Measuring the Attitude and Motivation of Rosetta Stone® Users

FINAL REPORT

Roumen Vesselinov, Ph.D.
Economics Department
Queens College | City University of New York
Roumen.Vesselinov@qc.cuny.edu
(718) 997-5444

July 2009

RESEARCH TEAM

- Roumen Vesselinov, PhD,
  Economics Department,
  Queens College, City University of New York
- John Grego, PhD,
  Statistics Department,
  University of South Carolina
- Brian Habing, PhD,
  Statistics Department,
  University of South Carolina
- Amy Lutz, PhD,
  Department of Sociology,
  The Maxwell School of Citizenship
  and Public Affairs, Syracuse University

CONTENTS

Executive Summary

This is the first scientific study measuring the attitude and motivation of Rosetta Stone® (RS) users. The study utilized three previously established and validated research instruments and multidimensional scales for measuring motivation and attitude. The three scales also passed the test for internal reliability based on our sample. The sample was stratified using propensity score calculations of the participants' language ability. The analysis was based on a stratified representative sample of 164 participants who used the Rosetta Stone® Spanish software for one full month. The motivation level of the participants was measured in the beginning of the study and at the end of the study.

MAIN FINDINGS

1. The level of intrinsic and extrinsic motivation for studying foreign language increased during the study and this increase was statistically significant. All six of the Academic Motivation Scale (AMS) subscales registered statistically significant increase at the end of the study. The average increase was about one point on a seven point scale, or 14.2% of the maximum score.

2. The level of the post experimental linguistic goals motivation was very high with an average of 5.6 on a 7 point scale, or 80% of the maximum score. The average scores for all five positive subscales of the Intrinsic Motivation Inventory (IMI) were above the middle point of the scale. The negative subscale had a very low score.

3. The level of the post experimental non-linguistic goals motivation was measured by the Attitude/Motivation Test Battery (AMTB) scale. The highest level of motivation was 5.4 on a 6 point scale, or 90% of the maximum score. These high levels were observed for "Attitude toward learning Spanish", "Interest in foreign languages", and "Desire to learn Spanish" and "Integrative orientation." The "Spanish course anxiety" and "Spanish use anxiety" scores were relatively low.

4. Two of the AMTB subscales were used to evaluate the RS language learning software and the process of learning Spanish. Users evaluated the RS software very favorably with average score of 5.3 on a 6 point scale (or 88.3% of the maximum score). The one month language course was also highly valued by the participants. The average score of this subscale was 5.2 on a 6 point scale or 86.7% of the maximum score.

5. Participants evaluated very favorably their experience using the RS software. About 90% of them considered the RS software to be easy, helpful, enjoyable, and satisfying. The majority of them (90.2%) stated that they would recommend the RS software to others who are interested in learning Spanish.

The overall conclusion of this study is that people using the RS language learning software were able to maintain and increase all aspects of their foreign language intrinsic and extrinsic motivation and their post experimental levels of linguistic goals motivation and non-linguistic goals motivation were very high. The participants had very favorable views of the software and the learning process.

Introduction

Creating and maintaining motivation and positive attitude for studying foreign languages is a very important matter. This research study is about measuring the level of motivation and attitude towards studying foreign language (Spanish) when using the language learning software Rosetta Stone®. The study was commissioned by Rosetta Stone® but the analysis was conducted independently by the research team.

The analysis is based on a sample of people who responded to an advertisement in Washington, DC area newspapers. To be eligible for the study, respondents had to be between the ages of 18 and 65 and neither native speakers of Spanish nor advanced speakers of Spanish.

The selected respondents were given the opportunity to use Rosetta Stone® Spanish (Latin American version) for one full month. They were given access to the web version of the software and the opportunity to use it at home. In the beginning of the study and at the end of the study information was collected about the level of their motivation and attitude. There were incentives for the participants who completed the study. They received a 6-month free subscription to the Rosetta Stone ® Spanish web version software and a bonus of $100.

Part 1. Methodology

SAMPLE

The original pool consisted of 4656 people that responded to a newspaper advertisement looking for participation in a Spanish language study. We had basic demographic information on everybody in the pool and we excluded people who were below the age of 18 or above 65, or were advanced or native speakers of Spanish.

The study was conducted in April-June 2009. We randomly selected 300 of the eligible respondents and invited them to participate in the study and required them to complete a new demographic survey, initial motivation measures and Spanish evaluation test. The study design was to allocate approximately 200 participants for this study and 100 for another exploratory follow-up study.

In order to address the differences in language skills we performed a propensity scores matching and stratified the sample before assigning the people to the different samples. In the propensity score model the outcome variable was the participants' level of Spanish ability evaluated by the web-based test WebCAPE. The independent variables in the propensity score model were the major demographic characteristics: gender, age, race, education, employment (full time or not), spouse speaks Spanish (Yes/No), respondent knows some Spanish (Yes/No), knowledge of another foreign language (Yes/No), reason for studying Spanish (work, travel, other).

