Translating and Localizing a Mathematics Tutoring System to a Spanish Speaking Country

An Interactive Qualifying Project Report
Submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE

by

_____

Andrew Euredjian

Approved: _____

Professor Ivon Arroyo, Project Advisor

Table of Contents

Abstract

MathSpring, formerly known as Wayang Outpost, is an intelligent tutoring system purposed towards assisting students with learning various topics in mathematics. After seeing success with students in the United States, we localized MathSpring to Argentinian schools through translating the system to Spanish in the hope of achieving the same or similar success. We analyzed the data retrieved from pretests and posttests as well as the time students in Argentina spent using MathSpring to determine that MathSpring can have a positive impact on these students, but did not show a successful result likely due to the time constraints of this study.

1. Introduction

MathSpring, formerly known as Wayang Outpost, is an intelligent tutoring system purposed towards assisting students with learning various topics in mathematics (Arroyo et al, 2014). Students learn and become proficient in topics such as operations with fractions, decimals, measurement units, etc. through MathSpring's interface, which boasts a virtual tutor character, hints, and adaptive problem difficulty. The software has seen use in US schools with measured success. In an attempt to expand the use of MathSpring outside the US and other English speaking countries, an effort was made to create a complete translation of the system into Spanish from the text on the interface that students interact with to the audio that they listen to while solving problems.

With the translations in place, the goal of this research project involves a study that was conducted with 6th graders from three different schools in Argentina that looks to answer the following question: "What is the effectiveness of the MathSpring tutoring system when localized to a Spanish speaking country in terms of both its interface and the problems themselves?" In gathering the data from this study, improvements can be made to the existing translations while also providing a process for additional translations to other languages.

2. Background Research

Researchers have begun to think about how developed and developing countries can collaborate to strengthen the effectiveness and potential of learning technologies worldwide. Such collaboration could take place with the implementation of an international standardized curriculum and standardized activities in subjects such as mathematics where the material taught is very similar in different countries. One concept that could be implemented is the use of high quality eBooks translated into different languages for different countries but utilize the same curriculum (Barr, n.d.).

The subject of research in this IQP builds on other studies on localizing tutoring systems and learning technologies to the needs of developing countries. This study uses the MathSpring online math tutor to repeat studies previously conducted in the United States on the effectiveness of the online tutoring system. In this previous study with MathSpring, which at the time was called Wayang Outpost, the research team set out to learn how effective online assistive tutoring tools could be on students using such tools versus those that don't use them. The team used the results from standardized testing as a point of comparison between students that used the tool and those that did not, and in some cases gave pretests and posttests to the students as additional data points for measuring students' progress.

This study also goes into how the MathSpring tool aims to help students learn math topics found in standardized testing. The tool specifically targets three different areas for improvement: cognition, meta-cognition, and affect. Cognition can be defined as a student's ability or inability to solve problems of varying difficulty and topic. Meta-cognition are "cognitive resources and

mechanisms that help students to regulate their own learning" (Arroyo et al., 2014). Lastly, affect can be described as the emotions or feelings experienced when using the software such as boredom, excitement, confusion, etc. In order to help a student improve in these areas, the system employs a variety of techniques and technologies both visible to the user as well as behind the scenes. Behind the scenes, MathSpring keeps track of the user's progress and adaptively assigns new problems based on previous performance measures such as the time to solve the problem, the number of attempts, and the number of hints used among others. Interaction with the user is handled through a dynamic animated character that acts as a teammate or classmate that the user can go to for help. Users also have access to basic progress data for problem sets they have worked on. A page is dedicated to showing the progress for each topic encountered and offers insight on how well the user understands individual problems in a set (Arroyo et al., 2014).

MathSpring was tested by a research team in the United States to measure its effectiveness in enhancing a student's learning of mathematics several times over several years. During one of these studies, the research team also used a variety of physical sensors to measure students' emotional states while using MathSpring in the hopes of improving MathSpring's learning companion to respond automatically to students' affective states. In studies in the United States, students have generally shown significant improvement in their mathematical capabilities after using MathSpring. These improvements were measured through monitoring MCAS scores for students using MathSpring and those not using MathSpring. Those that used the tool received higher scores on the exam (Arroyo et al., 2014). However, this system has not been evaluated in

other countries, such as in the developing world where internet connections could be slower, and different languages are required.

Another study looks into the feasibility of using learning technologies with low socioeconomic schools in Chile. The name of the learning technology used was not given in the paper but is described in a similar fashion to MathSpring and can be assumed to have similar features. This team, similar to the MathSpring team, had the goal of adding to the number of options that students have for practicing and improving their mathematics skills while providing real-time feedback to the students as they work on problems. The system, like MathSpring, also provides tools for teachers to keep track of individual and class progress going as far as to notify the teacher if there are certain students that may require personal attention (Araya, 2013).

As this study was conducted with low socioeconomic schools, there were some challenges that the research team had to overcome. Some of those challenges included only 90% of students having internet access at home, lower quality teachers, poor maintenance of school equipment, unstable internet connections, higher number of students whose skills were lacking for their grade and struggling with core concepts, and lower attendance rates. However, the research team was able to overcome these challenges and produce convincing results of the effectiveness of the system in question. In standardized testing, schools that used the system exhibited a large average point increase over other schools with similar socioeconomic status as well as other schools in general throughout the country. The study also concluded that feedback using video or audio was the most effective form of feedback to induce improvement in a student's learning (Araya, 2013).

