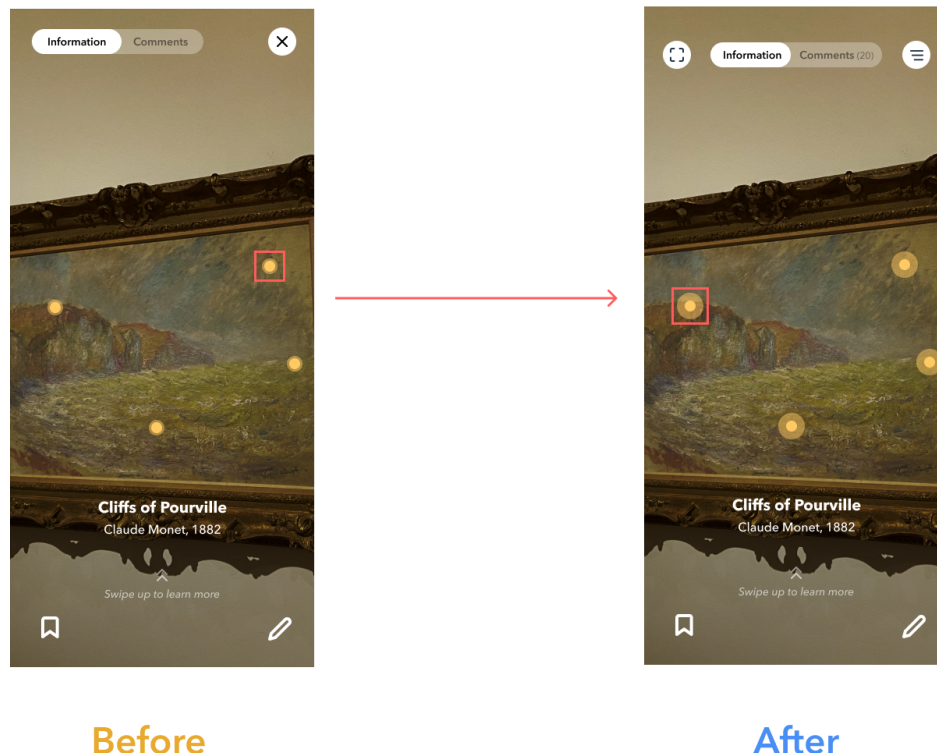
We also identified another **system status' visibility** issue from the feedback of our participants, which caused the majority of our participants' failure to have a try on the information tags. The previous visualization of the tags is not expressing enough information. For example, one of the feedback was the participant thought the "small yellow dots" was just the detection spot. In this way, we thought it was necessary to change how the information tags look and to give more information to the users of their affordance. We changed the tags by first increasing their size and the size difference between the inner circle and the outer circle, resembling more the conventional design of a tappable button. We also intended to add animation on the button to lead the users to tap on it and deactivate the animation once the user finishes exploring it.