The predicted propensity scores were based on the regression model described in the previous paragraph. The whole sample (N=300) was stratified in three equal size clusters/strata (N=100 each): low, medium and high ability according to the predicted propensity scores. Then we randomly selected from each strata with equal intensity to constitute the sample for this study. The starting sample size for this study was N=200. In addition, 3 more people originally assigned to other groups were added to the sample because they used Mac computers which were not compatible with the other exploratory study. Thus 203 people started this study.

Since we had to work with a stratified sample we used the appropriate statistical methods to account for the complex sampling. We applied the stratified estimator (Cochran, 1977) in the SAS® software procedure PROC SURVEYMEANS (SAS OnlineDoc, SAS Help and Documentation).

Part of the analysis incorporates the use of difference scores (DS). DS are computed for variables for which there are two values, one in the beginning of the study and one at the end of the study. This is the case with one of the motivation scales which was administered once at the beginning of the study and once at the end of the study. DS are calculated as the difference between the value of the variable at the end of the study and the initial value of the variable. Negative values of DS would mean that the variable level has decreased during the study, and positive values would mean that the level has increased. Zero and close to zero values of DS would mean that the variable level during the study has not changed.

SPANISH LANGUAGE EVALUATION TEST

In order to evaluate the level of Spanish ability we used the WebCAPE test (Web-based Computer Adaptive Placement Exam) developed by the Perpetual Technology Group (http://webcape.byuhtrsc.org).

This is a well established foreign language placement exam with good validity and reliability (test-retest = 0.86). According to their website, more than 500 colleges and universities use WebCAPE for language courses placement. Among them are Harvard University, Boston University, Vanderbilt University, Brown University, Queens College, CUNY, University of South Carolina, Cornell University, etc.

The maximum score for Spanish achieved empirically for this computer adaptive test had been 956. The scores are usually a positive number but it is possible to get zero or negative score because of the weights on the questions. A negative score can be interpreted in the sense that the participant did not take the test seriously or that there were other obstacles because the test is adaptive and every question depends on the answer to the previous question. In that respect, preferably negative scores should be set equal to zero. WebCAPE creators suggest the following cutoff points for placement in a Spanish college course depending on the length of the course measured by number of semesters.

A student at a college with 6 semester Spanish course will need at least 204 points on WebCAPE to move or be placed in Semester 2. Respectively a student at a college with 5 semester Spanish course will need at least 234 points; with 4 semester Spanish course – at least 270 points, and with 3 semester courses – at least 281 points.

Table 6. Suggested Calibration Scores

WebCAPE Suggested Calibration Scores							
Spanish: (3) Courses		**Spanish: (4) Courses**		**Spanish: (5) Courses**		**Spanish: (6) Courses**	
Sem 1	Below 280	Sem 1	Below 270	Sem 1	Below 324	Sem 1	Below 204
Sem 2	218 - 351	Sem 2	270 - 345	Sem 2	234 - 311	Sem 2	204 - 288
Sem 3	Above 351	Sem 3	346 - 427	Sem 3	312 - 383	Sem 3	289 - 355
		Sem 4	Above 427	Sem 4	384 - 456	Sem 4	356 - 434
				Sem 5	Above 456	Sem 5	435 - 497
						Sem 6	Above 497

1 - Personal correspondence with Dr. Jerry Larson, Professor of Spanish Pedagogy, Brigham Young University.

MEASURING THE MOTIVATION AND ATTITUDE

In this study we used three different sets of scales.

SCALE 1: ACADEMIC MOTIVATION SCALE (AMS)

This scale was developed first by Vallerand et al (1992, 1993). It measures motivation to study a foreign language and contains 7 subscales assessing intrinsic motivation towards knowledge, accomplishments and stimulation, as well as external, introjected and identified regulation, and amotivation. It contains 28 items assessed on a 7-point Likert scale. The items are rated on a scale, ranging from 1 (does not correspond at all) to 7 (corresponds exactly). Each subscale consists of four items so the score can range from 4 to 28. It can also be averaged so the score is between 1 and 7 which makes the interpretation easier. A high score indicates a high approval/support for that particular academic motivation. Subscales are related to intrinsic motivation (IM), extrinsic motivation (EM) and amotivation and are listed below.

IM-to-know

IM is performing an activity for itself, and the pleasure and satisfaction derived from participation. IM-to-know "… relates to several constructs such as exploration, curiosity, learning goals, intrinsic intellectuality, and finally the IM to learn. Thus, IM-to know can be defined as the fact of performing an activity for the pleasure and the satisfaction that one experiences while learning, exploring, as trying to understand something new. For instance, students are intrinsically motivated to know when they read a book for the sheer pleasure that they experience while learning something new" (Vallerand et al, 1992).

IM-to-accomplish

"Individuals interact with the environment in order to feel competent, and to create unique accomplishments. ... Thus, IM-to-accomplish things can be defined as the fact of engaging in an activity for the pleasure and satisfaction experienced when one attempts to accomplish or create something. Students who extend their work beyond the requirements of a term paper in order to experience pleasure and satisfaction while attempting to surpass themselves display IM toward accomplishments" (Vallerand et al, 1992).

