

Hello English Efficacy Study

FINAL REPORT

RESEARCH TEAM

ROUMEN VESSELINOV^{1,2}, PhD

Economics Department
Queens College,
City University of New York
roumen.vesselinov@qc.cuny.edu

JOHN GREGO, PhD

Statistics Department
University of South Carolina
grego@stat.sc.edu

March 2017

¹ Corresponding author.

² This report represents the individual opinion of the authors and not necessarily of the two institutions.

EXECUTIVE SUMMARY

The Hello English (HE) efficacy study was independently conducted by the Research Team from September, 2016 to January, 2017. The study was based on a random representative sample of 97 students in India from grades 8 to 12 from three schools: one government and two private schools.

The research goal for this study was to test whether the introduction of HE as an additional tool for learning English as a foreign language would significantly improve students' language skills compared to their classmates who did not use the new tool.

The students were randomly assigned to two groups: the HE group, which was given access to the HE app in addition to the regular school instruction and the Control group, which continued with their regular school classes with no access to the app. The participants took one English oral proficiency test in the beginning of the study, and the same test at the end of the study. The improvement in language abilities was measured as the difference in levels between the final and the initial language tests.

MAIN RESULTS

HE English Language Oral Proficiency Gain:

- Overall, 73% of the HE students improved their language proficiency by at least one level compared to 42% of the Control group. This difference was statistically significant.
- The 95% confidence interval for the improvement in language proficiency for the HE students was between 60% and 83% compared to an interval of 27% to 58% for the Control group.
- Truly novice English language learners (that is, those who were initially at a Novice-Low level) from the HE group improved the most with 88% gaining one or more levels. Among the more advanced, 45% improved by one or more levels. This is expected and consistent with the results from other language studies.
- From the HE group, 56% of students improved by one level, 15% of students improved by two levels and one student improved by three levels. None of the HE students decreased their language level.
- The language improvement of the HE group compared to the Control group remains significant even after controlling for demographics, school type and results of school based assessments. The HE app worked for everybody in the study regardless of their age, gender etc.

Conclusion. Students who used HE to study English significantly improved their language abilities and performed significantly better than their classmates who did not use HE.

Table of Contents

Introduction.....	4
Research Design	6
Initial Sample Description	7
English Oral Proficiency Test	7
Environmental Information	11
Analysis	13
Final Sample Description.....	13
Final English Oral Proficiency Test.....	15
Main Results.....	17
Factors.....	19
Conclusion.....	21
Cited Literature	22

Introduction

The importance of learning a foreign language —particularly English—is undisputable. This is especially true in India, where employability opportunities in urban areas have strongly been associated with an ability to speak English. Of late, a number of language learning apps have been launched in India, offering youth a self-learning opportunity. Schools in particular could benefit from such applications, especially if apps could help decrease the financial burden and make a good language education affordable for every student.

With this study, we are trying to evaluate the effect of introducing the language-learning app Hello English³ (HE) for school children in India and compare their progress in learning English as a foreign language to students at the same schools who did not use HE or any other language-learning app.

HE is an interactive, personalized, and contextual English learning application designed specifically for English as a second language learners. Launched in October of 2014, HE is Asia's most downloaded, and world's third most rated Educational application on Google Play Store (as of January 2017). According to HE, in just 2 years, more than 22 million users have accessed the app across the world, to learn English as a second language from 21 different vernacular languages. The app has consistently retained the top position under the free apps in education category in India, Sri Lanka, Bangladesh, Nepal, Qatar, and many others; with an average rating of 4.5/5 with more than 446,000 user reviews on the play store. The App was also awarded 'Best of 2016 - Made in India' by Google Play and was featured at Google I/O conference in 2016.

HE has the following features:

- Covers all four aspects of language acquisition: Reading, Writing, Listening and Speaking, with advanced voice recognition technology that allows learners to speak into the app and hold real-life, useful conversations
- Pairs interactive lessons with fun games and speaking practice for a complete learning experience
- Offers unique and engaging contextual learning tools that leverage news, sports and entertainment to help learners build their English vocabulary
- Makes learning seamless & saves data expenses for users as a majority of the app's features work offline

³ <http://helloenglish.com/>

* * *

This study was funded by the Central Square Foundation⁴, a grant making organization and policy think tank focused on improving the quality of school education and learning outcomes of children from low-income communities in India.

