IMA 2 - Usability Testing Report

DuoLingo

Maria Mercier
Lauren Lucernoni
Zachary Hill

[2/28/21]

Introduction

Duolingo is an American language-learning website and mobile app, as well as a digital language-proficiency assessment exam. The company uses the freemium model; the app and the website are accessible without charge, although Duolingo also offers a premium service for a fee. As of 10 February 2021, the language-learning website and app offered 106 different language courses in 38 languages. The app has over 300 million registered users across the world.

Three students in Dr. Park's EME 6208 Interactive Media graduate class conducted an on-site usability test using the mobile version of the Duolingo app. Mobile devices were utilized to collect data from a total of 6 target users/participants. An observation sheet and an end-of-session questionnaire were used to record the data. Users were asked to complete a total of seven tasks. The sessions captured each participant's navigational choices, number of clicks, task completion rates, comments, overall satisfaction ratings, questions and feedback.

Methodology

The test administrators recruited participants by asking people that they know directly if they would be willing to participate in a test that examines the usability of a foriegn language learning application. When the subject agreed, a time and location was verbally set according to availability of the participant. The test administrators worked one-on-one with each participant, in person to complete their test.

Each individual session lasted just under half an hour. During the session, the test administrator read the script from the <u>Usability Test Materials</u> Microsoft Word document provided within the assignment details. This script described the process that the participant would take part in. Participants were asked to download Duolingo to their mobile device and talk through each task to give live feedback to the test administrator, while the test administrator filled out an Observation Sheet that required the test administrator to time the participant through each task, count clicks, and make note of participant comments.

After the test, the administrator asked the participant to complete a survey rating their experience on a 5-point Likert Scale. This scale measured from Super Easy to Super Hard and included 7 questions regarding the tasks.

- How easy or hard was it to complete the steps necessary to create a new account?
- How easy or hard was it to find one language you like to learn and add it to your profile/account?
- How easy or hard was it to change the language to another language you would like to learn?
- How easy or hard was it to turn off and turn on auditory sound?
- How easy or hard was it to go to a setting/preference and change a setting option?

- How easy or hard was it to change your profile picture?
- How easy or hard was it to invite a friend to try Duolingo?

The next part of the survey asked the participant to complete a survey rating their experience on a 7-point Likert Scale. This scale measured from Strongly Disagree to Strongly Agree and included 7 questions regarding the participants overall satisfaction with the application. The questionnaire included the following 19 satisfaction related questions:

- Overall, if they were satisfied with how easy it was to use the language learning app.
- If it was simple to use the language learning app.
- If they could effectively complete the tasks using the language learning app.
- If they were able to complete the tasks quickly.
- If they were able to efficiently complete the tasks.
- If they felt comfortable using the app.
- If it was easy to use the language learning app.
- If they believed they could become quickly productive using the language learning app.
- If the app gave them error messages that told them how to fix problems.
- If they made an error, if they could recover easily and quickly.
- If the information provided with the language learning app was clear.
- If it was easy to find the information they needed.
- If the information provided for the language learning app was easy to understand.
- If the information was effective in helping to complete the tasks.
- If the organization of information on the app screen was clear.
- If the interface of the app was pleasant.
- If they liked using the interface of the app.
- If they thought that the app had all the functions and capabilities they expected.
- If they were satisfied overall with the language learning app.

After the participants had completed the survey, the test administrator asked the participants to study the language learning mobile application for a timed three minutes. Once they completed their study, the participants were asked to imagine their friend just emailed them to ask about Duolingo. What would they tell their friends? Participant responses were recorded on the Observation Sheet by the test administrator.

Based on the data, the following is a short analysis of Learnability, Errors, and Satisfaction of the mobile app:

Learnability
Duolingo has a very eye-pleasing and friendly user experience interface that allows our participants to walk through each task with ease. Participants were guided by the interface

as though they were familiar with the interface, though many had not used the mobile application before their experience. Most of the participants completed each task assigned to them with ease and efficiency throughout their experience. Most users agreed that they would recommend Duolingo to a friend because of how user friendly the application is and how easy it makes the hard work of learning a new language seem.