IM-to-experience stimulation

"IM-to experience stimulation is operative when someone engages in an activity in order to experience stimulating sensations (e.g., sensory pleasure, aesthetic experiences, as well as fun and excitement) derived from one's engagement in the activity. Research on the dynamic and holistic sensation of flow, on feelings of excitement in IM, on aesthetic stimulating experiences, and peak experiences is representative of this form of IM. Students who go to class in order to experience the excitement of a stimulating class discussion, or who read a book for the intense feelings of cognitive pleasure derived from passionate and exciting passages represent examples of individuals who are intrinsically motivated to experience stimulation in education" (Vallerand et al, 1992).

EM-identified

This type of motivation is external rather than internal. It is usually related to some other life goals like career success. One typical example is the following item: "Because I think that Spanish will help me better prepare for the career I have chosen."

EM-introjected

This subscale is related to other feelings of importance for life and career. A typical prompt is: "Because of the fact that when I succeed in studying I feel important."

EM – external regulation

This subscale is related to expectations for external guideline and directives, such as the following: "Because only with the language(s) I currently speak I would not find a high-paying job later on."

Amotivation

As the name shows it relates to the feeling opposite to motivation. A typical case is "Honestly, I don't know; I really feel that I am wasting my time studying Spanish"

SCALE 2: INTRINSIC MOTIVATION INVENTORY (IMI)

Scale Description :

"The Intrinsic Motivation Inventory (IMI) is a multidimensional measurement device intended to assess participants' subjective experience related to a target activity in laboratory experiments. It has been used in several experiments related to intrinsic motivation and self-regulation (e.g., Ryan, 1982; Ryan, Mims & Koestner, 1983; Plant & Ryan, 1985; Ryan, Connell, & Plant, 1990; Ryan, Koestner & Deci, 1991; Deci, Eghrari, Patrick, & Leone, 1994). The instrument assesses participants' interest/enjoyment, perceived competence, effort, value/usefulness, felt pressure and tension, and perceived choice while performing a given activity, thus yielding six subscale scores. Recently, a seventh subscale has been added to tap the experiences of relatedness, although the validity of this subscale has yet to be established. The interest/enjoyment subscale is considered the self-report measure of intrinsic motivation; thus, although the overall questionnaire is called the Intrinsic Motivation Inventory, it is only the one subscale that assesses intrinsic motivation, per se. As a result, the interest/enjoyment subscale often has more items on it that do the other subscales. The perceived choice and perceived competence concepts are theorized to be positive predictors of both self-report and behavioral measures of intrinsic motivation, and pressure/tension is theorized to be a negative predictor of intrinsic motivation. Effort is a separate variable that is relevant to some motivation questions, so is used as its relevant. The value/usefulness subscale is used in internalization studies (e.g., Deci et al, 1994), the idea being that people internalize and become self-regulating with respect to activities that they experience as useful or valuable for themselves. Finally, the relatedness subscale is used in studies having to do with interpersonal interactions, friendship formation, and so on."

IMI has 37 items grouped in 6 subscales. Each item is evaluated by a 7-point

2 - Source: University of Rochester: Self Determination Theory web page http://www.psych.rochester.edu/SDT/

Likert scale from 1 ("Not at all true") to 7 ("Very true"). The subscale score can be computed by averaging the items scores so the total score will be between 1 and 7 and thus is easier to interpret. The subscales are as follows:

Interest/Enjoyment
The subscale measures the interest and enjoyment while studying a foreign language. A typical item is: "While I was studying Spanish, I was thinking about how much I enjoyed it."

Perceived Competence
For learners of a foreign language, it is important to gain some confidence regarding their competence. A typical item is: "I am satisfied with my performance at studying Spanish."

Effort/Importance
A substantial part of the motivation is related to the perception of importance of the task at hand. A typical item is: "It was important to me to do well at studying Spanish."

Pressure/Tension
Undoubtedly the process of learning a foreign language is stressful and a certain amount of pressure and tension is inevitable. A typical item is: "I felt pressured while studying Spanish."

Perceived Choice
Even though the respondents for this study volunteered it is still important to evaluate whether they felt free to decide whether to study or not to study Spanish. A typical item is: "I did study Spanish because I wanted to."

Value/Usefulness
Last but not least, this subscale measures the value of learning a foreign language. A typical item is: "I believe studying Spanish could be beneficial to me."

SCALE 3: ATTITUDE/ MOTIVATION TEST BATTERY (AMTB)

AMTB was developed by Gardner (1985). The version used in this study was adapted from the English-language version of the AMTB for use by students studying English as a foreign language.

Scale Definition [3]
"The goals of any second language program are partly linguistic and partly nonlinguistic. The linguistic goals focus on developing competence in the individual's ability to read, write, speak and understand the second language, and there are many tests available with which to assess these skills. Non-linguistic goals emphasize such aspects as improved understanding of the other community, desire to contwinue studying the language, an interest in learning other languages, etc. Very few tests have been made available to assess these non-linguistic aspects.

The Attitude/Motivation Test Battery has been developed to fill this need. Its development follows more than 20 years of research…"

The scale has 104 questions defining 12 subscales. Each item requires an answer on a 6 point Likert scale ranging from 1 (Strongly Disagree) to 6 (Strongly agree). The total score for each subscale can be expressed as the average of the items' scores and it will also range between 1 and 6. The subscales are as follows:

1. Interest in Foreign Languages
This interest is obviously very important in the overall evaluation of the motivation and attitude. The bigger the interest is the bigger the motivation. A typical example is, "I wish I could speak many foreign languages perfectly."

