



duolingo usability study: duo not disturb 🌙

Ashlyn B. Aske, Michelle Pearman, Gene Wang, Mina Zavary

Table of Contents

Executive Summary	2
Study Objective	3
Participants	3
Method	3
Tasks	4
Findings, Areas for Opportunity & Recommendations	4
Positive Findings	4
Severity Rating for Areas of Opportunity	5
Area of Opportunity 1: Notification Navigation	5
Area of Opportunity 2: Lack of Confidence in Changing Notifications	7
Area of Opportunity 3: Unclear Terminology	8
Area of Opportunity 4: Branding & Gamification	10
Area of Opportunity 5: User Profile Completion	11
Next steps	12
Further Exploration	12
Future Learnings	13
Appendix	14
Task List	14
Test logistics	15
Screener data	15
Session data	16
Toolkit materials	16
Recordings	16
Citations	16

Executive Summary

Our study evaluated the usability of Duolingo, a popular language-learning app with “gamified” features and a public reputation for being pushy with its reminder notifications. We sought to evaluate the clarity of Duolingo’s systems to see if we could find inefficiencies that could be better streamlined. As such, much of our research focuses on the clarity of various processes and potential sources of cognitive overload.

Throughout our study, we came across [five major findings](#), each of varying severity. Duolingo’s notification settings were consistently complicated for users to navigate. Attempting to adjust the app’s notification frequency presented challenges that made the process downright unclear and inefficient. Our study also showed that participants received unclear feedback regarding their attempts to turn off notifications. Despite going through the proper process to adjust notification settings, participants found themselves unsure that they had done so properly. Additionally, we discovered that Duolingo contained a plethora of Duolingo-specific terminology that was often not clearly defined. As such, participants would often find themselves confused or second-guessing their interpretations. The app’s branding and gamified model generated mixed feelings from our participant pool, suggesting that the app’s presentation and educational model did not appeal to everyone. Our participants often cited receiving moments of cognitive overload that would distract them. The study also demonstrated that the user profile completion process lacked clarity in its completion status due to unclear language and inconsistent mental models.

To address the concern over notification settings our suggestion is to implement a system with less scrolling and clicking for easier navigation. We would also recommend adding a save button to the notifications settings and performing further research to improve feedback. Additionally, our research suggests better defining unique terminology during onboarding. To reduce cognitive overload, we propose toggle options for animations and gamified learning tools. Our research also suggests streamlining the onboarding process into one step and improving the verification of its completion. More details regarding our findings and our recommendations to alleviate these concerns can be found in the [Findings & Recommendations](#) section below.

Study Objective

Our own past exposure to the app and its media presence marred our image of the app: a potentially useful educational device hampered by inconveniences and design shortcomings. The app has an abundant amount of educational resources presented in a unique gamified manner, but our initial research indicated that many criticize the app's cognitive overload on its users. As such, we hoped to discover methods to streamline Duolingo's language-learning process to make the app more accessible to its users. To better focus our investigation, we focused our points of interest on the research questions below:

- What points of confusion do users encounter while navigating through the app?
- How can users better be able to customize their app usage to their liking comfortably?
- What obstacles do users encounter while trying to learn the language of their choice?
- How well can users navigate the setup process of Duolingo?
- What features are missing in the app?
- What features are distracting or superfluous to the app?
- Do users encounter cognitive overload within the app? How do they handle it?

Participants

Our study consisted of 10 users total; 5 new participants who had never used the app in any capacity and 5 experienced participants who had any amount of past experience using the app, regardless of the duration or recency of said experience. All participants had the following qualifications:

- An understanding of the Latin alphabet (ABCs)
- Access to Duolingo's free version, as this would be the version most new users are exposed to upon trying the app for the first time.
- Access to a smartphone
- Access to Wifi and Zoom (if remote)
- Located in the Seattle area (if in-person)

Method

We conducted 40-60 minute in-depth usability interviews, where users walked through 4 different [tasks](#) on the Duolingo App. These took place either in person or over zoom. During all interviews, participants were provided a consent form where they agreed to be recorded via audio and have their mobile device screen recorded. Two researchers were present for each session, one as the facilitator

and the other as the notetaker. We tracked **qualitative data** regarding participants' ability to navigate through the tasks and noted their struggles and frustrations. As such, keeping track of the user's expectations from the app enabled us to understand potentially more intuitive options.

