Harry's Habits: A Chatbot Investigating the Effect of Rewards on Habit Formation

Harry Mumford-Turner

School of Engineering
University of Bristol
SEPTEMBER 2017

Supervision from
Dr. Oussama Metatla &
Dr. Katarzyna Stawarz.

I hereby declare that this dissertation is all my own work,
except as indicated in the text:

Signature _____

Date _____/_____/_____

Acknowledgements

Thank you to my supervisors Dr. Oussama Metatla and Dr. Katarzyna Stawarz, for their patience and guidance through this project.

Abstract

Habits are automatic actions that require little effort. A simple action such as turning on the light to a room happens automatically, even if the light is already on. Forming new habits requires contextual cues, positive reinforcement and repetition. This research focuses on positive reinforcement in the form of rewards, experimenting with forming new habits using rewards from different modes. These rewards are delivered by a prototype chatbot, evaluated during a 28-day trial, testing the success of delivering rewards via a chatbot and the effectiveness of each modality on participants habit strength. 38 participants messaged a Facebook Messenger chatbot every day for 21 days, confirming or denying if they have performed their chosen habit (20 Press-Ups). After the evaluation trial, a 7-day follow up trial without interaction with the chatbot tests if participants continue performing their chosen habit. At the end of this 7-day period, participants evaluated the project success through an informal interview and a final validated habit formation questionnaire.

In summary this research:

1. Compares how rewards from different modalities effect the building of new positive habits
2. Evaluates a prototype chatbot using 28 real-users, revealing positive and negative aspects of the design.
3. Investigates the use of chatbots as vehicles for promoting behaviour change.

Definitions

Human-Computer Interaction (HCI) - Field of computer science that studies how people interact with computers.

Modality - In the context of HCI, a modality or mode is the classification of a single independent channel of sensory input or output between a computer and a human.

System with Different Modalities - A system that provides the user with multiple modes of interaction.

Chatbot - A method of communicating with a computer system via a conversation using natural language.

Contents

1 Introduction

1.1 Summary

Habits are automatic actions that require little effort. A simple action such as turning on the light to a room happens automatically, even if the light is already on. Forming new positives habit needs three elements: contextual cues, positive reinforcement and repetition. The contextual cue acts as a trigger and the positive reinforcement rewards the person encouraging them to perform the action again when triggered by the cue. A study measured repetition to show a habit takes on average, 66 days to form [1]. However, people still fail at forming new positive habits and give up because they don't stick to a routine [2, 3].

Mobile technology can help people stick to a routine by focusing on the three elements. Using reminders as triggers to repeat an action, helping us associate contextual information around the action and rewarding us with positive reinforcement. Yet most existing mobile systems aren't built on theory, leading to repetitive actions rather than habit automaticity. Therefore, people become dependent on technology and habit performance decreases when the system is eventually removed.

Building habit automaticity removes this dependency. We can increase automaticity by building motivation to complete the action [4, 5] We can encourage motivation by rewarding users with positive reinforcement by granting user's satisfaction after completing the action. However, reward type and delivery type are important. Giving money (extrinsic rewards) hinders motivation. Giving satisfaction (intrinsic rewards) benefits the person and should be preferred. The type of delivery should suit each individual user so a choice of delivery should be available. Different research [6] shows how interaction with users should span different modalities to suit the needs to users. This project combines knowledge from these two domains to focus on how intrinsic rewards from different modalities affect people's habit strength.

This project uses the three elements of habit formation to build a mobile technology tool to deliver rewards from different modalities. These rewards are delivered using a chatbot instead of a mobile app to present a novel method of interacting with users. Chatbots are a method of communicating with a computer system using natural language, providing deeper integration into users mobile phone and can hook into messaging services users are familiar with, such as Facebook Messenger. The chatbot design is based on a combined set of theory-based design requirements; one for building habit formation apps that aim to increase habit automaticity and one for increasing motivation with intrinsic rewards.

An evaluation trial tests the prototype during a 4-week controlled period to measure user's habit strength. 58 participants engaged in daily activity with the bot for 3 weeks, then for a further 1-week all bot interaction was suspended to test if users continued with their chosen habit after the bot is removed. The chatbot provides habit tracking using reminders as triggers, and three combinations of rewards, visual, auditory and visual-auditory combined. Participants are split into four groups, a control group without rewards and the remaining three groups recieve one combination each. The evaluation trial tested the success of the chatbot by measuring user's habit strength after the trail and comparing the results with the control group.

The project aims to deliver insight into how rewards from different modalities effect habit strength and opens up new research avenues for investigating the use of chatbots as vehicles for promoting behaviour change.

Background

People have goals they want to achieve that require repetitive actions, such as regular exercise or losing weight. Habits can be used to perform these actions with almost no conscious thought in a automatic-like way. Forming a positive habit increases the chance people can achieve these type of goals, by changing their behaviour [2]. Evidence [7] suggests there are three elements of habit formation: repetition, contextual cues and positive reinforcement. Associating the cue with performance and grounding the process with a reward encourages regular repetition, leading to automatic behaviour [8]. Building a new habit requires a contextual cue, to trigger the start of the habit (action), and a reward for positive reinforcement [7, 1]. For example, a reminder on your phone (trigger) reminds you to stretch (action), relaxing you and removing back pain (reward). Studies have shown that the process of creating a new habit takes on average up to 66 days of repetitive use [1]. But, anyone who has ever made a new years resolution knows the difficulties in changing behaviour. People try to create new positive habits only to drop them a few days into their new routine.

Motivation

Habit Formation

Technology can help solve this problem. Coaching us through a new routine until an action becomes a habit. Mobile technology provides us with an interactive platform that can help support habit formation. Issuing reminders and build motivation for repetitive tasks. Plenty of existing habit formation mobile systems use apps to guide users through a series of experiences to form a new habit. But the majority of these apps are unsuccessful because they don't ground themselves in habit formation theory [7, 9]. These apps create a dependency on the technology and don't build the automatic reaction to a trigger (habit automaticity) [10]. For example, the dependence is highlihgted whent he apps are removed as people stop the habit altogether.

Research [11] shows that routine-based remembering strategies are good for building habit automaticity. Stawarz et al. [12] produced a set of design requirements for building mobile apps grounded in habit formation theory that aim to build habit automaticity. Other literature [13] shows how intrinsic rewards build habit automaticity, producing a set of real-world implementation requirements. Combined, these two requirements form a new set focused on rewards. This project uses these to build a tool to help measure the effect of rewards on habit automaticity.

Chatbots

Interaction with current habit formation systems is often via a mobile app. This creates a notable difference in the person when the system is removed [14]. This is also the case with many mobile feedback systems that aid with behaviour change. When we remove the system any improved performance is lost [10, 15].

Chatbots are a method of communicating with a computer system via a conversation using natural language. They hook into messaging services users are familiar with, providing a better mobile phone integration for users. When removing the system, instead of removing an app, users stop messaging a person (the chatbot). This project shows how a chabot to deliver reminders and rewards to users, acts as a novel tool to interact with users.

Reward Modalities

Habit formation systems use reminders and rewards. Studies have shown that good reminder systems should use multiple modalities [16], that is providing alternative ways to interact with the user, either visual, auditory or both visual and auditory combined. This increases the likelihood that the delivery method is pleasant and satisfactory to the user. This project uses this idea for rewards, incorporating this technique into the chatbot by delivering rewards to users across multiple modalities. However, reminders are issued on a single mode (visual) to limit the scope of this project to test how rewards from different modalities effect peoples habit strength.

Aims

This project aims to evaluate how users habit strength is effected by rewards from different modalities using a chatbot as the tool to deliver these rewards.

Objectives

The chatbot provides habit tracking with reminder messages as triggers, and rewards as positive reinforcement in two modalities: visual and auditory.

The rewards provide the user with the satisfaction of completing the habit action and encourage them to build user habit automaticity. The visual reward is a gif and the auditory reward is an audio clip. Each reward mode content can be comparable with another, using mapping the content across each reward modality. For example, a visual picture of a bird would map to the sound (auditory) of a bird.