Change 8: Improve the related artwork

Before   **After**
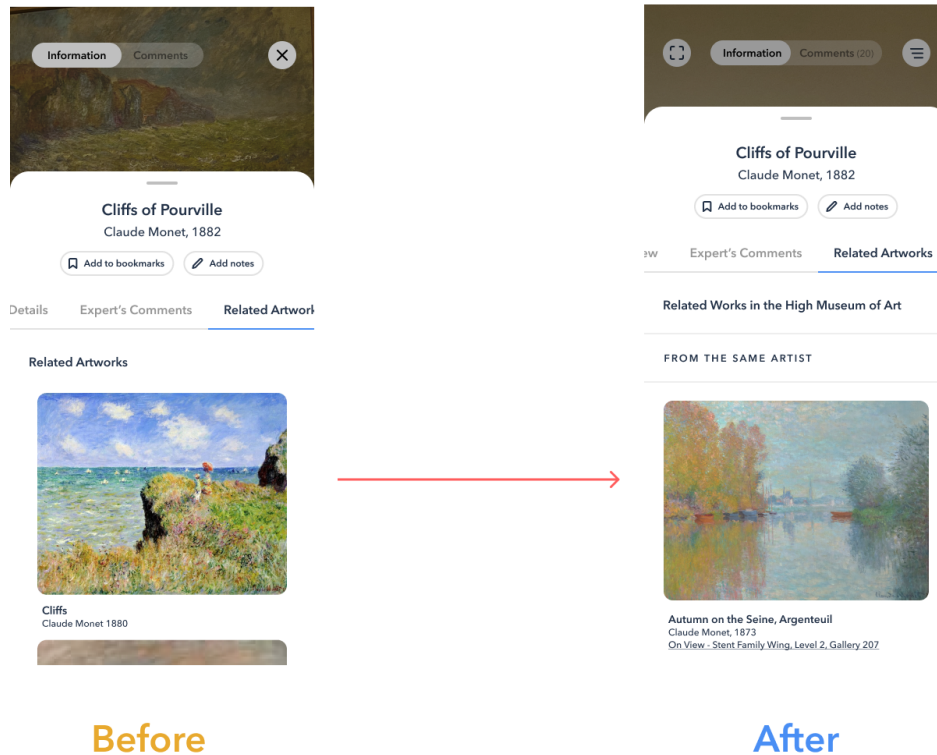
We have also received a cluster of feedback about how to improve the related artwork section we have in the information about the artwork. Previously we **did not have a very clear logic about how the related pieces of artwork would be presented or a definition of "related"**. After asking our participants' interpretation and feedback about this section, we first decided to limit the pieces presented in this section to those that are in the High museum currently because in this way the users could easily go and find them if they are interested. Secondly, we also categorize the artwork based on criteria that, for example, they are from the same artist, they are in the same style, they are having the same elements, etc. In this way, the "related artwork" section now looks more reasonable and organized, reducing questions and confusion.