2. Parental Encouragement
It is interesting to see what the role of parents was in the general orientation and attitudes towards learning a foreign language. A typical example is, "When I was in school, my parents felt that it was very important for me to learn a foreign language." For our sample this encouragement measurement may not be very closely related to the current situation because the average age is about 40 years of age but it is still valuable information.

3. Motivational Intensity
The understanding that foreign language is better studied every day or even more intensely is measured by this subscale. A typical example is, "I keep up to date with Spanish by working on it almost every day."

4. Spanish Course Anxiety
Anxiety is typically present in all types of study not only language study. It may have a negative connotation in the sense that it is better to have lower level of anxiety rather than high level. A typical example is, "It worries me that other students who study Spanish as much as I do, may speak Spanish better than I do."

5. Spanish Software Evaluation
This subscale is a validated measure for assessing language learning software. A typical example is, "My Spanish software is a great source of inspiration to me."

6. Attitudes toward Learning Spanish
Motivation can be obviously helped by a positive attitude toward learning a foreign language. A typical example is, "I really enjoy learning Spanish."

7. Attitudes toward Spanish-speaking people
This is a measurement of one of the "non-linguistic" goals. It may not be directly related to the ability to speak a foreign language but it is a very important component of the motivation. A typical example is, "I wish I could have many native Spanish speaking friends."

8. Integrative Orientation
Another "non-linguistic" goal that is very important for the overall characteristic of the learning process. Typical example is, "Studying Spanish is important because I will be able to interact more easily with speakers of Spanish."

9. Desire to Learn Spanish
A typical example is, "I have a strong desire to know all aspects of Spanish."

3 - R. C. Gardner, The Attitude/Motivation Test Battery: Technical Report (1985), University of Western Ontario.

10. Spanish Course Evaluation
This is another validated measure that helps with the overall evaluation of the foreign language learning process. A typical example is, "I enjoy the activities of my Spanish course much more than those of other courses I had before."

11. Spanish Use Anxiety
Anxiety is expected in the learning process but it would be preferable for a student to have a lower level of anxiety when using the foreign language. A typical example is, "I would get nervous if I had to speak Spanish to a tourist."

12. Instrumental Orientation
The reason behind the desire to learn a foreign language is very relevant to the level of motivation of the person. A typical example is, "Studying Spanish is important because I will need it for my career."

The three sets of scales AMS, IMI, and AMTB although related are used for slightly different purposes. AMS is used to evaluate the level of motivation to study foreign language before the actual study begins and is repeated after the study is completed. IMI is focused on the intrinsic and extrinsic linguistic goals motivation after the study is completed. It is specifically developed to evaluate the post-experimental motivation. AMTB measures a wide variety of non- linguistic goals motivation and attitudes.

Part 2. Sample Description

INITIAL SAMPLE

The full initial sample for this study (N=203) was predominantly female (67.1%) and the racial decomposition was Black/African American (24.6%), White (67.3%) and Other (8.0%). The majority of participants (82.3%) had a B.A. degree or higher. The average age was 41.1 years with the youngest being 19 years old and the oldest being 65 years old.

Part of the sample (40.4%) had some knowledge of Spanish and 11.8% of the sample had spouse who spoke Spanish. A third (36.5%) of the sample knew another foreign language. Only two people (1%) reported having parents, grandparents or great grandparents who speak Spanish.

The reasons for studying Spanish were diverse: work (33.0%), travel (15.3%), general interest (42.4%), and other reasons (9.4%). Some of the other reasons were:

"All of the above."

"Many Spanish people around me."

"To help my kids learn Spanish."

"To communicate with son-in-law and his family."

"Relocation."

"Enhance communication with students at my job and residents in my neighborhood."

"It's the second language spoken in the US. The need to learn it has become evident."

"To communicate with my neighbors better."

"I live in a large Spanish-speaking community."

"Daughter married Costa Rican."

Most of the respondents worked full time (71.9%) and their household income was significantly larger than the nation's median, with 76.4% of them having more than $50,000 annual household income.

There were 203 people who started this study. Of them, 39 dropped out. The dropout rate was 19.2% (39 out of 203) and the final sample for the full analysis was N=164. There were no statistically significant differences between the people who dropped out and the people who completed the study on the major demographic characteristics: gender, race, age, initial Spanish ability score, and initial motivation scores. The only significant difference (p=.018) was on education with less educated people being more likely to drop out of the study.

INITIAL EVALUATION OF THE SPANISH LANGUAGE SKILLS

The initial WebCAPE test before the study began, showed expected results. The average WebCAPE score was 104.8 points from a known maximum of 956 points. Almost a third (30.5%) scored zero on the initial test. These scores are well below the threshold level for semester two of Spanish.

Part 3. Analysis

3.1. SCALE RELIABILITY

We tested the internal reliability of the scales included in this analysis using the Cronbach's Alpha coefficient (Cronbach, 1951). Usually, a scale with Alpha close to 0.7 or above is considered reliable. The results of the reliability tests are presented to the right and below.