Methodology

The literature review enabled the construction of theory-based requirements that are focused on rewards for habit formation. Next a tool is built that uses these requirements, to ensure that the system is based on theory. A chatbot is constructed as a tool to track habits, deliver reminders as notifications and deliver rewards from two modalities, visual and auditory. The chatbot is built using a popular messaging platform, Facebook Messenger.

A 3-week evaluation trial tests the success of the chatbot by evaluating the tool and the effectiveness of each modality on users habit strength using a validated questionnaire. Afterwards, chatbot interaction is removed during a 1-week follow up trial to test if users continue with the habit. Participants split into four groups, all groups receiving reminders, three groups receiving rewards each from a different modality, and one group (control group) don't recieve any rewards.

Deliverables

To summarise, these are the following key project deliverables.

- A chatbot to serve as a tool that tracks habits by delivering rewards from two modalities: visual and auditory.
- Analysis about how rewards from different modalities effect habit strength.
- Design recommendations for building chatbots that deliver rewards in different modalities.

Added value

Evaluation from real-users reveals positive and negative aspects of the requirements. The follow-up trail in determines if the requirements were effective for building habit automaticity and tests the validity of the hypothesis. If user's habit automaticity does not increase, the project still presents a novel method of interacting with users and track habits. A system evaluation provides value on how to build a chatbot to deliver rewards from different modalities to support habit formation. Finally, the project opens up new research avenues for investigating the use of chatbots as vehicles for promoting behaviour change.

This thesis analyses literature around habit formation and rewards from different modalities, constructs a prototype to deliver rewards, conducts an evaluation trial and summarising with design guidelines and prototype analysis.

2 Habit Formation

TODO: Add snooze to requirements, look into Kathy research, find research about *why* I should have snoozable reminders

To understand how to build a system that supports habit formation, we must discuss how people fundamentally form habits.

Psychology defines habits as learned automatic cue-response actions, such actions that will perform automatically in response to another action or trigger that has been actioned repeatedly in the past [3]. Studies have shown people must keep to strict strategies and perform an action repeatedly before it turns into a action that occurs with little concious thought [2].

2.1 Three Elements of Habit Formation

Forming a habit occurs similarly to how a person changes their behaviour. A study into how habits are formed [8] shows that using the following three elements ensures an action becomes permanent.

1. Repetition
2. Contextual Cues
3. Positive Reinforcement

Repetition

Lally et al. [1] conducted a test on how long it takes for an action to become automatic, showing that the process of creating a new habit takes on average 66 days of repetitive use. The easier the action, the shorter the before the action turns into a habit, from drinking water (18 days), to going to the gym (254 days) [1]. However, applying the other two elements is needed before the action develops into a habit.

Contextual Cues

Context from information around the action, serves as a cue to trigger events to push the person onto performing the action. For example, if you wanted to adopt a stretching habit, you could attach it onto an existing habit like brushing your teeth. The contextual cue of brushing your teeth will trigger you to stretch.

Behaviour change literature [17] shows that attaching habits onto existing event-based cues are easier to remember, when compared with time-based habits, e.g. stretch every 4 hours. These help connect the contextual information with the action and builds habit automaticity [18]. Further research into the design implications of contextual cues shows how multi-cue routines are more effective that a single cue [11].

Positive Reinforcement

Self-efficacy, the belief in ones ability to succeed, plays a large part in forming habits and is a main part of behaviour change [4]. Rewards give people this experience by feeding back their success of their action. Rewarding a person with positive reinforcement, strengthens the habit by giving the

feeling of satisfaction [2]. However, the type of reward matters, as rewards that benefit the person with satisfaction (intrinsic rewards) should be used over monetary gains (extrinsic rewards), due to issues with extrinsic rewards hindering motivation [5].

2.2 Technology

Research into how mobile systems can support habit formation and behaviour change, shows a large number of habit forming systems are mobile apps. Studies into the effectiveness of these apps has been recently conducted [7, 10] revealing that although most of these apps are rated highly, they do not ground themselves in behaviour change theory. Research into some of these apps show that habits are not sustained when the app is removed, due to the lack of habit automaticity [7].

Two pieces of literature discuss different concrete strategies for building habit forming mobile apps that do ground themselves in theory [19, 13]. Stawarz [19] presented formal requirements for building habit forming apps, based on the above three elements of habit formation with the aim to build habit automaticity. Weiser et al. [13] shows that 'motivation is a key requirement for behaviour change' presenting five design principle requirements and six requirements about the implementation mechanics of habit forming systems that focus on rewards and motivational needs. This project builds upon this set of requirements, combing them into a new set of design requirements for habit formation systems that focuses on rewards.

2.3 Combined Requirements

Combining information from the above two sources [19, 13], produces the following list of methods to build a habit forming app.

The system must:

1. Help users define a memorable strategy
2. Give users small difficult tasks (challenges)
3. Give them Competition & Comparison & Cooperation
4. Give insights for strategy improvements and support changes
5. Remind them about cues and remembering strategies
6. Reward Users
7. Disable cue reminders when behaviour is routine
8. Check if the action has already happened

The following table shows how the two sources [19, 13] were combined.

REQ	Stawarz. [19]	Weiser et al. [13]	Combined Requirement
1.	REQ 1. Help users define a good remembering strategy. REQ 2. Provide examples of good remembering strategies	Design REQ 2. Support User Choice. Design REQ 4. Provide personalized experience.	Help users define a memorable strategy.
1a.	NA	Mechanic REQ 3. Offer challenges	Give users small difficult tasks (challenges).
1b.	NA	Mechanic REQ 5. Competition & Comparison. Mechanic REQ 6. Cooperation	Give them Competition & Comparison & Cooperation.
2.	REQ 3. Provide suggestions for strategy improvements and support changes.	Design REQ 1. Offer meaningful suggestions. Mechanic REQ 2. User Education.	Give insights for strategy improvements and support changes.
3.	REQ 4. Remind about cues and remembering strategies.	Design REQ 3. Provide User Guidance. Mechanic REQ 2. User Education.	Remind them about cues and remembering strategies.
4.	NA	Mechanic REQ 4. Rewards.	Reward users.
5.	REQ 5. Disable cue reminders when the behaviour becomes a part of a routine.	Design REQ 5. Design for every stage of behaviour change.	Disable cue reminders when behaviour is routine.
6.	REQ 6. Help users check whether the habit has already happened.	Mechanic REQ 1. Feedback.	Check if the action has already happened.

Figure 1: The combined table of requirements for designing mobile systems focused on rewards.

2.4 Requirements Overview

The combined requirements based on methods from [19, 13] create a list grounded in habit formation theory and focused on rewards. Each requirement provides detailed breakdown about why it's used and what mechanics it relates to from theory.

1. Help users define a memorable strategy

- Make personalized, well defined, structured multi-cue routines & also support users choice of not setting remembering strategies.
- Provide examples of some strategies to users.

1a. Give users small difficult tasks (challenges)

- Assignments: Turn the bigger habit into smaller assignments to make it more enjoyable, being careful to not make them forced.
- Quests: Same as assignments, but optional.
- Goals: User specified to support user autonomy. Should be specific and challenging to get better results.

1b. Give them Competition & Comparison & Cooperation

- Friends, teams & groups.
- Leader-boards and collections.

2. Give insights for strategy improvements and support changes

- Give users meaningful instant feedback & accumulated feedback based on their system usage.

3. Remind them about cues and remembering strategies

- Reminders can effectively support prospective memory in the short term.
- Educate them about what they should perform.

4. Rewards

- A good form of external motivation because they don't change the ability to perform a behaviour, unless the reward itself is a tool that increases ability.
- Provide a strong motivational source, but like all extrinsic motivators, these are less effective for changing behaviour in the long run, because externally motivated behaviour lasts as long as the external motivator exists.
- Identifying methods that enable internalisation of externally motivated behaviour.
- Achievements and badges.

