Accountable: Achieving Goals with Accountability

Levi Lian
Stanford University
Stanford, CA
levilian@stanford.edu

ABSTRACT

We describe the ideation, execution and data analysis of Accountable, a mobile application that helps people achieve goals through holding each other accountable. The idea sits in the broad productivity space but is informed by our literature review and generative study. The study shows a need for effectively building medium- to long-term habits and achieving goals with accountability. We then discuss the mission and detailed execution of the app, Accountable, and share our findings from our public beta. We conclude by describing key design implications for both our application and other behavior changing apps that incorporate social, communal or mutual elements.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords

Mobile; Productivity; Social Network; Personal Development

INTRODUCTION

In recent years, behavior change tools that help people achieve goals and build better habits have become increasingly popular. One main explanation for this change is that people spend more time on their phones. This both creates an opportunity for empowerment through technology and opens a floodgate to myriad problems with excessive phone usage. No wonder some behavior change tools now intend to help people avoid distractions by staying off their phones.

Another driving force is a natural desire from the developer and researcher side to use persuasive technologies to improve people's lives. Technologies should assist people, not get people addicted to their devices. This need is most prominently reflected in the establishment of several organizations such as Center for Humane Technology and Stanford Institute for Human-Centered Artificial Intelligence. To empower people to become better versions of themselves through technology is a major reason behind the integration of cutting-edge behavior change principles by Ariely and Fogg into mobile apps [1, 5].

However, one limitation is that most of the prior research focuses on changing one person's behavior through the app itself. The app is the agent for change, not the user. Not many research studies the interpersonal and intrapersonal impacts on habit building and goal completion. No academic exploration can be found on studying how family and friends can help drive behavior change through digital technologies.

Moreover, far more productivity tools in the app market focus on shorter-term goal completion – such as calendar apps that schedule events, and time management apps that help people stay focused on immediate goals – rather than building medium- to long-term habits (ones that take more than cramming at the last minute, often do not have a strict deadline, and are hard to complete in one sitting). Admittedly, there are apps that help achieve goals in one specific domain (e.g., body-building, sleeping habits, eating habits) abound. But they lack generalizability and do not exhibit the impact that a closed, private social network could have on habit building.

Do accountability systems increase medium- to long-term goal completion rate? What kind of accountability systems and what degree of accountability is needed for achieving a variety of goals? We want to answer these questions with our research. Through the disciples of the d.school design thinking approach, we explored the problem space with an open mind in order to develop research questions and come up with solutions.

We began by investigating the behavior change field as a whole and studying people's interactions with popular apps in the market. We then used the insights from our literature review and generative study to identify user pain points with existing systems. We were especially intrigued by how people could achieve short-term goals that have a strict deadline and do not take long to complete, but often failed to realize their longer-term goals. We were equally surprised that a healthy support system, however rudimentary, could help with building personal habits and achieving personal ambitions.

The main contribution of this paper is showing the possibilities of making habit building and personal development social (to a certain degree). The specific design insights can also help future research in this field enhance user experience and drive adoption of friendship within habit building. The following sections of this paper will outline the design process from conception to field study, and conclude with implications for design in this space.

RELATED WORK

We studied papers that explore different aspects of behavior change in physical and digital products. Apps that intend to

change behavior were also examined to extract cutting-edge industry methodologies.

Kummerfeld, Tang, Kay, and Yekeh [10] show that effective productivity tools need not be mobile. Ambient displays that require users to log in information daily can actually provide much easy access that a mobile app. This research helps us think more broadly about this implication of using a mobile app, i.e., whether it is always ideal to build a mobile solution and whether anything in real-life can be utilized to drive behavior change.

Stawarz, Cox, and Blandford [11] conclude that trigger-based events are better at facilitating habit formation than reminders. This goes in accordance with the tiny habits methodology proposed by Fogg [4], which illustrates the effectiveness of anchoring the habit that needs to be built to something more frequent and takes little or no effort to complete.

Kovacs et al. [8] find that with behavior change interventions on the focal goal (e.g., reducing time spent on Facebook), time spent on related apps (Gmail, Twitter and Youtube) is not increased (positive spillover effect). The implications of this research are far-reaching: if interventions have net positive influence on user's behavior, then researchers can put more emphasis on setting up the right trigger or nudging the user to the intended behavior without worrying about unintended consequences.

Kovacs, Wu, and Bernstein [9] studied whether a strategy of rotating behavior change interventions can improve effectiveness. The implications of this study are that users will adapt to static intervention methods over time and their effectiveness will decrease accordingly. Although tapping into a variety of interventions will increase effectiveness, caution needs to be taken if the fickleness of the app discourages users from using the app at all.

We also spent time looking to apps that already foster productivity and habit building in order to gain insights about what was already being done in our domain and how we could improve upon these apps.

Forest [1] is an app that makes people more aware of their phone usage and transfer time spent being unproductive on phone to productive activities off the phone. A user can set a goal of not using phones for a certain period of time, and the app gamifies the experience of staying away from phones by displaying a tree on the screen. If the user uses their phone during the period of time they set for themselves, the tree will die. It ups the stakes and plays to people's loss aversion so that people need to double think before they come back to their phones. However, this static behavior change intervention suffers from novelty and has a decreasing marginal utility over time. Moreover, it fails to make sure that the user does not find other distractions to kill time so that they don't have to work.

Calm[2] provides users with calming, pre-recorded audiovisual sessions to meditate. It provides soothing music and a environment for users to calm down and meditate. It identifies the universal need for a sense of peacefulness in life and better sleep in a overly stressed world.

Moment[3] is a screen time tracker that makes people more aware of their excessive phone usage that often goes unnoticed. Users can measure their usage and see how much time they are spending on their phone and which apps they are using most over time. However, it's also frustrating because there are just too many numbers on the screen - they don't tell a story to convince me that something is wrong with my phone or give me tips to stay off my phone. Numbers usually fail at conveying a story of what's wrong with a user's phone use. Users can see their family's device usage, establish screen-free times and attend guided sessions to achieve goals and improve habits. These features sound rewarding because holding each other accountable within family is very important to lots of young people.

