External Evaluation Report

duolingo™

Education and Translation for Free

Formative Evaluation conducted by
Justin Schuyler and Samantha Geitz
College of Education, University of Missouri

May 12, 2013

Evaluation Target

Executive Summary

This document contains a formative evaluation of Duolingo.com, a free online language learning system. Duolingo utilizes crowdsourcing principles, providing translation services to businesses while teaching users a new language. Our intended purpose is to analyze whether Duolingo is meeting its current goals and whether it has the potential to be effective in the long-term.

In this report, we provide detailed information about Duolingo's background, goals, and pedagogical methods. We also explain our purposes for performing this evaluation, including the major stakeholders in the project as well as the decisions they'll have to make about the Duolingo system. In addition, we list important questions that must be asked as part of the formative evaluation processes, then detail our methodology, sample population, and instrumentation for gathering data that answers these questions and aids the stakeholders in making important decisions about the system. We also identify possible limitations to our planned data collection techniques.

As a result of our evaluation, we determined that Duolingo is overall quite effective at meeting its goals. Through our data collection, we determined that the majority of the site's users find Duolingo to be fun, easy to use, and highly instructional. Our recommendations target a few interface features that can be improved, especially relating to the audio exercises, along with suggestions to improve language immersion, skill appreciation, and degradation.

Client Assumptions

This evaluation was conducted to answer questions about the implementation and effectiveness of Duolingo from our clients, Luis von Ahn and Severin Hacker. Our investigation is based on their need for recommendations to support decisions concerning the design, development, and feasibility of their education models. They assume that the results of this evaluation will provide evidence to drive changes and improvements to the system while also measuring value to determine its feasibility as a successful language-learning and translation-providing service in the future.

Introduction

In the last few years, a fairly large number of sites have offered language courses online; however, many of these -- such as the popular "Rosetta Stone" software -- can cost up to several hundred dollars. Even of the "free" services, many will allow users to take basic courses without cost and then charge users for more advanced offerings. Duolingo is a new online language learning system that claims to be entirely free for beginning to advanced courses, teaching users a new language while having them translate content on the web.

The goal of this study was to evaluate Duolingo's overall effectiveness as a learning system. We have performed a formative evaluation and, at the client's request, we have developed recommendations for the system and reviewed the effectiveness of existing elements. In the future, we will also be appraising the viability of Duolingo's crowdsourcing-based business model.

This evaluation was conducted by Justin Schuyler and Samantha Geitz, two master's degree students at the University of Missouri - Columbia, as part of Dr. Julie Caplow's Formative and Summative Evaluation course in Spring 2013.

Background

Duolingo is a language-learning system that provides free education and translation services through the power of crowdsourcing. It is distributed through the web and also available on iOS mobile devices with an Android version coming in the next few months. Duolingo is built on the model that people want to learn a new language and that both individuals and businesses want fast, accurate, and inexpensive content translation. So, people pay with their time, learning a language while providing valuable translations.

Learning a language through Duolingo is free, and at the moment upload accounts for users who want content translated are also free, although a disclaimer on the site warns that "in the future we may charge for speed and accuracy." Because they do not charge any of their users, their funding is almost entirely external; notable sources include venture capital firms NEA and Union Square Ventures, actor Ashton Kutcher, and author Tim Ferris.

As a company, Duolingo is fairly small, employing about twenty people in a variety of roles. In addition to the CEO and CFO, the team consists of several designers, developers, language experts, and researchers. Their main office is in Pittsburgh, PA.

A variety of language courses are available, including Spanish, English, French, German, Portuguese, and Italian, with a Chinese course coming soon. At the moment, translation services offered include English to Spanish as well as German, Spanish, and French to English;

no translation services are supported for Portuguese or Italian yet. The site claims that a recent independent eight-week study assessing Duolingo as a language learning tool found that the service is more effective than college language courses, teaching the same material in a shorter amount of time. (Vesselinov & Grego, 2012)

The system incorporates principles of gamification in order to drive user progress; users earn "experience points" to gain levels, "lose hearts" for incorrect answers, and also have the option of timed practices for extra experience points.