5. Disable cue reminders when behaviour is routine

- Relying on reminders in the long term can hinder habit development.
- Ease off from reminders later.

6. Check if the action has already happened

- Easy to forget whether a task done automatically, completed.

Summary

These requirements act as the basis for this project. They are guidelines for building habit formation systems that focus on delivering rewards to build habit automaticity and motivation. This project tests how different types of rewards from different modalities effect users habit strength.

3 Modalities

Habit formation research [11] shows how different contextual cues better support behaviour change, compared with a single mode. If we combine this with research [20] into crossmodal interaction, i.e. the process of signals we receive through a single sense affecting how we process information perceived through a different sense. We can map habit rewards across different modalities, enabling us to present users with the same reward type from a different modality. This allows us to test the different types of modalities and how they effect behaviour change. The next section discusses methods of interaction in practice.

3.1 Why Interaction from Different Modalities

'Multi-modal systems are required for user interaction' [6] states research into comparing unimodal reminders systems with multimodal. They suggests a need for 'highly flexible and contextualised multimodal and multi-device reminder systems' [6]. Although this study focused on the elderly, so the need to multiple modalities was important because some peoples sensory modalities decline with age, this principle still holds true for general case reminder systems. The study presents design guidelines for reminder systems. These are mainly focused on users needing a choice of modalities for interacting with users, as users want a highly configurable system. These aspects are implemented into our project and adapted for delivering rewards.

Research into designing reminder systems for the home that interact with users in different modalities, states that 'Good reminder systems should use different modalities, because they provide alternative ways to interact with a user' [16]. Using different modalities for interaction increases the likelihood that the information users are receiving are more pleasant to them, and decreases the chance the interaction will be disruptive or annoying [16]. Habit rewards should not be annoying, they should give users a feeling of satisfaction, therefore reducing the chance of disruption is another justification for using multiple modalities.

3.2 Habit Rewards

Why habit rewards?

3.3 Modality Types

Next we look at literature discussing the three main modality types and how they change peoples behaviour.

Visual

One study looked at improving habit consistency for how often patients took medication, by using a visual display device that gave constant feedback [15]. They found that this feedback improved consistency of the habit and increasing rating of self-efficacy. But when the device was removed, their performance dropped (from a 2-month follow up study), because users integrated the feedback display with their routines. This habit-forming system used visual feedback to encourage consistency, however, this system shouldn't integrate visual cues into the system, otherwise users become dependent on the technology. Users should instead build these cues outside of the system to build performance longevity after removing the system.

Auditory

Another key paper, discussed their need for different modalities when designing for the elderly [21], combing different sounds with high visual contrast to suit their needs, given deteriorating senses due to age. The study showed that using different modalities for interaction gives a means of communication to people with varying levels of sense ability. But studies have also shown people need a choice of mode when designing for interaction [6], and thus the design requirements produced from this study can be applied to general applications that use multiple modalities, such as this project. The design guidelines discuss the need for interaction consistency, such as using similar audio interaction. Therefore visual and auditory rewards are mapped identifying a pattern across the modalities.

Vibration

The majority of electronic activity monitors have behaviour change techniques and these monitors present a medium which behaviour change interventions could occur [22]. This provides us with a final modality to explore reward delivery techniques, implemented with a wearable device using vibration. A survey on activity monitors [22] ranked Fitbit devices 'Good vehicles for behaviour change techniques'. Thus, the Fibit, would be a good primary platform for integrating vibration for rewards. However, due to technical limitations, as discussed in (TODO reference tech limitations chapter), we are unable to implement them.

3.4 Reward Types

Rewards are delivered from each modality, with the content based on requirements created in (TODO reference chapter 2). Visual rewards present the user with a gif to give them satisfaction after completing the action and auditory rewards provide a similar result but via the auditory mode. The next section discusses design recommendations for delivery methods for these rewards.

3.5 Pairing Reward types

Talk about how you can match Visual with sound, and breifly vibration. Disucssion about types of rewards.

A Modal, setting the same frequency for sound and vibration, look into how mapping patterns TODO.

Experiment with gif to audio to vibration

4 Prototype Design

To verify the effectiveness of each modality, an experiment was needed to test a reward on a large sample. To accompany the experiment a tool was constructed to send users rewards from different modalities and help users track habits. The platform for the tool needed to be highly available for participants, interactive and time effective to build. A system on a mobile device grants us access to a 'highly available, contextually aware and interactive platform' [23] and based on previous design recommendations for behaviour change mobile systems.

4.1 Design Considerations

Current mobile systems use apps to interact with users, but a recent study showed apps fall into a low behavioural theory adherence scale [23]. Research into how to build systems that are theory-based, suggest four main stages for designing mobile health solutions. Conceptualisation, Formative Research and Pretesting, Pilot trials and Evaluation trials [23]. The first two stages use the Behaviour Change Wheel [24] (a framework for planning health behaviour interventions), to understand the behaviour, better define the characteristics and form the mobile concept into a prototype. The pilot trials tests the prototype before it moves to a finalised app with the commitment of a full trial. The final stage tests the finalised app with a wider range of participants. This project uses these four stages as design steps to build the prototype.

4.2 Mobile Concept to Prototype

This first stage explores different mediums to develop a prototype upon and to decide if there is a better method to interact with users than a mobile app. We conclude with a series of implementation options and choosing the path of a Chatbot.

When it comes to mobile phones, users have plenty of options for interaction. A popular choice that has revisited the market are Chatbots—applications that parse questions using Natural Language Processing (NLP) to provide a response. Bots act as a user interface to expose data and would use online services to parse the response, such as Amazon lex `https://aws.amazon.com/lex/`. These programs have conversations with users to achieve a goal and are not new inventions. Since 1966 [25], Eliza by Joesphs Weizenbaum, used simple expression matching to return a certain response for user trials. In the present day, these applications (commonly referred to as bots, sometimes chatbots), are found integrated into many different apps on the majority of users mobile phones. For example, Facebook Messenger (a popular messaging application) encourages developers to create bots to interact with their users. These bots act as a real person with similar interaction flow, plus a few additional features, such as *Quick Replies* for revealing a list of options to a user. 'Quick Replies provide a way to present buttons to the user in response to a message.' [26]. However, these bots would not reply like a real person, but rather would only reply if that question was pre-trained using machine learning algorithms. This technology requires the bot to be trained on a large set of data and the majority of use cases would have to be accounted for.

NLP would enable users to chat to the bot and get a friendly understandable reply. But, would this interaction develop a dependence on the user-chatbot interaction? Would it lead to losing automaticity if the chatbot supported habit formation? NLP won't be used to avoid these potential problems, and instead of natural language processing, the location of the bot (i.e. inside of an existing messaging app), ease of interaction and the additional features (quick replies) helps us easily communicate with a user.

Another option is a Web App. But a Web App cannot send users reminders, unless it is paired with an app or SMS platform, but this is hard to get a response from a user.

The below table summarises the available choices for developing a system for our requirements (TODO cite chapter). A mobile app can supply notifications, but for each platform a completely separate app would need to be built and users would need to download the app before it would be available to them.
A single cross-platform app could be constructed to reduce development time and complexity, but still users would need to download the app to start using it.
A web app has the advantage of being available to all users with a web browser (with users being able to save the site to home screen), but without notifications on all platforms (iOS), it won't meet our requirements.
Finally, a chatbot integrated into a popular messaging platform is easily available (if you have the messaging app already installed), simple (the user interface is already supplied), works on any platform the messaging app is available on and has notifications built in.

	Mobile App	**Cross-Platform App**	**Web App**	**Chatbot**
Notifications	✓	✓	✗	✓
Development Time	Long	Medium	Short	Short
High Availability	✗	✗	✓	✓
Simplicity	✗	✗	✓	✓

Figure 2: Comparing different prototype mediums.

From the table above the best method forward is a chatbot. It is already integrated into an existing app users are familiar with and (as long as the messaging app is available) works on all platforms.

