Project Part 1 – Proposal & Usability Test Report

CMPT 363 @ Simon Fraser University
Spring 2024

Team 9
Michael Kwan (301541356)
Ryan Vincent Mitchellin (301540810)
Joel Tai (301347412)
Cyrus Ng (301553946)
Huy Bui (301575130)

****Note that with the cover page and table of contents, this document should not exceed 11 pages ****

Adapted from Usability Test Plan Template from Usability.Gov
(https://www.usability.gov/how-to-and-tools/resources/templates/usability-test-plan-template.html)

Table of Contents

Part 1: Proposal

Target AI-based Application

Our target AI application for this project is Siri. While it certainly has some shortcomings (as will be discussed later), Siri is overall a readily available tool (at least for Apple users) and can perform many tasks without the user ever having to pick up their device. For users with multiple Apple devices, Siri's deep integration with the ecosystem allows it to function across these devices seamlessly. With the ability to program your own Siri shortcuts and workflows, Siri is more capable than ever and comes pre-installed on most Apple devices.

Competitors

Two notable competitors in this space are Google Assistant and Amazon Alexa. Feature- and reliability-wise, most would argue that these two competitors are way ahead of Siri, from being able to handle more complex queries to generally being more capable of processing natural language. However, as these aren't the default AI assistant (nor can they be set as one) on Apple devices, Siri is still the go-to personal assistant for many users.

Known User Interaction Issues

1. Siri is only able to process queries in the language(s) set by the user, and the currently supported list of languages are fairly limited, meaning many won't be able to interact with the tool at all.
2. Siri lacks contextual awareness (at times). In these instances, Siri won't be able to keep track of the conversation and require the user to repeat the context.
3. Siri quite frequently misinterprets queries (e.g., set a timer for "fifteen minutes" can sometimes be interpreted as "fifty minutes").
4. Siri doesn't perform certain queries and requires the device to be unlocked, which can be dangerous when driving.
5. Google Assistant (Siri's competitor) supports searching through Google Search, while Siri does not support searching capabilities.

These were only a few user interaction issues we were able to find during our brief experimentation with Siri. We are hoping to uncover several more through usability testing.

Part 2: Usability Test

This document describes a test plan and results from conducting a usability test for Siri. The goal is to conduct the usability test with at least 4 representative users. Data will be used to assess whether usability goals regarding an effective, efficient, and well-received user interface have been achieved. Insights from this test will be used to establish baseline user performance and user-satisfaction levels of the user interface for future design iterations.

Executive Summary

Siri is a primarily voice-based digital assistant that comes with a variety of Apple devices. This test focuses on evaluating Siri's usability through integration with other features or applications present on these devices, by inviting users to perform 4 tasks ranging from setting an alarm to ordering food through Siri. The usability test objectives are to identify potential usability issues, including misinterpretation of user speech and/or prompts, inconsistent behaviour, obscured features, unhelpful failure responses etc.

Methodology

Participants

There will be four participants recruited for the usability testing, all of whom are students or seniors and either Apple or Android users. Apple users may already be familiar with Siri, whereas Android users might not be familiar with it.

The participants' responsibilities will be to attempt to complete a set of representative task scenarios presented to them in as efficient and timely a manner as possible, and to provide feedback regarding the usability and acceptability of the user interface. The participants will be directed to provide honest opinions regarding the usability of the application, and to participate in post-session subjective questionnaires and debriefing.

During the usability testing, we will refrain from providing assistance to the participants beyond the provided task instructions to ensure the accuracy of the results. Participants come from diverse backgrounds to ensure representation in each case; some may be familiar with Siri, while others may not be. By doing so, we aim to obtain more accurate results and mitigate potential biases.

The demographics that will be selected are age, what cell phone operating system they use and exposure to Siri. These will help determine how familiar and adaptable the user will be with Siri.

Procedure

The tools required for the experiment are iPhone, iPad, or Mac computer with Siri and screen recording capabilities.

Participants will take part in the usability test using an iPhone, iPad, or Mac computer with Siri. The participant's interaction with the application will be monitored by the facilitator seated in the same location. The session will also be recorded using the device's built-in screen recording feature and microphone for analysis; if these features aren't available, the facilitator will videotape the session or take notes instead.