A third study conducted with MathSpring, at the time known as Wayang Outpost, focused more on the requirement of localizing the system to a country past just simple translation in order to be effective in being a mathematics tutor for students, the main concern being adapting the system to the local culture. Previous studies to this one have shown that adaptive tutoring systems such as MathSpring have been as effective or more effective than 1-on-1 human tutors which opens the door to introducing such systems in developing countries where there are a lack of teachers in some cases, common student absences in other cases, and both in more extreme cases. On top of these issues, teaching methods and curriculum are not up to par with those of more developed countries and as such can lead to a loss of learning potential for the students (Zualkernan, 2013).

This study also looked to delve into other methods of attempting to improve students' learning other than providing an adaptive online tutor. The team looked at the effects of different teacher-to-student ratios on individual learning, the result of which being that there were no significant effects with varying ratios. They also looked at the effects of providing textbooks to the students to use and study with at home. What resulted from this was that higher achieving students benefited while others showed little to no improvement as was shown through little to no increase in testing scores overall (Zualkernan, 2013).

The team then introduced the MathSpring system to 9 mathematics teachers from the same school in Peshawar, Pakistan for them to interact with and provide feedback on how well the system could align with the curriculum and teachings in the classroom. The teachers estimated that they could dedicated around 2 hours per week to using the system in their classes and noted

that the system more or less aligns with the types of problems done in class and the methods of teaching they use with their students. However, when asked if they agree that the system aligns with the class, most responded with "Neither Agree or Disagree". This feedback suggests that while MathSpring is viewed by these teachers as a tool with potential to aid in the classroom as a complementary tool to in-person instruction, improvements can be made to the localization of the system in order to better align with the culture of the region and the schools (Zualkernan, 2013).

3. Research Question

The following is the research question that I am attempting in this study, Which have not been answered yet given the background research presented before:

What is the general effectiveness of the MathSpring tutoring system localized to a spanish speaking country? Is it conducive to student learning? How effective can english problems translated to spanish be for spanish speaking schools/students?

4. Methodology

This section will detail the methods by which the study was conducted through providing details about the participants and about the administering of the MathSpring tutoring system.

4.1 Participants

The subjects of this study were Argentine students in 6th grade (11 year old children) from three different schools in the city of Cordoba. The schools in question will be referred to as School A, School B, and School C All three schools were private schools with School C specifically being a catholic school. School A had three classes participating in the study: the first class (6A) had 22 students, the second class (6B) had 18 students, and the third class (6C) had 18 students. School B had two classes participating in the study: the first class (6A) had 34 students and the second class (6B) had 37 students. School C also had two classes participating in the study: the first class (6V) had 16 boys and the second class (6M) had 15 girls.

4.2 Procedure

This study was conducted over a period of six weeks in which students from School B and School C used MathSpring once a week in the classroom while students from School A used the system twice a week in the classroom, once in Spanish and once in English.

The mathematics topics that the students practiced with MathSpring were:

1) Fractions and equivalent fractions (conversions)

2) Basic operations with fractions (addition/subtraction)

3) Advanced operations with fractions (multiplication/division)

4) Decimals (conversions from fractions to decimals and identification)

5) Basic operations with decimals (addition/subtraction)

6) Advanced operations with decimals (multiplication/division)

7) Unit conversion using the metric system (weights, volumes, distances)

8) Unit conversion using time (hours, minutes, seconds)

On the first day for each class, the research team gave a basic introduction of the MathSpring tutor and administered a pre-test consisting of ten questions covering the topics listed above to evaluate the students' current capabilities in mathematics and to use as a base from which to measure improvement with MathSpring. These questions are not available in the problem sets that the students will eventually work through. In some classes, there were not enough computers for each student to have their own. For these cases, students that did not get their own computer were given a paper exam and later paired up with students that had taken the exam on a computer to work together. Each student, or pair of students depending on the ratio of computers to students, received login credentials that they would use for the entirety of the study. Students from School A received two logins, one for Spanish and one for English. Once finished with the pre-test, students used the remaining class time to begin using MathSpring and solved problems from the first topic.
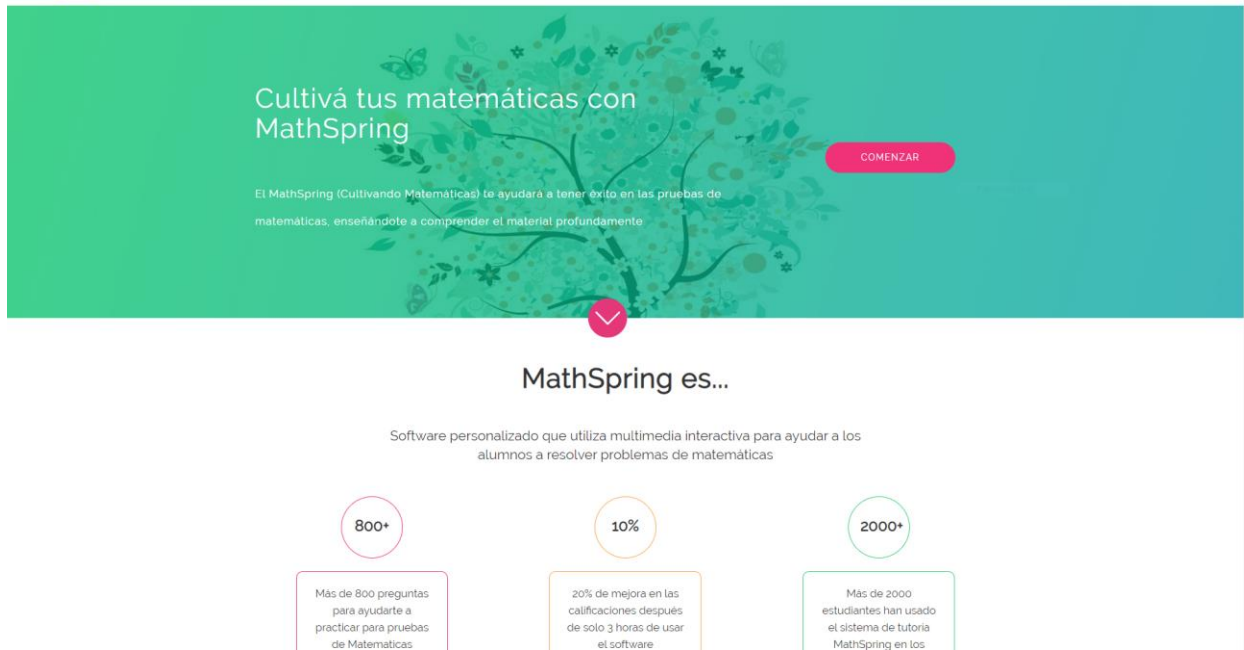
Figure 1: MathSpring homepage in Spanish