4.3 Chatbot Platform

Again, similarly there are lots of options about what platform to build the chatbot into. For example, Slack bots provide additional functionality, such as habit track, into the popular workplace communication service.

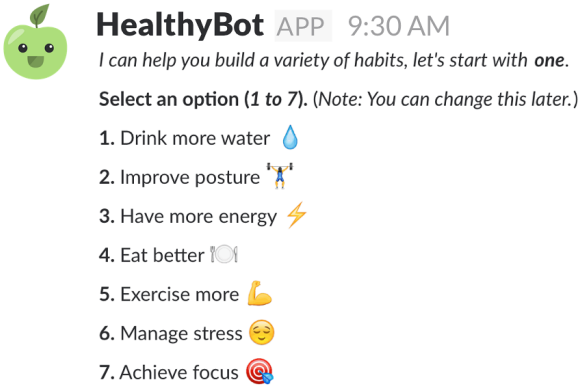

Figure 3: Healthy Bot: A Slack bot for forming new positive habits.

However, if our main aim is to interact with lots of people easily, we need to target existing platforms

that a lot of people are using.

	Facebook Messenger	WhatsApp	SMS	Telegram	Slack
High Availability	✓	✓	✓	✗	✗
Interactive	✓	✓	✗	✓	✓
Additional Features	✓	✗	✗	✓	✓

Figure 4: Comparing different chatbot platforms.

Facebook Messenger looks like the attractive option for user interaction with the ease of additional features, such as Quick Replies and with the benefit of:

- 1,200 Million active users per month (as of April 2017) [27]
- Embedded into a service users already use
- Quick replies allows for easy interaction

But, will users differentiate between a bot and a person? Will the text interaction put people off? This thesis aims to answer those questions.

4.4 Facebook Messenger

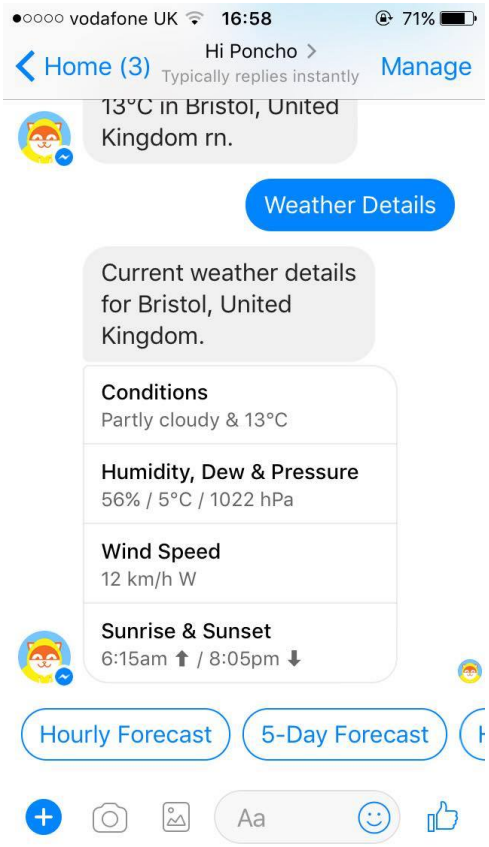

Figure 5: Poncho: An example of Facebook Messenger Weather Chatbot

Discussion about the additional features fb messenger has with discussion about previous chatbots.

¡photo of UI flow¿

¡photo of quick replies¿

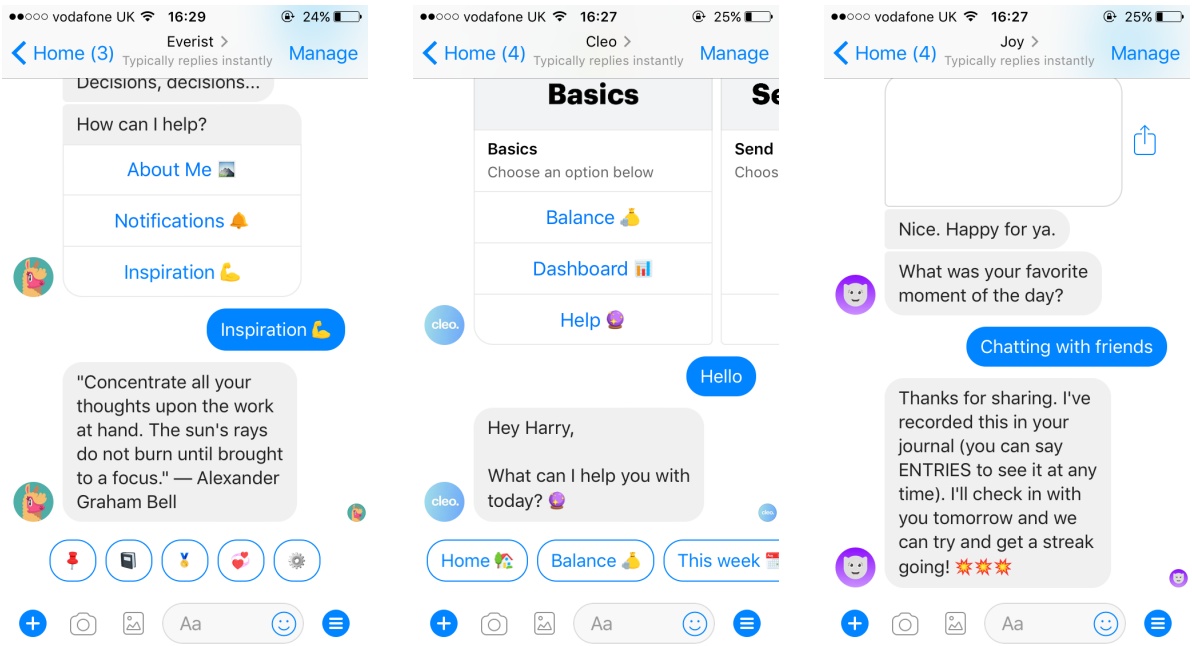

Figure 6: Examples of chatbots performing different actions

4.5 Supporting Habit Formation

Discussion about how we can support habit formation using a chatbot in reality.

Paragraph about Gamification elements: - Gamification elements [28] - Designing outstanding feedback loops [29]

Justify the type of habits chosen: Because the evaluation trial is shorter, the habits must be easier - as the easier it is, the quicker it will form, if you can easily do it, you are more likely to do it

4.6 Delivering Rewards

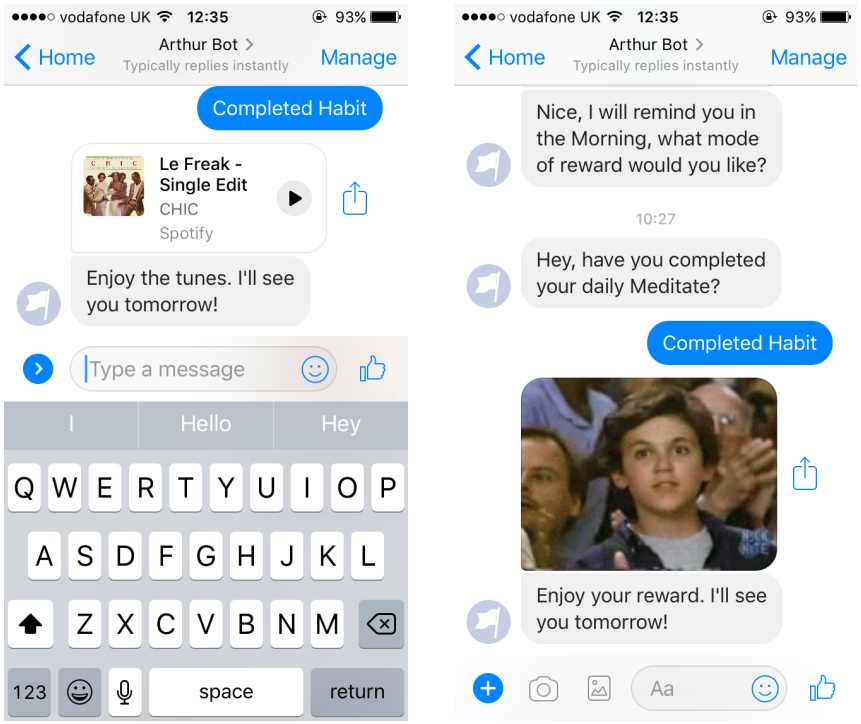

Figure 7: In-line rewards.

