



Comprehensive Usability Assessment for NEEK Language Acquisition App using Wearable Eye Tracking, Questionnaires and Face-to-Face Interviews

ارزیابی تجربه کاربری اپلیکیشن آموزش زبان
نیک با استفاده از ردیابی حرکات چشم کاربر،
پرسشنامه و مصاحبه‌ی رو در رو

Arghavan Aslani (aslaniarghavan@gmail.com)

BehPAS Cognitive Technologies

Niloofar Saeedi (niloofarsaeedi.ed@gmail.com)

BehPAS Cognitive Technologies

Parisa Navidi (parisanavd@gmail.com)

BehPAS Cognitive Technologies

Bahador Bahrami (bahador.bahrami@lmu.de)

BehPAS Cognitive Technologies

Crowd Cognition Group, Ludwig Maximilian University, Munich Germany

<http://crowdcognition.net>

Abstract

This white paper describes the details of usability assessment of NEEK Language Acquisition app by BehPAS Cognitive Technologies. A total of 24 users were recruited (12 female, 12 male; equally sampled in age groups 30-40, 40-50, and 50<). Usability was assessed with wearable eye tracking, questionnaires and face-to-face interviews. Quantitative and qualitative methods were used to interpret the data and key insights are provided that indicate a number of key strategic directions for NEEK's development.

Introduction

In our world today, private and public educational services are predominantly targeted at children and young adults. Lifelong education for adults is often dismissed on the assumption that adults are not good at learning or that they do not need to, or are not keen to learn. Defying these assumptions, NEEK Language Acquisition System has taken up the challenge of providing a digital platform (apps for Android and iOS) and customized content for learning English as a second language for grown-up adults i.e. 30 years and older.

To ensure evidence-based development of its UX and content, NEEK consulted BehPAS Cognitive Technologies to conduct a comprehensive Usability Assessment for the application. What follows is a detailed report of the methodology, the results and, most importantly, a set of strategic evidence-based insights for subsequent development of NEEK.

Methods

1. Head-Mounted Eye Tracking

Our eyes, often considered as the "windows to the soul," offer valuable insights into our inner world. Eye tracking is a potent technique for gathering data on gaze behaviors like where one looks, what we pay attention to and what we ignore. This method serves as a versatile tool for addressing various core inquiries, such as deciphering emotions, identifying interests, assessing levels of distraction or fatigue, and gauging overall mental and physical well-being. By measuring gaze indicators, in conjunction with the surrounding environment, we can understand and quantify the human experience.



Figure 1. Eye tracking

Eye movements contain valuable data about individuals' inclinations and preferences, rendering eye tracking a potent instrument in marketing and design realms. It facilitates the extraction of valuable insights into users' engagement with visual elements, telling us about user's areas of interest and focus of attention. Such insights can be leveraged to enhance a spectrum of elements ranging from UX design and point-of-sale exhibits to video content and online store arrangements.

We used Pupil Invisible (PI, [link](#), see Figure 2) wearable eye tracking device to assess user experience. Wearable eye tracking systems mount the video camera(s) onto a glasses-like frame, allowing the subject to move freely while their eyes are tracked from close up. Typically, they also include a forward-facing camera that records a video of the person's field of view, enabling eye-related signals to be analysed in the context of the changing visual scene. Pupil Invisible detects the appearance and/or features of the eye(s) under the illumination of ambient light to track the gaze location. The freedom of movement offered by Pupil Invisible vastly expands the range of tasks and situations in which eye-related signals can be recorded and analysed. The possibility of recording eye tracking data in open-ended environments yields rich, complex video datasets, with methods for analysing these data continuing to be an important topic of active research.