Each day between the first and last days, the students worked on problems throughout the class period in MathSpring. A new problem set was open to the students per day while problem sets from previous days were closed to make sure the students stayed on track. Students were expected to be finished or close to finished with a problem set by the end of each class period such that by the end of the study, all topics listed above would have been covered. In some classes two problem sets were opened with the expectation that the first would be completed by the students quickly and with time to spare to work on another problem set.
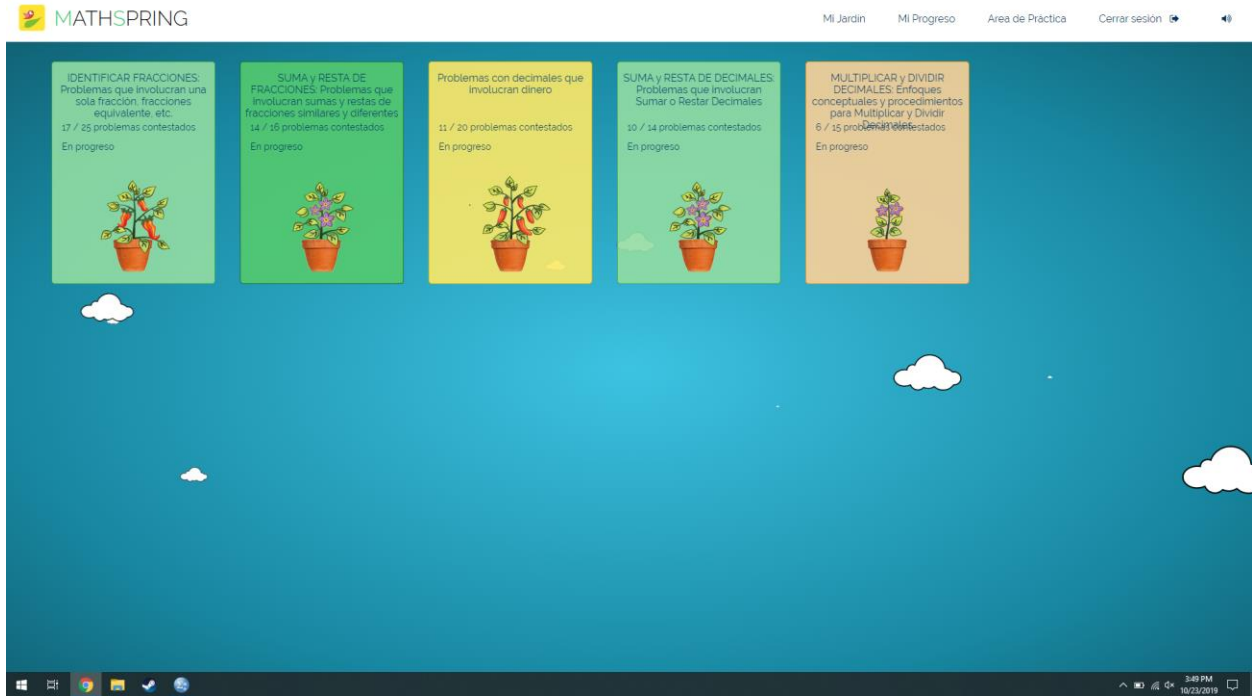
Figure 2: MathSpring dashboard page in Spanish displaying the different problem sets
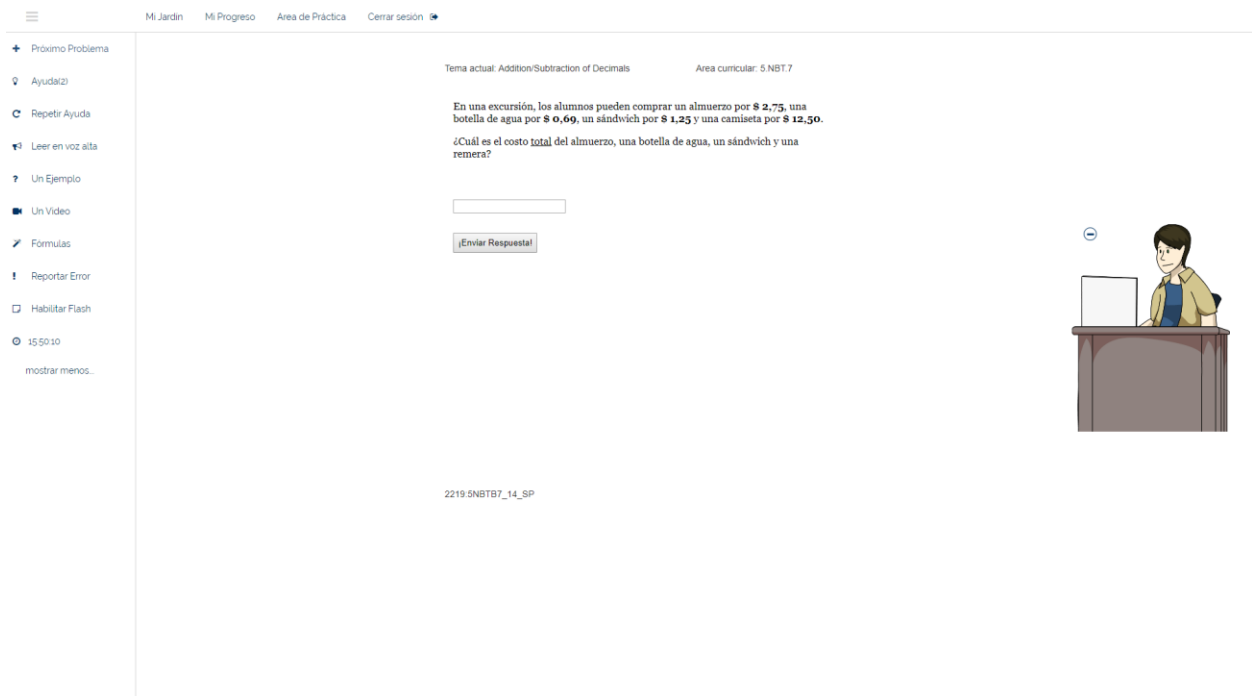