The facilitator will brief the participants on the application and instruct the participant that they are evaluating the application, rather than the facilitator evaluating the participant. Participants will sign an informed consent that acknowledges: the participation is voluntary, that participation can cease at any time. The facilitator will ask the participant if they have any questions. Participants will not be asked to complete a demographic and background information questionnaire prior to performing the tasks.

The facilitator will explain that the amount of time taken to complete the test task will be measured and that exploratory behavior outside the task flow should not occur until after task completion. At the start of each task, the participant will read aloud the task description from the printed copy and begin the task. Time-on-task measurement begins when the participant starts the task.

The facilitator will instruct the participant to 'think aloud' so that a verbal record exists of their interaction with the application. The facilitator will observe and record observations of user behavior, user comments, and system actions.

After completing the usability test, the participant will be given a questionnaire. The questions are used as a metric to assess how the user felt about the experience with Siri. It will also have questions to determine what demographic they are in.

Usability Tasks

Due to the range and extent of functionality provided in a selected application, and the short time for which each participant will be available, the tasks are the most common and relatively complex of available functions. The tasks are identical for all participants of a given user role in the study.

Tasks

1. **Ask Siri to set an alarm**: This task serves as an introductory task to familiarize users with using Siri. Setting an alarm is a common and straightforward action, therefore, by starting with this task, we can observe how easily users can interact with Siri for a basic function.
2. **Ask Siri to order your favorite meal**: This task introduces a little more complex scenario where the user interacts with Siri to perform an online transaction. Since ordering food involves multiple information, we can evaluate Siri's ability to handle more intricate commands and transactions by testing this task.
3. **Ask Siri to play your favorite songs**: Music playback is a commonly used feature of virtual assistant. However, Siri's interpretation of user commands can sometimes lead to unexpected results, such as playing the wrong song. This task aims to assess Siri's accuracy in understanding and executing 'vague' music-related requests.
4. **Ask Siri for directions to a location of your choosing**: Requesting directions is another common use case for virtual assistants, especially for users navigating unfamiliar places. This task tests Siri's ability to interpret location-based commands accurately, considering potential ambiguities such as multiple places with the same name.

Test Setup: The test setup requires an Apple device such as an iPhone, MacBook, iPad, or Apple Watch with Siri enabled. Users will need to activate Siri and provide instructions verbally to perform designated tasks. No additional accounts or equipment are necessary beyond the standard Apple device setup.

Conclusion: By conducting usability tests with these four tasks, we aim to gather insights into Siri's performance across various scenarios. The findings from this evaluation will help identify strengths and areas for improvement in Siri's functionality and user experience.

Usability Metrics

Usability metrics refers to user performance measured against specific performance goals necessary to satisfy usability requirements. Task completion success rates, error rates, and subjective evaluations via questionnaires, etc. can be used.

Assessment Metrics

- Completion rate: Measure how likely is Siri able to actually complete a task
- Completion time: Measure how long it takes users to complete a task with Siri and maybe compare it to performing the task themselves.
- Misinterpretation rate: Measure how likely Siri misunderstood our instructions
- Questionnaires:
    - Q1: Can the participants do the tasks easily without using Siri?
        - To see whether the tasks given are much more efficient to be done by the users itself or need Siri to help the users to do the tasks.
    - Q2: What is the most challenging part from these tasks?
        - For some users, it might be easier for them to do the task by themselves instead of using Siri.
    - Q3: What was your experience with Siri as a new user ?
        - To examine the user's initial experience with Siri.
    - Q4: How satisfied are you with the experience with Siri?
        - Use a scale from 1-10 to get a value.
    - Q5: What features/improvements would you like to be added to Siri?
        - To gain the user's perspective of what they wish that Siri could potentially relieve their burdens when doing certain tasks in the future.

Key Observations

Setting an alarm was the first task given to participants. Interestingly, we found that when users set an alarm in a 12-hour format without specifying AM or PM, Siri assumed AM – a safe assumption for most individuals. Notably, Siri completed this task flawlessly for all participants.

