Pimsleur Efficacy Study

FINAL REPORT

RESEARCH TEAM

ROUMEN VESSELINOV[1,2], PhD

Queens College, City University of New York

JOHN GREGO[2], PhD

University of South Carolina

April 2019

[1] Corresponding author: roumen.vesselinov@qc.cuny.edu .

[2] This report represents the individual opinion of the authors and not necessarily of their institutions.

THIS PAGE LEFT INTENTIONALLY BLANK

EXECUTIVE SUMMARY

The Research Team conducted an independent study of Pimsleur efficacy from February to April 2019. It was based on a random representative sample of 82 novice users of Spanish. The participants took one Spanish oral proficiency test in the beginning, then studied Spanish with Pimsleur for two months and took the same test again. We measured the improvement in language abilities as the difference between the final and the initial oral proficiency test results.

PIMSLEUR EFFICACY

- The efficacy of Pimsleur program and method is 83%.
  Eighty three percent of users who completed Pimsleur Level 1, i.e. completed 30 lessons, increased their oral proficiency level at least by one level and up to three levels.
  The 95% confidence interval for the efficacy is between 63% and 94%.
- The efficacy of Pimsleur for users who did not follow the program requirements exactly but have at least eight hours of study, is 73%.
  Seventy three percent of users who had at least eight hours of study increased their oral proficiency level at least by one level and up to three levels.
  The 95% confidence interval for this efficacy is between 61% and 82%.

USER SATISFACTION

- The majority of users thought that Pimsleur was easy to use (96%), helpful (97%), enjoyable (97%), satisfying (93%).
- Pimsleur received a positive Net Promoter Score of +62.8 from the users.
- Pimsleur efficacy was not affected by gender, age, education, native language, etc.
- Participants' motivation was very high and remained high after two-month study with average level of 72% of the maximum motivation score.

THIS PAGE LEFT INTENTIONALLY BLANK

CONTENTS

THIS PAGE LEFT INTENTIONALLY BLANK

INTRODUCTION

This is the 11th study of the Research Team testing the efficacy and motivation of language learning apps (Vesselinov, Grego, et al., 2009-2019). Our previous studies evaluated Rosetta Stone, Duolingo, busuu, Babbel, Hello English, italki and a new language app. The statistical design and methodology are practically the same for all 11 studies. The only deviation is the measurement of study time in the first Rosetta Stone study in 2009 when the time was self-reported versus recorded by server because participants were given CDs with the software and no objective measure for study time was technologically available.

This is a study designed to evaluate the efficacy of The Pimsleur® Language Program[3]. The company describes the Pimsleur program and method as follows:

"The Pimsleur Method™ was established nearly 50 years ago by Dr. Paul Pimsleur. "There exists an urgent national need for self-instructional materials in many of the world's languages," wrote Dr. Pimsleur in 1963.
With that goal in mind he developed and recorded Modern Greek, the first of the Pimsleur Language Courses. It was a test of his new theory of self-instruction which resulted from his years of teaching and his study in linguistics and the psychology of language, specifically of how memory could be triggered to best implant a new language.

Dr. Pimsleur's memory-training technique improves the language-learning centers in your brain. The Pimsleur programs are unique because of something they don't have -- and because of something they do have -- exclusively! No Mindless Repetition.

Many language courses try to teach a language by "drumming it in" with endless repetition of words and phrases, boring conjugations, irrelevant vocabulary, and mindless drills. Instead, Dr. Pimsleur focused on delivering the information at the right time to train your brain to respond easily. He called it: Graduated Interval Recall. Dr. Pimsleur's research into the psychology of linguistics revealed that unstructured mindless repetition (rote-memorization) has a dulling effect on the brain. It doesn't work. Through his studies at UCLA and Ohio State University, Dr. Pimsleur discovered that if words and phrases are repeated at specific time intervals, they are easily stored in long-term memory. This core principle is what makes the Pimsleur courses unique and highly effective.

The Pimsleur Language Program is one of the leading audio-based language learning programs, and is available in over 50 languages, including English as a second language."

This study was funded by Pimsleur, but the data collection and the analysis were carried out independently by the Research Team. The language test used in the study was designed and developed by an external independent testing company.

[3] www.Pimsleur.com

RESEARCH DESIGN

The random sample for this study was drawn from existing or new Pimsleur users residing in the U.S. There were some additional requirements for the potential participants, who had to:

- Be willing to study Spanish using only Pimsleur for two months;
- Take two sets of oral proficiency language tests;
- Be at least 18 years of age;
- Be novice or beginner learners of Spanish.

The test used in the study was the Oral Proficiency Interview by Computer® (OPIc)[4] created by Language Testing International (LTI). LTI is the exclusive licensee of the American Council on the Teaching of Foreign languages (ACTFL). The online test is recorded and evaluated by independent certified raters.

Table 1. OPIc Ratings

UR	Un-Ratable	**AL**	Advanced Low
NL	Novice Low	**AM**	Advanced Mid
NM	Novice Mid	**AH**	Advanced High
NH	Novice High	**S**	Superior
IL	Intermediate Low		
IM	Intermediate Mid		
IH	Intermediate High		

The specific definitions of the levels are presented on the company's webpage[5].