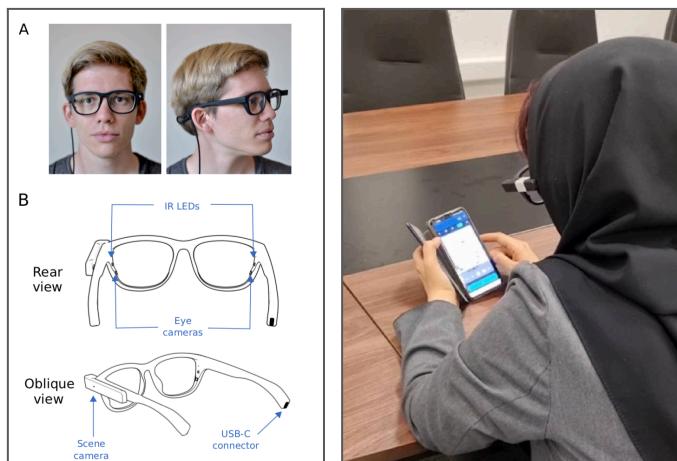


Figure 2. Pupil Invisible wearable eye tracking device. Left: schematic of components. Right: NEEK user wearing the device, allowing for unobtrusive, non-interfering recording and assessment of user experience.

To assess user experience and behavioural interaction with NEEK UX, we recruited our participants one at a time and had them wear the PI device for the duration of minimum 30 minutes of engaging with the NEEK application on their own cell phone (see Figure 3 and video link [here](#)). Each participant received a link (via SMS) for downloading and installing NEEK on their phone and proceeded to install and use NEEK. We recorded their experience and behavioural interactions with the app and submitted the data to quantitative analysis to answer a number of questions that have been listed in the Result section.

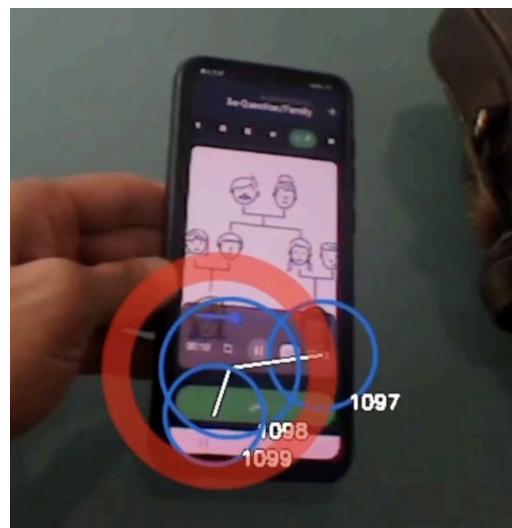


Figure 3. Example snapshot of eye tracking video data. Red circle: focus of attention. Blue: Three successive fixations. Lines: trajectories of transitions between fixations. Numbers indicate temporal order within the longer 30 minute video.

2. Questionnaires

At the end of the eye tracking procedure, once the participant had completed their interaction with NEEK for at least 30 min, they answered a detailed Questionnaire (see [LINK](#)). In this questionnaire we asked the participants to report how comfortable and pleasant they found working with NEEK, what they would change in it if they could, whether they would be keen to recommend NEEK to friends and family, and give a rough estimate of the price value of having the unlimited subscription to NEEK. The data was saved in a spreadsheet and subjected to quantitative assessment reported in the result section.

3. Interviews

During the interview, we posed open-ended questions to the participants such as “How did you find the experience of working with NEEK?” and “What aspect of NEEK did you find positive or negative?”. Participants' responses were noted down in writing and were then used for qualitative analysis.

Participants, Data collection, Handling and analysis

Data was collected from 24 people (male and female in equal number and in three age groups of 30 to 39, 40 to 49 years, and 50 years and above) living in Tehran. All categorised themselves as beginner level in English. All participants signed a written consent form before taking part in the study which was approved by the BehPAS board of ethics.