According to the Cronbach's Alpha all but one of the subscales had good internal reliability. Only one subscale with a Cronbach's Alpha of .588 did not reach the desired level. This was the Post- "Amotivation" subscale of AMS. For that reason the amotivation subscale will be excluded from analysis. This exclusion is not a problem for the analysis because it is based on the individual scales which are not affected by the presence or absence of one scale.

Table 1. Reliability of pre- and post level of motivation

Academic Motivation Scale (AMS)	Cronbach's Alpha	
	Pre AMS	Upper Limit
1. Intrinsic motivation - to know	.829	.819
2. Intrinsic motivation - toward accomplishment	.908	.916
3. Intrinsic motivation - to experience stimulation	.851	.854
4. Extrinsic motivation - identified	.884	.924
5. Extrinsic motivation - introjected	.876	.880
6. Extrinsic motivation - external regulation	.856	.900
7. Amotivation	.764	.588

Table 2. Reliability of Post AMTB

Attitude/Motivation Test Battery (AMTB)	Cronbach's Alpha
1. Interest in Foreign Languages	.674
2. Parental Encouragement	.953
3. Motivational Intensity	.687
4. Spanish Course Anxiety	.780
5. Spanish Software Evaluation	.860
6. Attitudes toward Learning Spanish	.830
7. Attitudes toward Spanish-speaking people	.819
8. Integrative Orientation	.773
9. Desire to Learn Spanish	.773
10. Spanish Course Evaluation	.839
11. Spanish Use Anxiety	.860
12. Instrumental Orientation	.716

Table 3. Reliability of Post IMI

Intrinsic Motivation Inventory (IMI)	Cronbach's Alpha
1. Interest/Enjoyment	.896
2. Perceived Competence	.859
3. Effort/Importance	.865
4. Pressure/Tension	.713
5. Perceived Choice	.653
6. Value/Usefulness	.850

3.2. MOTIVATION AND ATTITUDE

PRE- AND POST MOTIVATION

The AMS scale was used to determine the level of motivation before and after the completion of the study (pre- and post). The amotivation subscale did not reach the desired level of reliability and is excluded from the analysis. The results are presented below.

Table 4. Pre- and Post AMS Results

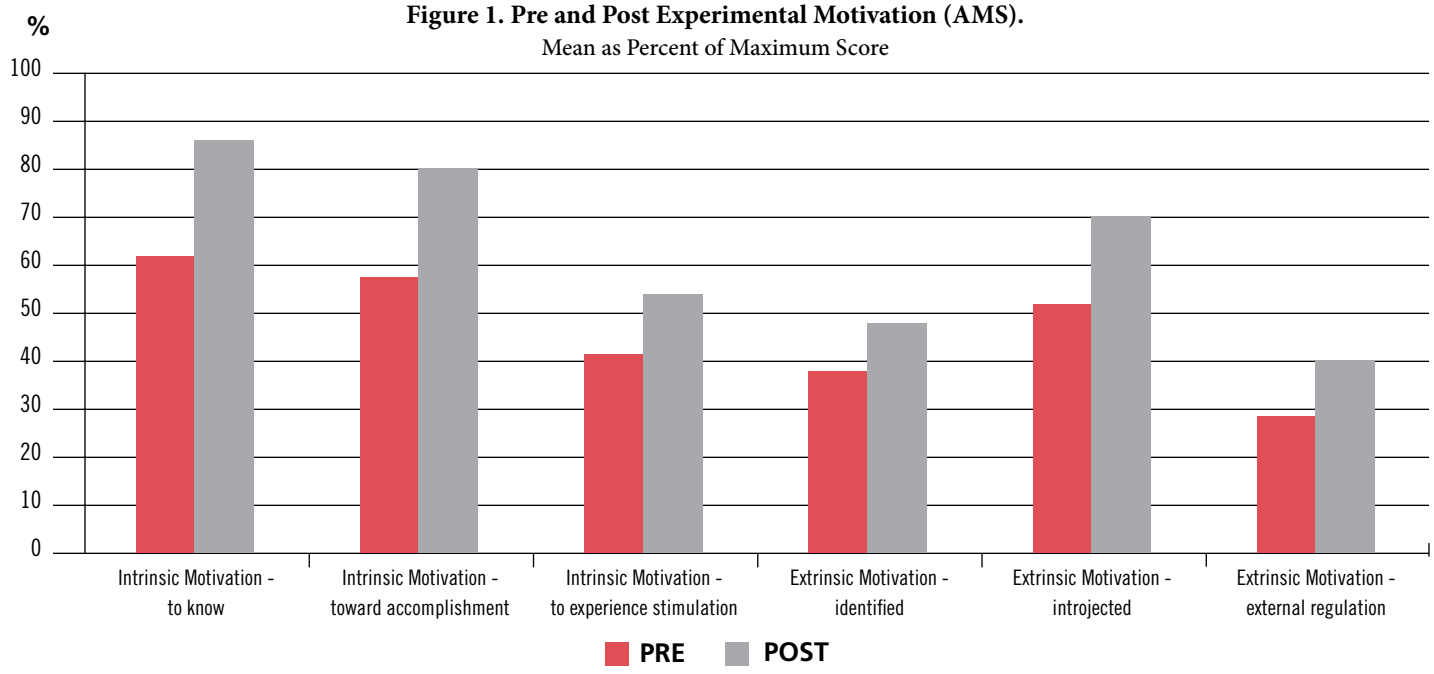
Academic Motivation Scale (AMS)	Pre			Post		
	Mean	Std Error*	Mean as % of max	Mean	Std Error*	Mean as % of max
1. Intrinsic motivation - to know	3.8	0.07	63.3	5.3	0.09	88.3
2. Intrinsic motivation - toward accomplishment	3.6	0.09	60.0	4.9	0.11	81.7
3. Intrinsic motivation - to experience stimulation	2.6	0.09	43.3	3.3	0.12	55.0
4. Extrinsic motivation – identified	2.4	0.09	40.0	3.0	0.13	50.0
5. Extrinsic motivation – introjected	3.2	0.09	53.3	4.3	0.11	71.7
6. Extrinsic motivation - external regulation	1.8	0.07	30.0	2.5	0.11	41.7
Average 1-6	2.9	na		3.9	na	