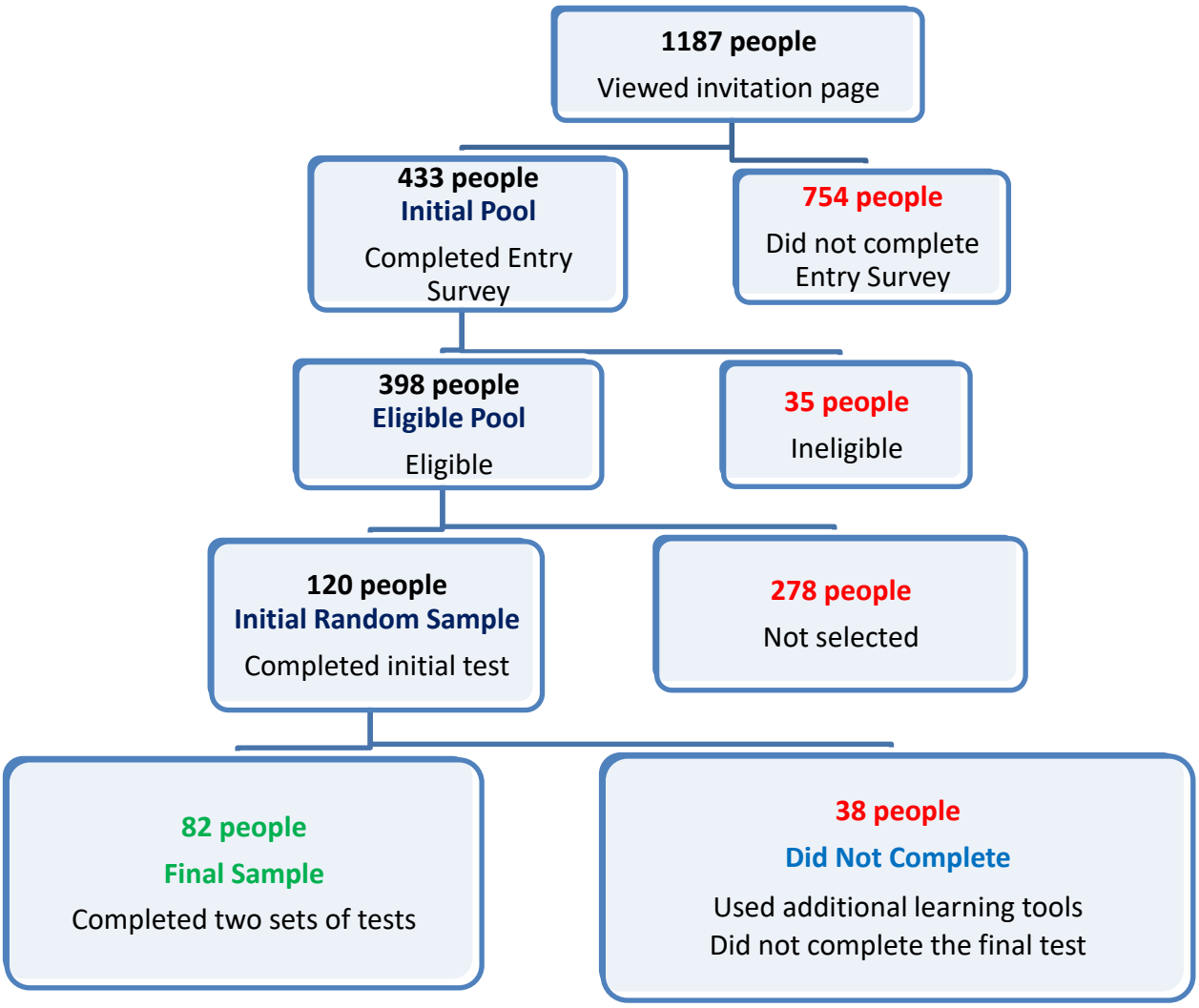
Sample Size and Power Analysis

We based our power and sample calculations on the typical results from our previous studies using OPIc. We designed the study to test the hypothesis for statistical significance of proportions of 0.6, 0.7 or 0.8, corresponding to the expected proportion of participants that improve their oral proficiency. We considered 0.05 statistical significance level (Alpha) and at least 80% statistical power. Under these assumptions we would need a sample of 60 people or more. Our initial sample size was set to 120 in order to reflect possible drop-outs from the study.

[4] http://www.languagetesting.com/oral-proficiency-interview-by-computer-opic
[5] https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012

Spanish language was selected as one of the most popular languages and because of the existence of previous research on Spanish for other language learning apps (Vesselinov & Grego, 2009-2019). The study lasted approximately 8 weeks and it was conducted between February and April 2019. Participants who successfully completed the study were given one additional level of Pimsleur Premium Spanish or one level of any language supported by Pimsleur. No monetary or other incentives were offered to the participants.

Figure 1. Sample Selection Tree



STUDY SAMPLE

In February 2019 emails were send to current and new Pimsleur potential users with an invitation to participate in Spanish language study for two months. They were directed to an online survey designed by the Research Team. This survey collected demographic information, and self-evaluation of their language proficiency level. We received complete responses from 433 people, and this constituted the initial pool for the study. From this pool we excluded 35 people who lived outside US, were younger than 18 years of age, or had intermediate or high level of knowledge of Spanish. This was the eligible pool (N=398) for this study.

We randomly selected 120 people of the eligible pool of participants and they completed the initial language test. They constituted our initial random sample (N=120).

Pimsleur study continued for approximately two months (8 weeks), starting in February 2019 and ending in April 2019. During the study the Research Team sent weekly e-mail reminders to the participants with information detailing the amount of time they had used Pimsleur each week. At the end of the study from the initial sample the following people were excluded:

- People who did not take the final test.
- People who regularly used additional learning tools during the study.

All participants were instructed at the beginning of the study that they could use only Pimsleur to study Spanish for the duration of the study. In the exit survey four people stated that they had regularly used other language apps and they were excluded from the study. Other people had occasionally used internet dictionaries, YouTube and translation websites and they were allowed to stay in the study.

The final study sample consisted of 82 people who had used only Pimsleur during the study and had valid initial and final OPIc tests. After the completion of the tests it was determined that two people had technical difficulties (cannot hear well the test questions) and their test results were not included in the analysis.

Final Study Sample versus Not Completed

From the initial random sample (N=120) 38 people (31.7%) did not complete the study for two different reasons: participants who did not take the final test and participants who used additional learning tools during the study. The 31.7% dropout rate is about average in this line of research.

We compared the two groups, the final sample of 82 people and the 38 people who did not complete the study by gender, age, education, employment status, initial knowledge of Spanish (initial OPIc) and reason for studying Spanish. There were no statistically significant differences (at 5% error level), which means that participants who did not complete the study were not very different from the ones that did.

Sample Description

In the final study sample 59.5% were female. The age varied from 19 to 80 years of age, with a mean age of 45.3 years. The pool was very well educated with majority participants having some college, undergraduate or graduate degree.

Figure 2. Age Distribution

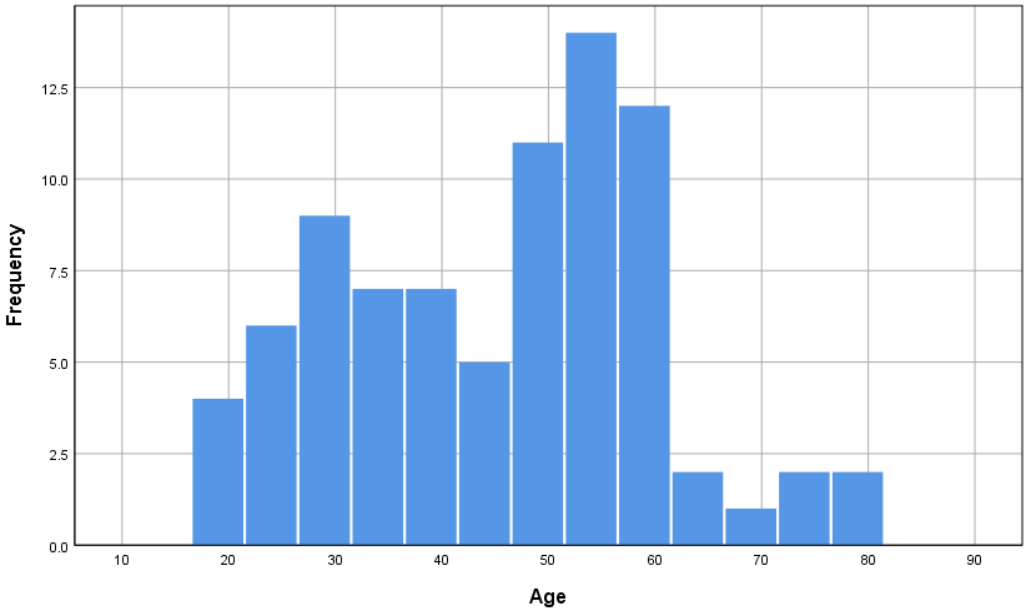


Table 2. Age and Gender Distribution

Age	Female (N)	Male (N)	Total (N)	Total (%)
18-20 years old	2	1	3	3.7
21-30 years old	6	9	15	19.5
31-40 years old	9	4	13	15.9
Over 40 years old	30	18	48	61.0
Total	**47**	**32**	**79***	**100**

* Some people declined to answer specific survey questions, so the number of answers can be less than 82 in some tables in this report. The number of answers is reported in the tables.

Table 3. Education

Category	N	Percent
2. High School diploma or equivalent	3	3.7
3. Some college but no degree	14	17.1
4. College graduate, BA or equivalent	23	28.0
5. Some graduate school but no degree	12	14.6
6. Master's degree (MA, MS)	22	26.8
7. PhD/MD/JD	8	9.8
Total	**82**	**100**

The majority of the people were employed full time (67.9%).

Table 4. Employment Status

Category	N	Percent
1. Employed full time	53	67.9
2. Employed part time	5	6.4
3. Homemaker	3	3.8
4. Student	4	5.1
5. Retired	7	9.0
6. Unemployed	1	1.3
7. Other employment	5	6.4
Total	**78**	**100**

Ninety-three percent of the participants were English native speakers and the rest included the native speakers of the following languages: Finnish, French, Japanese, Russian, Thai, and Turkish.

All participants in the final sample described themselves as beginner users or never studied Spanish. About 8% of the respondents' spouse, partner, or close friends spoke Spanish. About 4% of their parents, grandparents, or great-grandparents spoke Spanish.
About 82% of the final sample had formally studied a foreign language before (mostly at high school or college).

About 24% have lived outside U.S. in a non-English speaking country for more than 6 months. About 17% of the respondents were raised in a multilingual or non-English speaking household.

The primary reason for studying Spanish was personal interest (53.7%), followed by business or work (25.6%), travel (19.5%), and other reason (1.2%).

