gale: smart assistant for habit formation

Travis Chen - Timothy Lau - Rebecca Wong - Seth Pendergrass February 28, 2016

1 Roles

Travis	Artist, Digital Mockup Creator, Usability Tester (computer, recorder)
Timothy	Digital Mockup and Report Creator, Usability Tester (facilitator, recorder)
Rebecca	Artist, Usability Tester (facilitator, recorder, gale voice), Project Manager
Seth	Artist, Usability Tester (computer, gale voice), Editor

2 Problem and Solution Overview

There are a number of apps on the market that are designed to help a user set and work towards goals, both general and specific (e.g. fitness), but they are far from ideal. People still regularly fail to achieve their goals due to poor planning, lack of experience in goal-setting, forgetting, or just giving up.

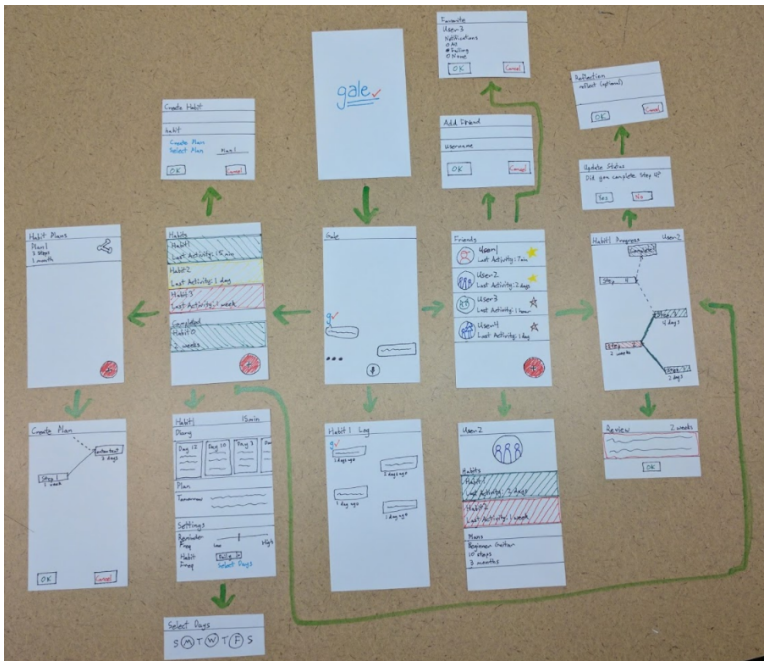
We propose the creation of gale, a next-generation voice assistant designed to make goal-setting easier and more effective than all currently available tools. gale will be a smartwatch and phone app, using a conversational AI to make goal-setting quick and painless for the user.

Additionally, gale will enable social accountability between users by notifying users of their friends' progress and offering a way to give positive encouragement. On the flip side, if a user's friend hasn't sent any updates about their habits recently gale will notify the user of this and encourage them to check in on their friend. One feature that we noticed was missing from many other goal setting apps was a place for people to reflect on the progress of their goals beyond recording raw data—we spent much of our time zeroing in on this issue.

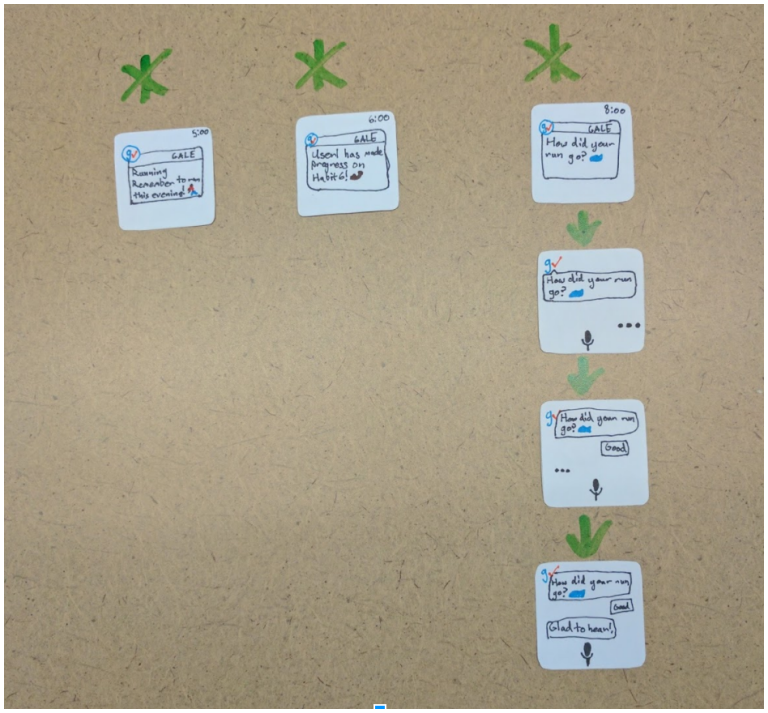
When gale inquires about user progress the AI pushes for deep reflection on their successes and failures. Advances in AI allow gale to talk to users like a friend would, allowing for natural moments of reflection. These factors will make gale the best available tool for a user to make, track and accomplish any type of goal.

3 Initial Paper Prototype

These are the screens for the phone app for gale. The phone app is paired with the watch app to provide richer functionality. Users begin most interactions in the conversational UI of gale seen in the center and navigate through the rest of the app as needed. For the initial prototype, we assumed that the system provided a back button that would allow the user to navigate back to the home screen (e.g. Android).

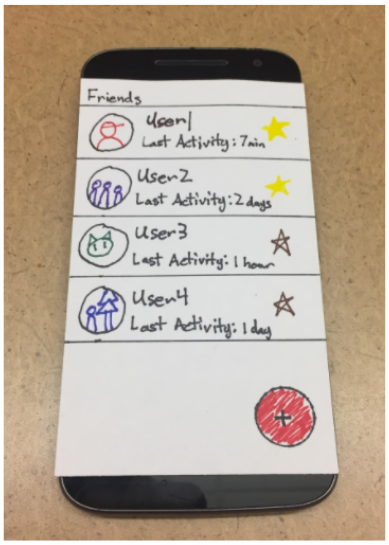


Smartphone application flow

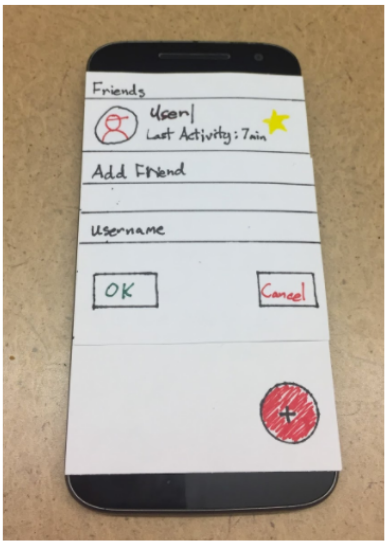


Smartwatch application flow (* indicates pop up notification)