The English oral proficiency language test used in the study was designed, developed and managed by Language Testing International (LTI), an independent US based language testing company⁵. The Research Team carried out the test data collection and statistical analysis independently.

* * *

⁴ <http://www.centalsquarefoundation.org/>

⁵ <http://www.languagetesting.com/>

Research Design

The sample for this study was drawn from three schools from India, one government school and two low-income private schools. Anecdotally, it is believed that in government schools, which are free of charge, the quality of education, including the teaching of English is worse than in private schools. The schools' characteristics are listed in Table 1.

Table 1. Schools Included in the Study

SCHOOLS IN THE STUDY*	
1. MGV Private School, Jaipur	
Location	Jaipur - capital city of Rajasthan, a large state in North India
Student Enrolment	250
Gender	Male and Female
Teachers	20
Grade	6 to 12
Income level of families	Average annual family income is less than INR 2 lakhs (USD 3,000)
Medium of Instruction	English
2. Government Senior Secondary Boys School, Old Faridabad	
Location	Faridabad - a city situated in the National Capital Region, bordering the Indian capital New Delhi
Student Enrolment	1350
Gender	Male only
Teachers	50
Grade	6 to 12
Income level of families	Average annual family income is less than INR 1.5 lakhs (USD 2,250)
Medium of instruction	Hindi and English
3. BMR Private School, Faridabad	
Location	Faridabad - a city situated in the National Capital Region, bordering the Indian capital New Delhi
Student Enrolment	800
Gender	Male and Female
Teachers	32
Grade	K to 12
Income level of families	Average annual family income is less than INR 2 lakhs (USD 3,000)
Medium of Instruction	Hindi and English

* The number of students and teachers are estimates.

Initial Sample Description

The students considered for this sample were in grades 8 to 12 and enrolled in one of the three schools. From each school, the selected students were randomly assigned to one of two groups: HE and Control. The HE group was offered free use of the Hello English (HE) app for smartphones. Students in the Control group had to continue regular learning of English at school, with no additional intervention. The incentive for the HE group was the free use of the HE app. Control group students were to each receive small packages of school supplies (pens, pencils, etc.) at the end of the study. Permission from the relevant authorities was obtained for conducting the study at each school. Parental permission was also obtained for the children to participate in the study.

97 students across grades 8 to 12 were selected to be part of the study, with 33 female and 64 male students. Hindi was the native language of a majority of students, with only 6 students speaking other languages in addition to Hindi, including Bengali, Bhojpuri, Maithili, Punjabi and Urdu.

All students had to take the online oral proficiency test in English, continue studying English at school for the duration of the study (3-4 months) and then take the same oral test again at the end of the study. Bi-weekly communication with the students was maintained — largely via visits to schools, phone calls and emails. The study took place between September 2016 and January 2017.

English Oral Proficiency Test

The test used in the study was the Oral Proficiency Interview by Computer® (OPIC) created by Language Testing International⁶ (LTI) based in the US. LTI is the exclusive licensee of the American Council on the Teaching of Foreign Languages (ACTFL). The online test is proctored and the recording of the test is reviewed and evaluated by independent LTI raters and the oral proficiency evaluation is provided by ACTFL. The specific version of the test for this study was OPIC English.

The OPIC tests have randomized selection of questions and situations. All questions are similar in nature but not identical.

⁶ <http://www.languagetesting.com/oral-proficiency-interview-by-computer-opic>

Table 2. OPIc Ratings

UR	Un-Ratable/Not Rated	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 definition of the levels is presented on the company's webpage⁸.