Table 5. Reason for Studying Spanish

Category	N	Percent
1. Personal Interest	44	53.7
2. Business/Work	21	25.6
3. Travel	16	19.5
4. Other	1	1.2
Total	**82**	**100**

Initial Language Test

All participants took an initial oral proficiency test and the results are presented below.

Table 6. Initial Oral Proficiency Results (OPIc)

Level	N	Percent
1. Novice Low	78	95.1
2. Novice Mid	2	2.4
3. Novice High	2	2.4
Total	**82**	**100**

As expected, almost all participants (95.1%) were placed at the lowest Novice Low oral proficiency level.

Initial Motivation

All participants completed a motivation scale in the beginning of the study to evaluate the role of motivation on efficacy.

As in the previous Rosetta Stone study (Vesselinov & Grego, 2019) we adopted a motivation scale approach largely based on the second language (L2) motivational self-system (Dörnyei, 2005, 2009) which stems largely from the concepts of possible selves and self-discrepancy theory. The model proposes that language learners are guided by visions of 'second language selves', one which attracts them toward becoming an idealized L2 user (ideal L2 self) and one which pushes them from societal obligation or a fear of failure (ought-to L2 self).

We adopted a specific 33 questions, 6 factors version of L2 Motivational Self System (see Appendix, Table A2) created by Kong et al. (2018).

Kong et al. (2018) offer the following descriptions of the motivation scale elements:

1. Ideal L2 self: "The ideal L2 self refers to a positive future image of the L2 self. For example, learners who have developed a vivid ideal L2 self are likely to endeavor to learn an L2 by imagining themselves communicating fluently using the L2 in the future."

2. Ought-to L2 self: **"(***This element*) pushes people from societal obligation or a fear of failure."

3. International posture: "It captures a tendency to relate oneself to the international community rather than any specific L2 group. The key characteristics of international posture are described as an interest in global issues or international affairs, a willingness to travel, stay, or work abroad, and a readiness to interact with foreigners or foreign cultures."

4. Competitiveness: "Competitiveness can be described as the desire to excel in comparison to others and contends that a learner constantly compares oneself with one's idealized self-image or with other learners, feels pressured to out-do other students."

5. L2 learning Experience or Attitudes: **"**L2 learning experience is related to the learners' environment including teachers, peer groups, curriculum, and their attitudes toward L2 learning."

6. Learners' Intended Effort or Motivated Behavior in L2 Learning: This motivation element evaluates how much effort are users determined to make and how hard they are ready to study.

As we can see from the table below and Figures 3 to 9 below, the participants vary a lot on initial motivation. The scale dimensions were re-coded, so the maximum motivation is equal to 100%.

Table 7. Initial Motivation Levels (%) Max=100

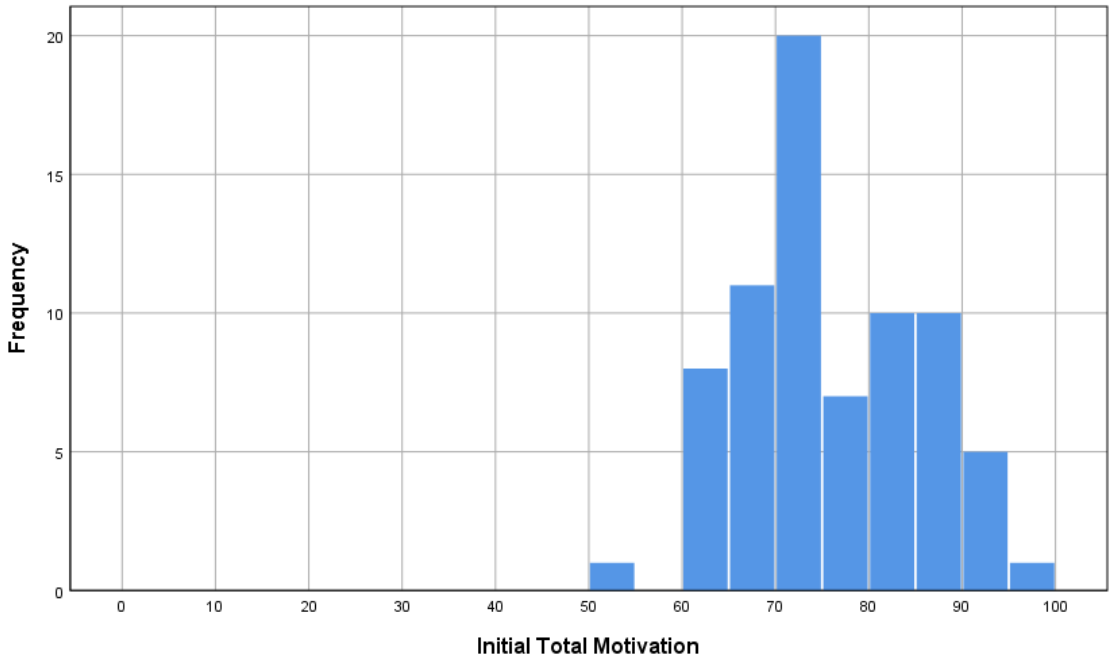
Motivation Dimensions	1st Quartile[6]	Median[7]	3rd Quartile[8]
1. Ideal Self	70.0	80.0	90.0
2. Ought-to-Self	42.9	55.7	65.7
3. International Posture	75.8	80.0	90.0
4. Competitiveness	72.5	80.0	85.0
5. Learning Attitude	80.0	85.0	95.0
6. Intended Effort	70.0	80.0	86.7
Total Motivation*	**69.6**	**74.4**	**84.5**

N=74

[6] First 25% of the sample.
[7] 50% middle point.
[8] First 75% of the sample.

The initial average level of total motivation was very high (Me=74.4%). From the motivation elements the highest level (85%) belongs to “Learning Attitude” which indicates that the participants were extremely eager to learn a new language. The element “Ought-to-Self” has the lowest level of all (55.7%) which suggests that the participants were not very afraid of failure or they were not that susceptible to pressure from societal obligation.

Figure 3. Initial Total Motivation Level (%) Max=100



As noted above the average initial level of motivation was very high and most people were overall highly motivated. Only a handful of people had motivation level less than 60%.

INITIAL MOTIVATION LEVELS:

Figure 4. L2 Motivation (Max=100)
M1 "Ideal Self"

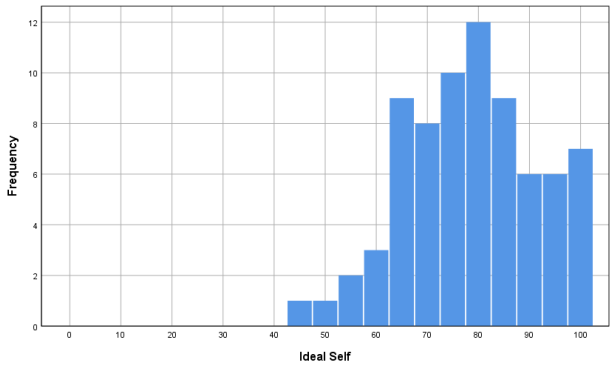


Figure 5. L2 Motivation (Max=100)
M2. "Ought-to Self"

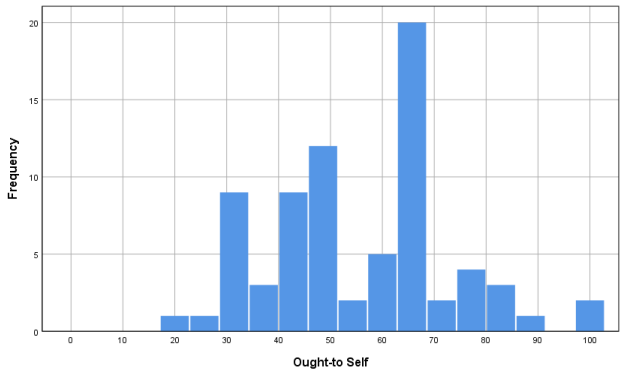


Figure 6. L2 Motivation (Max=100)
M3. "International Posture"