Figure 3: A Spanish problem in MathSpring on the practice page

The last day, the research team administered a post-test for each class. The post-test had the same

questions as the pre-test in order to find out if students could correctly answer questions that they

had struggled with or not known the answer to on the first day after having used MathSpring. Similar procedures to that of the pre-test were conducted where classes with an inadequate number of computers compared to the number of students had some students taking a paper exam. After completing the post-test, all problem sets were opened for the students to view their overall progress over the six week period.

Figure 4: Students of School A completing the pretest on day 1

4.3 Measures to answer the research question

There are a couple methods by which I planned to answer the research question previously stated. The first method is through administering a pre-test and post-test to the students at the

beginning and end of the study respectively. The pre-test acted as a way to find what skill level each student was at in mathematics prior to using MathSpring.The post-test was used to measure any and all improvement in the students' mathematics skills after having used the MathSpring tutor for six weeks. The two tests used the same questions that cover the 8 topics listed above.

The second method is through analyzing students' statistics for the problems they worked on. MathSpring collects analytical data such as time spent on a problem, number of attempts used, number of hints used, and other useful pieces of data that can help determine if a user is understanding the material, struggling with the material, or not making an effort and clicking randomly. Between the three schools, the results from this method may vary significantly due to the different situations. School A got more time per week with the system, School B had students working in pairs as there were not enough computers to accommodate all the students, and School C had a mix of students working individually and in pairs.

These two methods also work together to get a more detailed look at how certain students may have advanced more than others. We could also see if and when other students helped or provided answers to other students to an extent by analyzing their respective progress in MathSpring in correlation to their testing scores.

5. Results

The following sections will show various analyses of the three datasets retrieved during this study. The pretest data will tell us the starting points for each of the students with respect to their mathematical skill level. The problem set data will tell us how well each student was able to understand and complete practice problems. The posttest data will tell us the amount of improvement or change in each student's mathematical skills based on their scores from the pretest.

5.1 Pretest data

The pretest that all students took before beginning to work with MathSpring aimed to define where the students stood in terms of mathematics skills. The results from the pretest are displayed in the table below and categorized by school and math topic as well as displaying the averages of all students per school and of all students across every school. The average pretest score for all students across every school came out to around 44%. Between each topic, students scored best on problems involving decimals and decimal operations at around 66% and scored worst on problems involving measurements and unit conversions at around 27%. Students also scored an average of around 47% on problems involving fractions and around 38% on problems involving time.

When looking at the three schools individually, it can be observed that students from School A on average scored better than students from School B and School C. Schools B and C both scored similarly around 38-40% while School A scored a bit over 10% better at 52%. When looking at the scores for each mathematics topic, the three schools share some patterns. Students

16

from all three schools did best on problems with decimals where School A scored 68%, School B scored 70%, and School C scored 56%. Students from Schools B and C had scores within 10% of each other on average in three of the four topics suggesting the students from Schools B and C are at a similar skill level in mathematics in those topics.