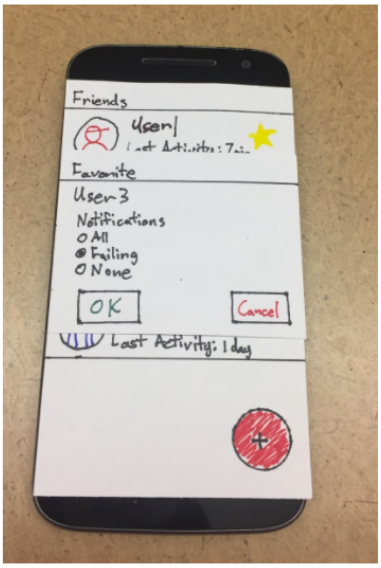
Task 1: Social Accountability



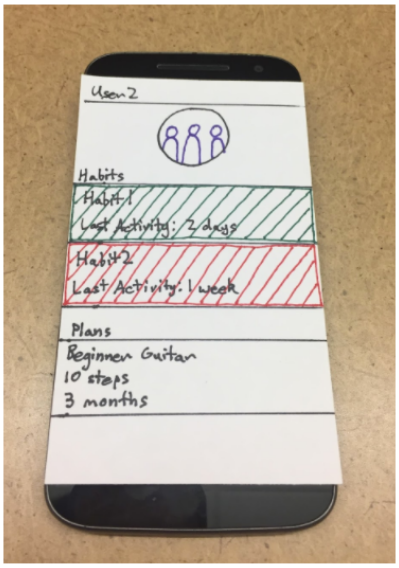
Friends list. Accessed by swiping from chat screen.



Add a friend by inputting their username (plus sign pressed).



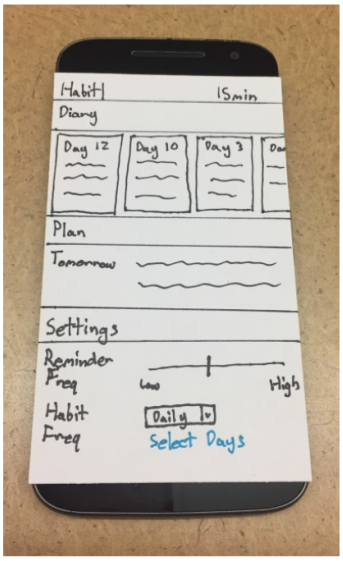
Favouriting a friend in the list (star pressed).



On the friend's profile the user sees their friend's habits and plans.

The wearable provides updates to the user when their friends make progress on their goals. The user can tap on the notification to view their friend's profile.

Task 2: Self Reflection

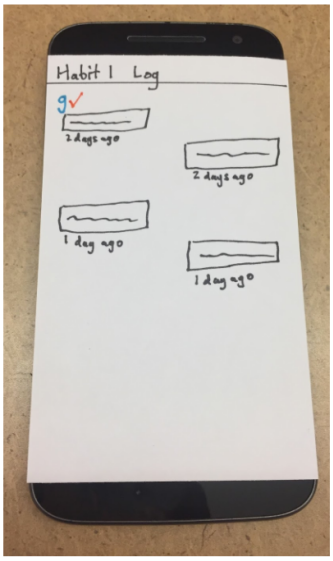


Detailed habit information.

Previous reflections can be accessed through the diary.

Plan shows the next step to be taken by the user.

Settings allows the user to change the frequency at which gale nudges the user and how often they expect their habit to be worked on.



A log of the conversations with gale about the habit, presented in a chat transcript. This will include the reflections made into the Diary.

A reminder from Gale.

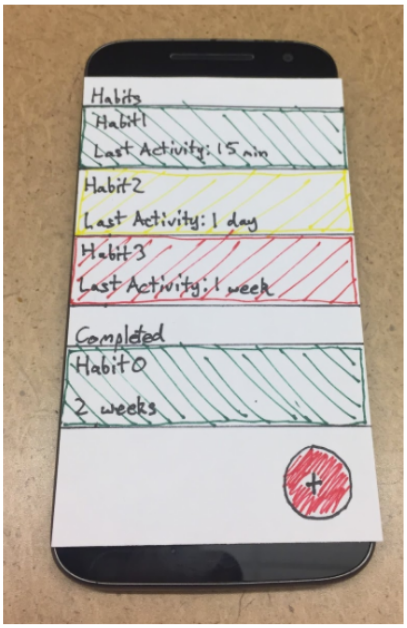
A prompt for reflection. The user can tap on the notification to reflect on their progress.

Inputting a voice note reflection.

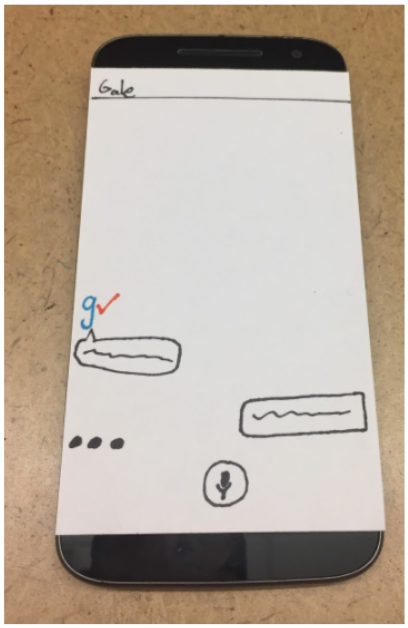
gale determining how to respond.

Feedback from gale. This conversation is stored in the chat log for this habit.

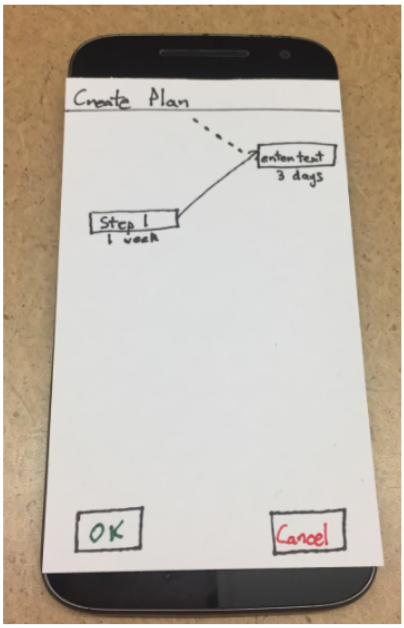
Other Screens



Users can view the list of habits they are currently working on or have completed.



gale's voice interface allows the user to quickly create new habits and log data. All app interactions can be performed from this screen.