Purposes

Evaluating Duolingo serves two primary purposes:

- To identify improvements to the design, development, and implementation of its teaching and translating services; and

- To determine the effectiveness and feasibility of its education and business models.

These purposes address potential changes to the system's future development. These formative elements include:

- To enhance the overall system, its interface, design, navigation, and accessibility;

- To map the pedagogical structure, analyzing educational artifacts like exercises, motivational tools, performance gauges, testing and sociability;

- To better the process of producing real-world translations; and

- To describe Duolingo's implementation and technical requirements.

Audiences

The primary audience includes Luis von Ahn and Severin Hacker, the CEO and CTO, respectively, of Duolingo. The two co-founders of the company requested the external evaluation from the authors of this report.

The following groups of people are also stakeholders in this evaluation:

- People interested in learning a new language. This may be based out of necessity (stuck in a foreign land) or curiosity. However, Duolingo's playful design is more indicative they are focusing on the latter group.

- People who have or are considering applying for a free Duolingo upload account in order to have content translated by the site's users.

- Duolingo designers and engineers: our recommendations can be used to improve the overall system and assist in making improvements to weaker areas of the application.

- Other businesses or institutions who are considering using a crowdsourcing model to drive content creation or another similar goal.

- Other businesses or institutions who are considering using similar gamification principles to motivate users in some way.

Decisions

After we gather information through our formative evaluation, the client will have several decisions to make about the Duolingo system. The following decisions have been

identified as important considerations for the stakeholders based on the results of our evaluations:

- Improvements to the website or the mobile interface

- Recommendations for implementation of new content, such as a new language course

Questions

The following questions will be answered to address the purposes of this evaluation and provide evidence to drive decision-making:

- How can the user interface, navigation, or design be improved?

- What pedagogical strategies are used to convey the system's material?

- What recommendations can be made regarding the structure and content of tasks, exercises, and assessment?

- What are the implementation steps and technical requirements of Duolingo?

Methods

To fulfill the purposes and address the questions outlined above, a collection of evaluation methods were employed. These methods were selected to provide multiple sources of data and insights into the design, development, effectiveness, and feasibility of Duolingo. They include:

- Focus groups with target users;

- Questionnaires;

- Usability analysis based on standards and heuristics;

- Review from subject matter experts in interface design, learner population, and content

The following table illustrates which of the four evaluation methods apply to each of our formative questions.

Formative Questions	Focus Groups	Questionnaire	Usability Analysis	Subject Matter Expert
How can the user interface, navigation, or design be improved?				Interface Design
What pedagogical strategies are used to convey the system's material?				Learner Population
What recommendations can be made regarding the structure and content of tasks, exercises, and assessment?				Learner Population and Content
What are the implementation steps and technical requirements of Duolingo?				

Sample

The primary sample for this evaluation focused on a large population of target users with various backgrounds. Nonconforming demographics, motivations, and experience levels provided a representative sample analogous to typical Duolingo users and allowed for more practical results. Data was collected from a large group using appropriate, easily distributed methods like questionnaires and online assessments.

Random individuals were then selected to give more detailed feedback in a focus group setting; these users attended a focus group session after using Duolingo for two weeks to a

month. Findings from the smaller samples were compared with trends in the larger population and serve as evidence for explaining those results.

An in-depth usability analysis was conducted by the evaluators, Samantha Geitz and Justin Schuyler. The analysis focused on Duolingo's German and French instructional materials.

In addition, subject matter experts were used for interface design, learner population, and content. Our interface design expert is Samantha Geitz, a professional web developer with two years of experience in the field. Our learner population expert is Jeremy Winn, M.A.Ed., an A.B.D. doctoral candidate in the field of Educational, School, and Counseling Psychology at the University of Missouri. Our content expert is Justin Schuyler, who has studied German for 10 years and earned a minor in the subject at the University of Missouri.