Errors

1. Two participants did not recognize that they needed to select a language to learn. They accidentally selected their mother tongue.

2. One participant first attempted to click the "speaker" icon when turning off the auditory sound, but then realized that it was the "toggle switch" instead.

3. Another participant made at least two errors (wrong clicks) while trying to find the button to change the language.

4. One participant kept clicking on the back arrow when trying to create the account. It took a while for her to understand the logistics of the 'placement test' prior to creating her account.

5. Another participant made up to 4 errors trying to figure out how to find and add a new language to his profile. He kept clicking on the wrong icons. This went on for six minutes. The rest of the tasks did not pose any major challenges for this participant.

Satisfaction

Most of the participants were very satisfied with the efficiency and usability of Duolingo. All participants were able to complete each task to 100% completion. Participants noted that they were satisfied with how easy the application is to use and they felt that they were able to complete their tasks quickly and efficiently. They all felt comfortable using the application without any training as the training is incorporated into the creation of a profile so users are trained as they are getting started. Each participant felt that the interface provided the information that the user needs to complete their tasks and it is enjoyable to use with functions and capabilities that they would expect in a productive language learning application. One of the participants liked the 'gamification' aspect of the application. Some of the participants also felt the application was great for overall speaking, vocabulary, and listening comprehension skills.

Participants

Maria's participants are both second and foreign language educators. Participant 1, a female, teaches English as a second language at the university level and Participant 2, a male, teaches French at the High School level. Both participants completed their usability testing tasks on Saturday, February 21st, but at different times of the day. The testing sessions took place in Maria's home. The participants used their own devices to complete the tasks. The participants were both born in the United States and their native language is English. They both enjoy using a variety of mobile apps, and they listed Stepcounter Pedometer, Disney+, and Google maps amongst their favorites.

Lauren's participants both completed their usability testing tasks on Sunday, February 22nd at consecutive times - 8:00PM and 8:30PM. The testing sessions took place in Lauren's home. Participant 1 is a male, whose occupation is in the financial services industry. He works in the middle office supporting financial advisors for a large firm.He is familiar with mobile applications, utilizing mostly social media, step counters, messaging, and money applications. He had never used a language learning application before participating in the usability test and quite enjoyed the process. Participant 2 is a female, full-time student in high school who uses mobile applications quite frequently - her favorites being Hulu, Instagram and Spotify. She had used both Duolingo and Drops mobile language learning applications before, but had uninstalled Duolingo. The participants were both born in the United States and their native language is English.

Zach's participants completed their Usability tests on February 24th, between 5:30 and 6:30 pm HST at Participant 1's house. The participants completed their tests on their own devices. Participant 1 is a female, part time adult ESL teacher, part time bartender. Participant 2 is also female and also a part time adult ESL teacher. Participant 1 is from the USA and her native language is English. Participant 2 is from Japan and is simultaneously bilingual in English and Japanese, having learned both languages in childhood. Both participants use mobile apps, such as Spotify, Instagram, and TikTok quite frequently in their daily lives, and have even used other language learning apps such as Babbel, Memrise, and Drops. Participant 2 had never used Duolingo before, while Participant 1 had used the app before but no longer had it installed on her device.

In total, 6 participants were scheduled over three separate testing dates. All 6 participants completed the test. Of the 6 participants, 2 were male and 4 were female.

Evaluation Tasks/Scenarios

Users were asked to complete a total of seven tasks. The sessions captured each participant's navigational choices, number of clicks, task completion rates, comments, overall satisfaction ratings, questions and feedback. The tasks were listed in the order as follows:

- Find one language you would like to learn and add it to your profile account.
- Complete the steps necessary to create a new account. Stop after your account has been created.
- Change the language to another language you would like to learn.
- Turn off and turn on the auditory sound.
- Go to a setting/preference and change a setting option.
- Change your profile picture.
- Invite a friend to Duolingo.

Please note: Task 1 and Task 2 were reversed as Duolingo requires the user to select the chosen language first.