Table 3. Initial Sample of Students (N=97)

School	N (%)		
	Control	HE	Total
1. MGV Private School	11 (45.8)	13 (54.2)	24 (100)
2. Government School	14 (40.0)	21 (60.0)	35 (100)
3. BMR Private School	14 (36.8)	24 (63.2)	38 (100)
Total (by group)	39 (40.2)	58 (59.8)	97 (100)

For each school, the Control group made up about 40% of the sample, and HE English comprised the remaining 60%. In all schools, the students from both groups received the same regular teaching of English. In addition, the HE group had free access to HE for smartphones. Students from the Control group were instructed not to use HE or other external language apps that they do not usually use in their studies. At the end of the study we asked all students about use of HE or other language apps. This was done to find out if there were any control group students who regularly used apps so that we should exclude them from the study.

⁸ <http://d2k4mc04236t2s.cloudfront.net/wp-content/uploads/2013/07/ACTFL-Proficiency-Guidelines-2012.pdf>

Table 4. HE vs Control Comparison on Demographics (N=97)

Characteristics	Control	HE	Significance*
Female N (%)	13 (33.3)	20 (34.5)	n.s.
Grade N** (%)			
8	5 (13.2)	1 (1.7)	n.s.
9	10 (26.3)	20 (34.5)	
10	13 (34.2)	16 (27.6)	
11	6 (15.8)	8 (13.8)	
12	4 (10.5)	13 (22.4)	
Age*** Mean (std****)	15.2 (2.1)	15.7 (1.8)	n.s.

* n.s. – the difference between the groups is not statistically significant at $p < .05$

** N=96.

*** Age estimation was based on year of birth only.

**** Standard Deviation.

We compared the HE and Control groups on their demographic characteristics. No statistically significant differences were found between the two groups on gender, age and grade.

Table 5. Past School Performance Test Levels (0 – 100%)

Mean % (std)

Evaluation	Control	HE	Significance
Overall School Performance (n=72)	62.9 (14.2)	62.2 (11.1)	n.s.
English Language School Performance (n=65)	59.9 (15.0)	61.2 (15.0)	n.s.

The students provided information about their most recent school scores, overall and specifically for English as well. There were no significant differences on these two measures between the two study groups.

Table 6. Performance Levels (0 – 100%) by School (N=97)

School	Mean % (std)	
	Overall	English
1. MGV Private School	65.0 (12.2)	72.5 (15.4)
2. Government School	62.1 (14.1)	62.4 (14.8)
3. BMR Private School	61.5 (10.0)	55.7 (13.1)
Total	62.5 (12.4)	60.8 (14.9)

Since there are no standardized tests across these schools (each school creates its own tests), direct comparison of the schools' past performance is not possible. However, we decided to use past performance tests' results as a control variable in the statistical models of oral proficiency improvement. Common performance test for all schools was not available.

All students took one initial OPIC English oral proficiency test. The tests were conducted under external supervision.

Table 7. Initial Oral Test (OPIC) Results (N=96*)

Characteristics	N (%)		Significance
	Control	HE	
0. Not Rated** (NR)	2 (5.1)	3 (5.3)	n.s.
1. Novice Low (NL)	27 (69.2)	34 (59.6)	
2. Novice Mid (NM)	10 (25.6)	16 (28.1)	
3. Novice High (NH)	0 (0)	3 (5.3)	
4. Intermediate Low (IL)	0 (0)	1 (1.8)	

* One student's test was not valid.

** Valid tests but students either kept silent, or did not speak any English or spoke in their native language.

As per the study design, the two groups were very similar in terms of demographic structure and initial level of knowledge of English. As expected, a majority of the students were at Novice Low and Novice Mid level. The difference in English language levels between the HE group and the Control group was not statistically significant.

Environmental Information

Students were asked to provide information about their access to computers at school and home and their experience with the internet and social media.

Table 8. Language and Computer Experience

Ns vary (63 to 97)*

Characteristics**	Percent "Yes"
Do any of your friends speak English well?	22.9
Do any of your parents speak English well?	8.4
Do any of your siblings speak English well?	45.1
Do you have easy access to computer at school?	23.2
Do you have full and easy access to computer or smartphone at home?	14.9
Have you ever used a smartphone?	69.1
Did you see an English-speaking movie or video last week?	37.8
Did you read a book in English (other than a textbook) last week?	17.1
Have you ever used a social network media?	58.5
Have you ever used Facebook?	47.2
Have you ever used Twitter?	12.3
Have you ever used Skype?	7.9
Have you done texting?	43.8

* There are missing data because some students chose not to answer some of the questions.