Users can create habit-forming plans that influence the timing of gale's notifications.

It is also possible for gale to auto-generate these plans by drawing on internet data, with no manual interaction from the user.

4 Testing Process

Methods and Participants

Heuristic Evaluation 1	21 / CSE / M, 21 / CSE / M
Heuristic Evaluation 2	20 / Informatics / F
Usability Test 1	20 / Math / F
Usability Test 2	21 / BioE / M
Usability Test 3	28 / Marine Bio / F

The first participant was selected because she owns a Apple watch and therefore is familiar with using a wearable. We conducted the testing in Allen library, because we were able to reserve a semi private booth where we wouldn't be interrupted. We had our participant speak aloud all her thoughts about using the app/wearable, whether it was her next step in using the app or confusion about what certain buttons meant. Seth played the role of "computer" by manipulating the prototype, while Rebecca was the facilitator who prompted the tasks that were to be achieved. Both helped record the participants critique and feedback.

Similarly, our second usability tester was chosen because he uses his Apple Watch daily. Rebecca facilitated the study and explained what tasks needed to be achieved. Travis performed the role of "computer" and manipulated the prototype. Seth provided the voice of gale and simulated gale's conversations with the participant. Tim took notes on the participant's behavior and suggestions. Rebecca introduced him to the idea of gale, and then prompted him through our two tasks using the Test Plan shown in the Appendix.

We chose our third participant because we had previously interviewed her for our design research, so she was familiar with the design problem that we were trying to solve. We chose to test at a UW coffee shop because it was relatively quiet and gave us a good work space. Tim facilitated the study and explained what tasks needed to be achieved. Seth performed the role of "computer" and manipulated the prototype. Rebecca provided the voice of gale and simulated gale's conversations with the participant. Travis took notes on the participant's behavior and suggestions. After each task, we asked the participant to give her feedback and suggestions.

Process Refinement

Our experience with usability tests gave us some insight on how to make the process effective and efficient and the preparation needed on our part. We needed to have a list of all the tasks and features we wanted the participant to interact with so we were not figuring things out on the fly. Everyone in the group needed to be on the same page on app functionality. We found it best to have four test administrators instead of two; this way one person could take notes, one person could be the voice of gale, one person could be the computer and the last person could be the facilitator.

During the actual test, it was important that we kept the pace of the testing slow so that we could ask thorough follow up questions and have time to write all the feedback down. Time budgeting was another consideration; our first usability test took an hour and a half, so we knew we had to schedule enough time for subsequent tests. We tried to choose participants for our usability tests who represented the average user, likely less technically inclined than a Computer Science or Informatics major. Overall our approach to testing was effective but improved as we smoothed out our process.

5 Testing Results

First Iteration (Initial Prototype)

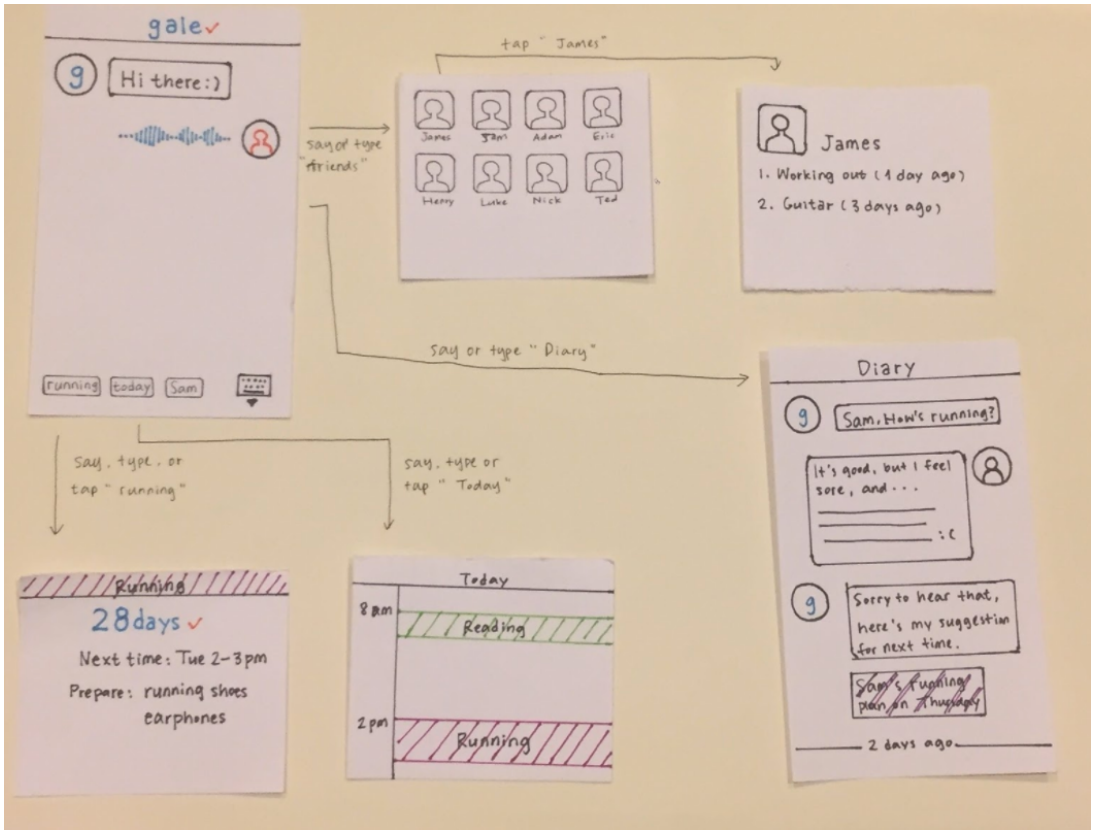
Heuristic Evaluations 1 & 2

- When presented with a prompt for reflection on the smartwatch, our evaluators suggested the use of a "snooze" function that would allow the user to respond to the prompt at a later time.
- Our evaluators suggested that receiving notifications about a friend and "favoriting" a friend should be separated. It was not immediately clear to our evaluators that "favoriting" a friend meant that that friend's progress would be easy to access.
- In the process of "favoriting" a friend, we presented users with different options for the types of notifications that they wanted to receive about their friends. This wasn't clear to our evaluators.
- We were recommended to allow the creation of new diary entries, and support voice dictation.
- We need to add buttons or a long-press menu to allow editing and deleting habits.
- We also might consider a "priority" habit, which stays at the top of the habit list akin to the favorited friends.
- We need to clarify how navigation is performed in our prototype.