Data Collection and Instrumentation

- Questionnaire (Appendix A): in order to evaluate users' opinions about the Duolingo system as well as how their previous experiences with online language learning systems may have impacted their thoughts on Duolingo's application of elements such as pedagogy, gamification, etc., an online questionnaire was given out to new users of Duolingo. The questionnaire is loosely based off of the Reeves Questionnaire and follo   ws the guidelines proposed in Robert Frary's "Hints for Designing Effective Questionnaires" (Frary, 1996).   Most of the items are in yes/no format or incorporate a Likert scale, but there are several open-ended questions as well for users to detail their thoughts.

- Subject Matter Expert checklists (Appendix B-D): for our subject matter experts in interface design, content, and learner populations/pedagogy, we created a checklist of relevant elements for evaluation. The SME evaluated whether the item in question is "Excellent," "Successful," "Developing," or "Deficient," then wrote comments in the space provided. Our interface checklist focused on system status, documentation, and functions; aesthetics; and accessibility. The learner population checklist included instructional methods, assessment, and motivation. Finally, the content checklist examined the learning material and delivery methods.

- Focus Group protocol (Appendix E): our first focus group protocol was developed for users who have only been using the Duolingo site for two weeks to a month. In it, we focused on their language learning goals, their tools and environments, their processes for accessing the site, and the incentives and motivations that keep them returning.

- Usability heuristic (Appendix F): Jakob Nielsen's *10 Usability Heuristics* (Nielsen, 1995) and a heuristic for video game evaluation (Pinelle, Wong, and Stach, 2008) inspired the development of a custom usability heuristic for gamified interface design. The evaluators used this heuristic in a usability analysis to identify potential problems and recommendations for Duolingo's interface, navigation, and overall design.

For our questionnaires, we targeted twenty-five people on Facebook who had at least some experience with the Duolingo system. We then sent them a questionnaire (Appendix A) to gather their feedback and received twenty-one responses. Of the users who submitted a

questionnaire, we asked eight to use Duolingo for several weeks and then held two in-person focus group sessions (asking each user to choose whichever was most convenient) based on our focus group guide (Appendix E) to gather their thoughts on learning a language with the system. These users' ages ranged between 17 and 51, and included several students, a housewife, and several people employed full-time in a variety of positions within the private sector. Five of our focus group participants were male and three of them were female.

We also performed subject matter expert evaluations for the interface design (Appendix B), pedagogy (Appendix C), and German language content (Appendix D) components of the Duolingo system. The evaluation guides for all three were designed before we performed our evaluation in order to target the effectiveness of Duolingo in these areas. Our subject matter experts used Duolingo over the course of several weeks to become familiar with how the system works and to form a thorough opinion on its successes and problems.

A usability analysis was conducted over a week of typical use through the lens of our gamification heuristic (Appendix F). Over that week, notes were collected regarding Duolingo's interface, performance, and accessibility.

Data Analysis

When we analyzed the data gathered by our questionnaire and focus groups, we looked at overall trends as well as individual factors such as age, background, etc. We looked at a summary of scores to look for possible areas of improvement and performed a thorough content analysis, reviewing user feedback for common themes, especially regarding problem areas and suggested solutions.

Our subject matter experts relied on their own extensive experience in the areas we asked them to evaluate along with their time spent using the Duolingo system to form their conclusions. They each completed the relevant evaluation guide, rating various components of the system on a 1-4 scale and writing comments in the space provided. To analyze the data, we looked at a summary of overall scores along with comments in order to target problem areas and make recommendations accordingly.

The heuristic notes were consolidated, reviewed, and explored for feasibility and priority. The analysis resulted in two lists: one describing how well many of Duolingo's interface elements matched principles from the heuristic and another detailing what could be improved.

Limitations

Several caveats must be observed before interpreting or generalizing the results of this evaluation. These limitations are primarily driven by scope and resource constraints. For one, Duolingo currently offers six language courses. Each course follows the same general topics and pedagogical strategies; however, the content is so deep that it is more feasible to focus on a specific language as a template rather than trying to evaluate the system as a whole. German was chosen for this purpose given the evaluators' experience with the language.

Another consideration involving resources was the limited amount of time available for formative evaluation. To fit the schedule, subject matter experts and participants had to be selected quickly. If the experts and Duolingo users had more experience, more detailed feedback might have been available.