* Stratified estimates

The initial level of motivation (average of 2.9) was below the midpoint of the 1 thru 7 Likert scale. The highest level was for the IM-to-know and the lowest for the EM-external regulation. The post-experimental level of motivation (average of 3.9) increased for all subscales on average by 1 point (or 14.3% of the maximum). The biggest increase was recorded for the IM-to-know, a 1.4 point increase (or 20% of the maximum). All of the increases measured by the difference score (post-pre score) were statistically significant at a 5% type I error rate (α).

Figure 1. Pre and Post Experimental Motivation (AMS).

Mean as Percent of Maximum Score



Also important is the fact that the average level of 3 of the 6 subscales moved beyond the midpoint of the scale (4 points) and thus got closer to the maximum score of 7 ("Corresponds Exactly"). These three subscales are very important characteristics of the motivation: IM-to-know, IM-toward accomplishment, and EM-introjected.

All elements of both intrinsic and extrinsic motivation did increase and the increase was statistically significant. This is a very important result. The participants were exposed to the RS language learning software for only one month and their average level of motivation increased by one point (or 14.3% of the maximum) on a scale of 1 to 7.

All the subscales except Pressure/Tension can be interpreted in a positive direction. In other words the higher the score of the subscales, the better the motivation and attitude of the respondent. For example, the higher the score of Interest/Enjoyment subscale, i.e. the higher the level of interest and enjoyment, the better the results are.

The Pressure/Tension subscale should be interpreted in the opposite direction; the lower the score, i.e. the lower the level of pressure/tension, the better the result is.

IMI has a 7 point scale (1-7) with a midpoint of 4 (neutral). All the results are above the midpoint and some are very close to the maximum score of 7. The highest score is for the perceived choice, followed by Interest/Enjoyment and Value/Usefulness. On average, the score was 5.6 (or 80% of the maximum). This is a very definitive result. After a one month study with the language learning software the level of post experimental motivation is very high and close to the maximum for the positive subscales and very low (2.6 average score) for the negative ones (Pressure/Tension). The 95% confidence intervals show that on average these levels will not fall below the midpoint of the scale (4 points).

Table 5. Pre- and Post Experimental Motivation Confidence Intervals (CI)

Academic Motivation Scale (AMS)	Difference score (DS)	Std Error	95% CI*
1. Intrinsic motivation - to know	1.4	0.07	(1.3-1.6)
2. Intrinsic motivation - toward accomplishment	1.3	0.09	(1.1-1.5)
3. Intrinsic motivation - to experience stimulation	0.7	0.09	(0.5-0.9)
4. Extrinsic motivation - identified	0.6	0.09	(0.5-0.8)
5. Extrinsic motivation - introjected	1.1	0.09	(1.0-1.3)
6. Extrinsic motivation - external regulation	0.6	0.08	(0.5-0.8)
Average 1-6	1.0	na	na

* Stratified estimates

POST EXPERIMENTAL LINGUISTIC GOALS MOTIVATION

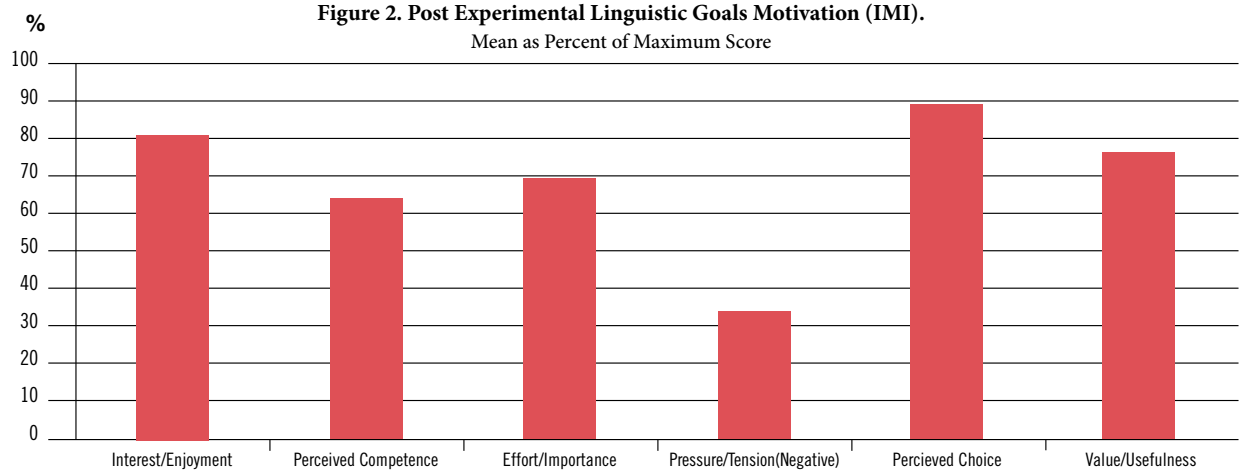
The level of post experimental motivation was measured by the IMI inventory.