Usability Test 1

- Navigating our app is too confusing and difficult because it feels like using an Excel spreadsheet.
- Voice interaction needs to be the focus instead of entering information by typing into a keyboard.
- gale should prompt more conversation during reflection, so as to avoid short and surface level answers. The amount that gale pushes for more reflection could be varied as a setting.
- Participants have no way of knowing how they can interact with gale and it was suggested that there should be some sort of help menu or conversation.
- Our participant felt uncomfortable having notifications of friends progress without a feature that let them communicate with that friend.
- Wanted the app to be more visually interesting (i.e. fun).

Changes



Revised Smartphone App Flow

In our inspections and first usability test, users found the mix of button and voice interfaces confusing. It was not immediately clear which commands were accomplished through which interface. We decided to refocus on the conversational aspects of gale. The home screen of

the app is a conversation thread with gale, where the user can ask gale to access functions of the app. gale supports voice or keyboard interaction, and presents requested information in card format in the conversation thread.

By swiping from the home screen, the user can enter into a conversation thread used specifically for journaling progress and reflections. Optionally, gale can be set to interact with the user through the journal entries. gale can prompt the user for their thoughts, and also can scan the user's text and provide smart suggestions and feedback.

From the main home screen, the user can type or speak the keyword "friends", at which point gale will open a menu with all of the user's friends. The user can tap on any of their friends to open up a profile that shows that friends' current status on their goals.

Second Iteration

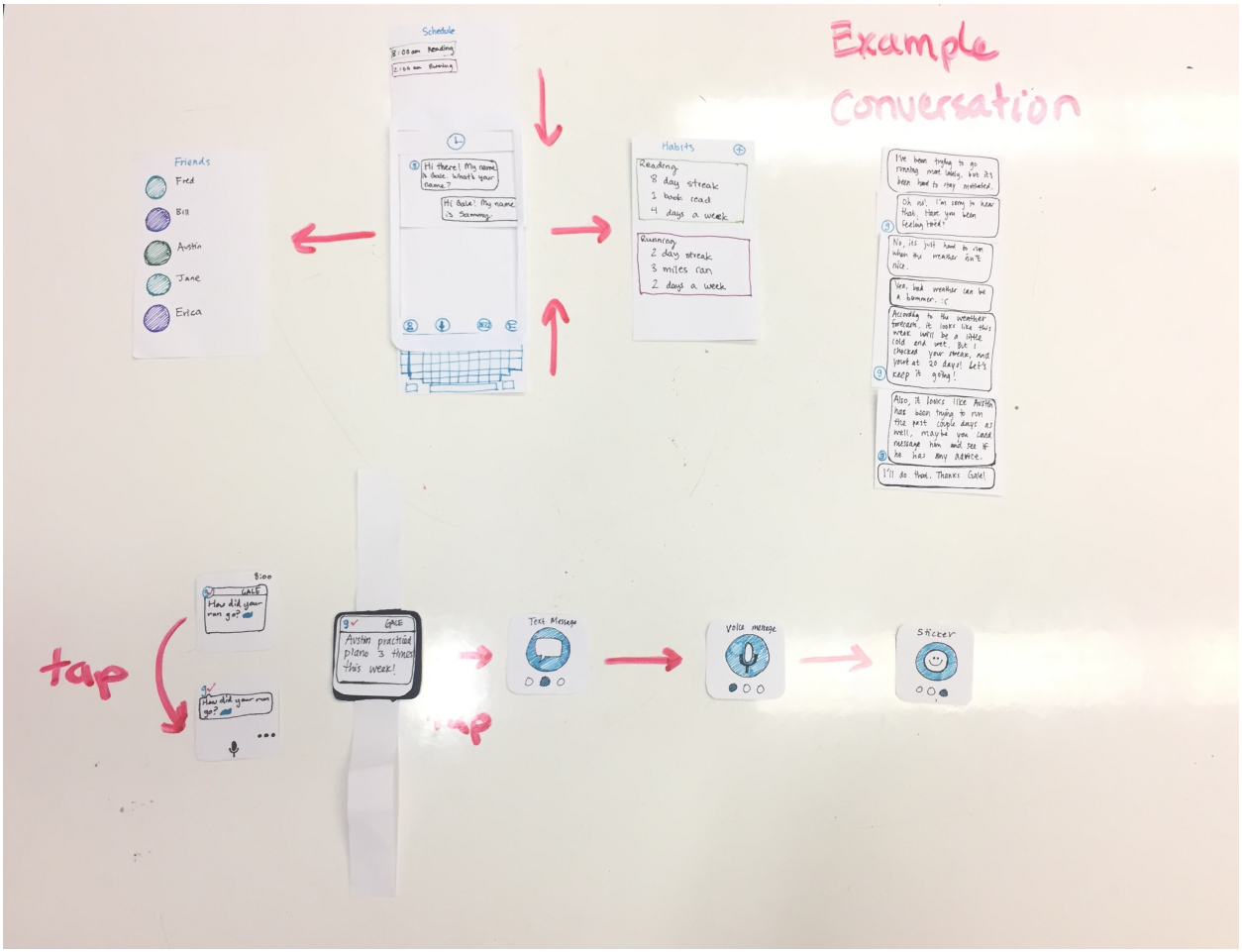
Usability Tests 2 & 3

- When the participant was asked to reflect on their run they felt that gale's responses were too passive. They thought that gale should be encouraging and pushing them to overcome whatever they found difficult instead of just acknowledging the issue. In response to this, we made gale have more active responses. If someone was having an issue, gale would try to offer advice or at least be more encouraging or recommend they stick with the habit because they had an X day long "streak" of continuing their habit.
- The participant also felt that gale sounded too much like a robot, which was what we were trying to avoid. For the next iteration of our paper prototype we tried to make gale more conversational and interact with the user of the app like how a friend would. Making gale push for deeper reflection mentioned in the above point helped make gale seem more human as well.
- When our participant pulled up the schedule of her habits for the day she suggested that the schedule also include other things she had to do that day that were in her google calendar. She thought that would help her so she doesn't schedule her habits when other events are happening. We decided to ask users when they first open the app if they wanted to sync their existing digital calendar with gale. This allows gale to not double book habits with other activities. We also provide a visual on the schedule screen of what times the person is already busy.
- In response to the smartwatch notification of a friend's progress our participant suggested that there should be an option to send immediate encouragement possibly in the form of an emoji. We changed the prototype so that users can interact with a notification about friend's progress by either sending an emoji, a voice message or a text message.
- When asked how to get back to the main chat screen our participant wanted to tap on the gale logo at the top of the screen since she was used to clicking on logos to go back to home screens. While we didn't implement this exactly it showed us that people probably wanted some more conventional ways to navigate our app instead of purely relying on

talking to gale. As a result we chose to make the friends list, daily schedule of habits and list of that user's habits accessible by swiping different directions from the home screen.