Finally, the volunteers used for questionnaires and assessments were not compensated for their participation. This may have affected their level of engagement or motivation in providing quality answers. The large population used to collect this data may help to lessen this effect.

Results

The general feedback from all of our data collection methods was very positive. Questionnaire and focus group participants all enjoyed their experience with Duolingo and expressed interest in continuing to use it. Of our eight focus group participants, seven said that after several weeks of using Duolingo to learn a language they would feel comfortable having a basic conversation with a native speaker. The other participant said, "I'd probably feel comfortable in another week or two, but I'm not sure I could come up with full sentences on the fly." Three of our users had previous experience with learning a language at a university level, and all three said that they found Duolingo to be a more effective learning system than what they found in the classroom. One commented, "I got a much better balance of reading, writing, and speaking than when I was in a classroom with forty other students."

Our subject matter experts found Duolingo's design and content to be very well structured and thoughtfully presented. Our learning systems expert, on a scale from 1-4 -- with 1 being "Deficient" and 4 "Excellent -- rated Duolingo's use of pedagogy to be a 3.75, calling it "especially well done... I don't know a lot about the language content, but it seems to do an excellent job teaching the material to [learners.]" Most of his suggestions for improvement related to the user's educational progress and immersion strategy. Using the same scale, our

interface expert rated Duolingo's design as a 3.5; she said that "it's easy to use and pleasing to the eye... I think most users could pick it up without much difficulty, and the iOS app is very convenient for users who want to interact with the system on the go." On the 1-4 scale, our content expert rated the system as a 3, saying, "Its formal components are lacking (which is very much intentional)... content-wise, they have a good mixture of basic and complex grammar, vocabulary, and even colloquial phrases. I never ran into poor German."

The usability analysis also confirmed the effectiveness of Duolingo's gamified elements. Following the heuristic, we found that the system behaved consistently, allowed a good level of customization, and provided the proper level of notification and guidance to learners. The comments of our focus group participants also support this evidence; several of them were acquaintances and had actually independently decided to "compete" with each other for the most points. These users pointed to this feature as something that kept them highly motivated.

The problems we discovered from our methods pointed primarily to existing features that could be revised and improved. The usability analysis uncovered some interface design and organizational issues, while the subject matter experts explored how certain learning materials, like the immersion exercises, and effects, skill appreciation and degradation, could be re-envisioned.

Recommendations

- To address inaccuracies in Duolingo's audio recording exercises, we recommend emphasizing the audio playback option, which is used automatically in other exercises to demonstrate correct pronunciation. Duolingo's long term goal should either be

upgrading or replacing the current audio processing system with one capable of detecting proper pronunciation and inflection as well as distinguishing between similar words and word endings.

- The recording button should be disabled and a loading icon should be added to provide users with feedback while their audio is being processed.

- Skill appreciation and degradation should be revised to reflect real-world, educational progress. Appreciation should be built over time, whereas degradation should be tailored to each user's individual learning profile.

- Real-world translation should be reintegrated into Duolingo's primary instructional material and revised to better fit its students' skills.

- Iteratively fix Duolingo's interface design issues. In the "mark all as correct" exercises, reposition the checkboxes. Upon module completion, allow the user to exit the mastery dialog with the "enter" key. In the exercise interface, modify the "Tips and Notes" button so it stands out.

- Duolingo should expand their immersion strategy by incorporating educational audio, video, and interface elements.

- A tagging system should be introduced to organize Duolingo's discussion forum resources.

- Formal instruction should be integrated into Duolingo's exercises where appropriate.

Works Cited

Frary, R. B. (1996). Hints for designing effective questionnaires. *Practical Assessment, Research, & Evaluation*, *5*(3), Retrieved from http://pareonline.net/getvn.asp?v=5&n=3

Nielsen, Jakob. (1995). "10 Usability Heuristics." *10 Heuristics for User Interface Design: Article by Jakob Nielsen*. Nielsen Norman Group. Retrieved from http://www.nngroup.com/articles/ten-usability-heuristics/.