Change 9: Make the menu button global

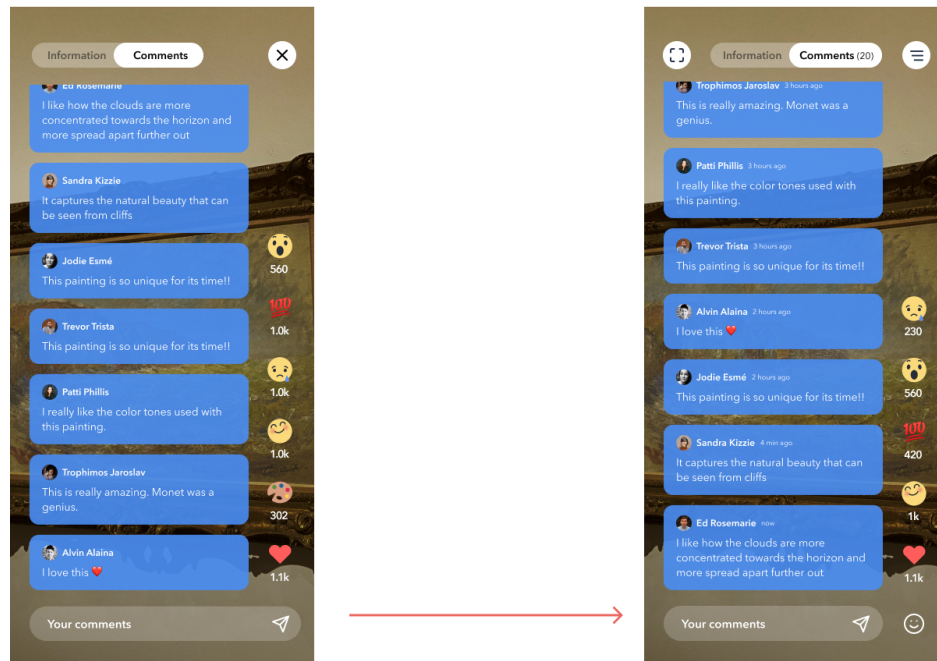**Before**   **After**
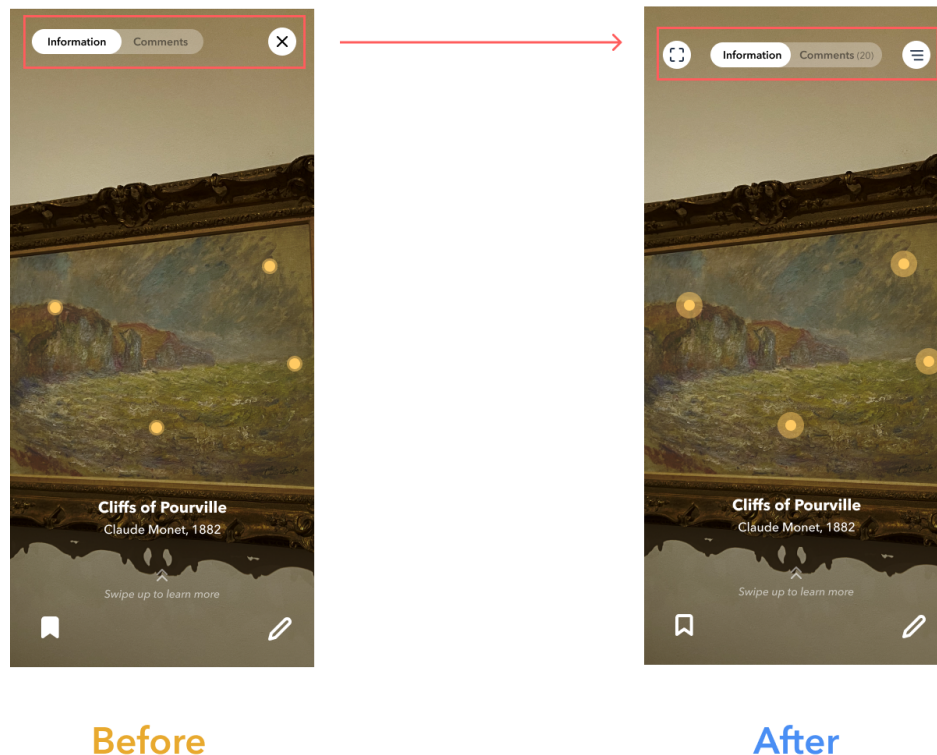
We also got repetitive needs for having the menu button as a global button which would appear on all the pages. We previously did not do it because we would want to let the users stay on the information page to do the exploration. However, we found that not showing the menu button on the page that the users spend the most time on would lead the users to **ignore the menu button or even be unaware of its existence**, which then brings barriers to the users when they try to find features in the menu page, for example, their bookmarks. Therefore, we decided to have the menu button as a global button and we also added the scanning button to allow users also be able to get back to the scanning page in just one tap no matter which page they are on.