- She also made the point that some people might want to create too many habits so it would be good if gale notified that person that they should consider focusing on fewer habits. We chose to incorporate this feature so if the total time a week a person is dedicating to all their current habits is high, then when that person goes to create another habit gale would warn them that they should consider not making more habits.

6 Final Paper Prototype



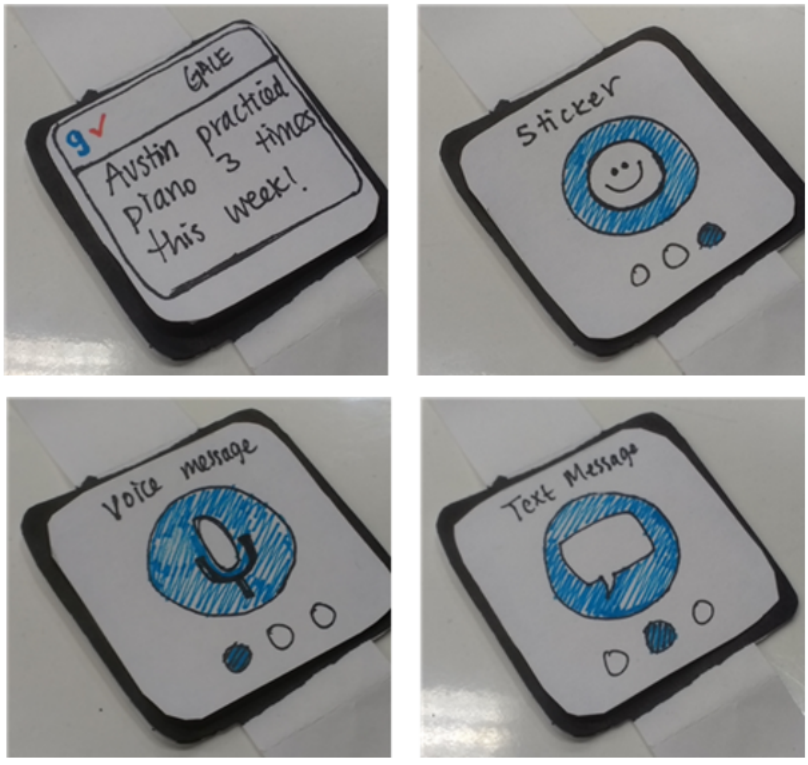
Final Smartphone and Smartwatch App Flows

For our third iteration of the paper prototype, we fixed major navigational issues and found a happy medium in balancing conversational aspects and intuitive tap/swipe functionality. We also focused on ensuring that gale's conversational capabilities were more visible in the prototype. Finally, we refined our views on the most important aspects of our design, with a greater emphasis on keeping people on track with social accountability.

After our first usability test, we realized that our app did not rely heavily enough on the conversational interface. We decided to redesign our app prior to the next two usability tests to focus almost all interactions on keeping things conversational. This shift is very well reflected in our third iteration of the paper prototypes, which showcase the chat aspect heavily.

Our third usability tester found using the app only by voice challenging; she would attempt to use tapping and swiping as ways to navigate. This showed us that we needed to support users who were not yet comfortable with conversational UIs. We decided to separate out the few key non-conversation screens (schedule, friends and habits) into pages that the user can swipe to akin to Snapchat. While this reverses some of the changes from our first revision (after the first usability test), we believe that this converges the design to a comfortable medium that should both focus on conversation with gale while making the app easy to learn.

Task 1: Social Accountability

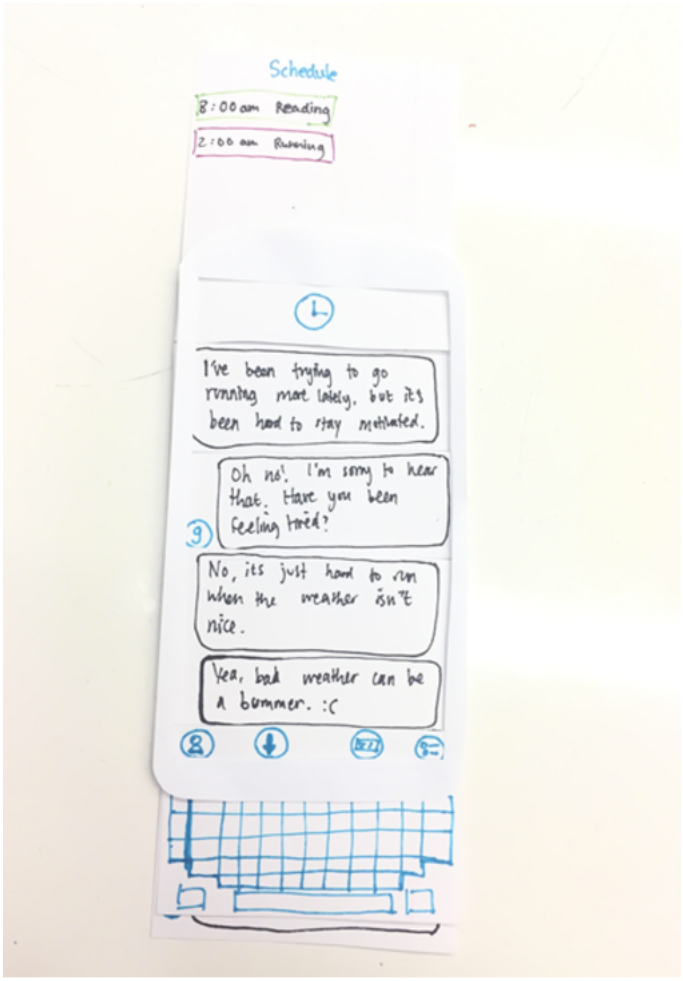


Revised smartwatch app, allowing for quick replies to status notifications

The smartwatch app provides a simple way to receive and respond to notifications from friends. The user receives the notification, which they can either swipe to ignore, or tap on to engage. Tapping will give them the choice of sending their friend a voice message, text message, or a fun sticker. We implemented these features in response to feedback from

usability testing that being notified about a friend with no way to respond was uncomfortable. We anticipate that this redesign will allow for positive feedback from friends to motivate users.