Along with the lowest mean score of the three schools, School C also had the highest standard deviation at 21.5% while Schools B and A were at 19% and 21% respectively. This suggests that there is a wider range of skill level in mathematics across students in School C than the other two schools while School B shows the most consistency between the skill levels of individual students despite having the most students of the three schools. The hope is that in using MathSpring, the average scores will be raised or the standard deviations will be lowered or a combination of both.

Table 1.
Means and Standard Deviations of Pre-test math ability for each school in the Argentina study

School		pre_test_score	pre_fractions_score	pre_decimals_score	pre_measurements_score	pre_times_score
Academia Arguello	Mean	.5231	.5817	.6827	.3910	.3654
	N	52	52	52	52	52
	Std. Deviation	.21015	.21985	.34341	.32820	.48624
Bernardez	Mean	.4059	.4338	.7059	.1863	.3529
	N	68	68	68	68	68
	Std. Deviation	.19152	.25414	.31428	.21068	.48144
Maria de Nazareth	Mean	.3872	.3910	.5641	.2479	.4359
	N	39	39	39	39	39
	Std. Deviation	.21543	.28559	.34754	.27272	.50236
Total	Mean	.4396	.4717	.6635	.2683	.3774
	N	159	159	159	159	159
	Std. Deviation	.21079	.26232	.33516	.28184	.48626

5.2 Posttest data

At the end of the study period, a posttest was administered to the students containing the same questions as the pretest. The results can be found in the table below in the same format as the pretest results shown previously. The average pretest score for all students across every school came out to around 49%. Between each topic, students scored best on problems involving decimals and decimal operations at around 67% and scored worst on problems involving measurements and unit conversions at around 31%. Students also scored an average of around 55% on problems involving fractions and around 44% on problems involving time.

When looking at the three schools individually, it can be observed that students from School A scored the highest of the three schools as they did in the pretest, but by less margin than the pretest. Students at School C scored an average of 45%, School B students an average of 49% and School A students an average of 51%. Once again, students from all three schools scored best on the section of the test involving decimals with all averages being above 60%. Schools B and C once again scored very similarly with scores being within 10% of each other in each topic. An interesting variation to the pretest results to note is that all three schools had very similar averages for problems involving fractions. This could be a result of the standardized problem sets for fractions in MathSpring that students from all three schools practiced on.

In contrast to the results of the pretest, School A ended up with the least average standard deviation of the three schools at 20%. School B had the most at 22% and School C had 21%. This data suggests that the skill level in mathematics across the students in Schools A and C

became more consistent while those of School B became less consistent. This could be due to some kids obtaining better understanding of the materials quicker than their classmates.

Table 2.

Means and Standard Deviations of Post-test math ability for each school in the Argentina study

School		post_test_score	post_fractions_score	post_decimals_score	post_measurements_score	post_times_score
Academia Arguello	Mean	.5154	.5529	.6442	.3526	.5962
	N	52	52	52	52	52
	Std. Deviation	.19841	.21770	.34777	.29084	.49545
Bernardez	Mean	.4912	.5515	.6985	.3235	.3382
	N	68	68	68	68	68
	Std. Deviation	.21973	.26821	.31314	.27602	.47663
Maria de Nazareth	Mean	.4548	.5323	.6774	.2258	.3871
	N	31	31	31	31	31
	Std. Deviation	.20951	.27189	.30409	.29043	.49514
Total	Mean	.4921	.5480	.6755	.3135	.4371
	N	151	151	151	151	151
	Std. Deviation	.21025	.25119	.32248	.28605	.49768

When comparing the means of all students' pretests and posttests as well as the scores on each topic in the tests, it can be observed that the students improved over the course of the study while using MathSpring. The overall test score average increased from 43% to 49% which, for such a short amount of time students were exposed to each topic, is a sizeable increase. When looking at the topics individually, students were found to have improved most with problems involving time, going from 35% to 45% with problems involving fractions a close second in amount improved going from 46% to 54%. Students also improved a decent amount on the topic of measurement and unit conversions going from 26% to 32%. However, an interesting piece of data to note is the lack of improvement on problems involving decimals which stayed constant at 66.3%. Although this result may suggest an ineffective system for improvement, it is important to look at the standard deviations as well to find if the students' scores became more consistent after the study period.

When comparing the standard deviations for the overall tests, we found there was a very small amount of change which turned out to be an increase of about 0.2%. This would suggest that the overall pool of students all improved on their posttests, as the mean shows a 5% increase, but maintained the same or similar skill gaps between the least performing students and the most performing students. Looking at the individual topics once again, we observed that there were not large changes in standard deviation among the separate topics either. Problems involving fractions or measurements only experienced a less than 1% change in their standard deviations. Meanwhile, we saw the standard deviation for problems involving time increased by about 2% suggesting that, despite the mean increasing by nearly 10%, some students obtained better understanding while others remained at or around their pretest level of understanding. Going back to the decimals section, we found that its standard deviation saw the most improvement in decreasing about 2% from 34% to 32%. This result implies that, although the mean did not experience a large change between pretest and posttest, the students scores became more consistent among one another which is the goal of providing standardized problem sets to all students through MathSpring.