Discussion about how the rewards will be delivered to the users. In-line verses same screen. How interaction is handled. Concludes with exactly how rewards will be handled in design way, brief discussion about vibration and how it would not work.

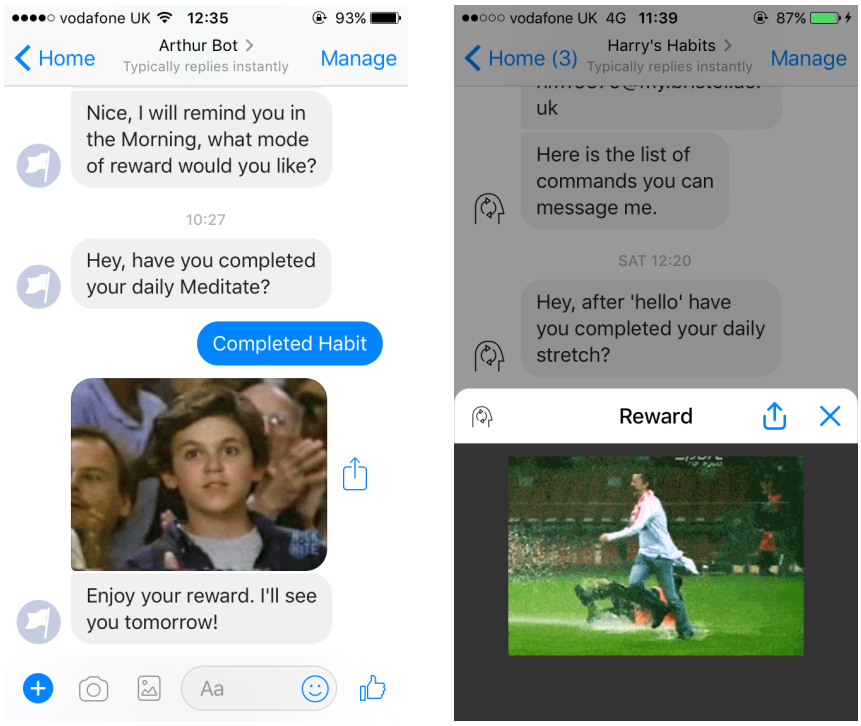

Figure 8: In-line visual rewards verses separate.

The types of rewards are separated into three categories, visual, auditory and combined visual-auditory. Different types of these rewards will be experimented with, and test if they provide user satisfaction. These rewards will be displayed to the user within the chatbot after they complete their habit. For a playful prototype, visual rewards will be light-hearted gifs and auditory rewards were selected to match the gifs.

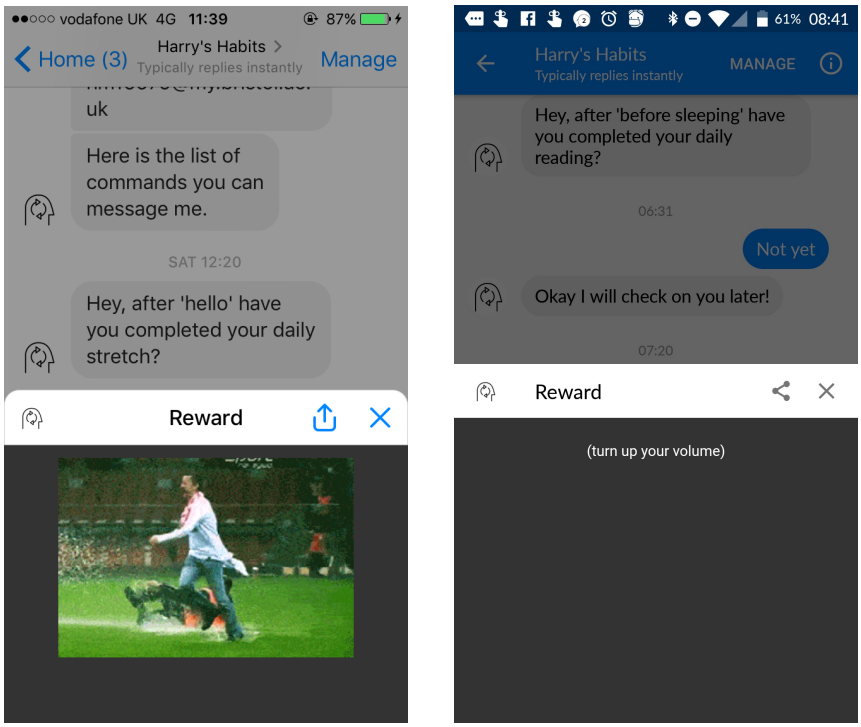

Figure 9: Consistency between every reward, even across devices.

To ensure consistency with the reward types, a standard method was deigned to deliver the reward to users. Instead of sending the rewards in-line with the messages, a *webview* was used to display a website where users opened their reward. This meant we could get out of the sandbox that the bot lives in, and use HTML5 elements to display the content in the same way, ensuring consistency across devices. Although this was not without its limitations.

4.7 Design Limitations

Auto-playing an audio track was not available when sending the audio in-line. But using a webview we could use the *HTML5* audio element to enable auto-play. But, for auto-play elements, the HTML5 standard needs a button press before it starts (TODO footnote, this is for a good reason, to stop people spamming you with audio when visiting sites, see more mozilla reference autoplay). This required another button to create a JavaScript hook to auto-play. However, during tests on low mobile data speeds, users found that they would have to press the button multiple times before the audio played. This was because the audio would only play after it had loaded and created a lengthy delay, along with seemingly broken display. To create a better user experience, the button disappeared when pressed, using CSS that would execute even if the audio hadn't loaded, and even on a poor connection. Then JavaScript would execute after the page had fully loaded and play the audio if the button had already been pressed from the CSS, but if it hadn't then it would create a hook to play the audio after it had been pressed. This ensured a seemly experience when using rewards for all levels of connection.

4.8 User Flow

Full user interaction from setup to completing a reward.

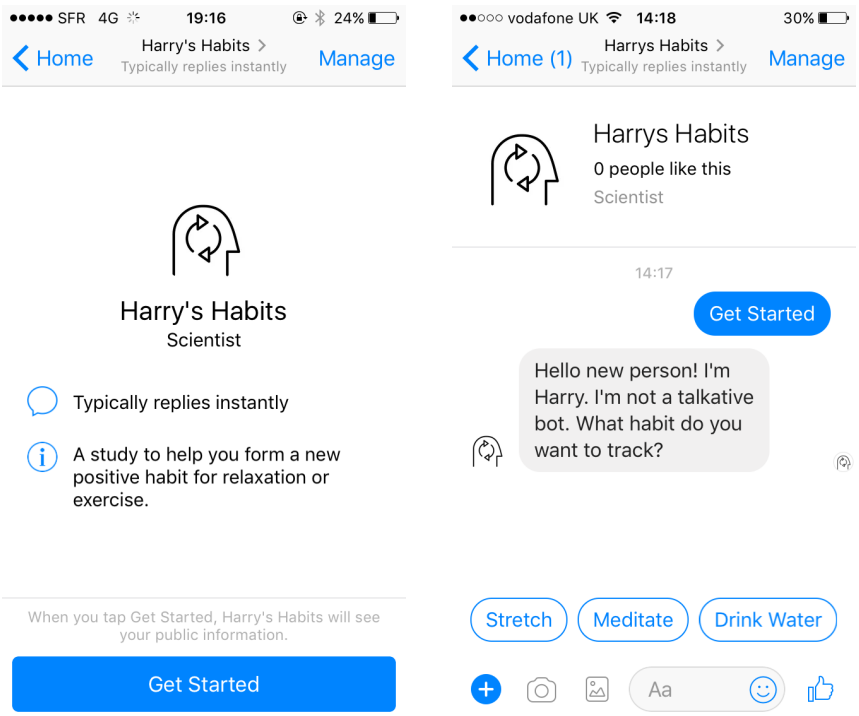

Figure 10: The screens initial users will see.