The next task required participants to order their favorite meal with Siri. While Siri offered helpful suggestions like nearby restaurants provided directions and even offered to call the place, it couldn't place the order itself. Additionally, specifying an app (e.g., DoorDash, UberEats) only resulted in Siri opening the app without further action. Overall, Siri was not able to complete food order requests and required additional actions from the participant.

Next up, participants were asked to play their favorite song with Siri. This required users to have an active music subscription (Spotify or Apple Music). Without this, Siri could shuffle a playlist (Spotify) or play a station (Apple Music), but also declined the request in a few instances (see Image 1 in appendix). Notably, some participants successfully requested playback on alternative apps like SoundCloud (see Image 2 in appendix). Additionally, Siri was able to accurately identify the preferred music platforms for playback (Spotify or Apple Music) without explicit user input. Despite this impressive feature, Siri completed this task only 41% of the time, with failures commonly stemming from misunderstood requests or limited integration with certain apps.

Lastly, participants were tasked with getting directions to a destination of their choosing using Siri. Impressively, Siri successfully guided users to their chosen destinations using the default Maps app (Apple Maps) for most participants. Users who preferred alternative navigation apps were also able to get accurate directions by specifying their preferred app.

Interpretations of Results

Siri's strengths are seen while completing simple tasks such as setting an alarm or playing a song, with the stipulation that the user has spotify or apple music, Siri will perform the task with ease. Users find that using Siri for these simple tasks is much easier than manually doing it themselves. Users also found using Siri as a new user was simple and easy to start.

Where Siri struggles is manually completing the task is still the preferred choice for some users to do such tasks. Lack of details, such as setting an alarm that does not state am or pm, is one of the reasons users prefer manually, even though the task was easier with siri.

One of the highlights of the observation is that some users who don't have an account with either Spotify or Apple Music can't use Siri to play music. In the beginning, before doing the testing, we thought that it was an easy task and everyone should do it easily. But, in fact, some people failed to do the task due to account problems that made Siri not work.

Changes to improve Siri would be first, increase the feedback from Siri when completing a task, currently Siri will confirm it has completed the task but additional details are left out, for example the alarm. Next, would be to tell users that a task is not possible, currently a certain task like ordering food is not possible, but Siri continues the interaction as if it will order food for the user. Lastly, there is no option for confirming the task for Siri, this causes Siri to do the wrong task at times, adding an option for users to require Siri to confirm will decrease errors for those who have a difficult time getting Siri to understand them.

Summary & Conclusions

Siri is a virtual assistant made by Apple that is built into their products such as the iPhone, Mac devices and iPad. Siri allows users to interact with their device with just voice commands and queries. These commands can range from setting an alarm to playing music. To do this, say "Hey Siri" or "Siri" and then the command or query and Siri will try to complete the command or answer the query. From testing, apps that have integration with Siri are easily completed, such as playing a song or setting an alarm. Where Siri struggled was with commands that have not been fully implemented, such as food ordering, which requires previous setup or only provides directions or a phone number.

Appendix

Consent Forms

SFU SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Thank you for agreeing to help us assess our software. We are taking a course at Simon Fraser University that involves the design computer software, and would like you to help us evaluate our design. In this session, we will ask you to work with an existing application or a paper or interactive software mock-up of our design and to help us evaluate its clarity and simplicity. We are evaluating the software, not you. Any difficulties that you may encounter while using the software will help us locate problems with our interface design.

This usability study is completely voluntary. You may decline to answer any question or stop the study at any time and for any reason. Any data gathered up to the point of stopping the study will be destroyed. If you are a student at Simon Fraser University, this interview will have no effect on your grades in any courses. The only data we will gather will be written notes. These notes will only be seen by our team members, our instructor, and our teaching assistant and will be destroyed at the end of the course. Your name will not be attached to any gathered data, but rather a unique ID code such as "Participant A".

Concerns or Complaints
If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, you may contact Dr. Jeffrey Toward, Director, Office of Research Ethics at jtoward@sfu.ca or 778-782-6593.

Do you agree to the following: I have read and understood the subject information and consent form and freely consent to participate.

(Yes)
No

Selecting yes indicates that you consent to participate in this study and that you are 19 years of age or older. Select no if you wish to decline or are under 19 years of age.