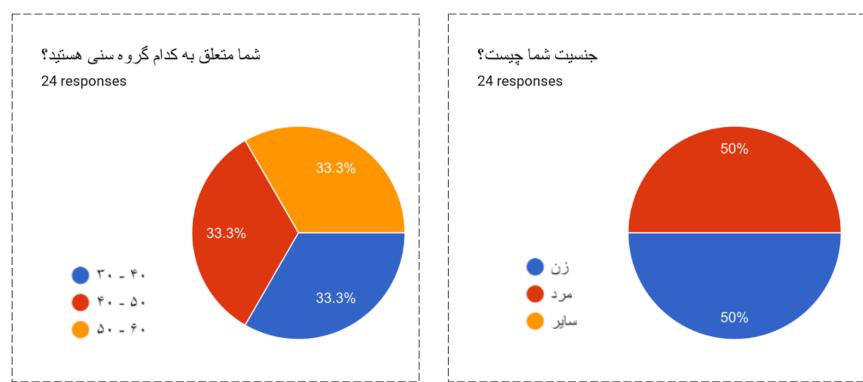


Figure 4. Demographics of recruited participants.

Results

Eye Tracking

The participants spent an average of 33 minutes (Figure 5) working with the Nik application and on average went to lesson 7 (Figure 6).

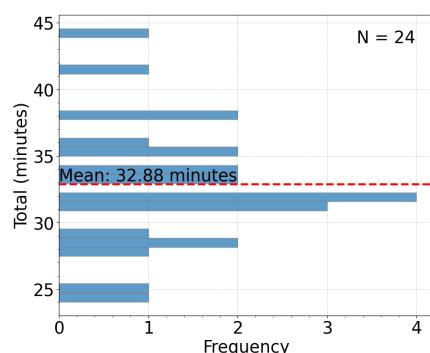


Figure 5. Frequency histogram showing the distribution of duration of engagement with NEEK.

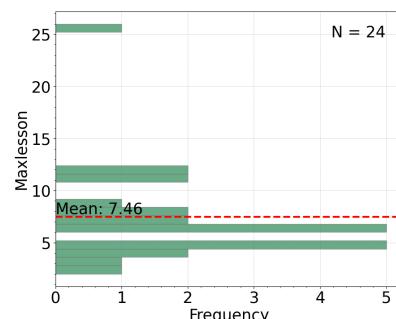


Figure 6. Frequency histogram showing the distribution of number of NEEK lessons covered in the session.

On average, participants were distracted from NEEK about 3 times (Figure 7) during a half hour session. More than half of people received at least one notification, and on average, each of them received up to two notifications while working with the app (Figure 8).

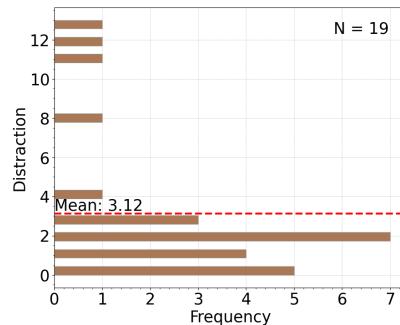


Figure 7. Frequency histogram showing the distribution of number of distractions during engagement with NEEK.

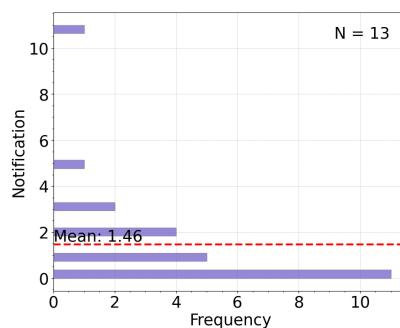


Figure 8. Frequency histogram showing the distribution of number of notifications received during engagement with NEEK.

To answer the question whether it was easy for people to learn how to work with NEEK from zero previous experience, we recorded how many times people had to resort to asking questions from the experimenter while using the app in the session. The results showed that about half of the people asked at least one question and on average each of them asked two questions about working with the program (Figure 9).

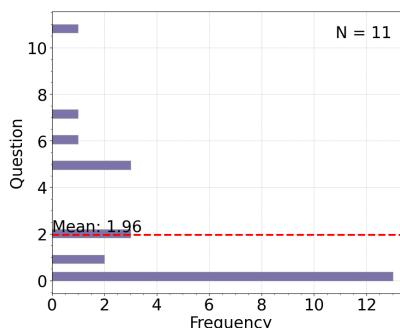


Figure 9. Frequency histogram showing the distribution of number of questions asked during engagement with NEEK.