Task 2: Self Reflection

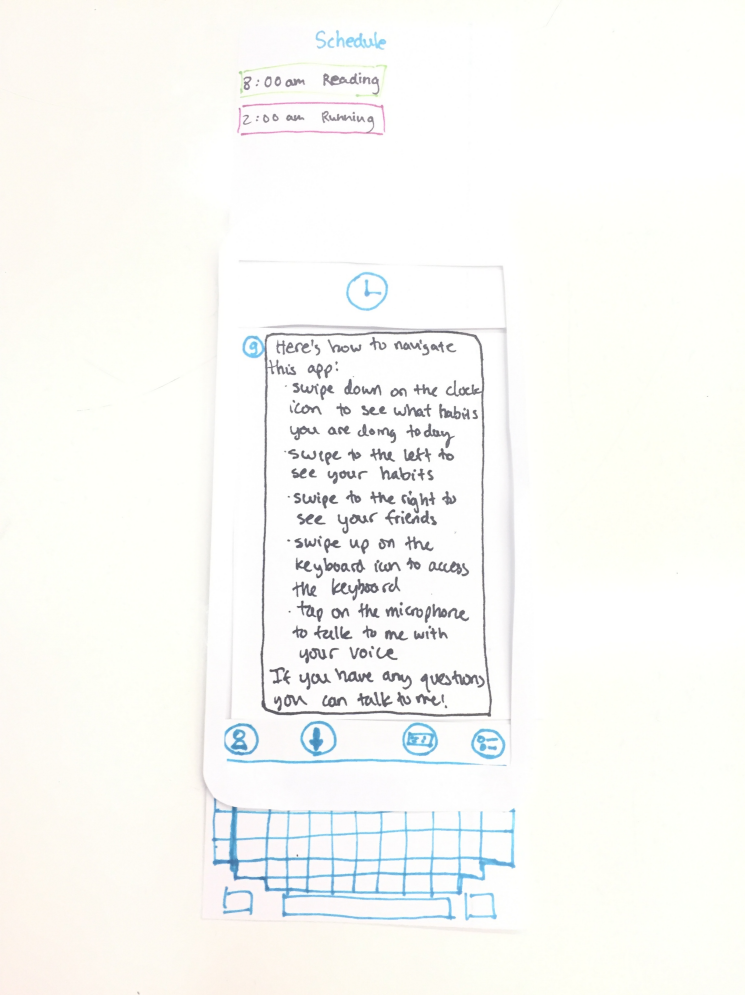


Primary chat interface, showing off screen keyboard and calendar

Reflection is accomplished in the phone app by talking to gale. The user can initiate the conversation, or gale can if she notices the user has not discussed their habits for a while. During the conversation, gale tries to encourage and help the user achieve their habits. Since gale is aware of the user's calendar, she can assist the user in scheduling time for their habits. By having a more natural conversation than in our initial prototype, gale will encourage the user to reflect with more depth and frequency. This should allow the user to better understand their progression.

Additionally, we eliminated the Diary screen, opting instead to have the user ask gale to filter past conversations based on topic.

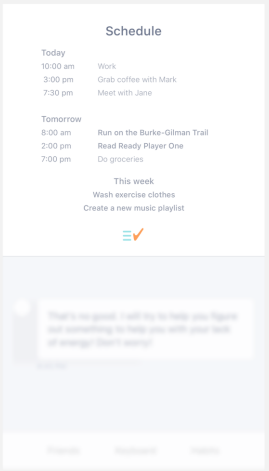
Onboarding

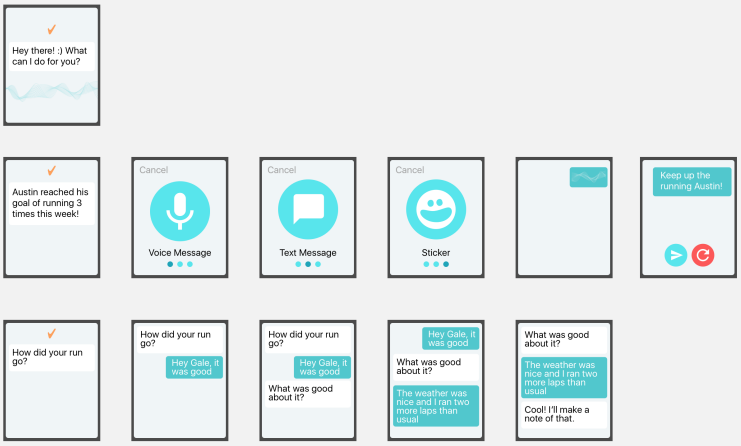


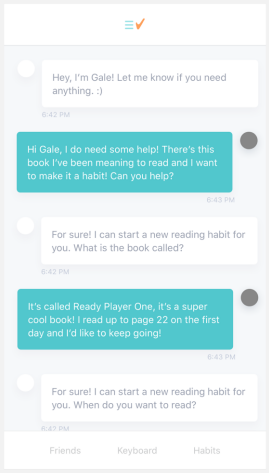
gale onboarding message, which indicates to the user how to use the app

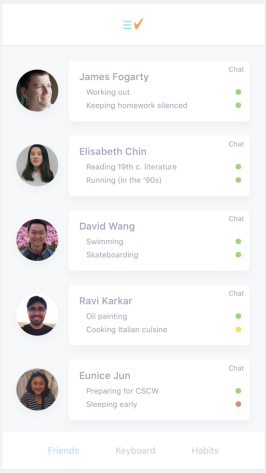
gale now includes introductory messages to acclimate the user to app usage. Our users had difficulty in navigating our app during our usability tests. Thus, we decided to have gale introduce herself and how the app is operated when first opened. This both gives gale a friendlier personality and explains all of the available functionality.