CMPT 363 Online Usability Testing Consent Form (v1.0 August 17, 2016)

Name: Yejin Lee
Signature: Yejin Lee

CMPT 363 Usability Testing Consent Form (v1.0 August 18, 2014)

SFU SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Thank you for agreeing to help us assess our software. We are taking a course at Simon Fraser University that involves the design computer software, and would like you to help us evaluate our design. In this session, we will ask you to work with an existing application or a paper or interactive software mock-up of our design and to help us evaluate its clarity and simplicity. We are evaluating the software, not you. Any difficulties that you may encounter while using the software will help us locate problems with our interface design.

This usability study is completely voluntary. You may decline to answer any question or stop the study at any time and for any reason. Any data gathered up to the point of stopping the study will be destroyed. If you are a student at Simon Fraser University, this interview will have no effect on your grades in any courses. The only data we will gather will be written notes. These notes will only be seen by our team members, our instructor, and our teaching assistant and will be destroyed at the end of the course. Your name will not be attached to any gathered data, but rather a unique ID code such as "Participant A".

Concerns or Complaints
If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, you may contact Dr. Jeffrey Toward, Director, Office of Research Ethics at jtoward@sfu.ca or 778-782-6593.

Do you agree to the following: I have read and understood the subject information and consent form and freely consent to participate.

(Yes)
No

Selecting yes indicates that you consent to participate in this study and that you are 19 years of age or older. Select no if you wish to decline or are under 19 years of age.

CMPT 363 Online Usability Testing Consent Form (v1.0 August 17, 2016)

Name: Kenzie Jovanka
Signature: Kenzie Jovanka

CMPT 363 Usability Testing Consent Form (v1.0 August 18, 2014)

SFU SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Thank you for agreeing to help us assess our software. We are taking a course at Simon Fraser University that involves the design computer software, and would like you to help us evaluate our design. In this session, we will ask you to work with an existing application or a paper or interactive software mock-up of our design and to help us evaluate its clarity and simplicity. We are evaluating the software, not you. Any difficulties that you may encounter while using the software will help us locate problems with our interface design.

This usability study is completely voluntary. You may decline to answer any question or stop the study at any time and for any reason. Any data gathered up to the point of stopping the study will be destroyed. If you are a student at Simon Fraser University, this interview will have no effect on your grades in any courses. The only data we will gather will be written notes. These notes will only be seen by our team members, our instructor, and our teaching assistant and will be destroyed at the end of the course. Your name will not be attached to any gathered data, but rather a unique ID code such as "Participant A".

Concerns or Complaints
If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, you may contact Dr. Jeffrey Toward, Director, Office of Research Ethics at jtoward@sfu.ca or 778-782-6593.

Do you agree to the following: I have read and understood the subject information and consent form and freely consent to participate.

(Yes)
No

Selecting yes indicates that you consent to participate in this study and that you are 19 years of age or older. Select no if you wish to decline or are under 19 years of age.

CMPT 363 Online Usability Testing Consent Form (v1.0 August 17, 2016)

Name: William Han
Signature: William Han

CMPT 363 Usability Testing Consent Form (v1.0 August 18, 2014)

SFU SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Thank you for agreeing to help us assess our software. We are taking a course at Simon Fraser University that involves the design computer software, and would like you to help us evaluate our design. In this session, we will ask you to work with an existing application or a paper or interactive software mock-up of our design and to help us evaluate its clarity and simplicity. We are evaluating the software, not you. Any difficulties that you may encounter while using the software will help us locate problems with our interface design.

This usability study is completely voluntary. You may decline to answer any question or stop the study at any time and for any reason. Any data gathered up to the point of stopping the study will be destroyed. If you are a student at Simon Fraser University, this interview will have no effect on your grades in any courses. The only data we will gather will be written notes. These notes will only be seen by our team members, our instructor, and our teaching assistant and will be destroyed at the end of the course. Your name will not be attached to any gathered data, but rather a unique ID code such as "Participant A".

Concerns or Complaints
If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, you may contact Dr. Jeffrey Toward, Director, Office of Research Ethics at jtoward@sfu.ca or 778-782-6593.

Do you agree to the following: I have read and understood the subject information and consent form and freely consent to participate.

(Yes)
No

Selecting yes indicates that you consent to participate in this study and that you are 19 years of age or older. Select no if you wish to decline or are under 19 years of age.