The analysis involved parsing through interview notes and recordings to complete a visual thematic analysis in FigJam software. Both groups were analyzed together without much regard for comparative analysis due to discovering our experienced participants tended to have a limited or outdated experience with Duolingo.

Tasks

Test participants attempted the completion of the following tasks:

1. Walk through the steps to set up Duolingo as a first-time user
2. Start their first lesson with a language of choice that they have never studied before.
3. Go through the process of setting up a user profile.
4. Adjust the amount of push notifications that users receive.

Findings, Areas for Opportunity & Recommendations

Positive Findings

The most salient positive comments participants had revolved around the branding and aesthetic of the app. Six out of ten participants mention enjoying the app's aesthetic, often referring to it as "cute." As such, it could be argued that the app succeeds at drawing people in with its aesthetic; making the experience fun and enjoyable. Participants made the following comments in appreciation of the app's branding design:

- Animations make the app "feel more fun and interactive" (P10)
- "Having a character is very good for something that's very educational and technical that can [otherwise] be really boring" (P3)

Second, we also found that, in general, aside from the specific finding in Area of Opportunity 1, participants were able to navigate the app with ease. Throughout the onboarding process, lessons, and profile creation components, the interface provided a clear and unidirectional path making it easy for users to get from one step to the next with minimal confusion. One metric we collected was the number of "wrong turns" participants made when attempting to complete a task (e.g. a participant clicking on a button that does not get them closer to the goal state

because the button was confusing). There were very few wrong turns in any of the tasks, with no task prompting more than 1 wrong turn per participant (and many of those wrong turns are attributed to a particular finding we will discuss later). This indicates that the app is generally very navigable, and it is clear which buttons users should select to advance toward each task's goal state.

Severity Rating for Areas of Opportunity

The following 5 major findings were ranked from 1-5 on severity, with 1 being the least severe finding and 5 being the most severe finding. For ease of use, the following findings have been organized from highest to lowest severity.

Severity 1: Lowest Severity

Severity 2: Low Severity

Severity 3: Mid Severity

Severity 4: High Severity

Severity 5: Highest Severity

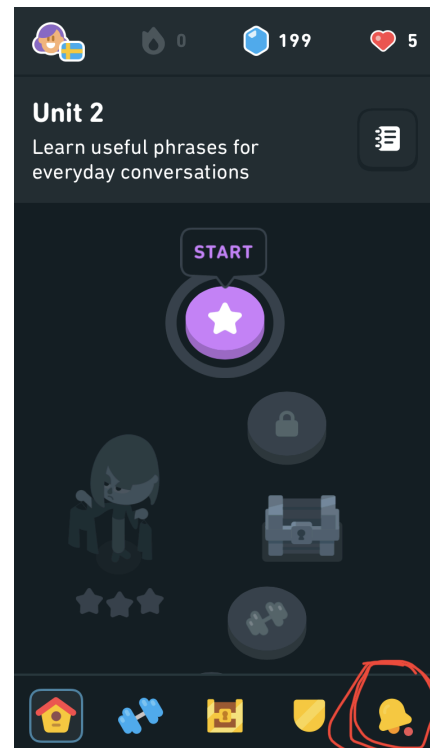
Area of Opportunity 1: Notification Navigation

Severity 5 (Highest Severity)

All ten participants discussed some areas for opportunity with this section of the study.