To compare engagement with different elements of NEEK, Figure 10, we see the relative share of time that participants spent working with each component. The graph shows us clearly that there was a lot of diversity and different participants chose to dedicate variable times to different parts. But some helpful inferences can be made. First, about 20% of the user's time is spent doing things other than using the content. This is shown in the rightmost column under "wasted time". Second, English Grammar was the least popular form of content for users. People also visited the Podcast and Video sections much less than other parts. Investigation of the video recordings, however, showed that users often had problems downloading the content in these two sections. Interviews (see below) showed that users complained that when waiting for videos and podcasts, the lack of a downloadbar to indicate how much more they have to wait was an important demotivator.

The most popular content, we can see in Figure 10, were the Tests, Vocabulary, Listening, and reading. This result was very surprising for us since, at the beginning of the study, we had anticipated that people would be unwilling to take tests. The interactive element of tests and their gamified design structure turned out to be very popular and successful in attracting the users. Many users later expressed their positive experience of working through the tests to collect COINs and use them in other parts of the NEEK.

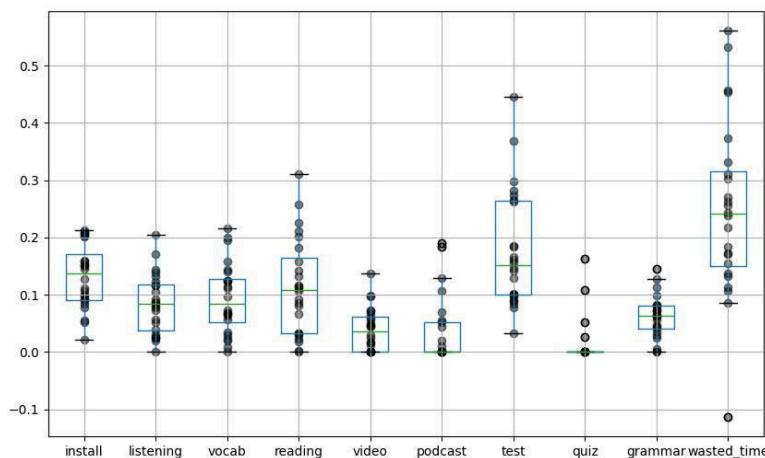


Figure 10. Relative share of time that participants spent working with each component of NEEK.

Questionnaires

The questionnaire's detailed results are available in the powerpoint report attached to this paper. For the sake of brevity, here we focus on a number of key findings.

When asked about the amount of time people were prepared to spend learning English with NEEK, users expressed a wide variety of opinions (see the pie chart in Figure 11). By weighted averaging of these data we can conclude that the optimal time for the duration of a daily lesson for NEEK users would be about 22-23 minutes.

اگر این برنامه در دسترس شما باشد، روزانه حدوداً چند دقیقه را به آن اختصاص می‌دهید؟

24 responses

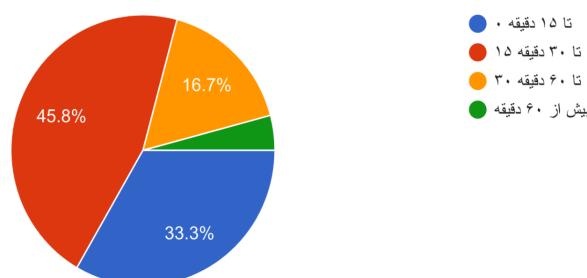


Figure 11. Pie chart indicating the distribution of times users would dedicate to using NEEK for learning if it were available to them.

To assess how much money the users would be willing to pay for the unlimited subscription of NEEK, we employed a well-known method in behavioural economics. Rather than asking people to estimate a number for the price of NEEK, we offered them a range of economic goods (see Figure 12) to the user to compare to NEEK subscription and tell us which product they would consider has the same value. This method has a number of advantages. Most importantly, it is independent of economic concerns such as inflation and value of national currency and therefore could be used in future versions of NEEK to make accurate comparisons. Second, by refraining from asking directly about money, we avoid any biases that different users may have about spending money such as loss aversion.