CMPT 363 Online Usability Testing Consent Form (v1.0 August 17, 2016)

Name: Stephen
Signature: Stephen

CMPT 363 Usability Testing Consent Form (v1.0 August 18, 2014)

SFU SIMON FRASER UNIVERSITY
ENGAGING THE WORLD

Thank you for agreeing to help us assess our software. We are taking a course at Simon Fraser University that involves the design computer software, and would like you to help us evaluate our design. In this session, we will ask you to work with an existing application or a paper or interactive software mock-up of our design and to help us evaluate its clarity and simplicity. We are evaluating the software, not you. Any difficulties that you may encounter while using the software will help us locate problems with our interface design.

This usability study is completely voluntary. You may decline to answer any question or stop the study at any time and for any reason. Any data gathered up to the point of stopping the study will be destroyed. If you are a student at Simon Fraser University, this interview will have no effect on your grades in any courses. The only data we will gather will be written notes. These notes will only be seen by our team members, our instructor, and our teaching assistant and will be destroyed at the end of the course. Your name will not be attached to any gathered data, but rather a unique ID code such as "Participant A".

Concerns or Complaints
If you have any concerns or complaints about your rights as a research participant and/or your experiences while participating in this study, you may contact Dr. Jeffrey Toward, Director, Office of Research Ethics at jtoward@sfu.ca or 778-782-6593.

Do you agree to the following: I have read and understood the subject information and consent form and freely consent to participate.

(Yes)
No

Selecting yes indicates that you consent to participate in this study and that you are 19 years of age or older. Select no if you wish to decline or are under 19 years of age.

CMPT 363 Online Usability Testing Consent Form (v1.0 August 17, 2016)

Name: Van Tran
Signature: Van Tran

CMPT 363 Usability Testing Consent Form (v1.0 August 18, 2014)

Questionnaire Responses

Google form to receive response:
https://docs.google.com/forms/d/e/1FAIpQLSdT_BANjwnI9L5Jnl_oDdD40mtxTTHWb_EMTS1J0n1VDLCbOQ/viewform?usp=sharing

Results:
https://docs.google.com/spreadsheets/d/1kncfvUnjkqDcBJ2nCQMMJKwXEIpW830C55_GAw-Kb04/edit?usp=sharing

Screenshots/Screen Recordings

https://drive.google.com/file/d/17oFrjYk9mJfi91cN13wuXUXTzEfkMoJR/view?usp=drive_link

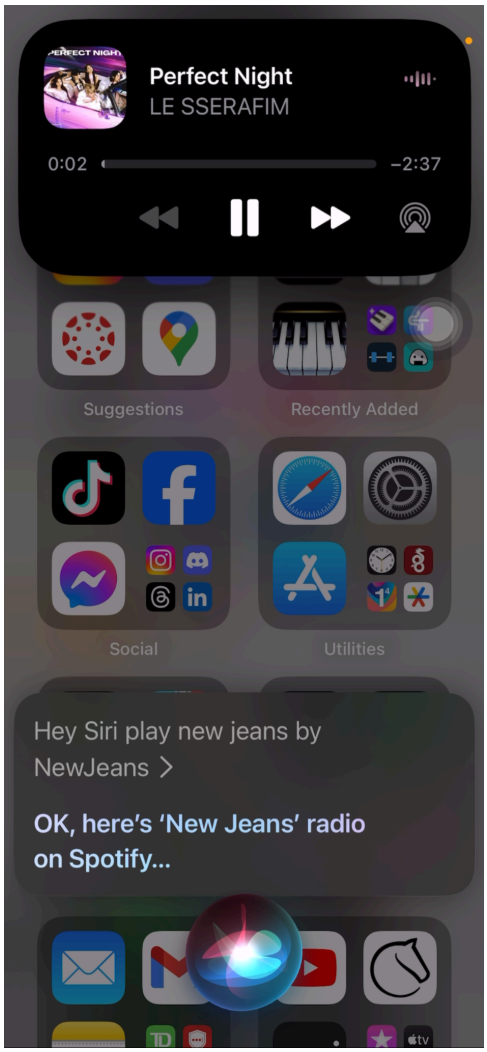

Image 1 - Siri plays a radio station if an active subscription is unavailable.