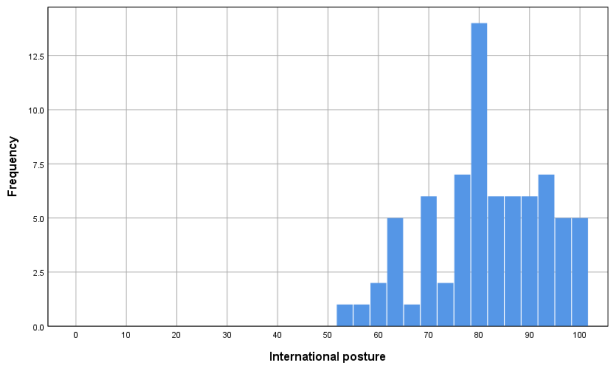


Figure 7. L2 Motivation (Max=100)
M4. "Competitiveness"

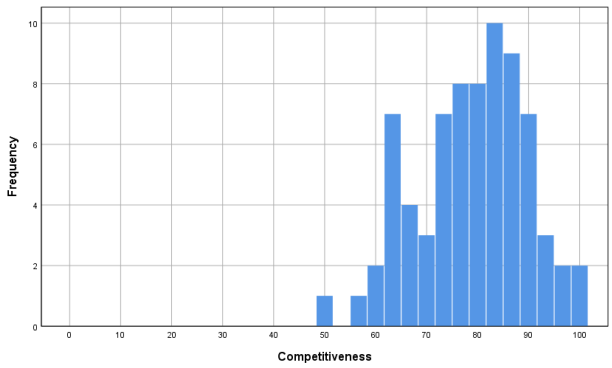


Figure 8. L2 Motivation (Max=100)
M5. "Learning Attitude"

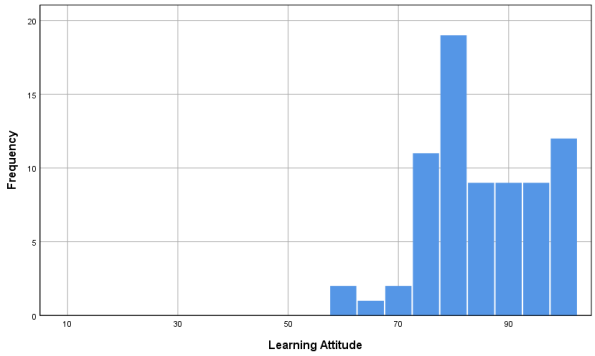
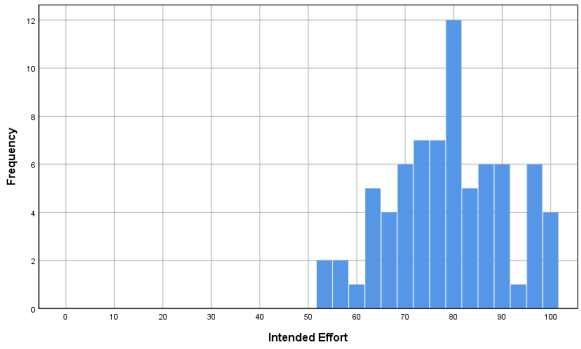


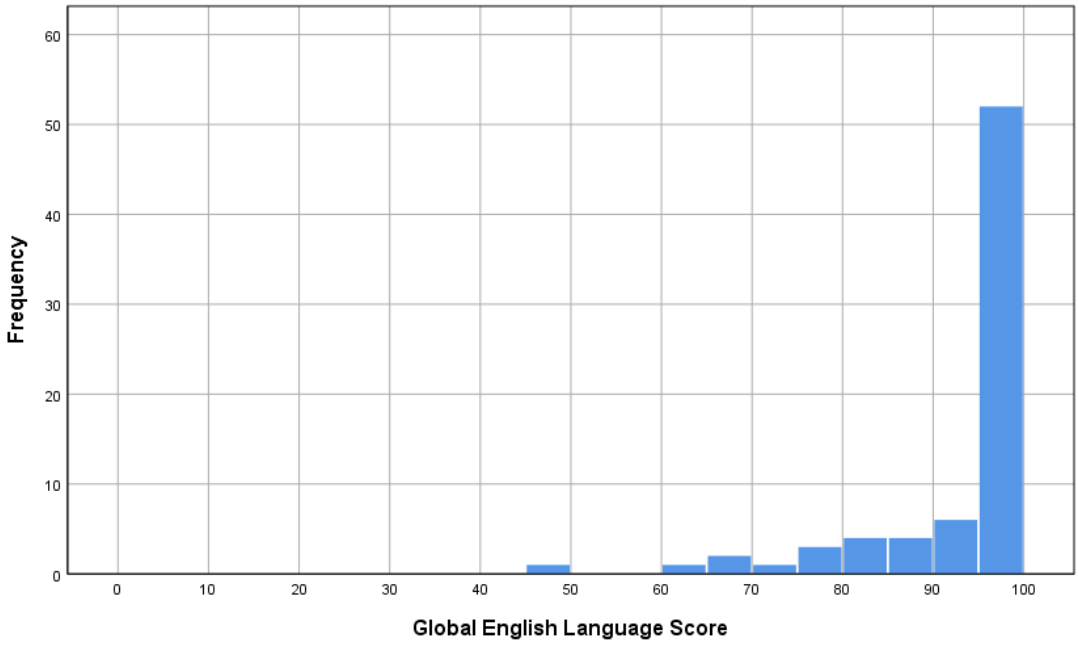
Figure 9. L2 Motivation (Max=100)
M6. "Intended Effort"



Language Profile

We asked participants to complete an adapted version of Bilingual Language Profile (Birdsong et al., 2012). We used the Global Language Score (GLS) for English and GLS for a second language for participants with a second language in use. GLS is based on separate modules for evaluating language history, language use, language proficiency and language attitudes. GLS can vary from 0 to 218, and we re-coded it, so the maximum is equal to 100.

Figure 10. GLS score for English Language (N=74) Max=100



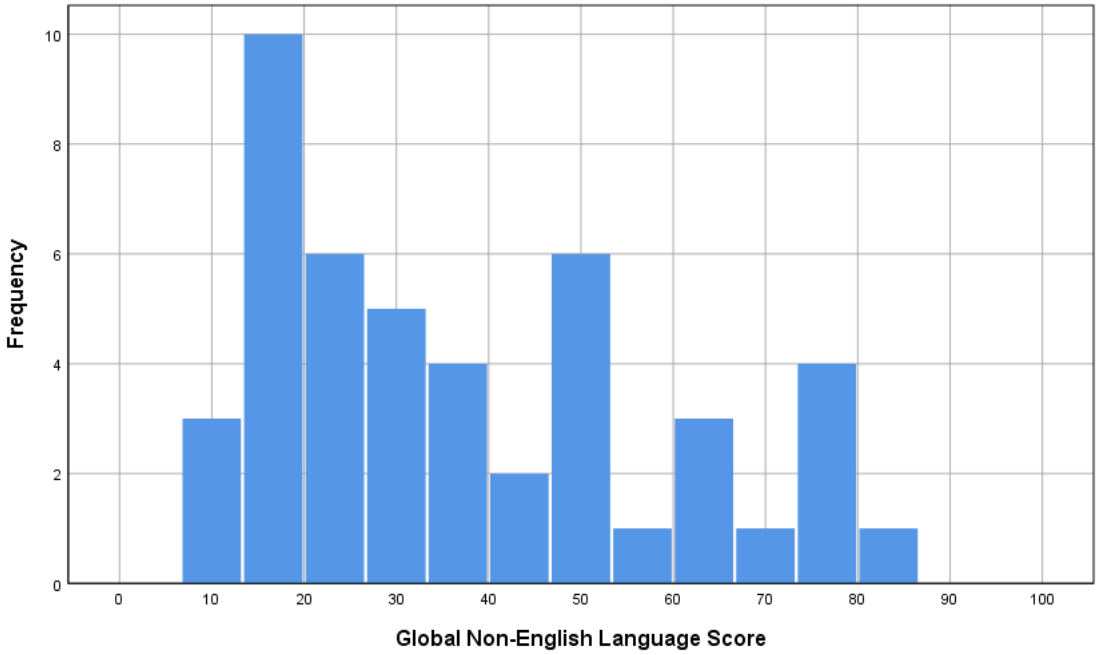
The median GLS percent was 97.7 (IQR[9]=5.7) which corresponds to initial sample of very strong English native speakers. Forty-six participants or 56% of the sample, felt comfortable enough to complete GLS for their second non-English language.