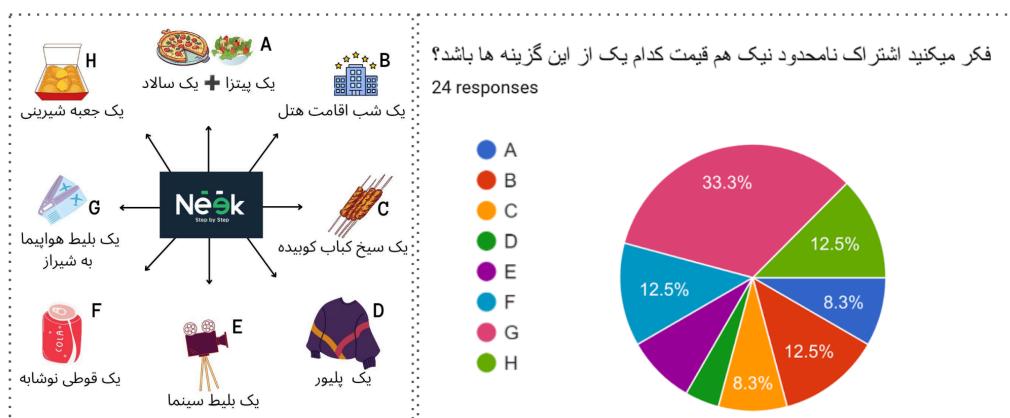


Figure 12.Pie chart indicating the distribution economic value that users would give to unlimited subscription of NEEK.

The largest group of the participants considered the unlimited subscription of NEEK to be equivalent to a plane ticket to Shiraz (33%) which at the time of testing was 1.5 Million Tomans. The weighted averaging of the data showed us that the user's valuation of NEEK subscription was 1.0 Million Tomans.

This measurement also allowed us to see if different age and gender groups of users would value NEEK similarly or differently. We found that among the three age groups of users, the oldest group (i.e., 50 and more) gave the highest price for NEEK. The results of the

regression model confirmed that people in the older age group estimate a higher price for unlimited subscription.

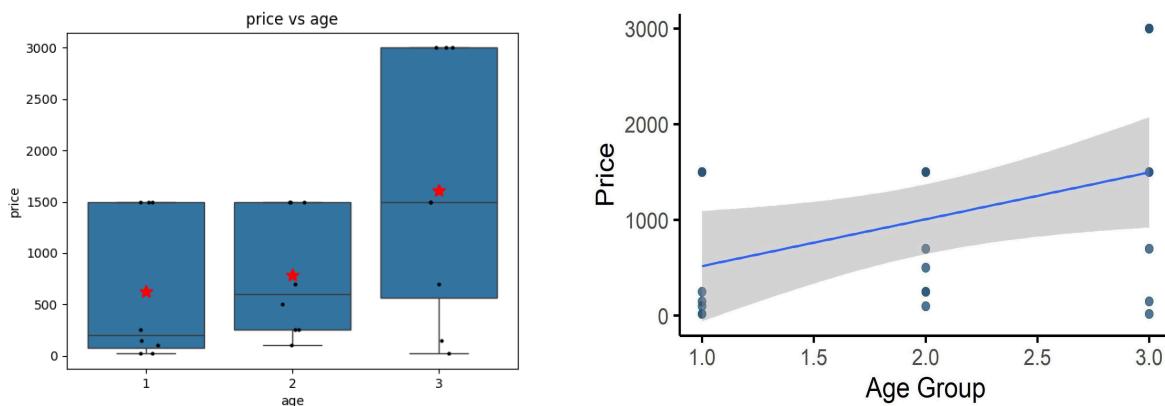


Figure 13. Oldest age group values NEEK subscription highest.

We also discovered that women value NEEK higher and are more eager to recommend it to their friends and family.