** Due to the missing values problem, comparison between the HE and the Control groups could not be performed.

On average, about a quarter of students' reported that their friends could speak English well. While less than 10% of students reported that their parents could speak English well, almost half the respondents felt their siblings knew English.

About a quarter of the students reported having easy access to computers at school. The emphasis was on access rather than availability, as computers could be present but students may not be able to use them. At home very few students (15%) could use a computer or smartphone with full and easy access.

About two thirds (70%) of the students said they had used smartphones.

Students were asked whether they had seen an English-speaking movie or video the previous week before the first test and about 40% of them reported doing so. Very few (17%) of them said they had read a book in English (other than a textbook) the week before.

More than half of the students (59%) had previously used social media including Facebook (47%) and Twitter (12%). Nearly half of the students (44%) reported an experience with texting and very few (8%) with Skype.

Analysis

Final Sample Description

At the end of the study all students were invited to take the final oral proficiency test designed by LTI. Six of the students (6.2%) dropped out of the study for various reasons. The final sample consisted of 91 students, with 55 in the HE group and 36 in the Control group. One of the students in the HE group did not use HE at all but the student remained in the group in accordance with the “Intent to Treat” principle used in experimental studies.

The students who dropped out of the study were uniformly distributed between the HE and control groups, as well as by gender, grades and schools. Specifically, 3 of the drop-outs were from the HE and 3 were from the Control group; 2 were female and 4 were male. Students were from grades 8, 9, and 10 —from all three schools. Details of the final sample are given below:

Table 9. Final Sample of Students

School	N (%)		
	Control	HE	Total
1. MGV Private School	9 (40.9)	13 (59.1)	22 (100)
2. Government School	14 (41.2)	20 (58.8)	34 (100)
3. BMR Private School	13 (37.1)	22 (62.9)	35 (100)
Total (by group)	36 (39.6)	55 (60.4)	91 (100)

Table 10. Final Sample: HE vs Control Comparison on Demographics

Characteristics	Control	HE	Significance
Female N (%)	13 (36.1)	18 (32.7)	n.s.
Grade N (%)			
8	4 (11.1)	1 (1.8)	n.s.
9	9 (25.0)	20 (36.4)	
10	13 (36.1)	13 (23.6)	
11	6 (16.7)	8 (14.5)	
12	4 (11.1)	13 (23.6)	
Age Mean (std)	15.3 (2.1)	15.5 (1.8)	n.s.

We compared the final HE and Control groups on their demographic characteristics. No statistically significant differences were found between the two groups on gender, age and grade.

Table 11. Final Sample: School Performance Test Levels (0 – 100%)

Mean % (std)

Evaluation	Control	HE	Significance
Overall School Performance (n=69)	62.3 (14.0)	62.0 (11.3)	n.s.
English Language School Performance (n=62)	59.9 (15.0)	61.5 (15.0)	n.s.

There were no statistically significant differences between students past exam scores across the two groups. Do note that each school uses their own test and thus the results are not entirely comparable among schools.

Table 12. Final Sample: Initial Oral Test (OPic) Results

Characteristics	N (%)		Significance
	Control	HE	
0. Not Rated (NR)	2 (5.6)	3 (5.5)	n.s.
1. Novice Low (NL)	25 (69.4)	32 (58.2)	
2. Novice Mid (NM)	9 (25.0)	16 (29.1)	
3. Novice High (NH)	0 (0)	3 (5.5)	
4. Intermediate Low (IL)	0 (0)	1 (1.8)	

There was no statistically significant difference between the two groups on the initial level of English language knowledge.