For example, a GLS score of 218 (or rescaled as 100) would be appropriate for participants born in English speaking family, in an English-speaking country, who started studying English immediately, for whom all classes at school were in English, who speak only English all the time with family, friends, and at work. Their language history and language use are entirely English-based. They feel totally proficient in English, and they identify themselves with an English-speaking culture.

[9] Interquartile Range = 3rd Quartile – 1st Quartile

GLS score will lose some points if the participants started learning English at older age; or some of the school teaching is in another language; or they use another language to speak with some of their friends and family, etc.

Figure 11. GLS score for Language other than English (N=46) Max=100



The median GLS percent for language other than English was 33.4 (IQR=32.2).

Final Language Test

The participants in the study completed the same language test at the end of the study.

Table 8. Final Oral Proficiency Results (OPIc)

Level	N	Percent
1. Novice Low	33	41.3
2. Novice Mid	38	47.5
3. Novice High	6	7.5
4. Intermediate Low	2	2.5
5. Intermediate Mid	1	1.3
Total	**80***	**100**

* After the study completion it was determined that two people had technical problems during the final test and their results were excluded from the analysis.

Final Motivation

Participants completed the same motivation scale (Kong et al., 2018) at the end of the study.

Table 9. Final Motivation Levels (Percent) Max=100

Motivation Dimensions	1st Quartile	Median	3rd Quartile
1. Ideal Self	65.0	77.5	85.0
2. Ought-to-Self	40.0	48.6	62.9
3. International Posture	70.0	76.7	86.7
4. Competitiveness	70.0	76.7	83.3
5. Learning Attitude	78.8	80.0	90.0
6. Intended Effort	66.7	73.3	80.8
Total Motivation*	**68.2**	**72.4**	**78.0**

*N=78

The total motivation level after two-months of studying remains remarkably high at 72.4%, although it is slightly below the initial level of motivation of 74.4%. The champion is "Learning Attitude" (80%) with "Ideal Self" as close second with 77.5% motivation level. Still, at the end of the study the participants were not very afraid of failure or societal obligations with "Ought-to-Self" 48.6% level of motivation.

Figure 12. Final Total Motivation Level (%), N=78 Max=100

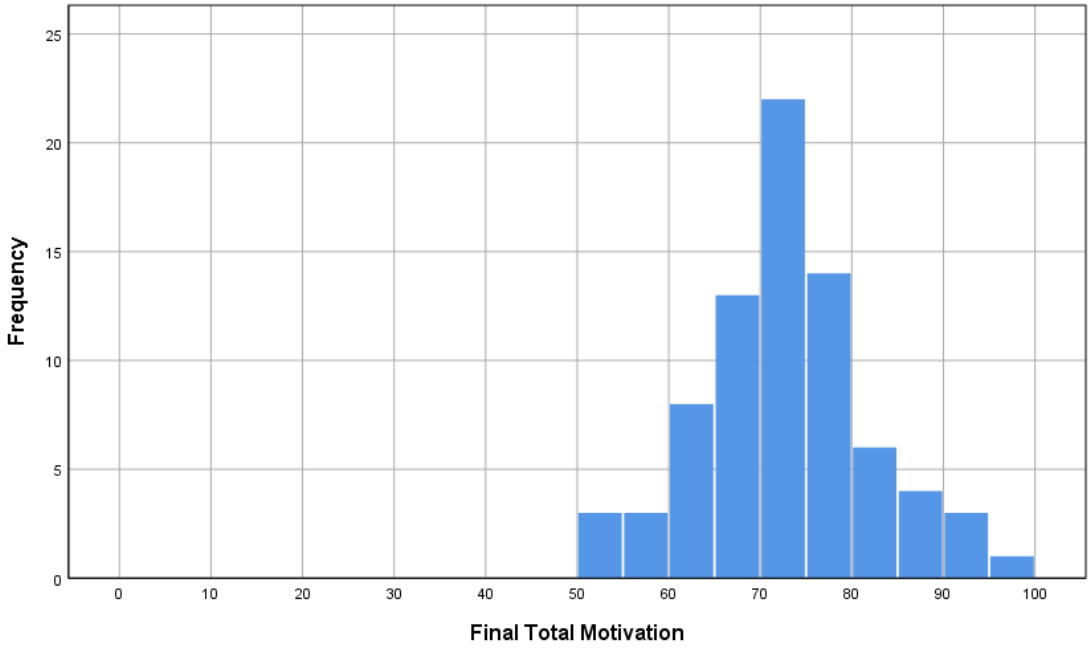
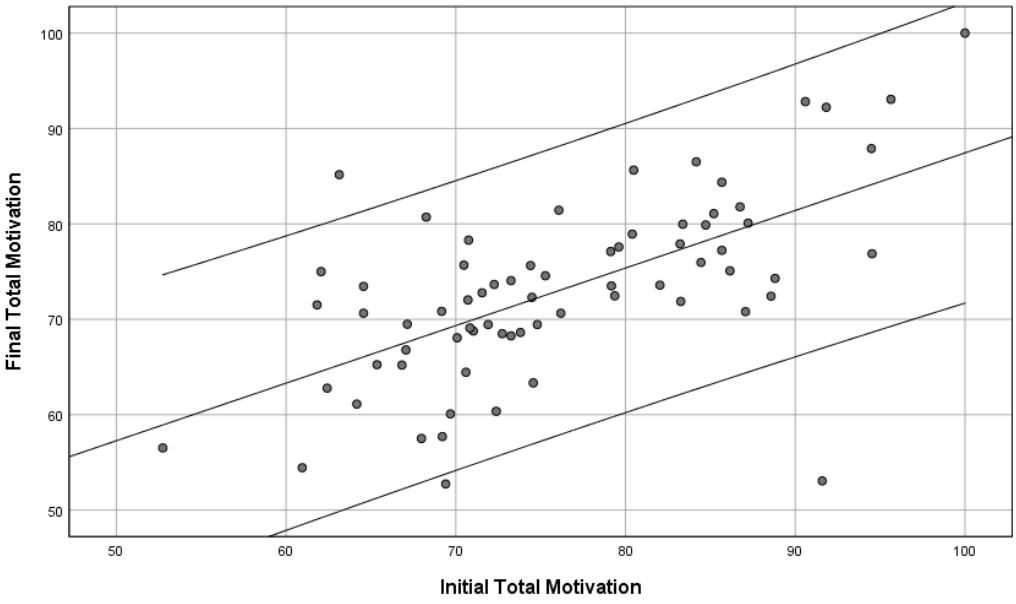


Figure 13. Initial and Final Motivation



The initial and final motivations are strongly correlated as the graph above shows. It includes the linear regression line and 95% confidence interval.

Study Time

We measured the study time objectively by the actual server time on a weekly basis and we reported the time to the participants regularly via e-mail in order to encourage them to keep studying. Several of the participants studied mostly offline by downloading the lessons and their time was estimated based on the number of lessons covered. The median study time for the final study sample (N=82) was about 16 hours, or about 2 hours of study a week. The total study time varied from about two hours to 140 hours during the two-month study.

We established a cut-off point of 8 hours to be consistent with the requirements for the OPIc test-retest (2-3 months of study) and comply with the established threshold from our previous studies (italki 2018 and busuu 2016).

In addition, the Pimsleur method requires half an hour a day study and Level 1 consists of 30 lessons. Overall 63 people (76.8%) covered the requirement of 8 hours of study and 24 people covered the Pimsleur requirement for 30 lessons or completed Level 1 of Premium Spanish.