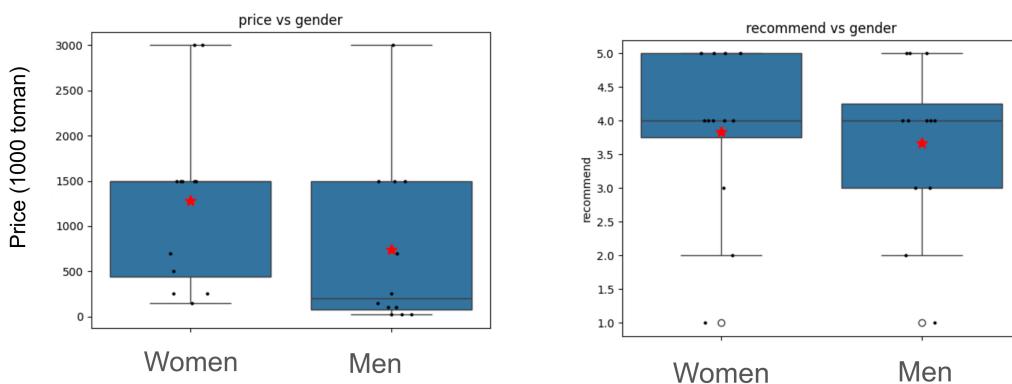


Figure 14. Women value the subscription of NEEK higher than men (left panel) and are more eager to recommend NEEK to their friends and family.

When asked “Which part of the NEEK would you change if you could?” change it, a considerable number of participants mentioned the colour-scheme and graphical design. A similar but less consistent opinion was observed across all age and gender groups.

Face-to-Face Interviews

When we put open ended questions to the users to tell us what they liked or disliked about NEEK, many users mentioned that the app’s gamification with coins gave them a positive experience that they would like to see more of in NEEK. Many suggested that having a more gamified environment would encourage them to use NEEK more frequently.

Interview results showed that people had the most negative opinion towards the grammar section which was also consistent with the results from eye tracking (see Figure 10) and with questionnaires (see Figure 15).

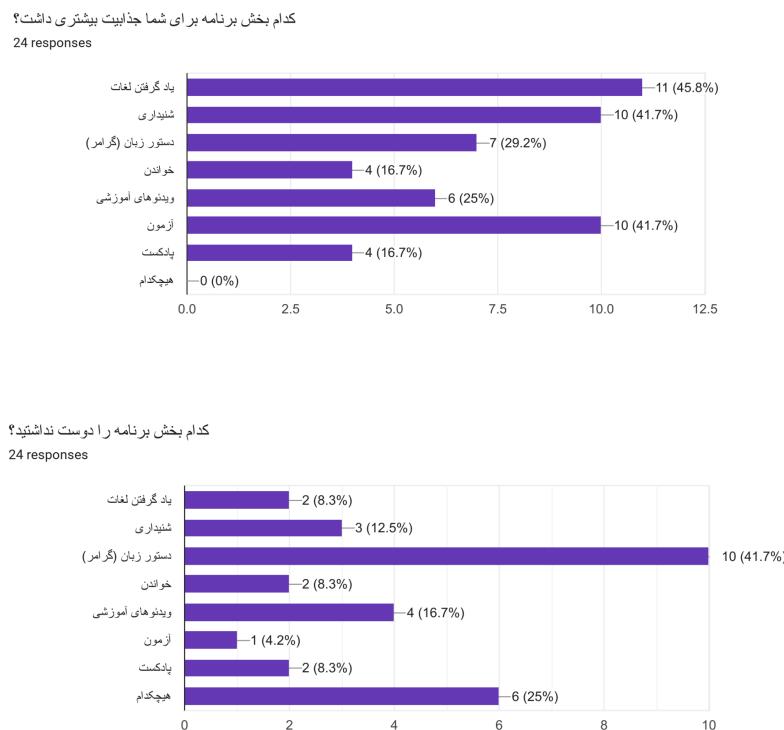


Figure 15. Favoured (top) and disfavoured components of NEEK according to Questionnaires. These results were broadly consistent with the interviews and eye tracking (figure 10).