MOTIVATION

Persuasive technologies have helped people achieve goals as varied as topics such as sleep [6], exercise [2], and productivity [7, 9]. However, each person's goal is different, and usually personal ambitions fall outside of the narrow domain of exercise, sleep and productivity.

Moreover, while there are a number of academic frameworks for behavior change systems, notably B=MAT [5], which demonstrates that (M)otivation, (A)bility and (T)rigger are three elements for effective behavior change, few mention the impact of a close-knit support network. Another framework, habit loop [3], shows that an iterative system where a trigger prompts an action, which leads to a variable reward, will ultimately allow the user to make an active investment and build habits. Such a self-contained system also ignores the role that family and close friends play in achieving goals and building habits.

We gleaned this insight specifically from our generative study, which is presented below.

Generative Study

Over the course of one week, we conducted a generative study to study people's interactions with mobile productivity tools. Our generative study consisted of four interviews preceded by contextual inquiries and one home tour preceded by an interview.

Research Questions

The main research questions we had at this stage were: How do people use their phones versus other instruments (e.g., physical tools, friends) to facilitate productivity? What environment (time and place) facilitates people's productivity? What prevents people from using productivity tools (e.g. Screentime)?

Methods

Contextual Inquiry

We recorded audio and took photos of our subjects as they used the think-out-loud protocol to describe what they were doing while working on digital device(s).

[1] Forest: https://www.forestapp.cc/

[2] Calm: https://www.calm.com/

[3] Moment: https://inthemoment.io/

Interview
We went through a list of questions we prepared for all interviewees in 20 minutes, and then collected enough information so that we could build a persona later.

Home Tour
We recorded our subject as he used the think-out-loud protocol to describe how he used hi home workspace to get work done. We also asked him to act out how he would get distracted in his home workspace and act out how he would get out of it.

We recruited five people for our generative study. The first person we recruited was Stanford student and was recruited from Facebook. We met him in his dorm room to conduct an interview and a home tour. Four other people were found by going to Blue Bottle Coffee in Palo Alto on an average Saturday afternoon. We interviewed each person and observed how each of the participants worked with their phones and laptops in Blue Bottle.

We decided to do a home tour because we thought it would be the best place to study productivity tools being used in a non-digital context, helping answering our second research question. We also thought that doing a home tour could help us determine why a person may like working from home rather than an outside space, answering our third research question on how environment plays a role in productivity.

In our interviews we could ask participants more about if they procrastinate with their mobile phones and why or why not. We could also ask them about what types of tools and how they use those tools in more detail with our observation (context inquiry). And because we chose the time and place to interview participants, we could ask them why they were at Blue Bottle in the afternoon to study.

Participants
The recruitment criteria we chose were the following: Participants had less than 3 years of professional experience Participants had little to proficient experience with productivity tools and methods Participants were moderately to extremely passionate about developing their career

We used a screener in figure 1 to make sure our recruited participants and interviewees fit the specific criteria.

Analysis of Data
We recorded audio using Otter.ai, which auto-transcribes the interviews for us. We then used a post-it generation Excel macro to produce 70 meaningful, standalone quotes. This allowed us to do an affinity analysis of the content (see figure 2). We also built a profile persona for each interviewee and created storyboards from their perspectives. We identified several major themes and launched one design sprint before narrowing down to the solution, Accountable.

Design Implications
Below are the three main insights we gathered from our affinity analysis and hoped to facilitate within our app.

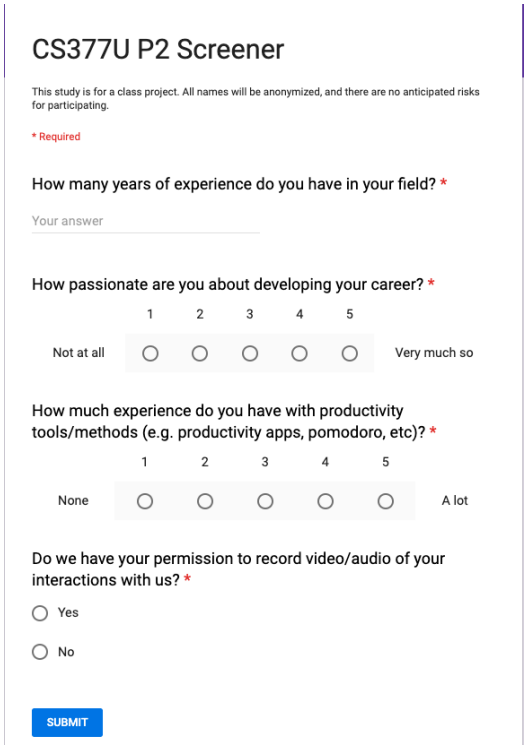

Figure 1. Participant Screener

Figure 2. Affinity Analysis of Interviewee Quotes

Facilitating accountability can help achieve medium- to long-term goals
Several participants in our study said that they usually completed short term goals with a deadline by turning off their phones or going to a quiet space. One user says, *"When I'm ready to work, I turn off my phone..."*

However, they usually had issues achieving medium- to long-term goals. An extrinsic incentive or push is often needed in order to achieve these types of goals. One person says,