Final English Oral Proficiency Test

Table 13. Final Oral Proficiency Test (OPic) Results

Characteristics	N (%)		Significance
	Control	HE	
0. Not Rated (NR)	0 (0)	0 (0)	.006
1. Novice Low (NL)	17 (47.2)	5 (9.1)	
2. Novice Mid (NM)	15 (41.7)	29 (52.7)	
3. Novice High (NH)	4 (11.1)	20 (36.4)	
4. Intermediate Low (IL)	0 (0)	1 (1.8)	

While the initial oral tests showed no difference between the HE and Control groups the final tests showed statistically significant difference between the groups ($p=.006$). At the end of the study there were no students with “Not Rated” level. But about 47% of the Control group was still at Novice Low level while only 9% from the HE group was at that level. More than a third of the HE group (36%) was at Novice High level compared to 11% of the Control group.

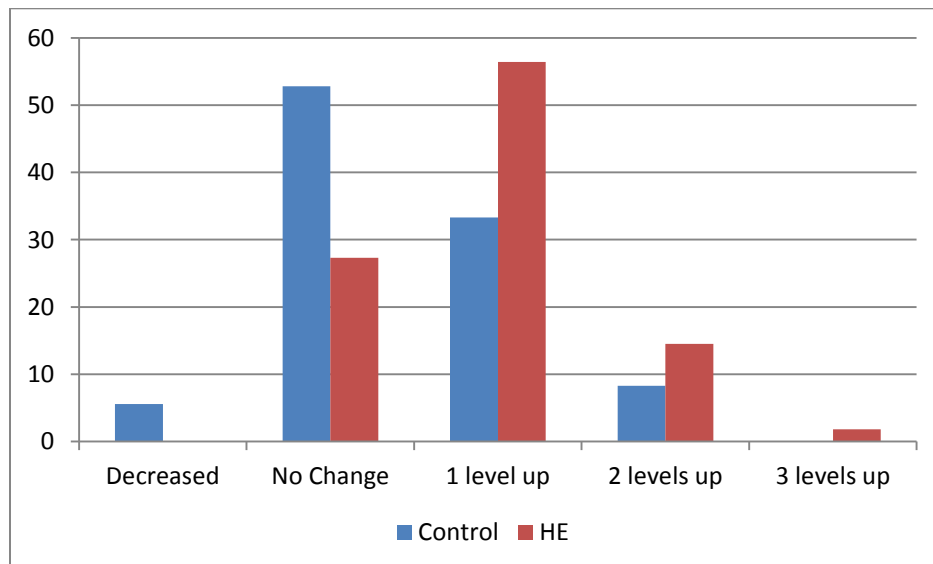
In previous studies measuring the efficacy of language learning apps (Vesselinov and Grego, 2012, 2016) time spent by participants on the app was part of the evaluation process. HE has the advantage that it can be used offline as well. While this is extremely beneficial for users, it makes measuring usage time challenging. In this context, usage time was not measured in the study.

Table 14. Final Oral Proficiency Test (OPIc) Change Results

N (%)

Change in Levels	Control	HE	Significance
-2. Decreased 2 levels	1 (2.8)	0 (0)	.05
-1. Decreased 1 level	1 (2.8)	0 (0)	
0. Same Level	19 (52.8)	15 (27.3)	
1. Increased 1 level	12 (33.3)	31 (56.4)	
2. Increased 2 levels	3 (8.3)	8 (14.5)	
3. Increased 3 levels	0 (0)	1 (1.8)	

Figure 1. Oral English Proficiency Improvement in Levels (%)



We recoded the cases in three major categories: decreased level, same level, or increased level.