Users were directly prompted to raise any negative points they had experienced during the session and the following complains were the most commonly observed items:

- A few users complained that the interface of NEEK was a bit confusing and perhaps a tutorial video guide or an onboarding screen (both for the registration process and the main interface) would be a good starter for them. (N=9)
- Texts were too long (N=8).
- Fonts were too small (N=5). This particular complaint was very relevant to participants who used glasses and large fonts in their default settings. Given the age group that NEEK targets as their key customer, this is a very important point to consider.
- A number of people complained about colour schemes and red on grey was not pleasant for them. Although frequent, there was no systematic structure to this complaint other than interindividual differences in taste.

The complete text of the interview points is also attached to this report in Appendix 1.



Observations made by the BehPAS team

- The appearance of the Download bar is inconsistent, appearing for some users and not for others, and this causes frustration and disappointment.
- Many users did not realise that they could turn the pages in the Vocab session.
- Different sections of lessons interfere with one another. A number of users who started an audio lesson and then opened the vocab tab and audio lesson continued to flow. Similar interference was observed in using videos.
- When working with Podcasts and Videos, which require downloading big chunks of data before starting, many users lost hope and abandoned the lesson because the app did not give them any information about the download progress.
- Many users did not notice the function of the icons at the top of the tabs and their toggle function was not self-evident to people.
- A repeatedly observed bug was that the app crashed (white screen) after a wrong response.
- Many users asked for a BACK button to revise their answers and return to what had just happened in the last screen.



Key Strategic Insights

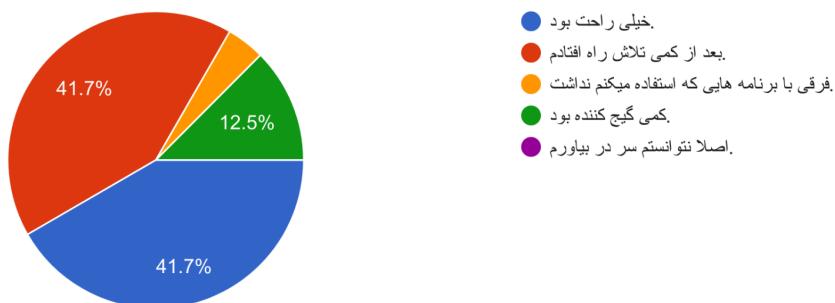
In this document we described the details of the usability test performed with 24 users (12 women, 12 men; three age groups of 8 people in each). This test was performed to assess the NEEK interface and content in order to converge to a number of guidelines and insights that could shape the next stage of development of NEEK.

Here we list these guidelines and insights that reflect the combined synthesis and aggregation of the data and statistical and qualitative analysis that has been detailed above.

1. **The target customers of NEEK are middle-aged and senior female users.** Women of 50 yrs of age and older were the ones who expressed the highest enthusiasm and financial value for NEEK. They also indicated that very few digital products target them as primary users, especially in domains that have self-development and education in their content. Important note: this group also stated that they are not technically skilled with their phone and they often have the younger members of their family install programs and handle their phone's technical issues for them (see insight 2).
2. **NEEK users need more and clearer instructions for how to use it.** Video guides, onboarding screens, and a simpler interface would be crucial for NEEK's target group. About 20% of study time per session was wasted mostly in figuring out how to navigate the app. This time will never be reduced to zero but simplifying the design and providing more sign-posting (eg BACK button, larger fonts, etc) can do a lot.
3. **More gamification and interactive play is highly useful.** Users of all ages and gender spent a lot of time in the TEST section and rated this section as most pleasant to work with.
4. **Waiting should be facilitated to avoid abandonment of the lesson.** Given the limits in download speed in most internet services in Iran, helping the users endure the waiting time for video and podcast files by a download bar/clock could prevent frustration and distraction.
5. **Ideal time for a daily lesson is around 20 minutes.** If a daily lesson takes about 20 minutes to complete, that gives the users a strong sense of accomplishment.
6. **Unlimited subscription to NEEK is valued around 1m Toman (roughly, 12-15\$) by the users.** This number could be usefully employed in setting NEEK's economic plans and targets.