- **Setup**
  - Setup the bot via the Facebook Messenger platform.
  - User chooses an existing habit they would like to develop.
  - User supplies an existing context to integrate their new habit into.
  - User supplies a time the existing context occurs.
- **Trigger**
  - At the time of the existing context, the user should perform their chosen habit.
  - The user receives a notification after the chosen time, that asks if they managed to complete their habit, or if they need more time.
  - If they need more time, the notification will *snooze* for about an hour and be sent again.
  - If users regularly snooze they will be asked if their existing context time would like to change.
  - If users say they have completed their habit, they will receive a reward.
- **Reward**
  - Users will press a button that will take them to a website so for each modality they can consistently experience their reward.
  - User will receive a reward (which represents success) from one of the following modalities:
  - Visual: A video with no sound

– Audio: A soundtrack

– Visual and Audio: A video with sound

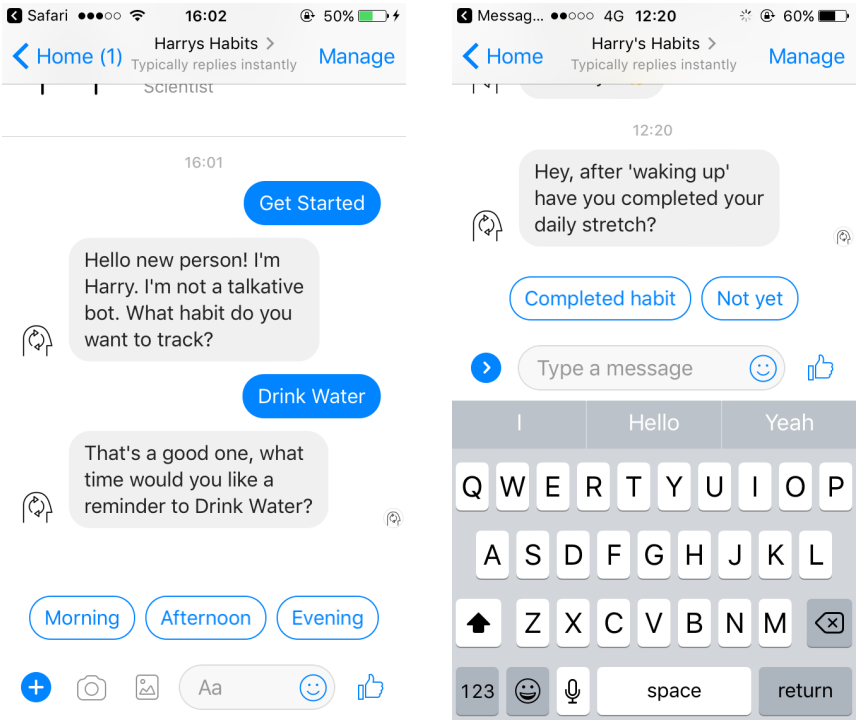

Figure 11: Example of the setup flow and a habit reminder.

Cite logo [the Noun Project by Yu luck]`https://thenounproject.com/term/custom/402041/`

5 Prototype Implementation

Abstract implementation overview of user and chatbot discussion.

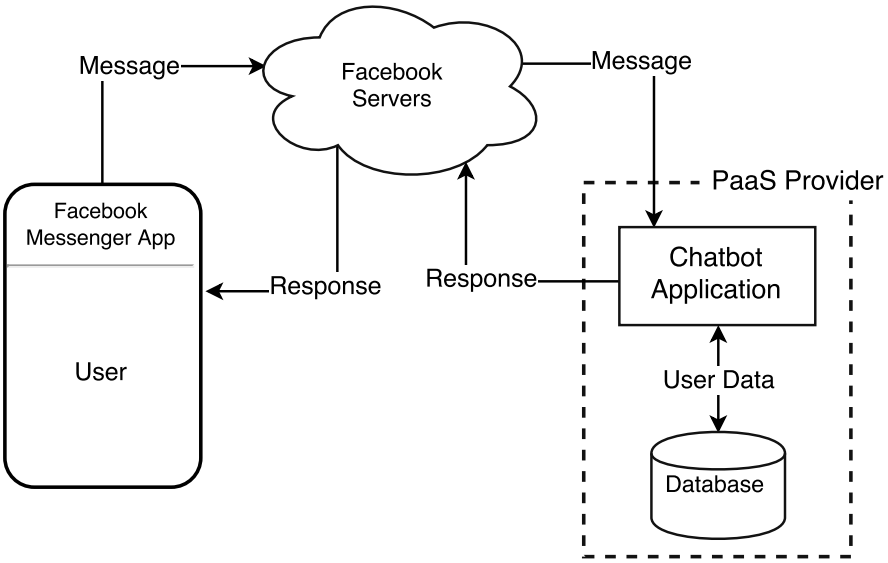

Figure 12: Prototype Component Overview

5.1 Platform

Language consideration, nodejs, java, other types, talk about heroku, hosting provider. Hosting provider. Database provider discussion, integration, airtable,

5.2 Detailed Overview

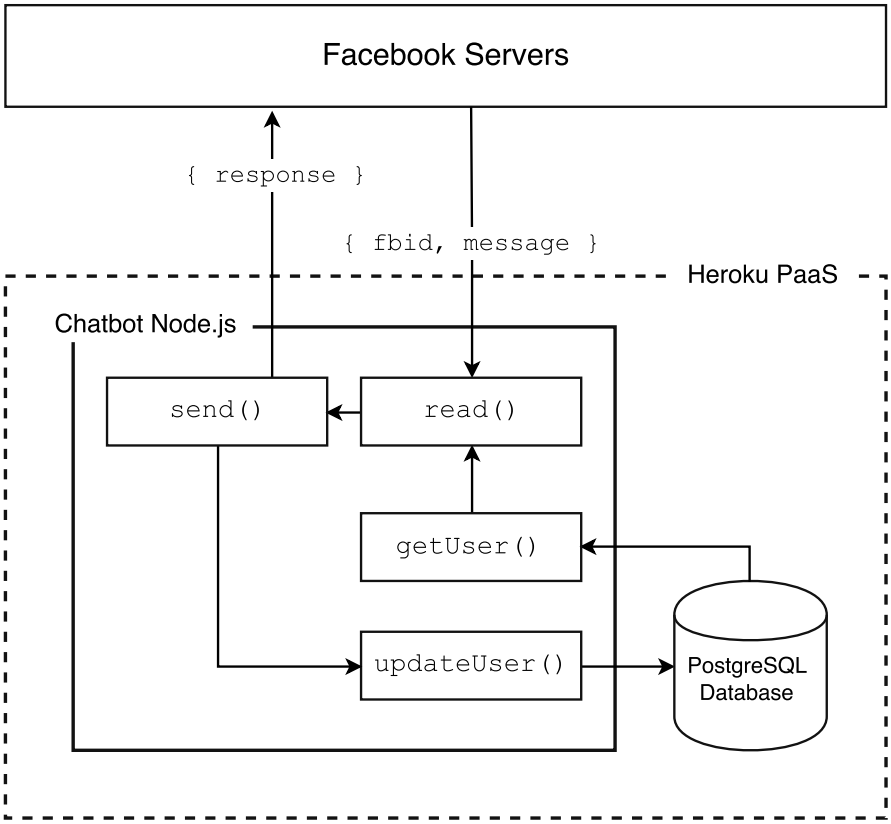

Figure 13: Detailed Overview

Scalable architecture because heroku

Implementaiton limitations about timezone with the scheduler
Scheduler
Free text input issues

5.3 Implementation Issues

(Go through all meetings and list implementation issues here.)