Results

Task Completion Success Rate

There were a total of six participants. Observing those participants were three test administrators who directed each participant without promoting through the tasks required. Participants 1 and 2 were observed by Maria. Participants 3 and 4 were observed by Lauren. Participants 5 and 6 were observed by Zach. All participants successfully completed all tasks at 100% completion: Task 1: Select a language, Task 2: Create a profile, Task: 3 Change the language, Task 4: Turn on and off auditory sound, Task 5: Change any setting, Task 6: Add a profile picture, Task 7: Invite a friend to Duolingo.

Task Completion Rates

Participant	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	Task 7
1	√	√	√	√	√	√	√
2	√	√	√	√	√	√	√
3	√	√	√	√	√	√	√
4	√	√	√	√	√	√	√
5	√	√	√	√	√	√	√
6	√	√	√	√	√	√	√
Success	7	7	7	7	7	7	7
Completion Rates	100%	100%	100%	100%	100%	100%	100%

Learnability: Time to complete a task successfully

Each team member recorded the six participants spent on each task. Some tasks were slightly more difficult to complete than others as reflected by the average time on each of the tasks.

Task 1 required participants to find a language they would like to learn and add it to their profile and it took an average of 3.12 minutes to complete. Task 1 completion times ranged from 0.5 minutes to 7 minutes, with half of the participants completing the task in less than a minute.

Task 2 required participants to complete the steps necessary to create a new account and it took an average of 2.5 minutes to complete. Task 2 completion times ranged from 1 minute to 3.96 minutes with most times less than 3 minutes.

Task 3 required participants to change the language to another language they would like to learn. This task took an average of 1.68 minutes to complete. Task 3 completion times ranged from 37 seconds to 3 minutes with most times less than 1.5 minutes.

Task 4 required participants to turn off and on the app's sound. This task took an average of 1.13 minutes to complete. Task 4 completion times ranged from 0.5 minutes to 3 minutes with most times less than 1 minute.

Task 5 required participants to go to a setting or preference and change one of the options. This task took an average of 42.5 seconds to complete. Task 5 completion times ranged from 7 seconds to 1 minute with most times less than 1 minute.

Task 6 required participants to go to their profile and add an image for their profile. This took an average of 54.5 seconds to complete. Task 6 completion times ranged from 27 seconds to 2 minutes with most times less than 1 minute.

Task 7 required participants to invite a friend to Duolingo. This task took an average of 1.04 minutes to complete. Task 7 completion times ranged from 47 seconds to 2 minutes with most times less than 1 minute.

Time on Task

	P1	P2	P3	P4	P5	P6	TOT	Avg.
Task 1	420	360	33	41	240	30	1124	187.3
Task 2	120	180	206	238	120	60	924	154

Task 3	180	180	91	37	60	60	608	101.3
Task 4	60	180	34	49	60	30	413	68.3
Task 5	60	60	8	7	60	60	255	42.5
Task 6	60	60	27	30	120	30	327	54.5
Task 7	120	45	49	41	60	60	375	62.5

Errors

Each team member captured the number of errors participants made while trying to complete the task scenarios.

Participant 1 made at least two errors (wrong clicks) while trying to find the button to change the language (Task 3). She also clicked on the back arrow several times when trying to create the account. It took a while for her to understand the logistics of the 'placement test' prior to creating her account, so this was her most challenging task. Participant 2 struggled with Task 1. He made at least 4 errors while trying to figure out how to find and add a new language to his profile. He kept clicking on the wrong icons. This went on for six minutes. This was the most challenging task for Participant 2. Nevertheless, both participants were able to complete all seven tasks and enjoy the overall application testing experience.

Participant 3 and 4 both selected English as the language they wanted to learn instead of a language foreign to them in Task 1, both making 4 clicks in error. Participant 4 accidentally went out of the settings to turn off auditory sound on and off in Task 4, making 5 clicks in error. Participant 5 first clicked the "speaker" icon associated with the "News Feed" when attempting to turn off auditory sound before finding the "toggle switch" under "Settings". None of these errors were critical and did not prevent successful completion of tasks.