Change 10: Saving the final postcard

Before   **After**
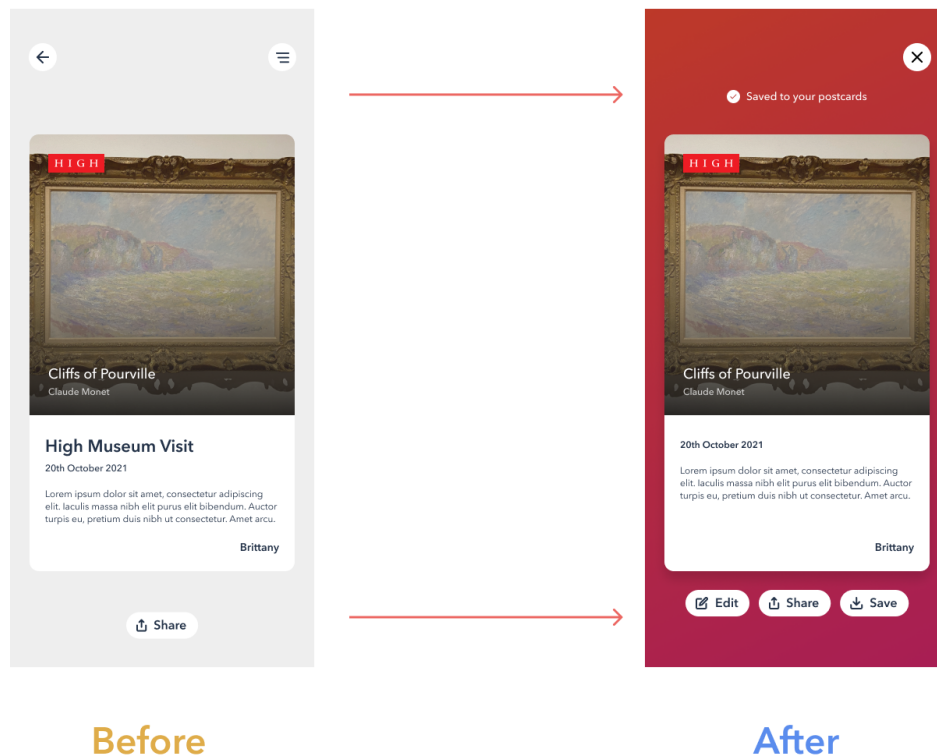
There were several points of confusion among participants in comprehending the final screen for the postcards task. Firstly, users were uncertain about the purpose of the final screen since they expected that they will be able to edit the postcard in this view. Secondly, the users were confused about where the postcard was being saved. Lastly, finding the "Save to Photos" option within the share button was difficult for most users. To obviate the purpose of this page, we made UI changes such as the background color and minimizing navigation options, highlighting the final presentation-view for the postcard. We also added a smooth move-in from the bottom pop-up animation to further enhance the view only and share the effect. We also added a small caption on the top indicating that the postcard has been saved within the app's postcards section. Finally, we added buttons for "Edit" and "Save" along with the "Share" button for more obvious user control in this view.

Change 11: Distinction between emoji reactions and comments

Before   **After**

Post our design change that incorporated static emoji reactions, we realized **users' confusion in distinguishing between the comment box and the emoji reactions**. With the comment box extended to the full width of the screen, as well as the keyboard containing the option to type emojis, users were uncertain whether their typed emoji in the comment box would reflect as their reaction. To distinguish between the two, we made a minor modification to the comment box, aligning it with the comments. We then added an emoji button for the interaction with the reactions which helped users visually understand the difference between the emojis that are part of the comments and the ones expressed as reactions.

Change 12: Saving information in notes

**Before**   **After**
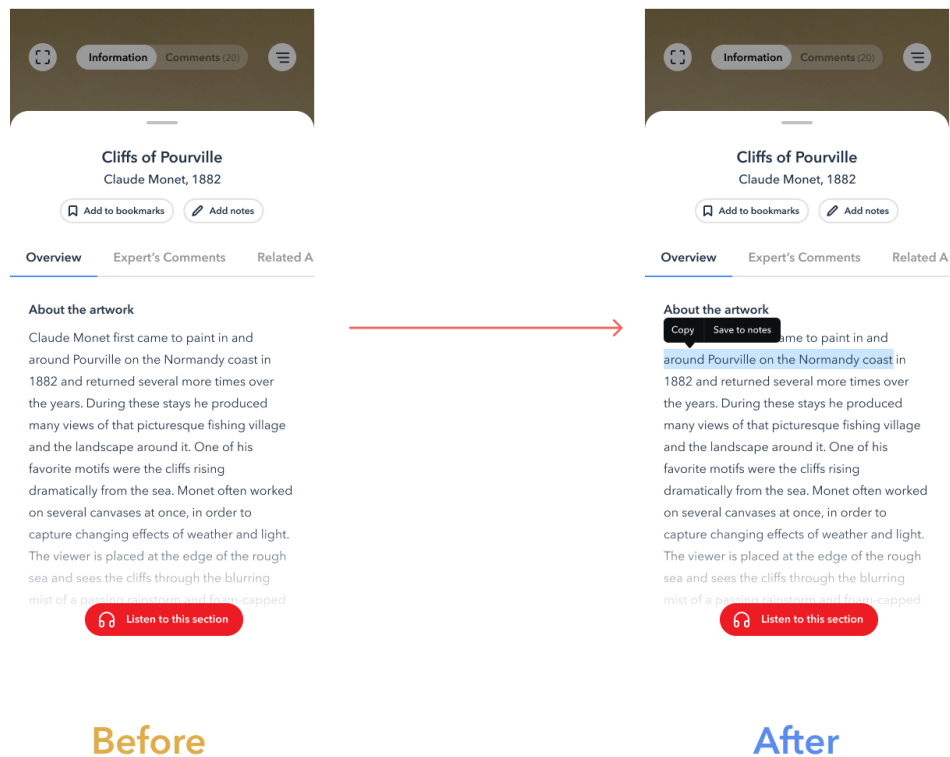
During the user tests, several participants suggested that they would want a simpler way of adding information to their notes. In order to save text from the information section to the notes section, the user would first select the text they wanted to store, copy the text, close the information tab, open the notes section, and finally paste the text within notes. To simplify these steps, we added tooltips as a part of text selection in the information section of the artwork. The tooltip contains a button called "Save to notes" which would act as a shortcut to perform the same action of copy and pasting the information in the notes section with a single tap of a button. This helped improve the design system's **flexibility and efficiency of use**.