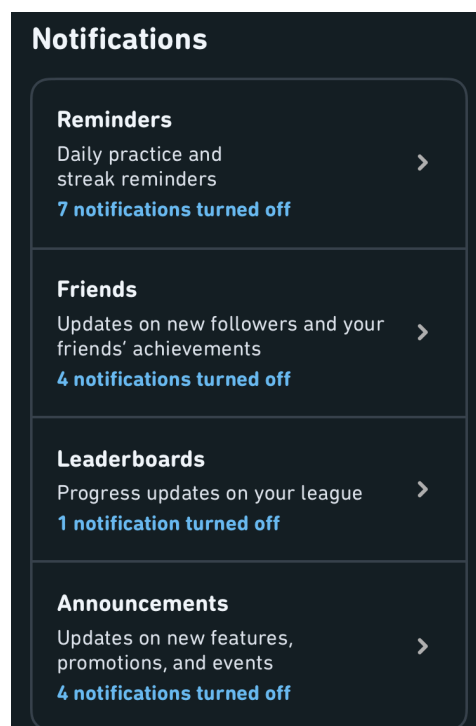
Navigating the notification settings was a point of confusion and inefficiency for nearly all of the participants in our study. First, seven of the ten participants struggled to find where the notification settings were located. This confusion stemmed from a bell icon leading to an unrelated page [see image to the right, the circled bell icon caused confusion]. In many users' mental models, bells are associated with notifications, leading to a wrong turn in finding the settings. All of the app's settings are located under the profile, which was unexpected for most of the users. One participant was quoted:

- "That's a weird place for it to be" (P5).



After participants determined that the bell icon did not relate to notifications and located the app settings under the profile, many found navigating the settings page to be inefficient. This impacted four of the ten participants. Due to the long list of settings, some scrolled past the notifications section, and others simply mentioned that they felt there was too much scrolling.

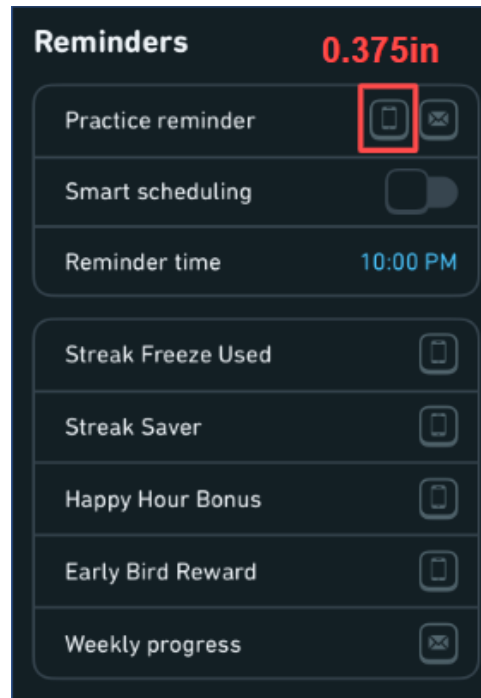
Furthermore, after the poor discoverability of the notification settings made locating them long and tedious, the process of actually adjusting was also inefficient. For context, there are 16 different types of notifications that can be adjusted, and they are divided across four categories. For each category, users must open a separate page to make changes, then return to the page of options and open the page corresponding to the next category. When the category pages are opened, individual buttons must be pressed for each of the 16 notifications. This was extremely inefficient for our participants, and three of ten mentioned they would appreciate one button to instantly silence all notifications.



Four categories, 16 notifications.

Finally, the process was further hindered by the small size of each of the 15 buttons that must be selected. In our study, participants mentioned that the small size made the already-long process even more difficult since they had to tap the buttons very carefully. Our team also speculates that the small size may hinder accessibility for people with some disabilities, which we elaborate on in the 'Next Steps' section.

In [Touch Targets on Touchscreens \[1\]](#), Nielsen Norman Group recommends touch targets should be at minimum 0.4x0.4 inches. We measured these buttons to be 0.375 inches across the smallest dimension. We were not able to examine the size of the touch target (which could be slightly larger than the button). Based on the information we do have, it seems that these buttons are too small.



The button size to turn notifications on and off is marked.