Errors

	P1	P2	P3	P4	P5	P6	TOT	Avg.
Task 1	1	4	4	4	0	0	13	2.1
Task 2	2	1	0	0	0	0	3	0.5
Task 3	2	0	0	0	0	0	2	0.33
Task 4	1	0	0	5	2	0	8	1.3
Task 5	0	0	0	0	0	0	0	0
Task 6	0	1	0	0	0	0	1	0.167
Task 7	2	0	0	0	0	0	2	0.33

Summary of Data

The table below displays a summary of the test data. Low completion rates and satisfaction ratings and high errors and time on tasks are highlighted in red.

Summary of Completion, Learnability (Time on Task), Errors

Task	Task Completion	Learnability (Time on Task)	Errors
1	6	187.3	2.1
2	6	154	0.5
3	6	101.3	0.33
4	6	68.3	1.3
5	6	42.5	0
6	6	54.5	0.167
7	6	62.5	0.33

Task – end of session questionnaire table

Participants	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Total (per participant)	Mean (per participant)
P1	1	1	3	1	1	1	3	11	1.57
P2	1	2	2	1	3	1	1	11	1.57
P3	1	2	2	2	2	1	3	13	1.85
P4	2	2	2	3	2	2	2	15	2.14
P5	2	1	1	3	2	1	1	11	1.57
P6	1	3	2	2	2	2	2	14	2
Total (per item)	8	11	12	12	12	8	12	75	10.7
Mean (per item)	1.33	1.83	2	2	2	1.33	2	12.5	1.78

* Scale: 1 (super easy) – 3 (not easy or hard) – 5 (super hard)

* Mean rating per participant = (total per participant from Q1 to your tasks) / # of items (above example used 5 items only)

* Mean rating per item = (total per item) / 7 (# of participants)

Satisfaction

In terms of participants' satisfaction using the Duolingo mobile app, three out of the six participants *agreed* that it was easy enough to use this language application. Two of them *strongly agreed* with that remark. The majority of the participants *agreed* that it was simple to use the app, whereas one of the participants *strongly agreed* with the statement. Two of the participants stated that they strongly agreed they could effectively complete the tasks using the Duolingo language learning app, while the remaining four participants stated they simply *agreed* with this statement. In terms of being able to quickly and efficiently complete the tasks using the app, the majority of the participants said that they *strongly agreed* with those statements.

Moreover, three of the six participants *agreed* that they felt comfortable using the app, as opposed to the single participant who stated he/she felt *very* comfortable with the app. Three of the six participants *strongly agreed* that it was easy to learn to use the application, though one of the participants remained *neutral* about it. This shows that the app was not always easy to use for at least one of the participants. To add, the majority of the participants were in slight disagreement with the statement that they could become quickly productive while using the app. Four out of the six participants felt *neutral* or in *disagreement* with the idea that the application gave adequate error messages that could quickly help them to quickly fix the mistakes (a lack of adequate error correction feedback) and easily recover from the errors. Half of the participants *strongly agreed* that the information provided with this app was clear, whereas the other half remained *neutral* about this statement. Most of the participants *slightly agreed* that it was easy to find the information they needed and that the information was *fairly easy* to understand and effective to achieve the completion of the tasks.

The item that received the *lowest satisfaction score* was item #10. Whenever participants made a mistake while trying to complete the tasks, they felt it was not always easy nor quick to recover. To conclude, four out of six participants *strongly agreed* that the Duolingo mobile application was effective and user friendly. They enjoyed using the language learning app and expressed overall satisfaction with it.

Satisfaction Questionnaire table

	1 Strongly Disagree	2	3	4 Do not agree or disagree	5	6	7 Strongly Agree
1. Overall, I am satisfied with how easy it is to use this language learning app.					3	1	2
2. It was simple to use this language learning app.					3	2	1
3. I could effectively complete the tasks using this language learning app.					2	2	2
4. I was able to complete the tasks quickly using this language learning app.					2	2	2
5. I was able to efficiently complete the tasks using this language learning app.					2	2	2
6. I felt comfortable using this language learning app.					1	3	2