Issues We Still Need to Work on

We've discussed changes we've made based on the feedback during the testing, and we will also present other issues that we discovered but could not address in our current stage of the prototype. We will talk about the issues, provide the evidence, and list some possible design solutions.

Issue 1: Sharing bookmarks, notes, and postcards

Our participants expressed that they would also like to share their bookmarks and notes in addition to the postcards. Some participants mentioned that would prefer sharing their bookmarks and notes over the postcards to share information about some of the more interesting artworks they saw at the
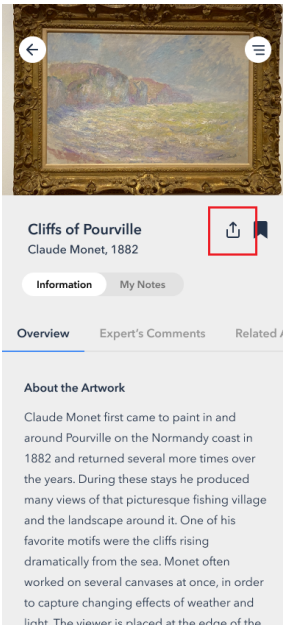
museum. In the current prototype design, sharing details about the artwork is represented with a single share button on the artwork's page. There is still some confusion regarding how the share button behaves when the user is on the "Information" tab v/s the "My Notes" tab. We would like to discover the best way to deal with issues have not been addressed it yet. A possible design solution is using a pop-up sheet with options selected between information and notes along with sharing options for different social media or messaging platforms. Alternatively, we could reposition the share button within the tabs section to differentiate their function as the user switches between the two tabs.

Issue 2: Categorizing information about the artwork
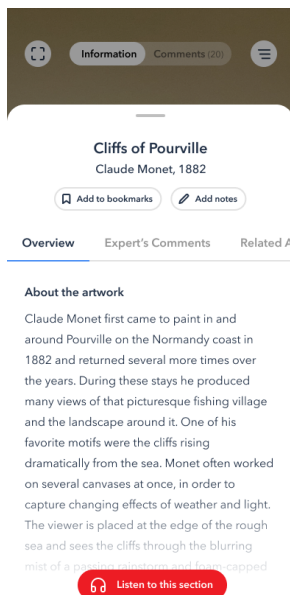
The current design uses a text and audio-based approach in presenting information to users. During our user testing, we observed that reading through paragraphs of information about the artwork and the artist requires effort from the user in terms of parsing and finding relevant details. Visually, we discovered a need of summarizing this information into quick glanceable categories for the user. Some of the categories could be time-period or era, art style, movements, etc. Identifying appropriate categories and information summaries could not be addressed with the current design, however, this issue presents interesting design solutions for us to explore. A possible solution could be to present these categories within the main view of the scanned artwork along with its title and artist. We could also simply incorporate these categories within the overview section with the detailed information presented below.