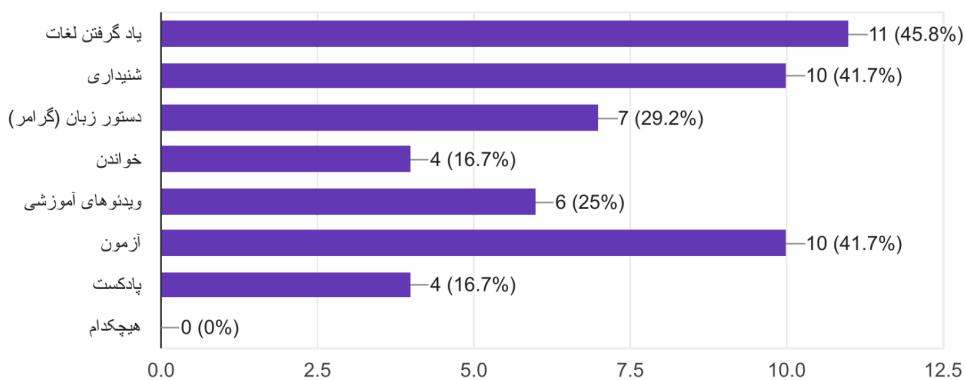
استفاده از این برنامه چگونه بود؟

24 responses



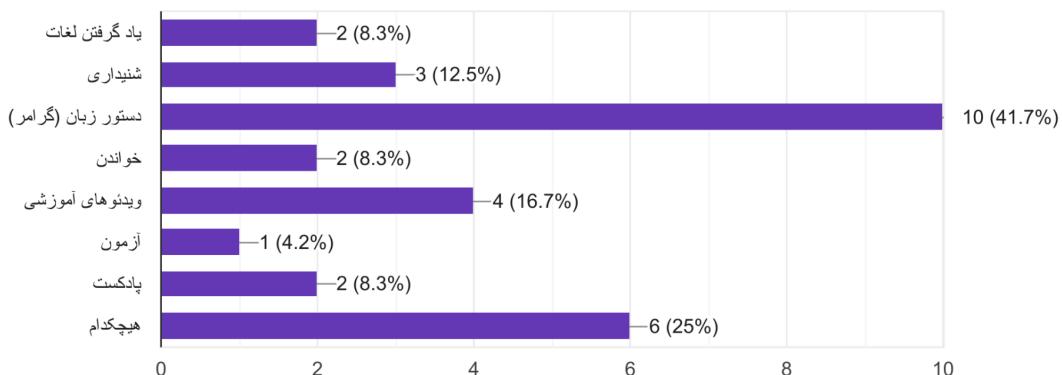
کدام بخش برنامه برای شما جذبیت بیشتری داشت؟

24 responses



کدام بخش برنامه را دوست نداشتید؟

24 responses



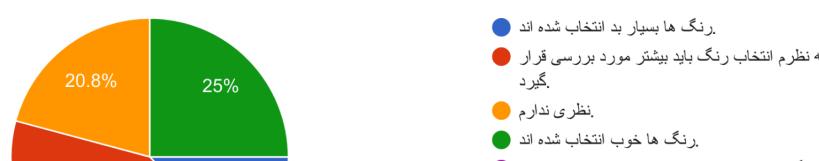
آزمون انتهایی هر بخش را چگونه ارزیابی می کنید؟

24 responses

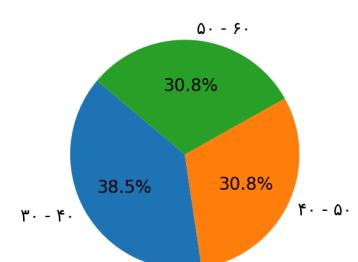


رنگ های به کار رفته در برنامه را چطور ارزیابی می کنید؟

24 responses



رنگ های انتخاب شده بسیار ترغیب کننده و جذاب بودند.



اگر این برنامه در دسترس شما باشد، روزانه حدوداً چند دقیقه را به آن اختصاص میدهید؟

24 responses