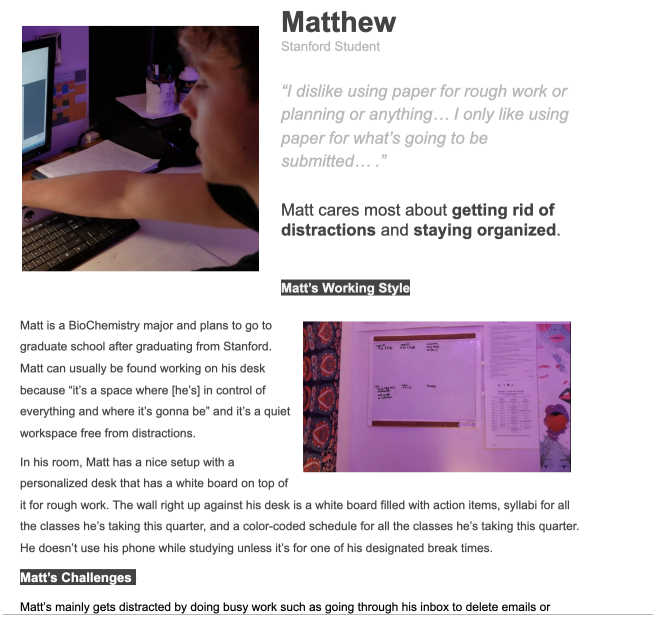

Figure 3. Sample Profile Persona

"... I feel motivated to finish long term goals when there's something to celebrate." Another interviewee stated, *"when I did my senior thesis, I sent a email to everyone and texted some close friends saying don't talk to me this month."* It doesn't take us long to realize that the importance of friends in helping us get things done.

Providing actionable steps to steer users toward completion

One participant from our home study and two participants from our interviews usually got goals completed by writing down what they needed to do step by step in order to complete their goal. We realized that to achieve goals is to break them down into actionable steps. This insight goes in line with the apps we've studied because it shows the action needs to be simple enough and followed by a reward. Therefore, our app should provide a space for users to turn a goal into a regular habit they want to build.

Generating a positive reinforcement loop for people to celebrate each other's success

One participant in an interview said, *"Asana is really cute. If you complete every three tasks, like a unicorn or narwhal jumps."* The playful and fun environment of the app made her want to engage in the app more and complete her goals. Therefore, working in a positive environment that celebrated her success made her more likely to complete her goals and should be implemented in our app. This insight also directly supports the hooks model and makes it possible for people to celebrate not only success of their own, but also of their friends. We think that it can complement really well the goal of facilitating accountability.

SYSTEM DESCRIPTION

In the following sections, we will present our initial paper prototype and the final design that was used in field study. We will also discuss the design process, including usability testing and the changes we made to the final design.

Initial Paper Prototype

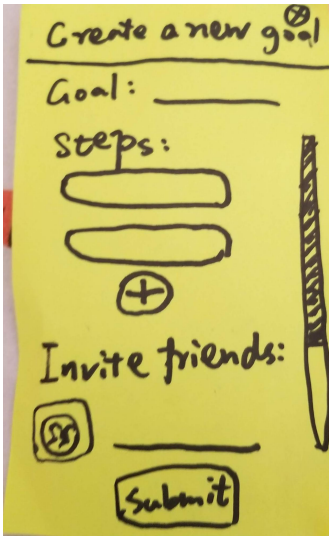

Figure 4. Create a Goal Screen

Our first main use case is creating a new goal. Within this use case, our participants are expected to name their goal, set how often they want to be reminded to do that goal, and invite a friend via a phone number to help them keep track of their goal. To see how this looks in our paper prototype, see figure 4.

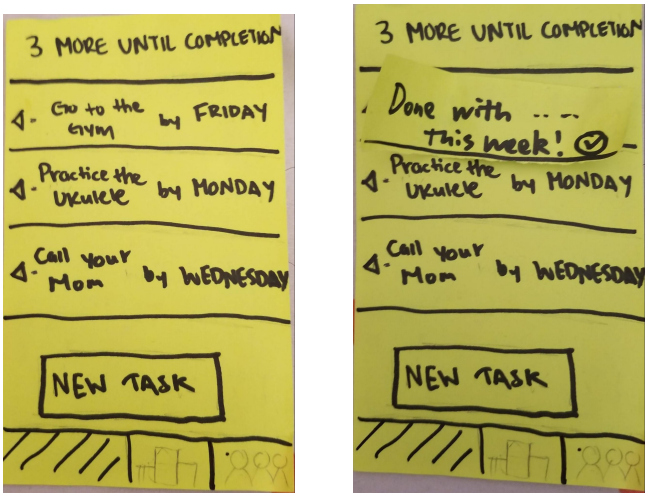

Figure 5. Incomplete and Complete Goal Screen

Our second main use case is tracking one's progress in completing one's goal. This screen displays information about each goal and is ordered by the most recent goal completed. In our initial usability study, users could swipe left to indicate

that they had completed their goal. To see how this looked in our paper prototype, see figure 5.

Figure 6. Messenger Screen with Auto-Populated Text

Our third main use case is a tricky part of the design, and we wanted to test our hypothesis that that asking friends to hold the participants accountable would lead to them more successfully completing their goals. In our app, the friend responsible for your progress will receive in-app messages about your missed task or you being on a streak. We will prompt them to reward or punish you via their default communication tool (Messenger in our case). We will auto-populate encouraging or motivational messages so that your friend can start a conversation with you in hopes of making you feel supported and proud or steering you on the right track. To see how this screen looks in our paper prototype, see figure 6.

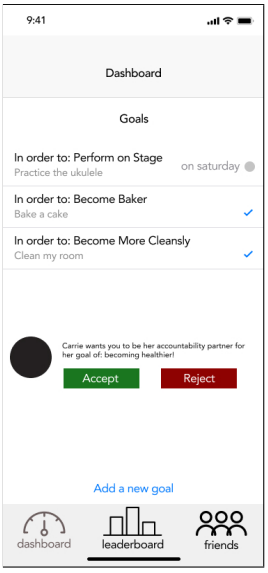

Figure 7. Figma Screen for Accepting or Rejecting Request To Be Accountability Partner

Finally, our fourth use case was added in our medium-fi prototype. The invited user was able to reject or accept the request to hold their friend accountable before being linked to their friend's progress data. To see how this looked on our medium-fi prototype, see figure 7.