Table 6. Level of Post Experimental IMI

Academic Motivation Scale (AMS)	Mean	Std Error*	95% CI*	Mean as % of Max
1. Interest/Enjoyment	5.9	0.07	5.8-6.1	84.3
2. Perceived Competence	4.7	0.09	4.6-4.9	67.1
3. Effort/Importance	5.1	0.10	4.9-5.3	72.9
4. Pressure/Tension (negative)	2.6	0.09	2.4-2.8	37.1
5. Perceived Choice	6.5	0.05	6.4-6.6	92.9
6. Value/Usefulness	5.7	0.08	5.6-5.9	81.4
Average 1-3,5,6	5.6	na	na	80.0

* Stratified estimates

Figure 2. Post Experimental Linguistic Goals Motivation (IMI).



POST EXPERIMENTAL NON-LINGUISTIC GOALS MOTIVATION

The level of the motivation and attitude related to the non-linguistic goals was measured by the AMTB scale.

AMTB works with a 6 point scale (1-6) with no neutral point so a score equal to or below 3 (Strongly Disagree to Slightly Disagree) will be interpreted as a negative finding, or lack of motivation for a particular subscale. And, accordingly, scores above 3 (Slightly Agree to Strongly Agree) will be interpreted favorably or as presence of motivation.

Overall none of the 12 subscales averages was below 3, i.e. overall all post linguistic motivation average levels were on the "Agree" side which by itself is a very important fact. The 95% confidence intervals also reveal that 11 out of 12 subscales, lower confidence limits will be above the "Disagree" level of 3 points. The only exception is "Parental Encouragement" for which the average level is also low (3.1 points). What we can conjecture is that on average, the parents of our participants when the latter were in school were divided between encouraging and not encouraging their children to learn foreign languages.

Table 6. Level of Post Experimental IMI

Academic Motivation Scale (AMS)	Mean	Std Error*	95% CI*	Mean as % of Max
1. Interest in Foreign Languages	5.3	0.04	5.2-5.4	88.3
2. Parental Encouragement	3.1	0.12	2.9-3.4	51.7
3. Motivational Intensity	4.8	0.05	4.7-4.9	80.0
4. Spanish Course Anxiety	3.2	0.06	3.1-3.3	53.3
5. Spanish Software Evaluation	5.3	0.05	5.2-5.4	88.3
6. Attitudes toward Learning Spanish	5.4	0.04	5.3-5.5	90.0
7. Attitudes toward Spanish-speaking people	4.5	0.06	4.4-4.6	75.0
8. Integrative Orientation	5.3	0.05	5.2-5.4	88.3
9. Desire to Learn Spanish	5.3	0.04	5.2-5.4	88.3
10. Spanish Course Evaluation	5.2	0.04	5.1-5.3	86.7
11. Spanish Use Anxiety	3.5	0.07	3.3-3.6	58.3
12. Instrumental Orientation	3.7	0.09	3.5-3.8	61.7

* Stratified estimates

LANGUAGE LEARNING SOFTWARE EVALUATION

Two of the AMTB subscales are geared toward evaluation of the learning process and the learning tool. These two scales are validated instruments and their reliability was again evaluated in this study and found to be very high (Cronbach's Alpha of 0.860 and 0.839 respectively).

Rosetta Stone® Spanish software was evaluated very favorably by its users. The average level of the subscale was 5.3 out of 6 (or 88.3% of the maximum) with a 95% CI of (5.2-5.4). This is one of the highest scores of all 12 subscales. The score of 5.3 is between 5 "Moderately Agree" and 6 "Strongly Agree".

SPANISH COURSE EVALUATION

This subscale includes items like, "I like my Spanish course so much, I look forward to studying more Spanish in the future."

The average score was 5.2 (or 86.7% of the maximum) or between 5 "Moderately Agree" and 6 "Strongly Agree". This is a very favorable opinion particularly given the relatively short time (one month) the participants spent studying Spanish.

Figure 3. Post Experimental Non-Linguistic Goals Motivation (AMTB) Scale 1-6.