Pinelle, D., Wong, N., and Statch, T. (2008). "Heuristic Evaluation for Games: Usability Principles for Video Game Design." Retrieved from http://www.irit.fr/recherches/ICS/projects/twintide/upload/447.pdf.

Vesselinov, R., & Grego, J. (2012). *Duolingo Effectiveness Study*. Unpublished manuscript, Queens'College, New York. Retrieved from http://static.duolingo.com/s3/DuolingoReport_Final.pdf.

Appendix A

**Duolingo.com Interface Design
Questionnaire**

The purpose of this survey is to determine usability problems found on the interface of the Duolingo application. All questions are optional¸ and all responses are anonymous. Thank you for your participation!

Background Information

The following questions will determine your prior experience with other language learning websites. Please place a check mark next to your selected answer.

1. How long have you been using Duolingo?
      ___ LESS THAN 1 MONTH
      ___ 1 - 6 MONTHS
      ___ 6 – 12 MONTHS
      ___ 12+ MONTHS

2. Have you previously studied a language in an academic setting?
      ___ YES
      ___ NO

3. Have you ever used another online language course such as Rosetta Stone, LiveMocha, Babbel, etc.?
      ___ YES
      ___ NO

4. If you answered **"YES"** to question 3, please select any of the following systems that you have used in the past.
      ___ LIVEMOCHA
      ___ ROSETTA STONE
      ___ BABBEL
      ___BUSUU
      ___VERBALPLANET
      ___LINGQ

5. How many languages do you currently speak with at least intermediate-level skill?
   ___ 1
   ___ 2
   ___ 3
   ___ 4+

6. How would you respond to the following statement? **"I am good at learning languages."**
   ___ STRONGLY DISAGREE
   ___ DISAGREE
   ___ NEITHER AGREE NOR DISAGREE
   ___ AGREE
   ___ STRONGLY AGREE

Learning System

The following questions are intended to help gather your thoughts about Duolingo's interface and instruction. Unless otherwise directed, please place a check mark next to your selected answer.

7. How would you respond to the following statement? **"I am willing to pay for quality online language instruction."**
   ___ STRONGLY DISAGREE
   ___ DISAGREE
   ___ NEITHER AGREE NOR DISAGREE
   ___ AGREE
   ___ STRONGLY AGREE

8. How would you respond to the following statement? **"I feel comfortable using the internet."**
   ___ STRONGLY DISAGREE
   ___ DISAGREE
   ___ NEITHER AGREE NOR DISAGREE
   ___ AGREE
   ___ STRONGLY AGREE

9. How would you respond to the following statement? **"I feel that, overall, Duolingo is easy to use."**
   ___ STRONGLY DISAGREE
   ___ DISAGREE
   ___ NEITHER AGREE NOR DISAGREE
   ___ AGREE
   ___ STRONGLY AGREE

10. How would you respond to the following statement? **"I think that Duolingo's color scheme and graphics are pleasing."**
      __ STRONGLY DISAGREE
      __ DISAGREE
      __ NEITHER AGREE NOR DISAGREE
      __ AGREE
      __ STRONGLY AGREE

11. How would you respond to the following statement? **"I am easily able to navigate through the Duolingo.com application."**
      __ STRONGLY DISAGREE
      __ DISAGREE
      __ NEITHER AGREE NOR DISAGREE
      __ AGREE
      __ STRONGLY AGREE

12. How would you respond to the following statement? **"I find that Duolingo's use of game elements, such as 'leveling up' and 'hearts', motivate me to try harder."**
      __ STRONGLY DISAGREE
      __ DISAGREE
      __ NEITHER AGREE NOR DISAGREE
      __ AGREE
      __ STRONGLY AGREE

13. How would you respond to the following statement? **"I think that Duolingo is addicting."**
      __ STRONGLY DISAGREE
      __ DISAGREE
      __ NEITHER AGREE NOR DISAGREE
      __ AGREE
      __ STRONGLY AGREE

14. How would you respond to the following statement? **"I think that Duolingo is more effective than college courses that I have taken."**
      __ STRONGLY DISAGREE
      __ DISAGREE
      __ NEITHER AGREE NOR DISAGREE
      __ AGREE
      __ STRONGLY AGREE
      __NOT APPLICABLE

15. Have you accessed Duolingo from a mobile device?
      __ YES
      __ NO

16. For each of the items below, please circle "VC" if you are very comfortable with using that feature, "SC" if you are somewhat comfortable with using that feature, and "NC" if you are not at all comfortable with using that feature.