Findings From Usability Study

It is not clear how to check in a task

Participants one, two, and three could not figure out how to let the system know they have finished a goal. For our original design, we had the user make a left swipe when they have finished a goal. The dashboard screen would then switch to a screen with a checkmark by the completed goal. We provided arrows on our screen in order to let the user know they could swipe to say they finished a goal. However, this affordance was not good enough for users to understand that had to swipe to finish a goal. Most of the users clicked on the goal card in order to indicate that they were done. Therefore, in our second design implementation, we changed the action from a swipe to clicking a button that makes the check mark appear. This design change was successful in our user test with participants in the medium-fi user test. In our actual app, we build a calendar system to indicate completion for the day.

It is not clear what the last step in adding a new app is for

Participant one did not understand what our *"invite friends"* button did in the create a new goal portion of the app. Originally, clicking on the contact icon in this part of the app would open the default contacts app and send a text message to your friend become the user's accountability partner. In our actual app, We decided to use different wording to explain what it means to enter a phone number. We provided more context and onboarding in the beginning of the app so that the value proposition is made clear. In our actual app, we used Twilio to allow your friend to track your progress without signing up for the platform.

Do not understand the concept of reward and punishment

Participant three did not understand what it meant to reward or punish a friend in the app. Therefore, in order to fix this design problem, we suggested adding a way to list the acceptable types of punishments and rewards a user wants for their specific goal. If the friend has accepted the user's request to be the user's accountability partner, the friend can then see what types of rewards and punishments are offered with that goal and choose from there to give out a reward or punishment. In our actual app, we allowed accountability partners to initiate a text message with their friend about their missed task or successful completion within the app.

Affordances are not clear enough for users to know which friends have completed goals

Participant four struggled with finding which friend has finished their goal, which friend was still in the process of completing the goal, and which friend needed was past the due date for finishing their goal. This user test was done on our second round of user testing, so for affordances we had a check mark if the user was done, greyed out text if the user was still in progress of completing their goal, and red colored text if the user had passed their due date. In our actual app, we

dedicated a friends page where all the data about friends' goals are displayed in the same fashion as one's own goals. This would become clear once the user sees different sub-sections under different friends' names.

Thought that there was a difference between the goals screen and the dashboard screen

Our fifth participant struggled with telling the difference between the goals (which shows the most recent goal to be completed) and the dashboard page (which shows the analytics). He clicked on the Goals title when he wanted to create a new goal because he thought he had to go to a new page in order to create a new goal. This problem may be a cause of the create-a-new-goal being so far down on the screen. The user also could have a different mental model of a dashboard and a goal screen. In order to fix this problem, we should make the create a new goal button bigger and more prominent, and replace the dashboard page with the friends page completely.

Product Vision

The findings from our usability study proved to be invaluable in forming our own product vision and help create the entire user experience illustrated below in Final Prototypes section.

The product vision is to achieve goals effortlessly and playfully with friends holding each other accountable. This aims to solve the problem with achieving medium- to long-term goals.

We built out our app with React Native[4] and Expo[5]. We used Twillio[6] for sending phone messages and Firebase[7] as our database.

Final Prototypes

The following are the final screens that we designed for our app. All of these screens fall underneath our main four use cases we wanted to accomplish with our initila design.

Screen 1: Create a goal and invite a friend to hold them accountable (Figure 8)
Screen 2: Friend receives invite and downloads the app (Figure 9)
Screen 3: The goal owner checks in (Figure 10)
Screen 4: Friends viewing check-in history (Figure 11)
Screen 5: Friend receives notifications when the habit owner checks in, and the habit owner receives reminders to check in (Figure 12)
Step 6: Friend sending words of encouragement to the goal owner (Figure 13)

FIELD STUDY

For the duration of 10 days, we had a group of 60 people in our target audience use the app. These 60 people consist of working professionals in the Bay Area under the age of 26, college students that include all four years and grad students. They are the ideal testers because not only their demographics

[4]React Native: https://facebook.github.io/react-native/
[5]Expo: https://expo.io/
[6]Twillio: https://www.twilio.com/
[7]Firebase: https://firebase.google.com/

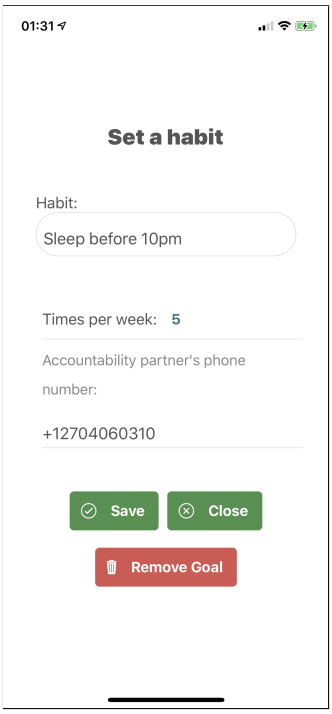

Figure 8. The goal owner can create a goal and invite an accountability partner to the app

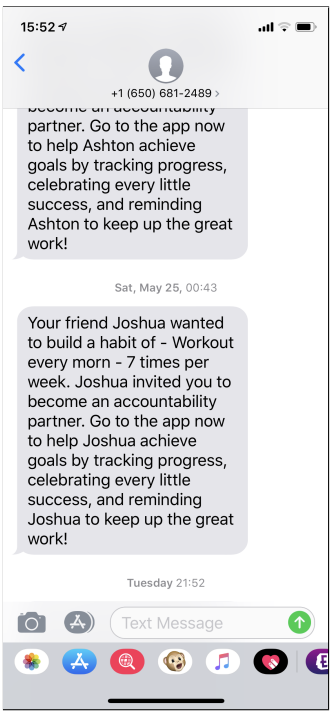

Figure 9. The friend will receive invites to download the app and understand the purpose of doing so, i.e., to hold their friend accountable for their success

fit in with the target audience identified in previous research, they proactively signed up for the usability study (some down-