Statement							
7. It was easy to learn to use this language learning app.				1		2	3
8. I believe I could become productive quickly using this language learning app				2	1	2	1
9. This language learning app gave error messages that clearly told me how to fix problems.			1	3	2		
10. Whenever I made a mistake using the language learning app, I could recover easily and quickly.			2	1	1	1	1
11. The information (such as online help, on-screen messages and other documentation) provided with this language learning app was clear.				3		3	
12. It was easy to find the information I needed.				2	3	1	
13. The information				1	2	2	1

provided for this language learning app was easy to understand.							
14. The information was effective in helping me complete the tasks.				1	1	2	2
15. The organization of information on the language learning app screens was clear.					2	1	3
16. The interface of this language learning app (buttons, menus, graphics) was pleasant.				1	1		4
17. I liked using the interface of this language learning app (buttons, menus, graphics).					2		4
18. This language learning app has all the functions and capabilities I expect it to have.				1	3	2	
19. Overall, I am satisfied with this language learning app.			1	1	1	1	2

Total Per Item	0	0	4	17	32	29	32
Mean Per Item	0	0	0.66	2.83	5.33	4.83	5.33

* Scale: 1 (strongly disagree) – 4 (do not agree or disagree) – 7 (strongly agree)
* Mean rating per participant = (total per participant from Q1 to Q19) / 19 (# of items) (above example used 8 items only)
* Mean rating per item = (total per item) / 7 (# of participants)

User Experience

Participant 1 found it odd and a bit annoying that she had to take a language placement test before she was allowed to create an account. She felt she had to go through too many steps to get to the point of creating the account and that this could be discouraging to other users. She did not like that if you get something wrong, the app shows you the correct answer, BUT the print is too small. She added that the fact that hearts need to be purchased in order to advance in some of the lessons or "be excused" from some of the lessons can be discouraging. She enjoyed the motivational messages in this app. As a language educator, she thought the constant 'cycling' of words, grammar points and phrases in the app is good pedagogy. She also appreciated the gamification aspect of the app. She added that, in her view, the app focuses on speaking, grammar and vocabulary. She would like to have the opportunity to do some typing/practice her writing as well. It would be great if the app included some images to go with the vocabulary words, she concluded.

Participant 2 stated that the app worked "pretty well." He added that the app does not provide a speaking assessment for pronunciation, but it is good for listening and vocabulary learning. This user is curious about how the app would work for languages that use symbols and characters as part of their writing system. He actually enjoyed doing the language assessment prior to creating an account. He thought the task which required selecting another language was tricky, and that the page did not refresh automatically. All, in all, he enjoyed completing the tasks and learning about the Duolingo app.

Participant 3 thought the app was very comprehensive and that there were a lot of language and courses on every language. She stated that DuoLingo is set up with a lot of features that bring all the lessons together at the end and the lessons are enjoyable.

Participant 4 thought that the application was good if you are interested in learning language because of how it starts the user off with the basics and allows you to build into learning sentences. However, he stated that when you make mistakes you lose heart health and have to wait four hours or wait for the heart health to refill, which works against productivity in the sense that an adult might only have 30-minute time frames here and

there and would be more productive to be able to work on language learning around their schedule instead of having the application set up like a game.

Participant 5 thought that the app was a very user-friendly free app for language learning and hoped that it would help her to improve her Spanish skills. On first impression, she found the owl to be very cute and the amount of white-space in the app to be very relaxing to the eye. She enjoyed the animated graphics and said that she would recommend the app to a friend.

Participant 6 also found the app easy to use, however, she did not like that certain languages were more developed than others. For example, Mandarin has audio for all words, but Hawaiian does not. She felt that the app didn't really help her to learn the language, just to memorize the words. On first impression, she felt that the owl looked "too happy." Interestingly enough, the app's interface appeared black on her screen, rather than white as on Participant 5's screen.

Recommendations

The recommendations section provides recommended changes and justifications driven by the participant success rate, behaviors, and comments. Each recommendation includes a severity rating. The following recommendations will improve the overall ease of use and address the areas where participants experienced problems or found the interface/information architecture unclear.