Figure 14. Study Time Distribution (N=82) Hours

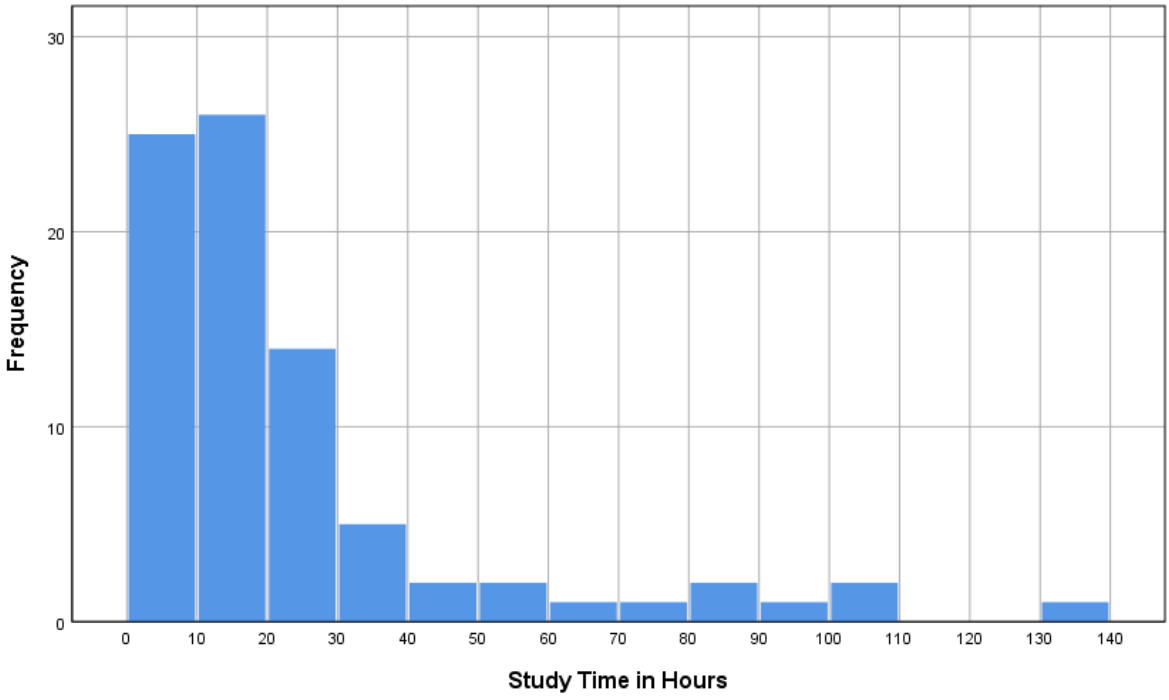
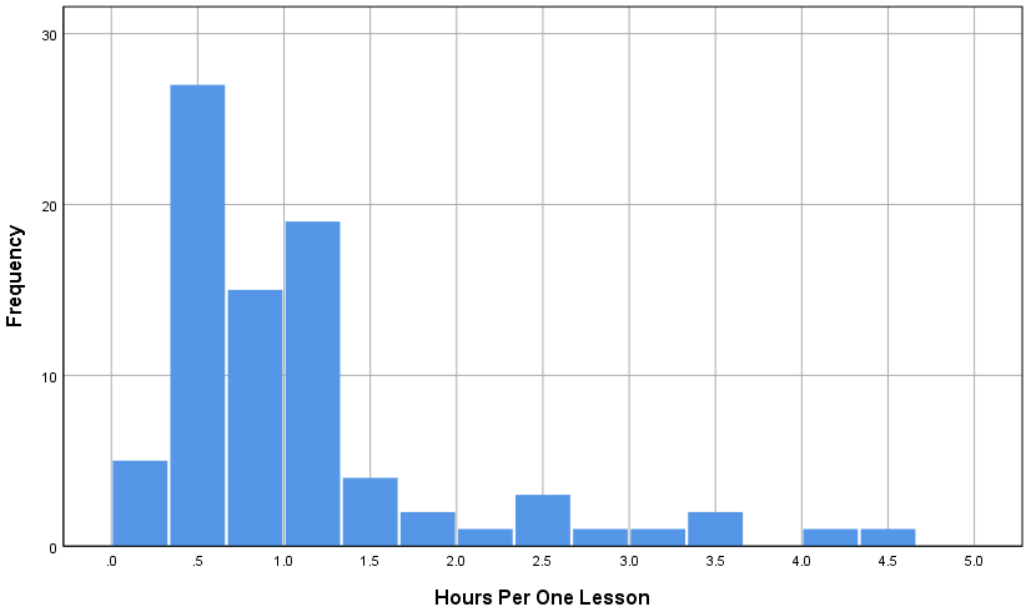


Figure 15. Study Time per One Lesson (N=82) Hours



The participants spent about 45 min (Me=45 min) per one lesson. The values varied from about 10 min to 4.5 hours per lesson, which included going several times over one lesson.

EFFICACY

Oral Proficiency Test Results

All participants took initial OPIc test before the start of the study and then again at the end of the study. We measured the progress or improvement as the difference between the final test level and the initial level.

Table 10. OPIc Oral Proficiency Level Placement

Level	Initial Test		Final Test	
	People (N)	%	People (N)	%
Novice Low	78	95.1	33	41.3
Novice Mid	2	2.4	38	47.5
Novice High	2	2.4	6	7.5
Intermediate Low			2	2.5
Intermediate Mid			1	1.3
Total	82	100	80	100

Participants at the lowest level decreased from 95.3% to 41.3%.

Table 11. Oral Proficiency Improvement with Study Time of 8 hours or more

#	Level Change	Improved		Study Time
		People (N)	%	Mean (Hours)
0	Same/No change	17	27.0	26.5
+1	One Level Up	40	63.5	29.3
+2	Two levels Up	4	6.3	36.6
+3	Three Levels Up	2	3.2	29.2
	Total	**63**	**100**	**29.0**

Overall, 73% of the users with 8 hours of study or more improved their oral proficiency level at least by one level and up to 3 levels. The 95% confidence interval is between 60.8% and 82.4%[10]. People who did not improve their level studied on average less than people who improved (26.5 vs 30 hours) but this difference was not statistically significant.

[10] 95% CI with Agresti-Coull correction (Agresti & Coull, 1998).

Table 12. Oral Proficiency Improvement with 30 Pimsleur Lessons or more

#	Level Change	Improved		Study Time
		People (N)	%	Mean (Hours)
0	Same/No change	4	16.7	37.3
+1	One Level Up	15	62.5	39.2
+2	Two levels Up	3	12.5	45.4
+3	Three Levels Up	2	8.3	29.2
	Total	**24**	**100**	**38.8**

Overall, 83.3% of the users with 30 Pimsleur lessons completed or more improved their oral proficiency level at least by one level and up to 3 levels. The 95% confidence interval is between 63.4% and 93.8%[11]. People who did not improve their level studied on average less than people who improved (37.3 vs 39 hours) but this difference was not statistically significant.

FACTORS FOR EFFICACY

Demographic Factors

We investigated the impact of main factors on language improvement (Yes/No), namely age, gender, education, employment, device used, native language, knowing another foreign language, presence of people around the participant who spoke Spanish (spouse, friend, parents, grandparents), and reason for studying Spanish.

None of these potential factors had a statistically significant effect on the efficacy (p=.05). In some instances, the number of cases by subgroups was too low to expect enough statistical power for the test of hypotheses.

This means that the Pimsleur app works similarly well for people with different gender, age, native language, education, employment status, etc.

[11] 95% CI with Agresti-Coull correction (Agresti & Coull, 1998).

Motivation Effect

Initial motivation level.