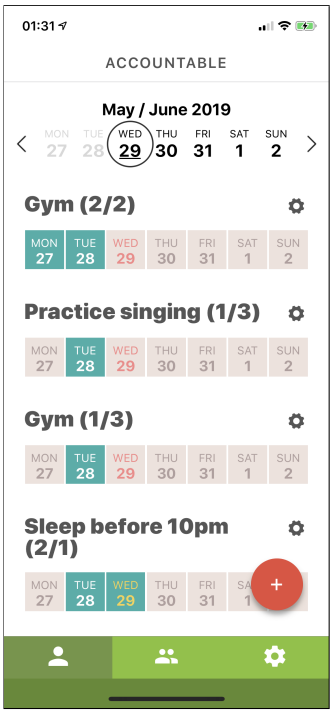

Figure 10. The goal creator can check in and view one's own progress

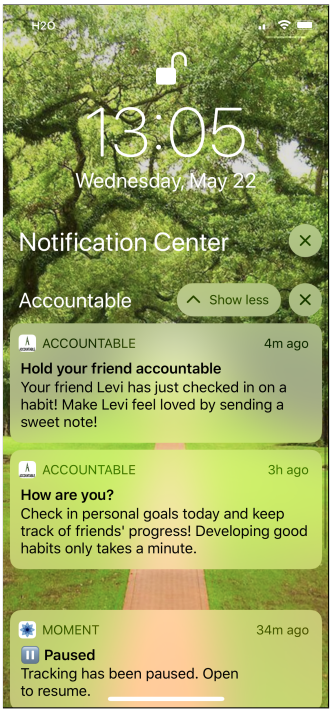

Figure 12. The friend will receive notifications when the person they need to hold accountable checked in in the app

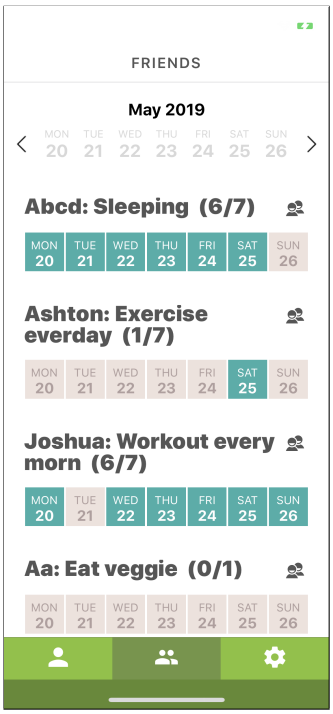

Figure 11. Friends can view your check-in history and celebrate success with you

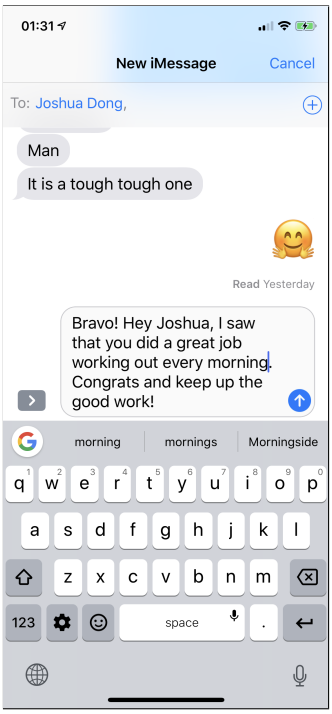

Figure 13. The friend, upon receiving notifications, can send words of encouragement to celebrate every single success their friend has achieved

loaded the app through TestFlight organically). Our hypothesis is that these people therefore have the motivation to build better habits and achieve goals, only lacking effective tools to help.

On the side, we did very small-scale Facebook marketing through our network across the globe and attracted a total of 81 downloads in three continents over the span of 10 days. Among them, 60 went through the sign-up process. Notably, 9 users (14%) were acquired through user referrals, i.e., being invited as accountability partners, with 8 other dormant users who, upon receiving the invite, never downloaded the app.

Research Questions

For the field study, we wanted to answer our overall research question. Do accountability systems help people achieve goals better? Does a closed, private network beget the right amount of accountability from a variety of goals? To do so, we need to look at how often people engage with the app, their retention and user funnels, and whether their behavior suggests that they get value out of finding accountability partners. Moreover, interviews will further help us understand the usefulness of the concept to each individual.

Methods

For our field study, we used Amplitude[8] to analyze the different events we tracked, interaction flows, user demographics, and retention analysis alongside 3 semi-structured interviews and a content analysis to learn more about our data.

Event Tracking, Interaction Flows, User Demographics and Retention Analysis

For the app, we instrumented 20 events to track with the analytics tool, Amplitude. Among them are such noteworthy events that affect the core user experience: *"gone through onboarding"*, *"began signing up"*, *"signed up"*, *"created habit"*, *"check-in notifications to friend"* and *"sendSMS"*. After the user has logged in/signed up, we create a unique user ID so that all future events can be traced back to the same person.

More important, other than noting down what event the user did, we also log in data relevant to the event. For example, when a user turns on a local notification, we note down the time the user sets for herself. By the same token, when a user creates or edits a habit, we register data from habit description, habit frequency, and phone number. This allows us to conduct extensive qualitative content analysis, which we will explore later.

While individual events can already tell the story about usage, Amplitude also provides us with the ability to look at user sessions, conduct funnel and retention analysis, and build a user demographic. The amount of data linked, correlated, and casually related allows us to dig into all areas of quantitative product analytics.

[8] Amplitude: https://amplitude.com/

Three Semi-Structured Interviews and Content Analysis on ten Users

The qualitative methods we chose are interviews and content analysis. These two methods are relevant because the former allows us to hear complaints about our product directly from our most vocal users, and the latter – given user consent – generates the most insight into improving our core services. We interviewed 3 users of different demographics and background on May 28th, 2019, each for 30 minutes, to discuss the usability of the app and the usefulness. On May 28th, 2019,we also conducted content analysis on 10 users' effective habits that we had been tracking through the system for at least a week.