EXAMPLE: (VC)   SC   NC

Website Element	Very Comfortable	Somewhat Comfortable	Not Comfortable
Choosing a language of study	VC	SC	NC
Signing up for Duolingo	VC	SC	NC
Selecting a lesson in the skill tree	VC	SC	NC
Completing assigned lessons in the language of study	VC	SC	NC
Translating real-world examples given on the site	VC	SC	NC
Accessing and completing lesson practice	VC	SC	NC

17. What is the one thing that you like **most** about Duolingo? Please explain in the space provided below:

18. If you could change one thing about Duolingo, what would it be? Please explain in the space provided below

Appendix B
Subject Matter Expert Evaluation – Interface

System status, documentation, and functions, based on Jakob Nielsen's "10 Usability Heuristics for Interface Design" (Nielsen, 1995)

Question	Excellent	Successful	Developing	Deficient
Is the system status always visible to the user? Comments:	☐	☐	☐	☐
Does the system follow real-world conventions, using words, phrases, and concepts familiar to the user? Comments:	☐	☐	☐	☐
Does the system support user control and freedom – can the user "undo" and "redo" tasks? Comments:	☐	☐	☐	☐

Does the system support consistency and standards? Do words, situations, and actions mean the same thing?	☐	☐	☐	☐
Comments:				
Are error prevention measures in place? Does the system present users with a confirmation option before they commit to an action?	☐	☐	☐	☐
Comments:				
Does the system rely on recognition rather than recall and make instructions either visible or easily retrievable?	☐	☐	☐	☐
Comments:				

Does the system cater to experienced and inexperienced users by allowing users to tailor frequent actions? Comments:	☐	☐	☐	☐
Do system dialogues frequently contain irrelevant information? Comments:	☐	☐	☐	☐
Are error messages expressed in plain language and offer users a solution? Comments:	☐	☒	☐	☐
Is documentation for the system available to the user and easily searchable? Comments:	☐	☐	☐	☐

Aesthetics

Question	Excellent	Successful	Developing	Deficient
Does the site have a consistent header/logo? Comments:	☐	☐	☐	☐
Is the navigation area consistent? Comments:	☐	☐	☐	☐
Does the site make good use of basic design principles - repetition, contrast, proximity, and alignment? Comments:	☐	☐	☐	☐
Does the site have a good balance of white space? Comments:	☐	☐	☐	☐

Does the site make appropriate use of chunking to separate large blocks of text? Comments:	☐	☐	☐	☐
Is the color scheme consistent and visually pleasing? Comments:	☐	☐	☐	☐
Is the typography consistent and visually pleasing? Comments:	☐	☐	☐	☐

Accessibility

Question	Excellent	Successful	Developing	Deficient
Is the website cross-browser compatible? Does it look the same in current versions of each of the major browsers? Comments:	☐	☐	☐	☐
Does the site work in older versions of Internet Explorer? (IE6-IE9) Comments:	☐	☐	☐	☐
Is the site responsive? Can it be accessed on mobile phone and tablet browsers? Comments:	☐	☐	☐	☐

Does the website validate with W3C coding standards? Is the HTML/CSS valid? Comments:	☐	☐	☐	☐
Do all significant images have a valid alt tag? Comments:	☐	☐	☐	☐
Are text-based alternatives in place to convey essential alternatives that are featured within images or multimedia files? Comments:	☐	☐	☐	☐
Is the typography easily readable? Comments:	☐	☐	☐	☐
Does the site load quickly on an average internet connection? Comments:	☐	☐	☐	☐