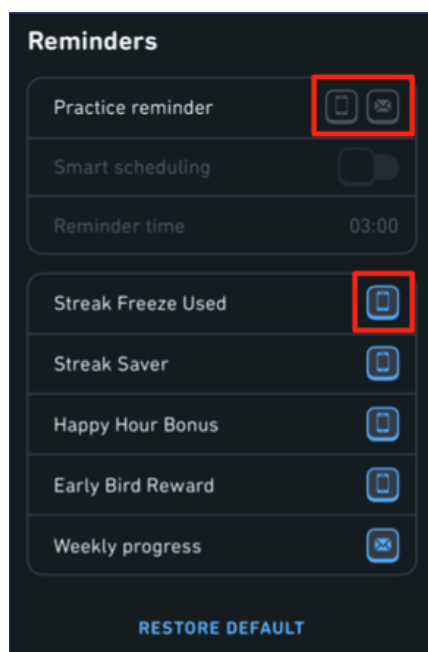
Since the causes of this finding were multi-faceted, we have multiple recommendations. To help discoverability, we recommend changing the bell icon to something else to avoid confusion and making sections of the settings collapsible to reduce scrolling time. To improve the efficiency of adjusting the notifications themselves, we suggest reducing the number of clicks by eliminating the requirement for opening a new page and also creating a “kill switch” for all notifications. We were unable to conclude whether or not the small button sizes cause accessibility issues, but we do think it would benefit all users to make these buttons larger and easier to select.

Area of Opportunity 2: Lack of Confidence in Changing Notifications

Severity 4 (High Severity)

6/10 participants mentioned that they aren't certain that they successfully turned off their notifications because there is a lack of enough verification/feedback from the app when they push buttons. Currently, clicking the blue phone/mail icons

next to each reminder makes them grayed-out as an indication that they had successfully been turned off (as seen in the figure below). However, participants desire stronger feedback from this process. This could be due to several factors: First, participants not being used to the type of buttons used. Some mentioned that they are more used to toggle buttons that are more traditionally used for turning notifications off.



Notification Settings Page, with examples of grayed-out and filled-in buttons.

Secondly, the preconceived notions that people have of Duolingo from social media / previous experiences. Some participants mention that they wouldn't expect to be able to turn Duolingo's notifications off easily because they see memes of Duolingo having "threatening and frequent notifications". As a way to work around this uncertainty, some mention that they would check their phone settings because doing so is a more sure way of determining whether or not they were changed: "I don't know if I turned off the notifications or not. I can go to my iPhone settings to turn it off - that's another way" (P1).

Our recommendation for this area of opportunity is to add a save button at the bottom that goes into an end screen similar to other screens across Duolingo to ensure user confidence. In future studies, it could be insightful to look into the type of button that promotes user confidence in changes.

Area of Opportunity 3: Unclear Terminology

Severity 3 (Mid Severity)

Another finding that stood out was the rampant unclear terminology on gamified language. The gamified nature of the app calls for features with specific terms such as “streak freezes”, “hearts”, “smart scheduling”, “personalized practice”, and “unlimited legendary”, among others. Six out of ten of the participants expressed confusion about the specific terminology throughout the studies. One participant remarked, “I don't understand— what is a streak freeze?” (P1). Multiple others commented on smart scheduling in the notification settings:

- “I assume smart schedule is just random whereas I could just turn it off and just do one time [notifications]. [...] not really sure what that is” (P3).
- “I’m not really sure what the smart scheduling is for” (P8).
- “What does [smart scheduling] mean?” (P9).

A specific page has lots of specific terms that are not defined prior to this page showing up: the Super Learners (premium subscription) page. A participant, while viewing the above page, stated the following: “Unlimited hearts--I still have no idea what that means. Personalized practice--I would want to know more about that. I can’t really click on these to learn more. Unlimited legendary--I have no idea what that means” (P2). Clearly, there is confusion around the specific terminology because there is no way of knowing what the terms mean on the spot.



Super Learners page during Onboarding.