Analysis of Data

For the analysis, we mainly focused on the effectiveness of accountability partners in achieving goals and whether a close friend network generated enough level of accountability. We used both quantitative and qualitative analyses to answer both questions.

FINDINGS

International audience helps increase diversity among users

Our 60 users during the study are from 4 different countries and reside in 27 different cities. This cohort of users has a diverse range of behaviors interacting with the app. For example, users from China were unable to authenticate with Firebase, a Google product, and had to be given test accounts to proceed. One user from Italy is the only user who sent us feedback directly through the app. California users usually set reminder times later in the day (after accounting for time zone differences) than people on the East coast. The wide range of experiences has allowed for cross-pollination of ideas and bodes well for international adoption.

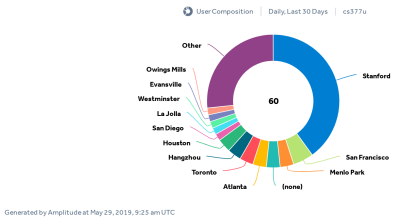

Figure 14. User Demographics By City

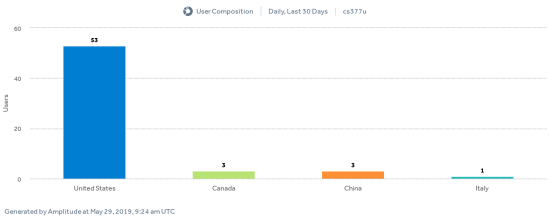

Figure 15. User Demographics By Country

Engagement can be regained if you roll out the right features

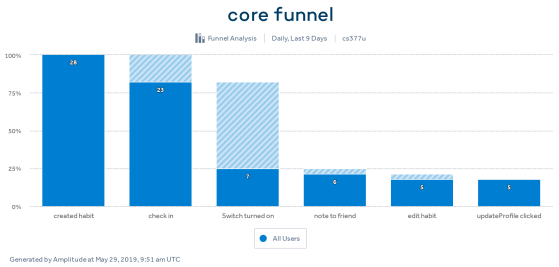

Figure 16. Most of the users who created a goal checked in. However, only 20% of them explored some form of advanced functionalities

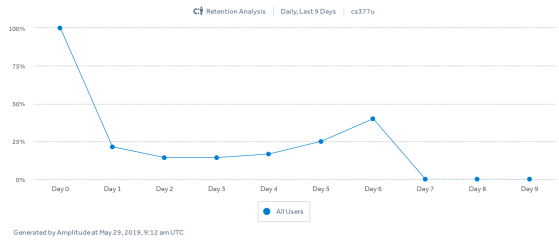

Figure 17. 7-day retention of users

The 7 day retention graph ends on a promising note. The retention averaged at 20% after the first day and is increasing over time. This is mainly due to bug fixes that happened on day 4 of the study, which brought people back to the app. This shows that while field study is hard to be perfect in one launch, quickly iterating can generate the value. More importantly, it shows that if the concept rings dear to users, they can tolerate bugs and still choose to use the app albeit less than ideal user experience.

Good ideas can enthrall people in the beginning

Check-in is such a key indicator of usage for productivity apps. For us, check-in is the achievement we're most proud of. This is partly due to the way we make check-ins easy (one click on the main screen), and introduce the accountability part. The conversion time from signing up to first check-in averaged at around 3 hours, which means that users immediately take action upon signing up. Moreover, 90% of logging in activities are followed by check-ins. Additionally, users had a median check-in frequency of 2 times a week.

Onboarding is a continual process of guiding people towards success with your product

People need to have an idea of what it is before they even sign up. Our app failed to do this as over half of users were lost soon after the blast of onboarding screens. People also spent less 3s was spent on each of our onboarding screens. Additionally, none of the interviewed users said they really looked at the onboarding screens. Finally, based on an in-app survey that reached 43 people and of the 25 people that responded, 64% said they had to figure out how to figure out how to create their own habit on their own. This shows that

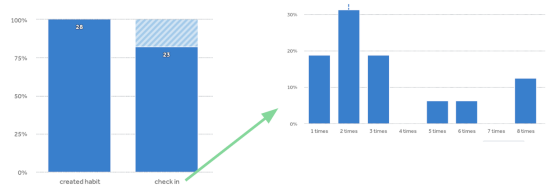

Figure 18. 80% of users who created one goal with one accountability partner checked in at least once over the study

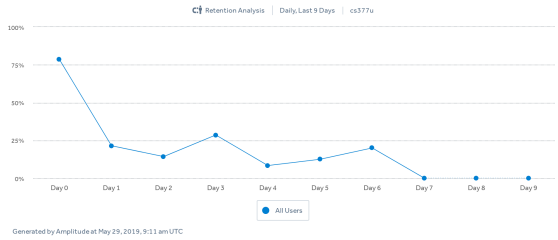

Figure 19. The 7-day retention of people who check-in. The graph has little difference from the log-in retention shown above, which suggests that most of users who log in are to check in for that day

our app isn't very intuitive or engaging and our onboarding screens need to be improved.

Figure 20. One of three current onboarding screens. There is too much text and graphic information, without reference to the actual user workflow

One user mentioned in the interview that *"I don't know why I should pull a friend in [by entering their phone number]"* while another user said *"I don't know what would a friend do via the app to hold me accountable upfront, so I left the app after onboarding"*.

Introducing accountability can be a beautiful user acquisition story

Intially, accountability is part of the research question we hope to get answers from. As the graphs show, persuasive technologies can have a positive spillover effect when a win-win situation is present – friends holding each other accountable can make themselves better people and cultivate friendship. Such a value proposition therefore proves a wonderful user acquisition story.