Appendix C

Subject Matter Expert Evaluation – Learner Populations

Instructional Methods

Question	Excellent	Successful	Developing	Deficient
Are the instructions for task clear? Comments:	☐	☐	☐	☐
Are the instructions concise or interesting/engaging? Comments:	☐	☐	☐	☐
Do the instructional tasks build upon prior knowledge and experience? Comments:	☐	☐	☐	☐
Are the tasks authentically meaningful to users, demonstrating clear real-world application? Comments:	☐	☐	☐	☐

Are cooperative, social learning strategies utilized? Comments:	☐	☐	☐	☐
Is learning primarily independent? Comments:	☐	☐	☐	☐
Are tasks based on user inquiry? Comments:	☐	☐	☐	☐
Are tasks dictated by the training program with little or no user input? Comments:	☐	☐	☐	☐
Are ample, high quality practice opportunities given prior to assessment? Comments:	☐	☐	☐	☐

Assessment

Question	Excellent	Successful	Developing	Deficient
Are the instructions for assessment clear? Comments:	☐	☐	☐	☐
Is the assessment interface easy to use? Comments:	☐	☐	☐	☐
Are the tasks appropriate for the ability level of the user? Comments:	☐	☐	☐	☐
Is feedback correct and prompt? Comments:	☐	☐	☐	☐
Is there a clear relationship between assessment feedback and meaningful learning? Comments:	☐	☐	☐	☐

Motivation

Question	Excellent	Successful	Developing	Deficient
Is the material appropriately challenging—not too boring or too frustrating? Comments:	☐	☐	☐	☐
Is the learning material interesting or novel? Comments:	☐	☐	☐	☐
Are users rewarded for their success, and are the rewards meaningful? Comments:	☐	☐	☐	☐
Is the learning experience autonomous? Comments:	☐	☐	☐	☐

Is the user able to make choices about their learning? Comments:	☐	☐	☐	☐
Is the learning tailored to meet the specific goals and values of the user? Comments:	☐	☐	☐	☐
Are positive social interactions facilitated by the program's tasks? Comments:	☐	☐	☐	☐

Appendix D
Content Subject Matter Expert

Learning Materials

Question	Excellent	Successful	Developing	Deficient
Does the progression of content transition smoothly and logically? Comments:	☐	☐	☐	☐
How well are rules and examples integrated? Comments:	☐	☐	☐	☐
How well is grammatical gender handled? Comments:	☐	☐	☐	☐
How well is syntax and sentence structure handled? Comments:	☐	☐	☐	☐

How well is case handled (nominative, accusative, dative, genitive)? Comments:	☐	☐	☐	☐
Is the vocabulary varied and relevant? Comments:	☐	☐	☐	☐
Are the students encouraged to extrapolate and experiment with the knowledge they gain? Comments:	☐	☐	☐	☐

Delivery Methods

Question	Excellent	Successful	Developing	Deficient
How clear are the exercise instructions? Comments:	☐	☐	☐	☐
How valuable are the social tools? Comments:	☐	☐	☐	☐
Are the strategies and vehicles for information delivery varied and rich in multimedia? Comments:	☐	☐	☐	☐
Do the exercises complement one another? Comments:	☐	☐	☐	☐
Do artifacts aid in the delivery of the learning materials? Comments:	☐	☐	☐	☐

Appendix E

Focus Group Protocol

Introduction – the facilitator should begin the session by reading the following:

"Thank you for coming today!

My name is **(Samantha Geitz, Justin Schuyler),** and I'm a master's degree student at the University of Missouri. **(Ad lib introduction.)** Why don't we get started by introducing ourselves? **(Ice breaker – participants give name, something interesting about themselves, and how long they have been using Duolingo.)**

We've been asked by Duolingo.com to conduct a focus group in order to hear your thoughts on the effectiveness of their language courses. All of you have been using Duolingo for at least a few weeks to learn a new language. We want to get feedback on how they're doing so that they can make improvements to their site and future language courses. This session is not going to be recorded, but I will be taking notes. However, everything you share with me will be confidential, so please feel free to be open and honest. No part of our discussion that includes names or other identifying information will be used in any reports, displays, or other publicly accessible media coming from this research.

This session shouldn't take more than two hours at most. Does anyone have any questions?

Let's begin!"

Vision	What is your main goal when you're using Duolingo?