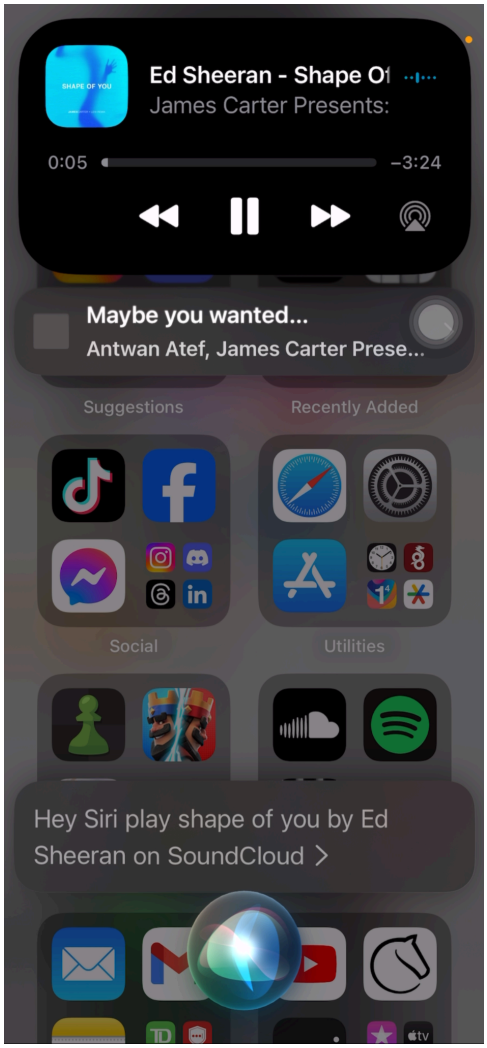

Image 2 - Siri plays a song on SoundCloud.

Usability Test Notes

Participants	Set alarm	Order food	Play fav music	Get directions
Van Tran	1/1 Siri was able to set an alarm.	0/3 Siri opened app (DoorDash & UberEats) but doesn't do much else	1/3 Siri was not able to play a song on Spotify/Apple Music if user doesn't have an active subscription and was able to play a song on SoundCloud only if explicitly stated;	3/3 Siri was able to get directions to various locations using both Google Maps and Apple Maps
Kenzie Jovanka	3/3 Quick and clear, Siri fully understands the task, she sets the alarm to "am" if we don't specify.	1/3 Siri opens Apple Maps by default, and gives a list of restaurants nearby. It asks the user which restaurant to choose, and offers an option to call the restaurant or give the direction.	2/3 By default, she opens spotify and plays the song from the liked album (would expect to play the most frequently listened song).	3/3 Straightforward, by opening the Apple Maps as the default application (if we didn't mention the map application) and search the direction from our current location.
Yejin	3/3 Sir will set the alarm with a single command, but lacks specifying am or pm	0/3 User wanted to order from chipotle but Siri did a web search for chipotle and only gave the user the option to call or get directions.	2 /3 The user searched for their favourite song but it took Siri 3 tries to get it right but next time it was able to do it for the first time.	3/3 The user was able to find directions to SFU easily and selected Apple Maps by default.
William Han	3/3 Siri directly goes	1/3 User need to	0/3 Siri by default is	3/3 Siri directly

	to the Clock and sets the right time.	turn on the current location and Siri can only give the recommended places to go, but not help to ordering	trying to open Apple Music and it doesn't work if you don't have any account yet. Users also try to guide Siri, by telling them to go to Safari and search for Youtube to play some music. However, Siri response by feature is not supported.	directed the user to the Maps and successfully did the job.
Stephen	3/3 Siri sets the alarm as directed ("Siri, set an alarm for 6am" → Alarm set and enabled for 06:00)	0/3 Siri brings up a list of nearby restaurants. When the user tries to click on one of them it just shows the restaurant's info in Apple Maps	3/3 Siri starts playing an Apple Music playlist filled with frequently played songs for the user right away	0/3 The user asked for directions to multiple locations, but Siri failed to identify the intended destination and gave directions to similarly named locations instead, with no options for ambiguous locations for the user to choose from