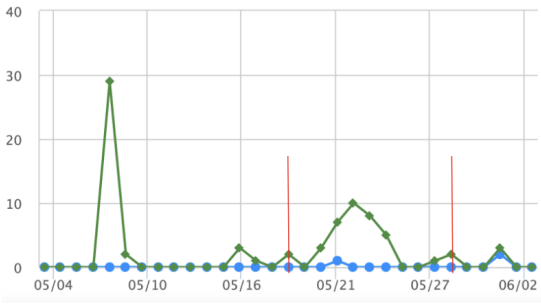

Figure 21. 28% of new users were acquired because their friends invited them

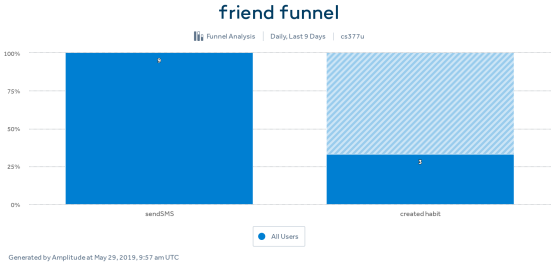

Figure 22. 33% of the acquired users invited their friends to the app, too

Default choice and playful design is crucial to driving behavior change

Only 19% of users turned on notifications. This issue may be due to the fact that by default the users have to go into their settings page to turn on notifications. This prevented accountability partners from receiving updates about their friends checking in, leading to lower retention rates. Users' existing session was not preserved if they exited the app completely which creates friction and lowers retention since signing in requires phone authentication. When users want to add a friend, they must enter their friend's phone number manually, which causes issues if a user can't easily remember their friend's phone number. Finally, from interviews, we gathered that users wanted something more playful when it came to rewarding or giving consequences to their friends rather than pre-populated text.

People loved the idea of tying personal development with cultivating friendship

The demographics of our users spanned 4 countries and 27 different cities, which proves that no matter where you are, friends care about each other's success. Several participants also noted how they reconnected with old friends by using this app as a conversation starter. Participants clearly love the idea of facilitating personal development through friendship.

Thoughtful product design, however subtle, will create lasting, delightful memories of user experience

One user, in response to a sweet reminder message the app sent, wrote to us: *"I feel like I am really treated as an actual friend first before as a user of the app"*. Our notification says "How are you" in the title, and then ask the user to check in for their goal and take a look at how their friend progressed. We also highlight how easy the action is by mentioning "the whole process will take less than a minute".

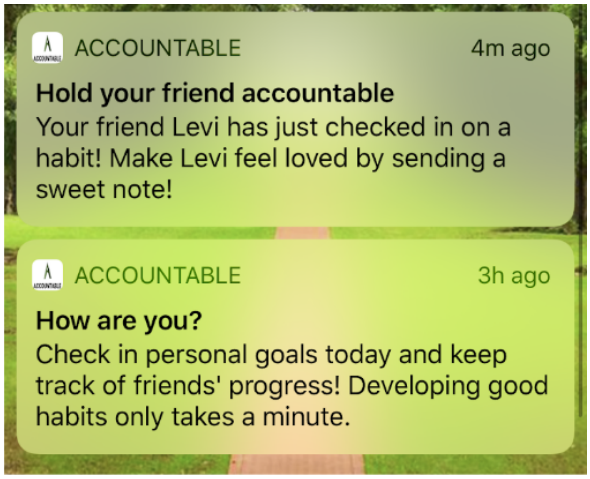

Figure 23. A push notification that delighted one user because of its personal touch

Content Analysis

For the content analysis, we went through 10 users (Will S., Minha, Ashton, Kai, Kate, Matt, Will C., Chin, Zheng, and Joshua) and analyzed the content of their goals. Five of these users had time-specific goals (eg. Go To Bed Early Every Night) and the other five users have atemporal goals (eg. Meditate). There was no specific difference between how likely these users completed their task or did not complete their task for the week based on whether their task was time-specific or atemporal. This most likely has to be the result of the user's willingness to come back to the app and check in rather than their goal not being specific enough.

Additionally, of the 10 users being studied, the 90% of them made self-improvement goals, such as working out or going to bed early. 20% of them made hobby-oriented goals such as dancing and reading more. And only one user made a group-oriented goal, *"rock climbing together"*. In fact, out of all 17 goals created, this was the only group-oriented one. One would expect that group-oriented goals would be more prevalent due to the fact that the goal of this app is to make habit-making more community focused, but this was not the case. This may have to do with the fact that users don't understand the point of the app from the value proposition not being stated clearly in the onboarding experience. Additionally, you can see from

the qualitative data, that users stopped talking to their friends via the app over the course of the week, and therefore, it leads to less discussion of the application and goal creation between friends.

	Categories of Goals			
Name	*Time Oriented*	*Self Improvement*	*Hobby Oriented*	*Completed*
Will S.	No	Yes	No	Yes
Minha	Yes	Yes	No	Yes
Ashton	Yes	Yes	No	Yes
Kai	Yes	Yes	No	No
Kate	No	Yes	Yes	No
Matt	No	No	No	Yes
Will C.	Yes	Yes	No	No
Chin	No	No	Yes	Yes
Zheng	No	Yes	No	Yes
Joshua	Yes	Yes	No	Yes

Table 1. Types of Goals Made by 10 Users

DISCUSSION

The overall data shows that people who understand the core experience are engaged with the app, and have gained value from it.

The app stands in between a productivity app and a social media app. It is similar to a range of habit building apps, where people can create habits and check in, and be reminded to complete their habits by the app. It is also similar to social media in that specific friends can see, comment, and receive notifications on another friend's status. However, it is also different from both fields because the product mission encourages people who care about each other to be on the app. The app simply serves as a platform, rather than an end to building habits. It is also diametrically different from social media apps that heavily rely on content creation and engagement. It actually hopes that lots of activities that happen between friends happen off-line, where accountability partners do activities or habits with the habit owner or encourages the habit owner in real-life. The app serves as being a trigger or a conversation starter.