Expectations	How long do you expect it will take you to meet your goal?
	How often do you use Duolingo? (Daily, weekly)
Feedback	How will you know when you are meeting your language learning goal?

	What is a good way for you to find out how well you are meeting your goal?

Focus Group Protocol

Tools	How do you most commonly access the Duolingo website? (Windows/OSX/Linux and IE/Firefox/Chrome/Opera)
	Have you ever had any trouble accessing their website?
	Have you ever used their iPhone mobile application? How often?

	Do any of you use Android? (If yes) If they developed an Android application, would you use it? How often?
Environment	Where do you most commonly access Duolingo? (If home, specific location of the house) Are there any specific challenges with accessing Duolingo in that environment?

Focus Group Protocol

Processes	During a typical session using Duolingo, what activities do you perform? (Courses, translations, practice)
	Do you have a microphone? Do you utilize the speaking section of Duolingo?
	Are you using any other resources to help you use a language? (Books, another site, native speakers, tutors, etc.)

Rewards	In what ways will you benefit if you learn a language through Duolingo?
	In what ways do you expect to use the language you learned through Duolingo?

Recognition	How do others view your use of Duolingo?
	Do you compete with any of your friends on Duolingo? (If yes) How do you feel when you're ranked higher than them on the leaderboard?
Incentives	In what ways does Duolingo encourage you to spend some time every day learning a language?

	What other incentives would lead you to do more?
	How do you feel when you get an incorrect answer and lose a heart?
	How do you feel when you gain a level?

Motivation	What led you to sign up for Duolingo?
	What is the longest participation streak you've achieved?

Self-Concept	If you learn to speak another language fluently with Duolingo, how would you feel differently about yourself?
Capacity	What is the most challenging part about Duolingo to you?

Knowledge/Skill	What skills do you believe are most important in order to learn to speak another language well? (Reading, writing, speaking, etc.)
	After using Duolingo for a few weeks, which of your skills do you think are strongest? Which of your skills are weakest?
	Would you feel confident having a basic conversation with a native speaker of the language you're learning?

Conclusion	Do you have any other thoughts about Duolingo that you'd like to share?

Conclusion – the facilitator should read the following aloud:

"Thank you so much for your participation in this focus group. I just want to say again that everything you've shared is confidential, and I appreciate your honesty. Does anyone have any questions for me about our discussion today or about our research on Duolingo?"

<div align="center">

Appendix F
<u>Gamification Heuristic: Evaluating Game Elements in Interface Design</u>

</div>

1. Respond Consistently

When a user performs an action, they should typically be provided with a predictable response. Consistent behaviors following task completion, events, and navigation establish a fluid interface. Unpredictable or inconsistent responses can lead to frustration and confusion.

2. Allow Customization

Users should be able to customize their experience. This is especially important for modifying difficulty and technical requirements. Resource -heavy, demanding audio and video features should be scalable to encourage participation on multiple platforms. Additionally, features should be modular, allowing users to add and remove components.

3. Unobstructed Navigation

The user should have a clear view of his or her surroundings. Navigating the interface should be simple, intuitive, and unhampered.

4. Skippable Content

If instructions or content are frequently displayed, they should be skippable. Not doing so interferes with the flow of the interface.

5. Harness Controls

Common controller elements like WASD, the number, and arrow keys should be taken advantage of to provide the user with alternative, quick, and predictable options for interacting with the interface. Controls should be intuitive and customizable.

6. Handle Responsively

User input should be responsive, affecting the interface quickly and accurately. The user should not have to wait between inputs and lag time should be minimized.

7. Show Status

User actions and tasks are often based on status, like handling notifications, reacting to a leaderboard, or strategizing a next move. Status, statistics, and other information relevant to the user should be displayed cleanly and clearly. More detailed information should be available upon inspection.

8. Provide Help

Instructions and help should be easy to find and consistently placed. Contextual help should assist users in finding solutions to activity-specific problems. Common issues should be identified, organized, and made searchable.

9. Represent Visually

Visual representations, like icons and avatars, should be used to encapsulate and convey information. Their function should be evident in their design.