Table 3.
Comparison of means between the overall Pre-test and Post-test and individual topics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	pre_test_score	.4319	141	.21021	.01770
	post_test_score	.4901	141	.21223	.01787
Pair 2	pre_fractions_score	.4628	141	.26548	.02236
	post_fractions_score	.5443	141	.25580	.02154
Pair 3	pre_decimals_score	.6631	141	.34067	.02869
	post_decimals_score	.6631	141	.32457	.02733
Pair 4	pre_measurements_score	.2624	141	.28118	.02368
	post_measurements_score	.3168	141	.28543	.02404
Pair 5	pre_times_score	.3546	141	.48010	.04043
	post_times_score	.4468	141	.49894	.04202

After a simple comparison of means, a paired samples t-test was conducted to find if there was any statistical significance to the test scores from pretest to posttest. The table below contains the result of this t-test. Looking at the overall test, we found that the change in average score was significant with p-value being less than 0.05 at 0.005. However, from the individual topics only the fractions topic was determined to have a significant change in average score from pretest to posttest with a p-value of 0.002 despite not having the highest improvement. The other three topics were determined to not have significant changes in average score with p-values of 1.0 for decimals, 0.079 for measurements, and 0.085 for times.

Table 4.
Result table of a paired-samples T-test on the overall Pre-test and Post-test and individual topics

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	pre_test_score - post_test_score	-.05816	.23969	.02019	-.09806	-.01825	-2.881	140	.005
Pair 2	pre_fractions_score - post_fractions_score	-.08156	.30103	.02535	-.13168	-.03144	-3.217	140	.002
Pair 3	pre_decimals_score - post_decimals_score	.00000	.41404	.03487	-.06894	.06894	.000	140	1.000
Pair 4	pre_measurements_score - post_measurements_score	-.05437	.36433	.03068	-.11503	.00629	-1.772	140	.079
Pair 5	pre_times_score - post_times_score	-.09220	.63133	.05317	-.19731	.01292	-1.734	140	.085

5.3 Problem set data

While the students used MathSpring, the system kept track of how much time each student spent

on the system. In theory, all the students should have similar times logged once the study

completed. However, due to unknown non-student days at the schools, there were significant

differences in the amount of time spent using MathSpring between students of each school and

even students of the same school from different classes. There were also cases of students

memorizing their login credentials and using the system at home. While this is good for the

students to have more exposure to the system and hopefully learning more, it creates more

variability in the amount of time spent on the system per student. As a result of the variability of

time spent on the system per student, an analysis was done on the relation between different

amounts of time spent on the system and the change in score from pretest to posttest. The result

of this analysis can be seen in the table below generated from a linear regression where the

posttest score was the dependent variable and total time spent in MathSpring was the

independent variable along with the pretest score. With a p-value of 0.324 being greater than

0.05, it can be observed that despite the varying time spent in MathSpring between students of

22

each school, the amount of time spent in MathSpring was not a significant predictor for

determining the scores on posttests.

Table 5.
Result table of a linear regression conducted on the Post-test using Pre-test and time spent in MathSpring as the independent variables

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.304	.049		6.174	.000
	pre_test_score	.356	.080	.352	4.437	.000
	minutes_in_MS_TOT	.000	.000	.079	.989	.324

a. Dependent Variable: post_test_score

Despite not being significant for the overall posttest, a different story is told when looking at the

individual sections of the test. Looking at the fractions section first, the time spent in MathSpring

turned out to not be significant at all with a p-value of 0.975 being greater than 0.05. For

decimals, the time spent in MathSpring was still not significant, but yielded a better result with a

p-value of 0.211. For measurements and unit conversions, the time spent in MathSpring was also

not significant, but just barely considering it had a p-value of 0.061. For times and time

conversions, the time spent in MathSpring turned out to be very significant, resulting in a p-value

of $0.009 < 0.05$.

Table 6.
Result table of a linear regression conducted on the Post-test fractions section using the Pre-test fractions section and time spent in MathSpring as the independent variables

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.397	.056		7.147	.000
	minutes_in_MS_TOT	-1.230E-5	.000	-.002	-.031	.975
	pre_fractions_score	.321	.078	.334	4.146	.000

a. Dependent Variable: post_fractions_score

Table 7.

Result table of a linear regression conducted on the Post-test decimals section using the Pre-test decimals section and time spent in MathSpring as the independent variables

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.582	.076		7.639	.000
	minutes_in_MS_TOT	-.001	.001	-.104	-1.256	.211
	pre_decimals_score	.221	.079	.232	2.814	.006

a. Dependent Variable: post_decimals_score

Table 8.

Result table of a linear regression conducted on the Post-test measurements section using the Pre-test measurements section and time spent in MathSpring as the independent variables

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.185	.055		3.344	.001
	minutes_in_MS_TOT	.001	.000	.156	1.887	.061
	pre_measurements_score	.171	.084	.168	2.034	.044

a. Dependent Variable: post_measurements_score

Table 9.