The findings validate our literature review. In particular, we find that particular anchors help a lot with goal completion. To complete a goal is to make certain behaviors regular enough so that people become good at it or no longer need to be reminded of it. Therefore, how users set goals is important. The app only attempts to nudge the user to turn a goal into regular behaviors without much guidance and handholding. We do find that those who have more granular habits have a much better check-in history and completion rate, and therefore engagement with the app.

The app showed us that gamification is a great way to make light of the unbearable weight of building habits. When difficult things look easy and playful, people will really give it a try. Showing support and having people play games while building habits and friendship is a valuable and should be facilitated more by our app.

IMPLICATIONS FOR DESIGN

The aforementioned findings suggest that for a novel system that integrates existing, well understood systems, the most important factor in driving adoption and useability is onboarding. This is because onboarding is the very first contact between the app and the user.

Build a coherent, and guided onboarding experience and make sign-up part of that process

Currently, 50% of users were lost during the onboarding. 15% of 7-day retention also indicates that the product value is not clear. This is mainly an issue with onboarding. During interviews, people mentioned how certain actions were hidden, e.g., left swipes, action hidden behind an icon, etc. While 80% of users completed the core user funnel, only 20% explored power features. Users need to know why they are here, where they are today and where they could be if they join us together to building better habits. We need to guide them through habit creation, accountability partner invitation, and show why it's necessary to sign up, before asking them to make an investment in the app.

Rethink all the default choices to optimize for value, intuitiveness, and engagement

Default choices are deeply related to whether users can opt in or opt out for reminders, which screen the user sees first when logging in, and whether they will be guided to create a new habit with friends upon logging in. We need to really dig into the optimal default level where users will understand how to use the app and what features they should already have access to without us accompanying them side by side to show them which functionalities they have access to.

Introduce gamification, specifically Badging, Randomization, and Consequences, to the core user funnel

Currently, 28% of users downloaded the app and signed up through friends' invitations. Of them, only 33% actually went through the core user funnel (create a habit $\rightarrow$ check-in). The in-app messaging 7-day retention plateaued at 10% and only 20% ever encouraged their friends within the app. This stands in contrast to the 82% check-in rate and 20% 7-day retention for check-in. During our interviews, we quickly realized that friendship is such an important part even for people to use the app in the first place. One interviewee said: *"It would be so much fun if I can roast my friends for not completing habits"* whereas the other mentioned how *"[accountability partners] choosing a truth-or-dare challenge for a friend who failed to complete a habit"* would bring people together. Therefore, we hope to introduce gamification concepts such as badging (for both the habit owner for completing habits and the partners for holding their friends accountable), randomization (choosing whom to invite for the habit), and consequences (which dumb truth-or-dare punishment question to send, and what reward to give to friend when they succeed).

Personal development systems with a social component have a promising future

The bigger picture that comes out of both quantitative and qualitative study is that people like the idea of making each other a better person. In the wake of Cambridge Analytica

event and other privacy breaches in the social media space, consumer software users are fond of any idea that counters such attention-grabbing systems. Therefore, it is a good timing for productivity and personal improvement research to venture into the social space. People love the idea of supporting each other and at the same time making themselves better.

While this app does not exclude the possibility of creating a shared goal where people can build toward it together, similar systems can further answer the question of how much of social support is needed for goal completion. Moreover, it is also interesting to see if this only works for close friends, or maybe a truly social model that involves strangers can achieve the same or possibly better effects. This insight is directly supported by the creation of some social goals where people can build toward together.

More importantly, the popularity of domain-specific apps in practicing mindfulness, yoga, cross-fit, cycling, and university homework shows that each domain needs intricately different approaches, and therefore it is hard to build a one-size-fits-all, general purpose habit building or goal completion app. This still leaves room for the possibility of validating different levels of support in habit building, and comparing and contrasting the spectrum of optimal amount of social element in habit building. This will have huge implications for behavior change research and shed light on the design principles for future HCI systems that intend to empower people to become better versions of themselves.

ACKNOWLEDGEMENTS

We would like to thank Frank Bentley and Geza Kovacs for their instruction of CS337U Understanding Users at Stanford University.

REFERENCES

1. Dan Ariely and Klaus Wertenbroch. 2002. Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science* 13, 3 (2002), 219–224.
2. Sunny Consolvo, David W McDonald, Tammy Toscos, Mike Y Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, and others. 2008. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1797–1806.
3. Nir Eyal. 2014. *Hooked: How to build habit-forming products*. Penguin UK.
4. BJ Fogg. 2019. Tiny Habits.(2015). *Retrieved June* 24 (2019), 2015.
5. Brian J Fogg. 2002. Persuasive technology: using computers to change what we think and do. *Ubiquity* 2002, December (2002), 5.
6. Matthew Kay, Eun Kyoung Choe, Jesse Shepherd, Benjamin Greenstein, Nathaniel Watson, Sunny Consolvo, and Julie A Kientz. 2012. Lullaby: a capture & access system for understanding the sleep environment. In *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, 226–234.
7. Young-Ho Kim, Jae Ho Jeon, Eun Kyoung Choe, Bongshin Lee, KwonHyun Kim, and Jinwook Seo. 2016. TimeAware: Leveraging framing effects to enhance personal productivity. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 272–283.
8. Geza Kovacs, Drew Mylander Gregory, Zilin Ma, Zhengxuan Wu, Golrokh Emami, Jacob Ray, and Michael S Bernstein. 2019. Conservation of Procrastination: Do Productivity Interventions Save Time or Just Redistribute It? (2019).
9. Geza Kovacs, Zhengxuan Wu, and Michael S Bernstein. 2018. Rotating Online Behavior Change Interventions Increases Effectiveness But Also Increases Attrition. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 95.
10. Bob Kummerfeld, Lie Ming Tang, Judy Kay, and Farahnaz Yekeh. 2015. SAL: a small, simple, situated, ambient logger. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*. ACM, 403–406.
11. Katarzyna Stawarz, Anna L Cox, and Ann Blandford. 2015. Beyond self-tracking and reminders: designing smartphone apps that support habit formation. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. ACM, 2653–2662.