Throughout the implementation process different techniques were explored to implement the design. Some of the research areas were not used in the final prototype due to technical issues and limitations with the approach.

Using vibration as a modality would've been a great additional modality to use. Unfortunately the chatbot sandbox meant that the vibration ability in the phone could not be used, so another device would be used in combination with the bot. Smart watches and fitness trackers were researched to test if they could programatically vibrate, with the pattern of vibration matching the frequency of

the audio. However, the majority of these devices did not have an API that exposed the vibration element. The best method was found to programatically set an alarm 1-minute into the future using a Fitbit fitness tracker. This would trigger the vibration when the alarm sounded. Although this would mean a 1-minute delay after completing a habit, a good user flow could've reduced the wait time with some additional dialogue. But, this approach relied on the fitness tracker to sync with the phone after the alarm was programatically set. Unfortunately forcing the tracker to sync wasn't available, so this modality was abandoned.

Another issue occured with stopping the audio after it had been played during a reward. If a user closed the reward box, there was no way to stop the audio, unless a user waited until it had finished. This limitation was very minor, but also showed how difficult it is to seemly connect a webview and a bot.

5.4 Testing the chatbot

Test harness, tested the full functionality programatically. Hooking into Travis CI and other continuous integration services. A pilot trial tested the basic chatbot functionality. Preparing for the full evaluation trails.

6 Evaluation

Many people agree about the importance of designing systems for health and behaviour change [23, 30, 31]. But each have varying opinions about how to evaluate these systems. Klasnja et al. [32] focuses on system usability and does it meet the needs of users. Whereas, Stawarz and Cox [31] argue evaluating a system of this type requires information from other fields to properly consider the systems effectiveness. The validated Behaviour Change Wheel Framework [24] does just this. Evaluating the system with validated behaviour change techniques from multiple domains. This project will use this framework to evaluate the chatbot with evaluation trials. These will test the long-term effect and efficiency of the bot, with information from two fields of study, HCI and health psychology.

6.1 Evaluation Trial Overview

Evaluation trials are the final part of the Behaviour Change Wheel Framework [24]. HCI research that focuses on health interventions [23], demonstrates the importance of evaluation trials for evaluating mobile health systems. These trials have three goals to test: objective-quantitative efficacy, subjective-qualitative feedback measures and real-world feedback about how the system is utilised [32]. I will conduct an evaluation trial for this project.

The length of the trial will be based on two factors, the time needed to form a habit [1] and the results of a previous habit formation trial [7]. First, the number of repetitive days required for an action to be considered a habit varies based on the complexity of the action [1]. Simple actions, such as drinking 2 glasses of water a day, can take a minimum of 18 days to form. The suggested actions used for this project will be considered as simple, e.g. stretching for 30 seconds. Second, a previous evaluation trial on habit-formation systems [1] showed an increase in habit automaticity after 4 weeks. This project will mirror that timeframe.

A 4-week evaluation trial will test the success of the chatbot by evaluating the tool and the effectiveness of each modality on users habit strength. Chatbot interaction will be removed during the follow up study to test if users continue with the habit. Participants will split into four groups, all groups will receive reminders, three groups will receive rewards each from a different modality, and one group (control group) will not recieve any rewards.

6.2 Testing Habit Strength

Habit strength will be measured using a validated 12-question questionnaire that specifically looks at automaticity [33]. Automaticity will also be measured using a validated subset of the questionnaire from [33] to test users habit behavioural automaticity index [34]. This will show the impact each modality has on habit automaticity and test the hypothesis. Participants will fill out the questionnaires [33, 34] at three stages: half-way through the trial (at 2-weeks), after the trial has finished and after the follow up trial.

6.3 3-Week Trial

Discussion of what a participant would actually do in this 3-week trail.
Bullet point list based on beyond self tracking paper
Ethics approval, Kathys PHD 5.3.
How participants were recruited, screenshots of all adverts
Creating the adverts

Photo of website
Photo of Screens users would see, facebook post, website, get started, Setup (Link to full setup screens in appendix), recieving reminder message, reward, snoozed lots messages, end of study messages.

6.4 1-Week follow up trial

Discussion of what the expected behaviour of the participant should be, bullet point list.
Transcript/screenshot of end of study messages.
Talk about asking for interview.

6.5 Interview

General overview of interview questions.
Informal chat.
Questions: - General questions firsts - Habit formation insight - Chabot interaction - Modality interaction - Are you still doing that habit? - Why did you pick that habit? - Do you want it back?

7 Results

General discussion about how the results were analysed. Technology used. Security.
'Does giving users too much personalisation effect the study?'

7.1 General Discussion

General results analysed, with discussion. Lots of charts and graphs.

- Feedback - Add a 'before' or 'after' option to avoid sythetic awkwardness (erasmo) - Liked music rewards. - If you tell the bot they haven't done their habit after their context, when they snooze it isnt going to be their context anymore!!! e.g. lunch at 5pm. The solution is to remove habit context from snoozed reminders

OLD STATS

- 58 Unique users
- 19: iOS
- 13: Mac
- 13: Android
- 7: Windows
- 5: Linux
- 5.4k total messages sent and recieved by the bot
- Average ¿100 messages sent/received per each user
- 55 users pressed got started (3 users didn't press get started)
- 38 people completed their setup (17 People failed to setup)
- 63%/24 Male, 32%/12 Female, 2 didnt say
- More people snoozed EARLY_MORNING (16), MID_MORNING (38 snooze presses),

LATE_AFTERNOON (14) and LATE_EVENING (17 snooze presses) - Most people chose Meditation as their habit (12 people), followed up Press ups (9) and
stretching (8) - Average age was 26.6 (3 outliers of 57, 63 and 53. Removing these gives us 23 avg)
- Only 7 people have used habit apps before
- Most people set a reminder time for the evening (21), most LATE_EVENING (10)
- Most people snoozed PRESS_UPS (49), MEDITATION (41) and STRETCH (22)
- Users combined streaks were higher for MEDITATION and PRESS_UPS (both 18 total)
- Highest streak was 10 (stretching)
- 18 daily active users :/
- Talk about male female split

7.2 Key results

Key questions answered, with discussion about what the key questions mean.

Was the chatbot successful at running a trial?
Did users habit automaticity increase after using the chatbot?
Did users habit automaticity increase when using a specific reward?

7.3 Issues

Issues with study results and general issues with implementation.

Discussion about the issues I had w the bot

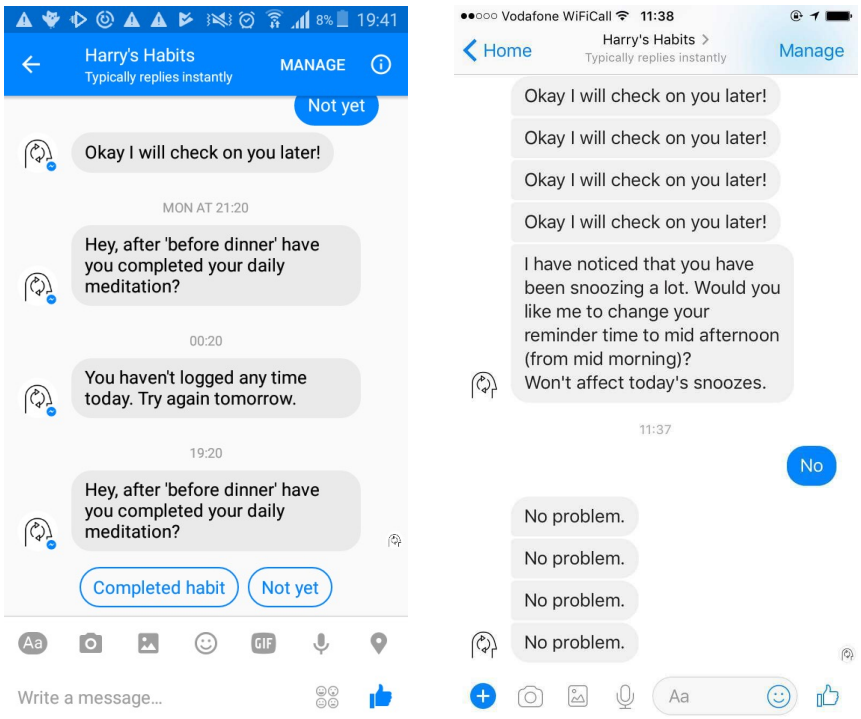

Figure 14: Example of 2 issues that occured, habit context using the word 'after' and the bot sending multiple of the same messages.