Table 13. Initial Motivation Levels (%) and Language Improvement Max=100

Motivation Dimensions	No Improvement	Improved	ANOVA
	Mean	Mean	p-value
1. Ideal Self	74.7	80.9	.12
2. Ought-to-Self	50.7	57.5	.19
3. International Posture	75.8	83.4	.023
4. Competitiveness	76.5	80.9	.15
5. Learning Attitude	85.9	85.4	.85
6. Intended Effort	77.1	80.7	.31
Total Motivation*	**73.4**	**78.3**	**.11**

N=57

Overall people who successfully increased their oral proficiency by one level or more were more motivated than those who did not improve. But these differences were not statistically significant except for International Posture motivation (p=.023).

Language Profile Effect

Table 14. Global Language Score (GLS): English and Non-English (%) And Language Improvement Max=100

GLS	No Improvement	Improved	ANOVA
	Mean	Mean	p-value
English (n=57)	93.8	92.7	.75
Non-English (n=38)	44.1	33.9	.20

The language profile (English and Non-English) did not have statistically significant effect on language improvement.

USER SATISFACTION

After the study the participants were asked for their opinion about Pimsleur, specifically how easy it was to use, how helpful, enjoyable, and satisfactory. The 5-point Likert scale was recoded into two categories: Strongly Agree/Agree vs Strongly Disagree/ Disagree/Neutral.

Table 15. User Satisfaction (N=78)

	Percent	
Do you agree with the following statement?	**Strongly Disagree/ Disagree/Neutral**	**Agree/ Strongly Agree**
"Pimsleur was easy to use"	3.8	96.2
"Pimsleur was helpful in studying Spanish"	2.6	97.4
"I enjoyed learning Spanish with Pimsleur"	2.6	97.4
"I am satisfied with Pimsleur"	6.4	93.6

After two months of study, the overwhelming majority of users (93% and above) agreed with the positive statements that: Pimsleur was easy to use, helpful, they enjoyed learning with Pimsleur and were satisfied with it.

All respondents in the exit survey declared that they will continue to use Pimsleur after the study ends.

In the exit survey a special question was included: "How likely are you to recommend Pimsleur to a colleague or friend?" with 11 possible answers, from 0 "Very unlikely" to 10 "Very likely". The answers to this question were used to compute the so-called Net Promoter Score (NPS). This is "a management tool that can be used to gauge the loyalty of a firm's customer relationships" (Wikipedia). It was developed by Reichheld (2003) and it categorizes users in three categories: "Promoters" (answers 9, 10), "Passives" (answers 7, 8), and "Detractors" (answers 0-6). NPS is equal to the difference between "Promoters" and "Detractors" and in general it can vary from -100 (all detractors) to + 100 (all promoters). As a rule, positive NPS is good news for the company and the higher the score the better indicator for the company.

From our exit survey the "Promoters" were 67.9%, the "Detractors" were 5.1% and "Passives" were 26.9%. The Pimsleur NPS was +62.8.

LIMITATIONS OF THE STUDY

The population of adult people who are seeking to study foreign language with language app is highly educated with majority of them having college level education and above. This is true not only for the U.S.[12], but also Europe[13] and the rest of the world[14]. This was confirmed by all our previous studies[15]. This population has higher education level than the general population. Our current sample for the 2019 Pimsleur efficacy study is representative of this population and it should not be compared to the general population.

This study measures the progress of novice/beginner users of Spanish. The study results cannot be generalized for intermediate or advanced users of Spanish.

The OPIc test used in this study is not tailored to any specific learning tool, including Pimsleur. On the one hand, some participants in the study complained that the test contained words or expressions that were not part of their regular course with Pimsleur. On the other hand, people insisted that they had learned a lot more than the test asked for. The test is valuable as an independent tool for evaluation which allows us to compare efficacy across different apps, however it does not provide a complete measure of the full progress of users. So, their progress evaluation of language proficiency is generally conservative.

The Research Team sent e-mail messages every week with individualized information about the study time for the previous week. This seemed to stimulate the study process. In normal settings when people work individually on their studies, this stimulation is not available. The results of the study should be valid in a setting where users study regularly for two months.

The study results could be generalized for studying Spanish with Pimsleur. For other languages the results could be markedly different.

The results of this study or the ten previous studies, cannot be compared to a standard college semester of Spanish for at least two reasons. First, progress or success in college is determined usually with one midterm exam and one final exam, plus some form of testing for oral proficiency and homework. The progress is measured very differently in a college setting compared to our 11 studies so far. Second, the study time at college is difficult to measure scientifically and it is not the same for everybody.

---

[12] Rosetta Stone (2009, 2019), Duolingo (2012), italki (2018)
[13] Babbel (Germany & US), Busuu (UK and US).
[14] New Language App, 2015 report, (world sample).
[15] Except Hello English (2017) where the participants were of high school age.

If there is a need to compare the progress of language app users and college language learners, a new study can be designed to satisfy the efficacy definition. To the best of our knowledge such study has not been done yet and until then the two sets of results cannot be compared scientifically.

There are limited number of studies with measure of efficacy available to compare with the results of this study. More help is needed from users, investors, and analysts to require the creators of language learning apps to provide independent efficacy measures. Hopefully in near future all language apps will be required to present independent efficacy studies to their customers.

CONCLUSION

The Pimsleur efficacy study is based on a random sample of 82 people, 18 years of age or older, residing in the U.S. All participants were self-reported novice/beginner users of Spanish.

The main goal of measuring the efficacy of Pimsleur was achieved with this study. The results show that, on average, 83% of the users who follow the Pimsleur method of half an hour per lesson and complete 30 lessons (Level 1) improve their oral proficiency level at least by one and up to 3 levels. The 95% confidence interval for this efficacy measure is between 63% and 94%.

In more general sense, comparable to our previous studies, 73% of the Pimsleur users who did not follow the Pimsleur method exactly but studied for at least 8 hours for a two-month period, improved their oral proficiency level by at least one level and up to 3 levels.

These results are not generally valid in different context: shorter than two-month period or for languages other than Spanish, and for intermediate or advanced learners of Spanish.

CITED LITERATURE

Agresti, A., Coull, B., 1998, Approximation is better than "exact" for interval estimation of binomial proportions, *American Statistician*, 52, pp. 119–126.

Birdsong, D., Gertken, L., & Amengual, M. Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism. COERLL, University of Texas at Austin. Web. 20 Jan. 2012. https://sites.la.utexas.edu/bilingual/ .

Dörnyei, Z. 2005. The psychology of the language learners. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z., 2009. The L2 motivational self system. In Z. Dörnyei, & E. Ushioda (Eds.), Motivation, language identity and the L2 self (pp. 9e42). Bristol, UK: Multilingual Matters.

Kong, J., Han, J., Kim, S., Park, H., Kim, Y., Park, Hy. L2 Motivational Self System, international posture and competitiveness of Korean CTL and LCTL college learners: A structural equation modeling approach, *System*, Volume 72, February 2018, Pages 178-189

Reichheld, F., 2003, "One Number You Need to Grow", *Harvard Business Review*, 2003 December.

Vesselinov, R., Grego, J., Sacco, S., Tasseva-Kurktchieva, M., 2019, The 2019 Rosetta Stone Efficacy Study . http://comparelanguageapps.com/documentation/The2019_RS_FinalReport.pdf

Vesselinov, R. and Grego, J., 2018, italki Efficacy Study. http://blog.italki.com/wp-content/uploads/2017/12/italki2018FinalReport.pdf or http://comparelanguageapps.com/documentation/italki2018FinalReport.pdf

Vesselinov, R. and Grego, J., 2017, Hello English Efficacy Study. http://centralsquarefoundation.org/grant/hello-english-efficacy-study/ , or http://comparelanguageapps.com/documentation/HelloEnglish_2017Study.pdf

Vesselinov, R. and Grego, J., 2016b, The Babbel Efficacy Study. http://comparelanguageapps.com/documentation/Babbel2016study.pdf , or http://press.babbel.com/en/releases/2016-09-29-Spanish-Study.html

Vesselinov, R. and Grego, J., 2016, The Busuu Efficacy Study. http://comparelanguageapps.com/documentation/The_busuu_Study2016.pdf , or https://blog.busuu.com/wp-content/uploads/2016/05/The_busuu_Study2016.pdf

Vesselinov, R. and Grego, J., 2015, Efficacy of New Language App, http://comparelanguageapps.com/documentation/LA_Final_Report.pdf .