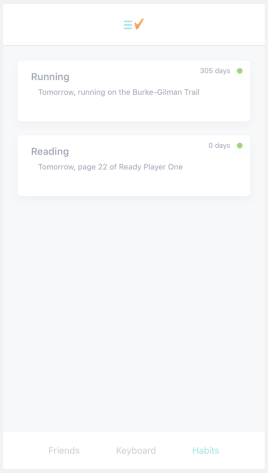
Digital Mockup

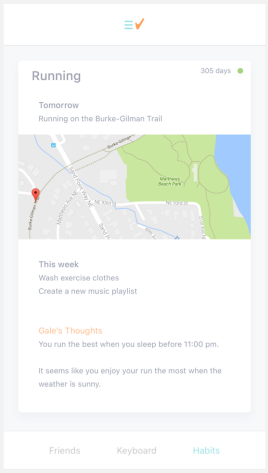


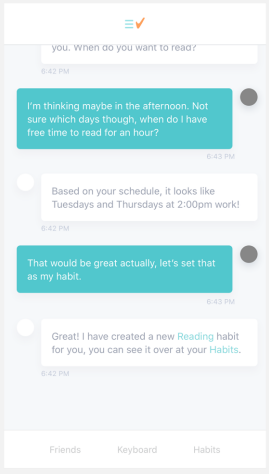






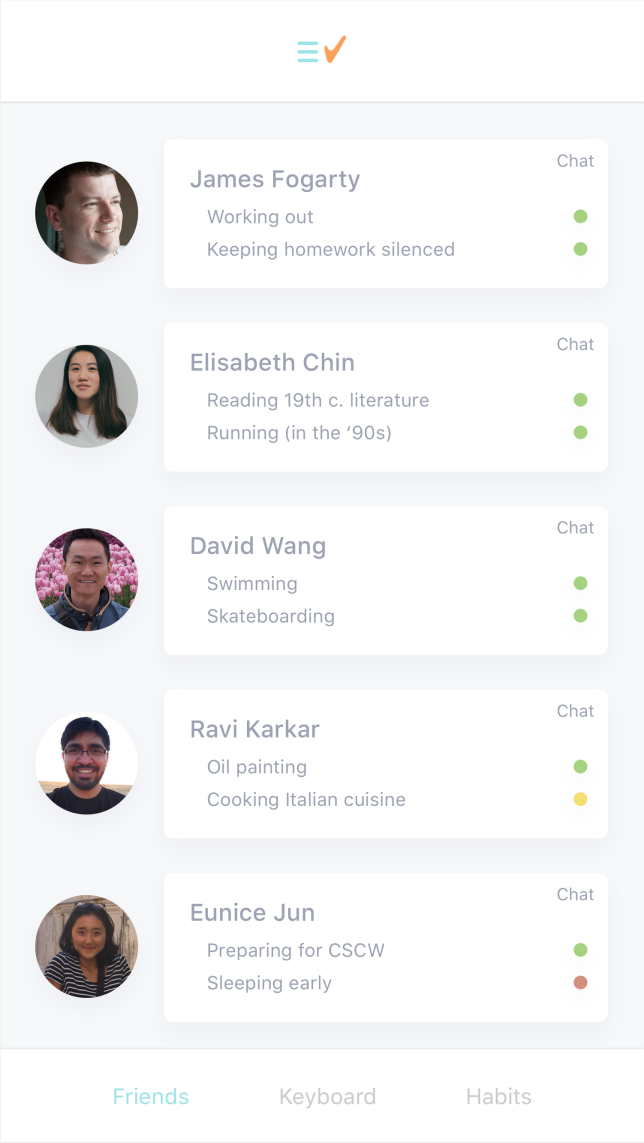






Moving from paper to digital was a surprisingly painless process. We had gone over (3!) iterations of paper prototyping so by the time it came for the digital mockups we had tested and solidified many of our design decisions. The above is the culmination of our efforts, in both wearable and phone formats.

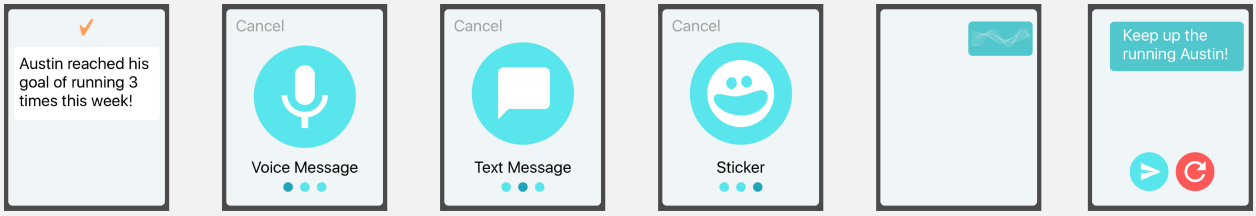
Task 1: Social Accountability



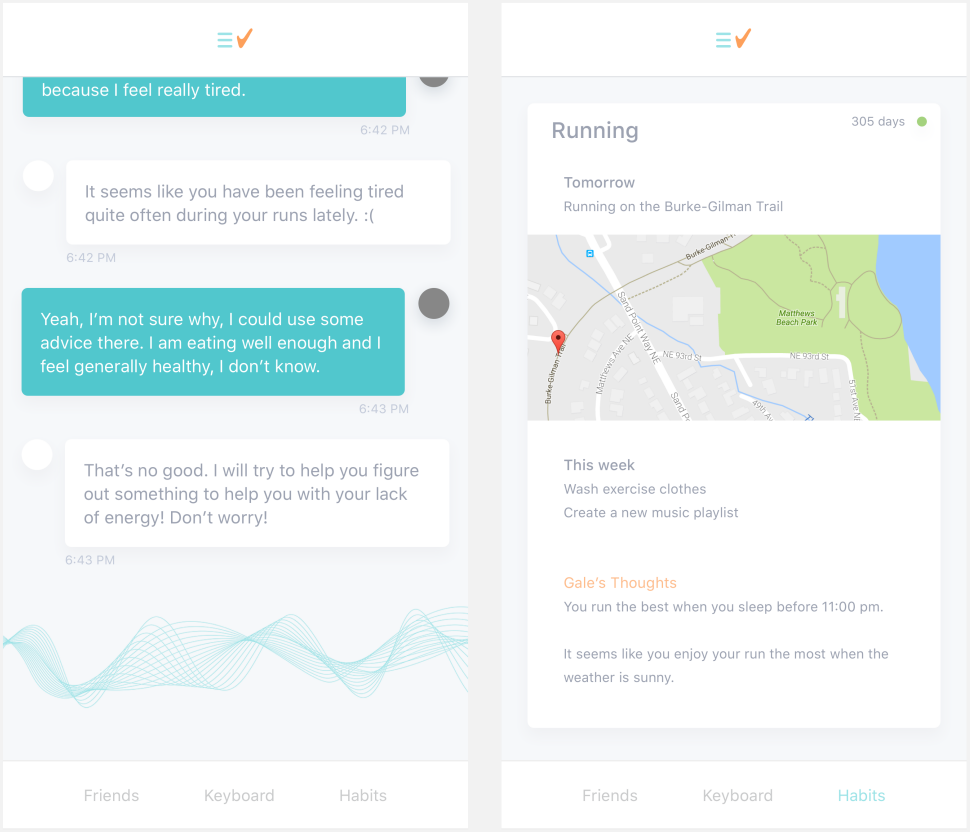
Social Accountability
A lightweight friends list on the phone app (left) provides a quick overview of friend progress and is linked to chat functionality.

Gale is meant to support only your closest connections, and users are not encouraged to be tracking more than a handful of people at a time.

Gale users can keep updated on their friend's progress through notifications on the wearable app (below). Users can select various types of responses or dismiss these notifications.



Task 2: Self Reflection



Self Reflection

Users can self reflect in both the phone and wearable apps (above). The wearable is context aware and so can prompt for reflection at relevant times. Reflections from both sources can be processed by the AI so that Gale can provide insights over time (Gale's Thoughts).