Result table of a linear regression conducted on the Post-test time section using the Pre-test time section and time spent in MathSpring as the independent variables

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	.166	.096		1.725	.087
	minutes_in_MS_TOT	.002	.001	.219	2.669	.009
	pre_times_score	.194	.085	.187	2.278	.024

a. Dependent Variable: post_times_score

6. Discussion

Despite the study being completed with some positive results, there are improvements that could be made in the next study should it occur. Such improvements can be made in all stages of the study from the preparation work to the study itself. Looking at preparatory work for the study first, most improvements here that can be made are mostly to make conducting the study and post-study analysis easier for the researcher. One such improvement could be to set up some sort of tracking system of users that have multiple accounts (for example, one account in English and one account in Spanish as was the case for one of the schools in this study) or accounts that have multiple users. A couple ways of approaching this could be to utilize a manual method of writing down the multiple users/accounts per account/user respectively or employ a solution in software to connect the accounts in some way or allow an account to have multiple users listed on it. Another improvement that could be made in preparation for the next potential study would be to trial and test the MathSpring system in its translated form extensively to find any potential issues. This would allow for issues to either be fixed before a study or for an opportunity to create a concrete process of getting around the issue during the study. The final improvement that can be made to the preparatory work for a new study could be to find out about any school holidays or non-student days during the period that the study will be conducted in order to build a schedule that works around those holidays to provide the most uniform experience with MathSpring to all participants.

There were a couple places where improvements could be made during the study as well. The first improvement would be for the very beginning and end of the study when the pretest and posttest are being administered. In this study, we observed that due to the seating arrangements

of the students, groups of students would work together and copy answers off one another to distribute the workload and hopefully get a better score. Although the ability to collaborate is a good quality for students to have, in this case we preferred to have each student working individually to obtain individual performance measures. As such, a new study could employ the use of dividers or specific seating arrangements that deter the sharing of answers during pretesting and posttesting. Another issue that was observed in this study was that the students were somewhat confused on how to navigate MathSpring in the beginning of the study. This can be improved on by simply giving a more in-depth tour of MathSpring and how to navigate the system, such as showing where each of the menu buttons go and how to navigate to different problem sets. With regards to the pretest and posttest, this study experienced an issue where the one problem in the tests meant to test students on time and time conversions was essentially a fraction multiplication problem. Future research teams can easily fix this by selecting a different problem under the topic of time and time conversions to be used in the tests. Lastly, this study encountered the issue of students being absent from school for the first day of the study or longer and as a result, starting somewhat behind from the rest of the group. A solution to this issue is not immediately clear, however one solution could be to take note of the student arriving late and cut that student's data out from the analysis after the study. At the same time, such an extreme measure can not be employed in the case of overwhelming absences on a single day. Therefore, this solution could be modified to consider how many classes an individual student was absent from as well as their progress with respect to the rest of the study.

Despite there being a number of areas for improvement in this study, there were also a few cases in which the study went smoothly. At the beginning of the study, the students were given papers

with their usernames to use for the duration of the study. However, these small pieces of paper had a tendency to get lost or misplaced and as a result, the student would not be able to access their account having not memorized their username. Fortunately, MathSpring's teacher tools had a way of finding out which usernames belonged to which student and we were able to provide students with their login credentials without having to provide new accounts. Additionally, our research team was able to quickly put in a process before each class of enabling the flash player for the Google Chrome browser and set the language when needed after encountering issues of students not being able to see problems that required flash to be enabled. With this process put in, time wasted fixing this issue during the class period was reduced to a minimal amount. Lastly, we quickly found that although new problem sets were being released for the students to explore and work through each class period, the students were opting to keep working on older problem sets. Needing to get data for all the math topics, we started disabling older problem sets while enabling newer ones in MathSpring's teacher tools for each class period to get the students to work on the newer problem sets. This capability in the teacher tools allowed us to get at least some data for each of the math topics.

7. Conclusion

The overall goal of this study was to introduce and administer a mathematics tutoring system called MathSpring to 6th grade students in Argentina to see if by translating the system to Spanish, the same or similar effect could be achieved compared to students in the United States using MathSpring. More specifically, does the use of a localized MathSpring tutor yield improvement in student's mathematical skills? When looking at the data shown in the results, the first answer to this question might be: No, the use of MathSpring, as shown in the linear regression of the overall tests with time spent in MathSpring, does not provide a significant improvement in learning in its localized form. However, when delving deeper into the results of the study and the results from the individual topics, a different conclusion can be observed. We find that the calculated significance for the overall test scores based on time spent in MathSpring are greatly skewed by a very high p-value for the fractions section of the test. Though the decimals and measurement sections also had p-values that suggested insignificance, they were far closer to being significant than the fractions section. Also, the improvement in the time section was calculated to be very significant, granted the one problem in the section was essentially a fractions problem. Therefore, due to the nature of the study having had to cover multiple math topics over such a short period of time, it's possible to conclude that time spent on MathSpring can have a significant impact on a student's development of their math skills, but due to the time restrictions in this study, we were not able to observe such an impact on the students of the participating classes.