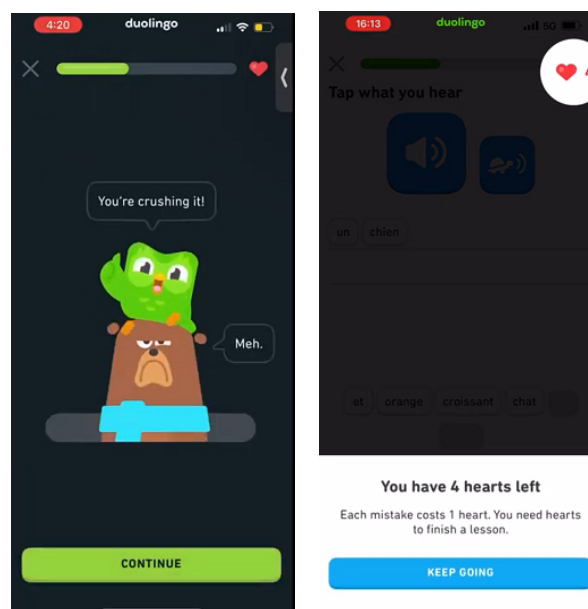
One recommendation to alleviate the confusion that comes with the vocabulary is to define all Duolingo-specific terms early on during onboarding. Another recommendation is to provide users with access to a “dictionary/information” feature where gamified, specific terms are explained to give users all the definitions in one easy-to-access place all the time. These fixes would

provide an experience that prioritizes learning rather than detracting from it while users wonder about the terms used.

Area of Opportunity 4: Branding & Gamification

Severity 2 (Low Severity)

While many users did express positive sentiments regarding the branding, we also discovered that Duolingo's branding, especially the more gamified aspects, were distracting for some. Many users cited cognitive overload from excessive animations, bright colors, and unclear gamified features. Three out of ten participants mentioned the app's aesthetic detracting from their appearance, and six out of ten participants had some form of criticism over the gamified features, such as the hearts system, XP, and streaks.



Example of Duolingo Animations and Heart System

The following quotes are from participants commenting on the branding and gamification of Duolingo:

- “It’s also kinda like a distracting app in general. The design isn’t necessarily the best [for] focus if you’re trying to learn” (P8).
- “This heart system is good to motivate people to actually pay attention. But I feel like it could also be very discouraging, especially if you’re new” (P10).
- “Too much gamification would mean you’re more invested in these streaks rather than the point behind it” (P3).

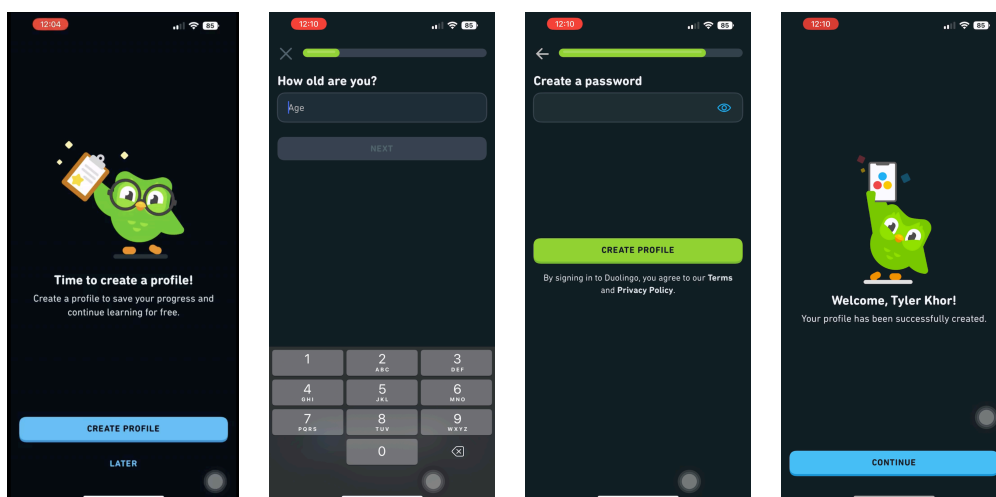
It's worth noting that this isn't to say these components were entirely unsuccessful, as discussed above. To address this finding, we suggest further research into the effectiveness of Duolingo's branding, given the conflicting viewpoints. Having a deeper understanding on what users enjoy about the app's aesthetic and what simply adds to cognitive overload has great potential for efficient redesigning should that be necessary. It's worth noting that this primarily applies to a wide-scale overhaul of the app's branding, as the aesthetic is more deeply rooted and important to understand before making large changes.