Issue 3: Museum Map

While our problem statement focussed on improving visitor engagement with the artworks at the High Museum of Art, participants of the user tests mentioned
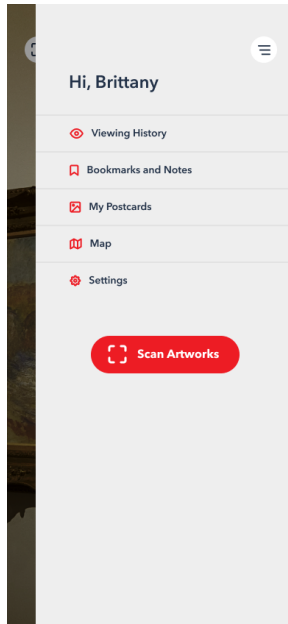
that they would need the museum's map within the app to better navigate through the museum. A tab for the map is available on our navigation panel, but with the current designs, we are yet to include an interactive version for the users to browse through it. The map feature becomes even more significant as we aim to connect it with our related artworks section. Our proposed solution is to include a view with zoomed-in floor plan layouts that the users can browse through to locate galleries and artworks. We also consider embedding a miniature version of this map within the related artworks section to help users better explore different artworks in the museum.

Issue 4: Moderation of comments
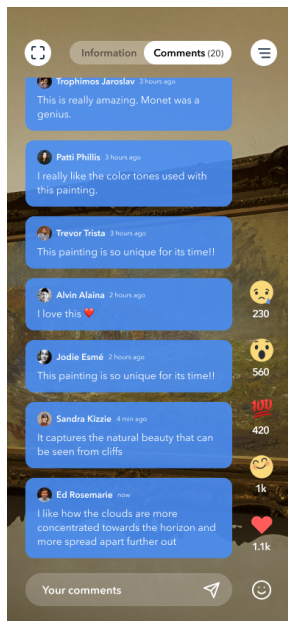
Though the comments section is designed as a medium to express visitors' thoughts and reactions, public comments on any digital platform require moderation. This concern was also raised during our expert evaluation and we were unsure how to express moderation as part of our design. Though not explicitly mentioned in our current design, we assume that these comments will have profanity/sentiment checks to ensure that profane or obscene comments are not shared with visitors. Additionally, we want to allow users to flag or report comments that seem to disrupt their viewing/reading experience.

Lessons Learned

We have learned a lot regarding how to better conduct feedback sessions and work in a group. There are many things that we could have done better. We will surely keep all of them in mind in future feedback sessions and group collaborations.

Firstly, we need to better polish the tasks before assigning them to participants. In the expert-based evaluation, we came to learn that some of the tasks are not as effective. As a result, we changed some tasks after the first and second expert-based evaluation sessions. This causes problems in our data since participants are given different tasks and they evaluate the system based on these different tasks. The result is thus not entirely standardized and leads to problems when analyzing the findings across different sessions.

In addition, we also realized that it is essential to check for the likeliness of use after users are done with the think-aloud during the user-based evaluations. For example, one may complete the task, point out some problems, and share recommendations. Everything seems pretty good, but when we ask them "will you use this function," the answer is a firm no. Having users complete the task and share their thoughts does not mean that users are willing to use certain functions in real life. It is thus very important to check the likeliness of use and the reason behind their answers.

We also noticed that we should keep a better record regarding changes to the prototype, especially in the case of group collaboration. We started building a checklist of changes to make halfway through the feedback sessions when we come to realize that we need to maintain a record of what to improve on after each feedback session. We met on occasions when two group members are fixing one problem separately because they did not realize others are also working on the same issue. It would thus be beneficial to keep a record of changes to make and have group members update the checklist to keep each other notified about any progress.