7.4 Summary

Are the results even valid? What did we find out about the trail? Did we answer our research question?

8 Conclusion

The survey on the relevant literature demonstrates two sets of design requirements for building systems that support habit formation. In particular building habit automaticity is key to building lasting behaviour change and multi-cue routines better support behaviour change than single cues. Rewards from different modalities are key to building habit automaticity with interaction from different modalities increasing user interaction. These combine to form new requirements that focus on delivering rewards from different modalities.

A chatbot is designed and constructed from these requirements to track habits and deliver rewards to users from different modalities. A 3-week evaluation trial tested the effectiveness of the chatbot implementation, the design requirements and if bots are a good method for collecting research data. Finally a 1-week follow up trial tested user's habit automaticity.

Evaluation from real users revealed important positive and negative aspects of the requirements. Following up after the study determined if the requirements were effective for building habit automaticity and testing the validity of our hypothesis.
The project found ....

Therefore, the project presents a novel method of interacting with users to track habits and a system evaluation provides value on how to build a chatbot to deliver rewards from different modalities to support habit formation. Finally, the project opens up new research avenues for investigating the use of chatbots as vehicles for promoting behaviour change.

9 References

[1] P. Lally, C. H. M. van Jaarsveld, H. W. W. Potts, and J. Wardle, "How are habits formed: Modelling habit formation in the real world," *European Journal of Social Psychology*, vol. 40, no. 6, pp. 998–1009.

[2] P. Lally and B. Gardner, "Promoting habit formation," *Health Psychology Review*, vol. 7, no. sup1, pp. S137–S158.

[3] W. Wood and D. T. Neal, "The habitual consumer," *Journal of Consumer Psychology*, vol. 19, no. 4, pp. 579–592.

[4] A. Bandura, "Self-efficacy: toward a unifying theory of behavioral change," *Psychol Rev*, vol. 84, pp. 191–215, Mar 1977.

[5] R. R. R. M. Deci, Edward L.; Koestner, "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation.," *Psychological Bulletin*, 1999.

[6] M. R. McGee-Lennon, M. K. Wolters, and S. Brewster, "User-centred multimodal reminders for assistive living," pp. 2105–2114, 2011.

[7] K. Stawarz, A. L. Cox, and A. Blandford, "Beyond self-tracking and reminders: Designing smartphone apps that support habit formation," *Proceedings of the 2015 Conference on Human Factors in Computing Systems (CHI-2015)*, pp. 2653–2662, 2015.

[8] P. Lally, J. Wardle, and B. Gardner, "Experiences of habit formation: A qualitative study," *Psychology, Health & Medicine*, vol. 16, no. 4, pp. 484–489. PMID: 21749245.

[9] L. T. Cowan, S. A. V. Wagenen, B. A. Brown, R. J. Hedin, Y. Seino-Stephan, P. C. Hall, and J. H. West, "Apps of steel: Are exercise apps providing consumers with realistic expectations?," *Health Education & Behavior*, vol. 40, no. 2, pp. 133–139. PMID: 22991048.

[10] I. Renfree, D. Harrison, P. Marshall, K. Stawarz, and A. Cox, "Don't kick the habit: The role of dependency in habit formation apps," pp. 2932–2939, 2016.

[11] K. Stawarz, M. D. RodrÃguez, A. L. Cox, and A. Blandford, "Understanding the use of contextual cues: design implications for medication adherence technologies that support remembering," *DIGITAL HEALTH*, vol. 2, p. 2055207616678707.

[12] K. Stawarz, A. L. Cox, and A. Blandford, "Don't forget your pill!: Designing effective medication reminder apps that support users' daily routines," pp. 2269–2278, 2014.

[13] P. Weiser, D. Bucher, F. Cellina, and V. De Luca, "A Taxonomy of Motivational Affordances for Meaningful Gamified and Persuasive Technologies," *3rd International Conference on ICT for Sustainability (ICT4S)*, vol. 22, pp. 271–280, 2015.

[14] L. Ventä, M. Isomursu, A. Ahtinen, and S. Ramiah, ""my phone is a part of my soul” - how people bond with their mobile phones," pp. 311–317, 2008.

[15] M. L. Lee and A. K. Dey, "Real-time feedback for improving medication taking," pp. 2259–2268, 2014.

[16] J. R. Williamson, M. McGee-Lennon, and S. Brewster, "Designing multimodal reminders for the home: Pairing content with presentation," pp. 445–448, 2012.

[17] B. Verplanken, "Habits and implementation intentions," in *The ABC of behavioural change.*, pp. 99–109, Elsevier.

[18] P. M. Gollwitzer, "Implementation intentions: strong effects of simple plans.," *American Psychologist*, vol. 54, no. 7, p. 493.

[19] K. M. Stawarz, *Towards better medication adherence apps: Preventing forgetfulness by facilitating the formation of routine-based remembering strategies.* Doctoral thesis, UCL (University College London.

[20] K. K. Evans and A. Treisman, "Natural cross-modal mappings between visual and auditory features," *Journal of Vision*, vol. 10, no. 1, p. 6.

[21] R. de Oliveira, M. Cherubini, and N. Oliver, "Movipill: Improving medication compliance for elders using a mobile persuasive social game," pp. 251–260, 2010.

[22] M. B. R. J. Lyons EJ, Lewis ZH, "Behavior change techniques implemented in electronic lifestyle activity monitors: A systematic content analysis," *J Med Internet Res 2014;16(8):e192.*

[23] A. Smith, K. de Salas, B. Schüz, S. G. Ferguson, and I. Lewis, "mhealth intervention design: Creating mhealth interventions for behaviour change," pp. 531–536, 2016.

[24] S. Michie, M. M. van Stralen, and R. West, "The behaviour change wheel: A new method for characterising and designing behaviour change interventions," *Implementation Science*, vol. 6, no. 1, p. 42.

[25] J. Weizenbaum, "Eliza - a computer program for the study of natural language communication between man and machine," *Commun. ACM*, vol. 9, no. 1, pp. 36–45, 1966.

[26] Facebook, "Quick replies," *Facebook for developers*, 2017.

[27] statista.

[28] W. Luton, *Free-to-Play: Making Money From Games You Give Away.* Pearson Education, 2013.

[29] Author, "How to design outstanding feedback loops," *Smashing Magazine.*

[30] S. Consolvo, P. Klasnja, D. W. McDonald, and J. A. Landay, "Designing for healthy lifestyles: Design considerations for mobile technologies to encourage consumer health and wellness," *Foundations and Trends in HumanComputer Interaction*, vol. 6, no. 34, pp. 167–315.

[31] C. A. L. Stawarz, K., "Designing for health behavior change: Hci research alone is not enough.," *Crossing HCI and Health: Advancing Health and Wellness Technology Research in Home and Community Settings, CHI 2015 Workshop.*

[32] P. Klasnja, S. Consolvo, and W. Pratt, "How to evaluate technologies for health behavior change in hci research," pp. 3063–3072.

[33] B. Verplanken and S. Orbell, "Reflections on past behavior: A self-report index of habit strength," *Journal of Applied Social Psychology*, vol. 33, no. 6, pp. 1313–1330.

[34] B. Gardner, C. Abraham, P. Lally, and G.-J. de Bruijn, "Towards parsimony in habit measurement: Testing the convergent and predictive validity of an automaticity subscale of the self-report habit index," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 9, no. 1, p. 102.