More actionable solutions would be providing settings to help users customize their Duolingo experience in a manner that can help cut down on cognitive overload. Giving users the ability to turn off animations or switch to using some more traditional language-learning methods (flash cards, definition pages, etc.) will enable them to balance having what features are enjoyable with which ones create superfluous load.

Area of Opportunity 5: User Profile Completion

Severity 1 (Lowest Severity)

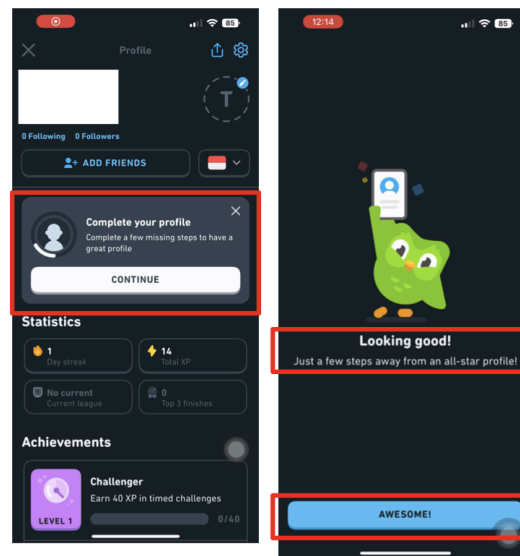
Overall, all participants find the profile onboarding process to complete their user profile intuitive as a result of the simple layout with a progress bar to indicate how far along users are, as well as the confirmation they receive from the end screen that says "Your profile has been successfully created".



Onboarding Screen Examples: Beginning, Some of the Middle, and End

However, 6/10 participants are confused/frustrated because they are met with a probe to "Complete your profile" in their user profile since they assume that they

had already finished the entire process (image below). Some mention that this confusion stems from having a mental model that setting up a profile should all be done in one place, so they assumed they were finished when they were onboarded the first time. Upon exploring this secondary profile completion probe, participants often opt to skip adding additional information (such as adding a profile photo and synchronizing their contact list), but are frustrated when the prompt does not disappear even though they finished clicking through it. Some mention feeling misled by the language used on the end screen (such as “Looking Good”, and “Awesome!”) that indicates that they successfully completed the task and the prompt should disappear.



Left: Profile Completion Probe in User Profile Page; Right: End Screen of Additional Profile Completion Pages

Two recommendations arose for this area of opportunity. We propose combining all profile completion steps into one task instead of two separate tasks to avoid confusion and support people’s mental models. Another recommendation is changing the language/text hierarchy on the end screen to give a stronger understanding to users of whether or not they completed their profile.

Next steps

In concluding our study, our next steps fall into two broad categories: further exploration and future learnings.

Further Exploration

Despite the thoroughness of our study, we found that several of the issues we discuss above could not be fully addressed through our research. These issues

concern Duolingo users' trust and perceptions of the brand; as well as certain accessibility issues we were not able to examine in our study.

Many of the issues our study demonstrated related to the notification settings of the app. To summarize, participants in our study struggled to find the settings, endured an inefficient process to adjust them, and then were not confident that they had actually successfully muted the notifications. While we were able to identify some issues with the interface that contributed to these problems, our team also speculates that Duolingo's brand identity may have impacted these interactions. In the past few years, Duolingo has reached internet infamy as memes circulated proclaiming the app's notifications are annoyingly omnipresent, invasive, and even threatening [1]. 6 out of the 9 participants who had heard of Duolingo before the study mentioned these memes throughout the study. Our team would be interested in further investigating how public perception of Duolingo notifications impacts users' trust in this part of the app.

Second, we also found that some aspects of the interface, while usable for able-bodied and neurotypical users, may not be accessible for users in other groups. While screening for ability status was outside the scope of our study, we believe that further studies looking at the following are conducted to ensure an inclusive experience for all:

1. We found that moving graphics and pop-ups were distracting and over-stimulating to a few of our participants. Our research team is interested in investigating how the many moving graphics and pop-ups impact neurodiverse users, who may be more severely affected by these aspects of the app.
2. We also found that some small buttons were tricky, but not impossible, for participants to select. Research on if these features are usable for participants with low vision or limited mobility in their fingers merits further research.