8 Discussion

In these weeks we learned just how much our design could change through the process of iterative design. We had several major revisions—we strayed away from the gale experience we imagined in our storyboards in early prototypes before taking things back in the right direction.

While we thought our initial paper prototype was well designed, our first usability test made it clear just how confusing it was. It was tap-centric, felt clunky and didn't rely heavily enough on voice commands. We set out to create a conversational UI; this usability test highlighted that we were not accomplishing that goal.

We then shifted our design to the other end of the spectrum, where there was almost no physical interaction and participants had to talk to gale for nearly everything. After two usability tests we decided that we had stripped away too many conventional interactions. Our final design became a compromise between the first two. The final design exposed traditional navigation for quick access and delegated everything else to gale's conversational interface.

We did not alter the two tasks that we started with, but the way they are implemented has changed. Initially, we prototyped the social accountability task by having a friends list where users could mark certain friends as favourites and adjust frequency of notification of friend updates. While our final paper prototype still has the friends list we decided to eliminate favourites because their sole purpose was to filter overload of notifications. In the final prototype we decided that users would only be tracking a handful of important connections, removing the need for favourites. With this iteration, the person sending out notifications adjusts their frequency, similar to how status updates on Facebook or Snapchat currently work.

As for the reflection task, our prototype separated out the reflection pages by habit so you had to select a habit before you would have access to that habit's reflection. We removed this unneeded friction in our final revisions by integrating reflections into the conversation.

The initial prototype also had the ability for gale to prompt reflection through the wearable, however, participants didn't think that gale pushed for enough reflection. We kept the ability to reflect through the smartwatch, but changed the AI design to push for more in-depth reflection through follow up questions. This also addressed the issue that gale felt too robotic or clinical to some of our testers.

Each iteration of our design got us closer to a *theoretically* ideal design, so we think that more iterations would still be beneficial. That being said, we are happy with the current design of gale after our testing and two full revisions.

Appendix

Test Plan

Phone App

1. Ask them to access their friends list.
2. Ask them to look at what James is doing.
3. Ask them to look at their running habit. What would they want to see on this screen?
4. Ask them to create a new habit (they should ask gale to create a new habit)
5. gale does back and forth convo to create the habit
    - name of habit
    - how often habit is performed
    - time of day habit is performed
    - what they need beforehand
6. gale fills out the habit card as the person gives them info about the habit
7. Once card is filled out it shows up, gale asks if the info is correct
8. Ask them to look at what habits they have scheduled for today
9. Ask them to try to move reading to 9:00 instead of 8:00
10. Ask them to open up their journal
11. Ask them to try to see what they reflected on 2 days ago.
12. Ask them to filter out reflections by one habit

Wearable

13. What would you do if you wanted to ignore this notification?
14. How would you engage with the notification
15. gale will converse with them about their habit
16. How has your habit been going?
17. Anything you have noticed that has made that habit better?
18. What problems with it are you having?
19. Ask them if they liked that gale asked specifically about the run or would you have preferred it said something more open ended like “Is there anything else you want to tell me?”
20. gale gives context aware notifications. Since it is the morning and you planned a run after work it is reminding you of what you need.
21. Prompt them that they want to cancel the run. What do they do?
22. Second scenario: How you you have gale remind you later in the day about their run?
23. For notification about your friend’s progress how would you want to engage with it? Would you want to “like” it or send a message over text or a voice message?

Questions to ask at the end:

- Ask for opinions on a reward system. Earning points to buy gifts for yourself or friends, or your friends set little rewards to motivate you to achieve your goals.
- Discuss account setup. Should we have gale when you first enter ask your name and if you want a picture for your account?

Critical Incidents

Heuristic Evaluation

User Control and Freedom: When presented with a prompt for reflection on the smartwatch, our evaluators suggested the use of a "snooze" function that would allow the user to respond to the prompt at a later time.	We decided that a snooze function would be best implemented by "swiping" the notification, while tapping it would allow the user to engage with the notification.
Match between system and real world, Severity 2: Our evaluators suggested that receiving notifications about a friend and "favoriting" a friend should be separated. It was not immediately clear to our evaluators that "favoriting" a friend meant that that friend's progress would be easy to access.	We decided that most users would only want to track a handful of their friends (< 10) in the first place, and that it would make more sense to remove this functionality.
Help and Documentation, Severity 1: In the process of "favoriting" a friend, we presented users with different options for the types of notifications that they wanted to receive about their friends. One of these options was "failing", meaning that the user would receive a notification when their friend was falling behind on their goal. This wasn't clear to our evaluators. Possible solutions included providing a help box, or including a short description of each option.	See above.

Flexibility and efficiency of use: We were recommended to allow the creation of new diary entries, and support voice dictation. We had intended for this to already exist; however, this highlights the difficulty of discovery in a voice controlled application.	We ended up removing the separated diary and allowed for reflection to be integrated into the main conversation with gale.
Consistency and standards: Our evaluator was confused by the ability to enter habits both by voice and by hand. Our paper prototype also failed to make clear that habit steps can be automatically filled.	Our second and third redesigns helped us realize that navigational aspects were best suited to tapping/swiping and that everything else could remain conversational.
Error prevention: We need to add buttons or a long-press menu to allow editing and deleting habits. We also might consider a “priority” habit, which stays at the top of the habit list akin to the favorited friends.	Habit editing is now controlled primarily through the conversational UI.
Flexibility and Efficiency of Use / Error Prevention: We need to clarify how navigation is performed in our prototype. We should also come to a decision as to whether we want a navigation bar or other visual indicator.	The home screen of gale serves as a chat bot, where users can describe the actions they want to perform which gale will carry out.

Usability Tests

Severity 3: Liked conversational interface / Navigation confusing and difficult / Voice should be input focus (e.g. friends screen).	We redesigned our phone app for the next two usability tests with a better focus on its conversational UI (e.g. friends card).
Severity 2: Reflection should continue conversation, not stop after first input.	Changed gale to continue conversation with users.
Severity 3: Users don't know how they can interact with gale.	Added suggestions at the bottom of the input screen (revision 2). Added on boarding conversations (revision 3).
Severity 2: gale wouldn't have had much impact on the user's ability to form habits.	Discussed how we could improve gale's ability to motivate the user. Our best plan was to improve the social responsibility felt by the user through voice/ text responses to notifications.
Severity 0: gale should be aware of the user's calendar so that it can help the user fit their habits in their daily schedule.	Added calendar integration and associated conversations with gale.
Severity 1: Desire to be able to respond to notifications from friends in a richer manner.	Voice responses and emoji responses.