Vesselinov, R. and Grego, J., 2012, Duolingo Effectiveness Study. http://comparelanguageapps.com/documentation/DuolingoReport_Final.pdf, or http://static.duolingo.com/s3/DuolingoReport_Final.pdf

Vesselinov, R., Grego, J., Habing, B., Lutz, A., 2009a, Measuring the Attitude and Motivation of Rosetta Stone Users. http://comparelanguageapps.com/documentation/MeasuringTheAttitudeandMotivationofRSUsers.pdf

Vesselinov, R., Grego, J., Habing, B., Lutz, A., 2009b, Comparative Analysis of Motivation of Different Language Learning Software. http://comparelanguageapps.com/documentation/ComparativeMotivationAnalysisofDifferentLanguageSoftware.pdf

Vesselinov, R., 2009, Measuring the Effectiveness of Rosetta Stone. http://comparelanguageapps.com/documentation/MeasuringTheAttitudeandMotivationofRSUsers.pdf, or http://resources.rosettastone.com/CDN/us/pdfs/Measuring_the_Effectiveness_RS-5.pdf.

APPENDIX

Table A1. Study Participants' Geographic Distribution: US States

Number of people

	State	ST	Initial Pool	Eligible Pool	Initial Sample	Final Sample
1	Alabama	AL	3	3	1	1
2	Alaska	AK	1	1		
3	Arizona	AZ	8	7	1	
4	Arkansas	AR	1	1		
5	California	CA	57	56	17	16
6	Colorado	CO	4	4		
7	Connecticut	CT	4	4	1	1
8	Delaware	DE	2	1		
9	Florida	FL	27	26	7	6
10	Georgia	GA	18	16	4	3
11	Hawaii	HI	6	6	3	3
12	Idaho	ID	1	1	1	
13	Illinois	IL	9	9	1	
14	Indiana	IN	6	6		
15	Iowa	IA				
16	Kansas	KS	2	2	1	1
17	Kentucky	KY	6	6	3	2
18	Louisiana	LA	2	2	1	
19	Maine	ME	1	1	1	1
20	Maryland	MD	17	16	5	3
21	Massachusetts	MA	15	15	6	2
22	Michigan	MI	16	16	3	3
23	Minnesota	MN	4	3	2	1
24	Mississippi	MS	2	2		
25	Missouri	MO	6	6	2	1
26	Montana	MT				
27	Nebraska	NE	1	1	1	
28	Nevada	NV	3	3		
29	New Hampshire	NH	1	1		
30	New Jersey	NJ	11	10	5	4
31	New Mexico	NM	1	1	1	1
32	New York	NY	26	26	11	9
33	North Carolina	NC	18	17	6	6
34	North Dakota	ND				
35	Ohio	OH	5	5	1	1

Table A1. Continued

	State	ST	Initial Pool	Eligible Pool	Initial Sample	Final Sample
36	Oklahoma	OK	3	3	1	
37	Oregon	OR	2	2	2	1
38	Pennsylvania	PA	3	3	1	1
39	Rhode Island	RI	2	2	1	
40	South Carolina	SC	3	1		
41	South Dakota	SD				
42	Tennessee	TN	10	10	4	3
43	Texas	TX	26	26	3	2
44	Utah	UT	4	4		
46	Virginia	VA	17	17	7	6
45	Vermont	VT	2	2	1	
47	Washington	WA	20	20	5	1
49	Wisconsin	WI	5	5	2	1
48	West Virginia	WV	2	2		
50	Wyoming	WY				
	District of Columbia	DC	2	2		
	Unknown state (but US)		25	25	8	2
	Outside US		23			
Total	**All**		**433**	**398**	**120**	**82**

Table A2. Motivation Scale

Developed by Kong et al., 2018.

A. Ideal L2 self (4 items)

1. I can imagine myself living abroad and having a discussion in Spanish.
2. I can imagine myself speaking Spanish with international friends or colleagues.
3. I can imagine myself speaking Spanish as if I were a native speaker of Spanish.
4. Whenever I think of my future career/life, I imagine myself using Spanish.

B. Ought-to L2 self (7 items)

1. I study Spanish because close friends of mine think it is important.
2. Learning Spanish is necessary because people surrounding me expect me to do so.
3. I consider learning Spanish important because the people I respect think that I should do it.
4. Studying Spanish is important to me in order to gain the approval of my peers/teachers/family/boss.
5. It will have a negative impact on my life if I don't' learn Spanish.
6. Studying Spanish is important to me because an educated person is supposed to be able to speak it.
7. Studying Spanish is important to me because other people will respect me more if I have knowledge of it.

C. International posture (6 items)

1. I want to make friends with foreigners visiting U.S.
2. I would feel somewhat uncomfortable if a foreigner moved in next door. (reverse-coded)
3. I want to participate in a volunteer activity to help foreigners living in the surrounding community.
4. I am interested in an international career/living abroad.
5. I often read and watch news about foreign countries.
6. I have thoughts that I want to share with people from other parts of the world.

D. Competitiveness (6 items)

1. I want to survive in the future.
2. I don't want to be an illiterate person.
3. I want to succeed in life.
4. Other people will consider me an elite if I have a good command of Spanish.
5. I don't want to place behind any of my friends.
6. I want to have a head start on other people.

E. L2 learning Experience or Attitudes (4 items)

1. I like the atmosphere of my Spanish classes with Pimsleur.
2. I find learning Spanish with Pimsleur really interesting.
3. I always look forward to Spanish classes with Pimsleur.
4. I really enjoy learning Spanish with Pimsleur.

F. Learners' Intended Effort or Motivated Behavior in L2 Learning (6 items)

1. If Spanish course were offered in the future, I would like to take it.
2. I expend a lot of efforts in learning Spanish.
3. I do my best to learn Spanish.
4. I spend lots of time studying Spanish.
5. I concentrate on studying Spanish more than any other topic.
6. Compared to other people I know, I think I study Spanish relatively hard.

Table A3. Language Profile

Developed by Birdsong et al., 2012.

I. Biographical Information

II. Language history

In this section, we would like you to answer some factual questions about your language history.

1. At what age did you **start learning** English?

   Since birth 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

2. At what age did you **start to feel comfortable** using English?

   As early as I can remember 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+ not yet

3. How many years of **classes (grammar, history, math, etc.)** have you had in English (primary school through university)?

   0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

4. How many years have you spent in a **country/region** where English is spoken?

   0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

5. How many years have you spent in a **family** where English is spoken?

   0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

6. How many years have you spent in a **work environment** where English is spoken?

   0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20+

III. Language use

In this section, we would like you to answer some questions about your language use.

7. In an average week, what percentage of the time do you use English **with friends**?

   0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

8. In an average week, what percentage of the time do you use English **with family**?

   0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

9. In an average week, what percentage of the time do you use English **at school/work**?

   0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

10. When you talk to yourself, how often do you **talk to yourself** in English?

    0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

11. When you count, how often do you **count** in English?

    0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%

IV. Language proficiency

In this section, we would like you to rate your language proficiency.

	0=not well at all						*6=very well*
12. How well do you speak **English**?	0	1	2	3	4	5	6
13. How well do you understand **English**?	0	1	2	3	4	5	6
14. How well do you read **English**	0	1	2	3	4	5	6
15. How well do you write **English**?	0	1	2	3	4	5	6

V. Language attitudes

In this section, we would like you to respond to statements about language attitudes.

	0=disagree						*6=agree*
16. I feel like myself when I speak **English**.	0	1	2	3	4	5	6
17. I identify with an **English-speaking** culture.	0	1	2	3	4	5	6
18. It is important to me to use (or eventually use) **English** like a native speaker.	0	1	2	3	4	5	6
19. I want others to think I am a native speaker of **English**.	0	1	2	3	4	5	6

THIS IS THE LAST PAGE

OF THE 2019 PIMSLEUR EFFICACY STUDY

********** **No text bellow this line** **********