Future Learnings

One of our biggest takeaways in the process of completing this study was the importance of interdisciplinary communication across product teams. During the study, the app underwent two major software/UI updates, which impacted the study tasks. These changes had an adverse impact on the consistency of our data across participants. As new researchers, this experience demonstrated that in future studies we should be intentional to communicate with engineering teams about planned updates and plan our studies accordingly.

Additionally, in the future, we would like to be more intentional about the data we collect and how we plan to analyze it. We collected so much data that we did not

have the time or resources to analyze it as in-depth as we would have liked. For instance, we surveyed participants with the System Usability Scale (SUS) but did not have a plan for how we would integrate analyzing that data with our other more qualitative findings.

Appendix

Task List

Task	Task Description	Starting State	Successful Completion Condition
Walk-through the setup process	Users go through the steps to set up Duolingo as a first-time user	Initial landing page when the app is first downloaded onto a device	Finishing setup process to see the Home page of Duolingo
Go through a section of a lesson	After setting up, users will start their first lesson on their language of choice that they haven't ever learned before. First, they should review the lesson tips, then go on to complete the first lesson. The user should tell the researcher when they believe they are done with the lesson.	Set-up ending page	Completion Page: "Lesson Complete!"
Set up user profile	Finish setting up a user profile with the additional information asked for. After adding a profile, accept the achievement for doing that.	Finished lesson ending page	Completion Page: "Looking good!"

Adjust the amount of push notifications	Find settings with the intention of reducing the number of practice reminder notifications that they get.	Home page	All “practice” notifications successfully turned off
---	---	-----------	--

Test logistics

- Operating system: To perform our user testing, we worked with both Androids and iPhones on the Duolingo mobile app.
- Equipment: We used the participant’s phone to complete each study. To take notes, we designated filled out a notetaking google form with our personal laptops that was attached to a spreadsheet that would collect all answers. For the remote option, both the participant and the researcher joined a zoom call where the participant was instructed to screenshare their phone screen.
- Environment: We provided both remote and in-person testing for closer observation and expediting the process. We conducted the remote option using Zoom and the in-person option anywhere that is mutually convenient and quiet enough to not interfere with the study.
 - In-person
 - Sit beside the researcher with their device screen visible to both people
 - If one of our “new user” participants is an Android user, Ashlyn, Michelle or Mina can provide their phone for the participant to use
 - Screen record the phone if the participant is okay with it
 - Have the participant send the screen recording to the researcher
 - Remote
 - Have the participant join a Zoom call from their computer and their phone, but don’t connect to Audio using the phone.
 - The participant will screen share their phone.
 - Record the Zoom call using the Zoom recording functionality.

Screener data

- Age of participant
- Level of familiarity with iOS Duolingo
- Participant smartphone OS of choice
- Familiarity with Latin alphabet

Session data

- Number of wrong turns for each task, meaning the number of times the user diverts from the expected pathway for each task
- Specific paths that participants take for each task
- Areas of confusion within each task
- Participants' struggles and frustrations

Toolkit materials

[Table of Contents](#)

[Note-Taking form \(internal use\)](#) & [Interview notes \(written from form\)](#): We used a google form to take notes during the interviews, and then we used the integrated form-to-sheet option in Google sheets to transfer our notes into a Google sheet. This allowed us to use database capabilities on the notes, such as quickly searching and filtering. We got this idea from the assigned readings!

[Screening Questionnaire](#)

[SUS Questionnaire](#)

[Discussion Guide](#)

[Debrief Notes](#)

[Recruitment/Onboarding email template](#)

[Consent Form](#)

Recordings

[Participant 1](#)

[Participant 2](#)

[Participant 3](#)

[Participant 4](#)

[Participant 5](#)

[Participant 6](#)

[Participant 7](#)

[Participant 8](#)

[Participant 9](#)

[Participant 10](#)

Citations

[1] [Touch Targets on Touchscreens](#)

[2] [How Duolingo built a successful \\$250 million brand by being kind of a jerk](#)