Table 15. Final Oral Proficiency Test (OPIc) Results (3 categories)

N (%)

Change in Levels	Control	HE	Significance
-1 Decreased	2 (5.6)	0 (0)	.006
0. Same	19 (52.8)	15 (27.3)	
1. Increased	15 (41.7)	40 (72.7)	

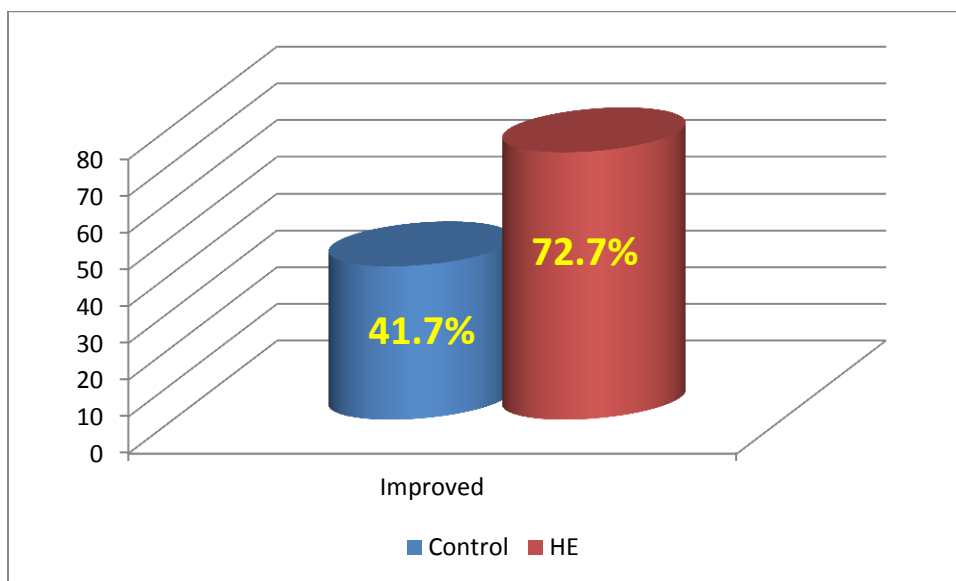
As seen in Tables 14 and 15, about 58% of the Control group did not increase their language level compared to only 27% of the HE group. From both groups there were some students (a total of 12) who increased their language level by two or three levels. Further, there were only two students who decreased their language level; both were from the Control group. Most importantly, a majority (73%) of the HE group showed an improvement in their language levels, compared to only 42% of the Control group.

Main Results

Table 16. Overall Improvement in Language Skills.

N (%)

Level Change	Control	HE	Significance
0. Same/Decreased	21 (58.3)	15 (27.3)	.003
1. Increased	15 (41.7)	40 (72.7)	

Figure 2. Oral English Proficiency Improvement (%)**Table 17. Main Finding: Overall Improvement in Language Skills.**

Group	Percent	
	Improved	95% Confidence Interval ⁹
Control	41.7	27.2 – 57.8
HE	72.7	59.6 – 82.7

The main finding of this study is that 73% of the students who used the HE app in addition to their regular school classes improved in their English proficiency by at least one level compared to only 42% improvement of their classmates who did not use HE. The advantage of the HE group was statistically significant ($p=.003$). The average improvement was about 73% with 95% confidence interval between 60% and 83%.

Overall the HE group students were 3.7 times more likely (Odds Ratio=3.7) than the Control group to improve their language level.

⁹ 95% CI with Agresti-Coull correction (Agresti & Coull, 1998)

Factors

The main study results indicate that using the HE app in addition to studying English at school has a statistically significant positive effect on the students' language education. We further investigated whether there were other factors that affected these results.

We built additional statistical models which included known possible factors like gender, grade, age, and school. The results remained unchanged and the HE group retained its advantage compared to the Control group.

Table 18. School Performance: Improvement in Language Level

School	N (%)	
	Control	HE
1. MGV Private School	6 (66.7)	10 (76.9)
2. Government School	3 (21.4)	12 (60.0)
3. BMR Private School	6 (46.2)	18 (81.8)
Total (by group)	15 (41.7)	40 (72.7)

As demonstrated in Table 18, the students in the two private schools did better than those in the government schools. Between 46% and 66% of the students in the control group in the two private schools improved their language level compared to only 21% in the Government school. Students using HE improved their language level in the three schools with 60% showing improvement at the Government school and between 76% and 80% showing improvement at the two private schools.

In a separate statistical model for language level improvement (Yes/No) even after introducing school as a factor the HE group still performed better than the Control group—this difference was still statistically significant.