References

Araya, R., & Molen, J. V. D. (2013). Impact of a blended ICT adoption model on Chilean
vulnerable schools correlates with amount of on online practice [PDF file]. Retrieved
from
https://www.researchgate.net/publication/262068933_Impact_of_a_blended_ICT_adopti
on_model_on_Chilean_vulnerable_schools_correlates_with_amount_of_on_online_pract
ice

Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A Multimedia
Adaptive Tutoring System for Mathematics that Addresses Cognition, Metacognition and
Affect [PDF file]. Retrieved from
https://www.researchgate.net/publication/266398598_A_Multimedia_Adaptive_Tutoring
_System_for_Mathematics_that_Addresses_Cognition_Metacognition_and_Affect

Barr, A., Richards, T., & Robson, R. (n.d.). Applying Standards and AI to Educational
Technology: The IEEE Actionable Data Book Project [PDF file]. Retrieved from
http://ceur-ws.org/Vol-1009/0606.pdf

Zualkernan, I., Arroyo, I., & Woolf, B. P. (2013). Towards Localization of Automated Tutors for
Developing Countries [PDF file]. Retrieved from
https://www.semanticscholar.org/paper/Towards-Localization-of-Automated-Tutors-for-
Zualkernan-Arroyo/613278210c21f54127fd8666734a589345600ba8

Appendix

Appendix A. Pre/Posttest for students at the beginning and end of the study

Nombre: _____   Grado: _____   Division: ____   Colegio: _____

**Ahora vamos a hacerte algunas preguntas de matemáticas. Respondé lo mejor que puedas.
Si no sabés, no hay problema: estimá la respuesta lo mejor que puedas.**

9) En esta recta numérica, ¿cuál de los puntos representa mejor la ubicación de la fracción 1/5?
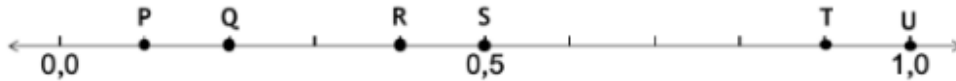
A. El punto P
B. El punto Q
C. El punto R
D. El punto S
E. El punto T

10) Horacio puso estampillas en algunas postales como se muestra en el dibujo.

¿Qué fracción de las postales tienen una estampilla?

Respuesta: _____

11) ¿Cual es el resultado de esta expresión?

$$\frac{2}{3} + \frac{7}{9} - \frac{5}{6}$$

A. 1/2
B. 3/4
C. 11/18
D. 5/9
E. 5/18

12) Juan quiere comprar una libreta que cuesta $ 1,64, una caja de plasticolas por $2,66, una caja de lápices por $ 2,81 y una goma de borrar por $ 0,78. ¿Cuál es el costo total de los cuatro artículos que Javier quiere comprar?

    A. 7,57
    B. 7,96
    C. 7,81
    D. 7,89
    E. 7,74

13) Valeria compra **3 metros** de cordón para hacer pulseras. Ella necesita **22 centímetros** de cordón para hacer **1 pulsera**. ¿Cuántas pulseras puede hacer Sofía si usa todo el cordón que compra?

Respuesta: _____

14) Un carpintero usa 3/7 de metro de una barra de madera para construir uno de los lados del marco de una ventana cuadrada. El carpintero necesita 4 pedazos iguales para enmarcar la ventana (porque la ventana es cuadrada). ¿Cuantos metros de madera necesitará el carpintero para completar el marco de la ventana?

Respuesta: _____

15) Lucía quiere comprar 5 aros de hula hula para jugar con sus amigas. Cada aro cuesta $1,25. Lucía tiene un cupón de descuento de $ 0,75 del precio de un aro. ¿Cuánto deberá pagar Lucía por los 5 aros de hula hula?

    A. $ 6,00
    B. $ 6,50
    C. $ 5,75
    D. $ 5,25
    E. $ 5,50

16) Carlos tomó un ómnibus para ir a visitar a sus tíos por cuatro días. En la estación de autobuses, esperó 2/3 de hora hasta que llegó el momento de abordar el autobús. ¿Cuántos minutos esperó Carlos para abordar el autobús?

Respuesta: _____

17) Jorge contó todo los pájaros en que vió en su patio en una semana. Los resultados están representados en la figura de abajo. ¿Qué fracción de los pájaros eran horneros?
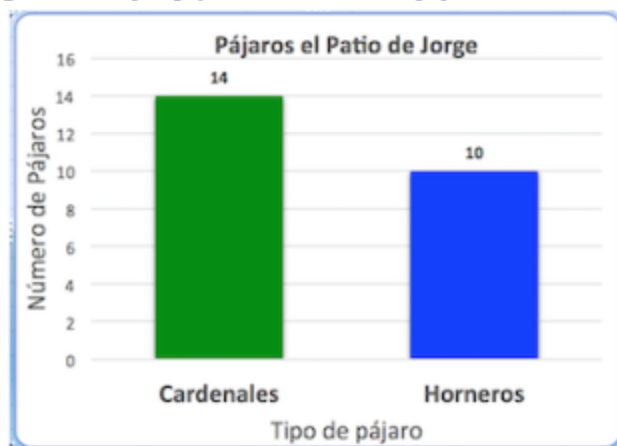
A. 10/24
B. 10/14
C. 14/24
D. 14/10
E. 24/10

18) Mariana llenó una regadera con 2,93 litros de agua. Usó 70 mililitros para regar su Jazmín y 220 mililitros para regar sus Margaritas. ¿Cuál es la cantidad de agua que queda en la regadera después de regar?

A. 3 mililitros
B. 264 mililitros
C. 587 mililitros
D. 1244 mililitros
E. 2640 mililitros