Based on the results of our "Usability Test," we found that our participants mostly found Duolingo to be enjoyable and easy to use. However, there were some issues that our participants mentioned. First of all, the "gamified learning" process does not work well for everyone, and one participant found it somewhat discouraging that in the free version you have limited "hearts," which are lost when you make a mistake. After losing all your hearts, you must wait several hours for them to regenerate before continuing with your learning. This could be very inconvenient for busy adults who may have limited time to dedicate to language learning in the app.

Recommendation
Other language learning apps, like Drops, simply give users of the free version a limited amount of time to learn in the app (5 minutes) and users are not penalized for their mistakes. When setting up a new account on Duolingo there are questions about user preferences. It would be an improvement if users could choose between limited "hearts" or limited time, depending on their needs.

Secondly, as Participant 6 mentioned, some languages are more developed than others. While this is likely a result of the fact that much of Duolingo's content is crowd-sourced from users (https://incubator.duolingo.com/), and some languages are more heavily-used than others, the creators of the app could make a greater effort to further develop less popular languages. In Hawai'i, for example, there are many efforts to revive and preserve the Hawaiian language (link) and it would likely not be difficult for the app's developers to tap into this existing body of knowledge to further develop its Hawaiian language lessons.

Surely the same could be said for other supposedly endangered languages featured on the app, like Navajo.

Recommendation
Duolingo's developers could engage in greater partnership with community members to further develop their lessons for endangered, or less widely-spoken, languages. Some participants felt that the app was less effective for teaching non-romanized or tonal languages, as it relies heavily on translating sentences or simply matching words from the target language with words from the users' native language. With the gamified learning process, it is sometimes easy to simply memorize the process to "win," rather than actually learn the language.

Recommendation
Other language learning apps, like Drops, allow you to actually "draw" non-romanized characters on the screen with your finger, and match new vocabulary words with pictures, rather than translate. Duolingo could adopt some of these methods, which create a stronger "tactile" or "visual" learning of the new language.

Recommendation
It would be useful if Duolingo had a speaking assessment feature. The app tests the users' grammatical, reading, and vocabulary knowledge and targets some listening skills, but it does not test speaking skills.

Learning a Language on Duolingo (Free Practice with the App)
Most of our participants were able to perform the seven tasks with little or no errors, however, when engaging in free practice with the app, some issues were encountered. The following chart shows recommended changes and the justification for making the changes. While these are problems that our users noted when using the app, they did not prevent participants from using the app, and so we have classified their severity as "Low."

Change	Justification	Severity
1. Include drawing or typing option for languages with non-romanized characters 2. Include pictures or visual aides to go along with the vocabulary and/or content	1. Some participants found the app to be less effective for learning languages with non-romanized characters. Making this change would make the app more effective for learners of languages with non-romanized alphabets, such as Arabic or Mandarin. 2. The way the app is currently designed relies too heavily on translation. Incorporating more visual elements would create a stronger cognitive connection between new vocabulary in the target language and corresponding pictures.	1. Low 2. Low

3. Further develop lessons for less widely-spoken languages.	3. Duolingo makes big claims about teaching languages that are in danger of extinction, like Navajo and Hawaiian. However, their lessons for these languages are not nearly as robust as the lessons for in-demand languages like English and Spanish. This is likely because much of their content is crowd-sourced from bilingual volunteers who contribute. If they want to make good on their claim of helping to preserve these "endangered" languages they should be more proactive about tapping into the knowledge of these communities that are already working hard to preserve their own languages.	3. Low
4. Further develop the app to include a speaking component.	4. Language learners need and want to develop their speaking skills when learning a new language in order to communicate effectively. It is a natural aspect of the language acquisition process.	4.Intermediate

Conclusion

All in all, our participants reported high levels of satisfaction and usability with the language learning app, Duolingo. For this reason, Duolingo is one of the most popular language learning apps on the market today, and its "gamified learning" process helps to motivate users to be consistent with their language learning and connect/compete with other language learners around the world. However, there is always room for improvement, and our recommendations offer some guidance, based on real-life user experience, for how the app could be improved to meet the needs of a diverse array of learners. To learn more about Duolingo and its online community of language learners, check out https://www.duolingo.com/info for more information about the mission and methodology of this popular language learning application.