Effect of School Performance Test Levels (0 – 100%)**Table 19. Students' School Performance and Language Improvement.**

Mean (std)

School Performance	Change in Language Level		Significance
	0. Same/Decreased	1. Increased	
Overall School Performance	63.3 (12.9)	61.3 (12.2)	n.s.
English Language School Performance	58.2 (13.3)	63.2 (15.9)	n.s.

Another interesting question was whether the students' success at school would affect their language improvement. Neither their overall exam scores nor their English language scores affected their language improvement. That is, any recorded differences were not statistically significant.

Initial Language Level as Factor**Table 20. Improvement in Language Level by Initial Level**

N (%) Improved

Initial Level	Control	HE	Total
0. Not Rated (NR)	2 (100)	3 (100)	5 (100)
1. Novice Low (NL)	14 (44.0)	28 (87.5)	39 (68.4)
2. Novice Mid (NM), Novice High (NH), Intermediate Low (IL)	2 (22.2)	9 (45.0)	11 (37.9)

The biggest improvement in language level occurred for students who had very little knowledge in the beginning of the study. This result confirms previous studies' findings of

similar nature (Vesselinov & Grego, 2012, 2016). HE seems to work well for both true beginners (Novice Low, 88% improvement) and those at a higher level (45% improvement), but the true beginners register the greatest progress.¹⁰ It is more difficult to increase your language level when you are a more advanced student.

Conclusion

The Hello English efficacy study addressed the question of whether the introduction of an additional learning tool to students from grades 8 to 12 would significantly improve their language skills. The study demonstrates that school-going students who use the HE language app in addition to their regular language classes show greater progress compared to their classmates who do not use the HE app. On average about 73% of the HE users can expect to increase their language knowledge by at least one level compared to only 42% of their classmates. The difference between the two groups is statistically significant. The 95% confidence interval for the progress of HE students is between 60% and 83%. The language improvement of the truly novice English language students (Novice Low level) from the HE group was the most impressive with about 88% of them improving by at least one level. Students with a higher initial level also improved, but at a lower rate (45%). This result is in line with previous language studies (Vesselinov & Grego, 2012, 2016). The progress of more advanced language learners is smaller compared to the progress of the true beginners.

The above findings are promising as they demonstrate that the Hello English app can be used by school students for learning the English language. With the availability of student usage data, we could garner more granular insights on patterns of use thus shedding light not only on whether children learn through the app, but also on how they do so.

¹⁰ The group of “Not Rated” students was too small to be analyzed

Cited Literature

- Agresti, A., B. Coull, 1998, Approximation is better than “exact” for interval estimation of binomial proportions, *American Statistician*, 52, pp. 119–126.
- Vesselinov, R. and J. Grego, 2016, 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 J. Grego, 2016, The Busuu Efficacy Study.
[http://comparelanguageapps.com/documentation/The busuu Study2016.pdf](http://comparelanguageapps.com/documentation/The_busuu_Study2016.pdf) ,
or [https://blog.busuu.com/wp-content/uploads/2016/05/The busuu Study2016.pdf](https://blog.busuu.com/wp-content/uploads/2016/05/The_busuu_Study2016.pdf)
- Vesselinov, R. and J. Grego, 2015, Efficacy of New Language App, report forthcoming.
- Vesselinov, R. and J. Grego, 2012, Duolingo Effectiveness Study.
http://comparelanguageapps.com/documentation/DuolingoReport_Final.pdf, or
http://static.duolingo.com/s3/DuolingoReport_Final.pdf
- Vesselinov, R., J. Grego, B. Habing, A. Lutz, 2009a, Measuring the Attitude and Motivation of Rosetta Stone® Users.
<http://comparelanguageapps.com/documentation/MeasuringTheAttitudeandMotivationofRSUsers.pdf>
- Vesselinov, R., J. Grego, B. Habing, A. Lutz, 2009b, Comparative Analysis of Motivation of Different Language Learning Software.
<http://comparelanguageapps.com/documentation/ComparativeMotivationAnalysisofDifferentLanguageSoftware.pdf>
- Vesselinov, R., 2008, 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.