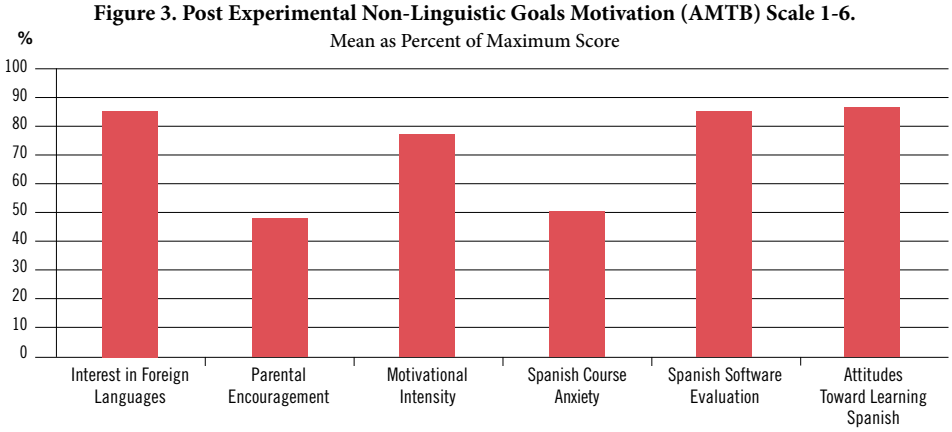
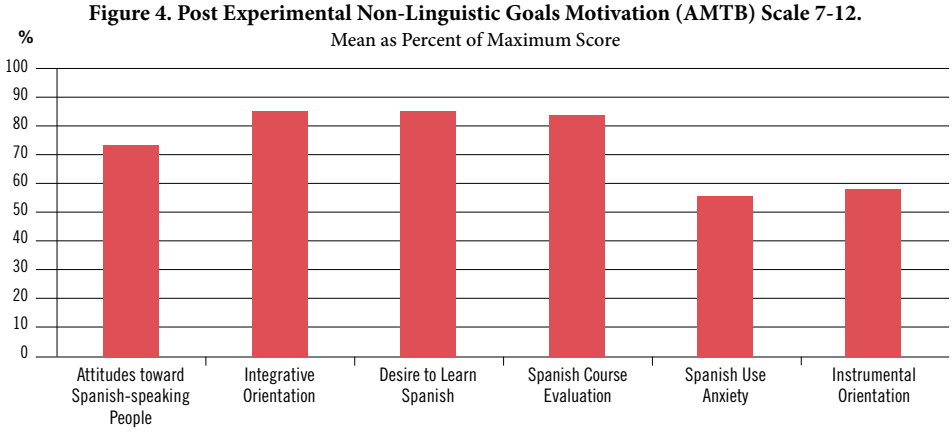


Figure 4. Post Experimental Non-Linguistic Goals Motivation (AMTB) Scale 7-12.



3.3. ADDITIONAL EVALUATION

In addition to the motivation scales the participants were asked a series of single questions which did not constitute a scale. The purpose was to extract additional information about their evaluation of the software.

EXPERIENCE

After using Rosetta Stone Spanish software for one month all participants were asked to evaluate 5 statements about their experience with Rosetta Stone Spanish with possible answers ranging from 1 ("Strongly Disagree") to 5 ("Strongly Agree").

Q1 (Easy). Rosetta Stone Spanish is easy to use.

Q2 (Helpful). Rosetta Stone Spanish is helpful in teaching me the language.

Q3 (Enjoyed). I enjoyed learning Spanish with Rosetta Stone.

Q4 (Satisfied). I am satisfied with Rosetta Stone Spanish.

Q5 (Recommend). I would recommend Rosetta Stones software to others who are interested in learning Spanish.

These results are very convincing that RS Spanish software is extremely easy to use, very helpful, satisfying, and enjoyable to work with. Finally, 90.2% of the users of the RS software will recommend it to others who are interested in learning Spanish.

Table 9. Software Evaluation

Scale	Easy	Helpful	Enjoyed	Satisfied	Recommend %
Strongly Disagree	2.4	1.8	1.8	1.8	1.8
Disagree	1.2	0	0.6	1.2	0.6
Neutral	4.3	3.7	6.1	11.0	7.3
Agree	43.9	46.3	33.5	43.9	36.6
Strongly Agree	48.2	48.2	57.9	42.1	50.6

If we consolidate the "Agree" and "Strongly Agree" categories, we get the following results.

Table 10. Software Evaluation

Rosetta Stone Software	Percent "Agree" or "Strongly Agree"
Easy	92.1
Helpful	94.5
Enjoyed	91.5
Satisfied	86.0
Recommend	90.2

Conclusion

The instruments and multidimensional scales used in this study of motivation and attitude worked remarkably well. We were able to register and distinguish between the levels of motivation and the change in the motivation levels even for the short period of the study.

The study had enough statistical power to test for significance and build the confidence intervals around the point estimates. Using a stratified sample allowed us to consider users of different Spanish ability: low, medium and high ability measured by the WebCAPE test. The propensity scores approach alleviated the influence of the demographics and other individual factors.

The main results of study were very encouraging and robust. It was very important to find out that the level of intrinsic and extrinsic motivation for studying foreign language increased during the study and this increase was statistically significant. All six of the Academic Motivation Scale (AMS) subscales registered a statistically significant increase at the end of the study. The average increase was about one point on a seven point scale, or 14.2% of the maximum score.

The level of the post experimental linguistic goals motivation was very high with an average of 5.6 on a 7 point scale, or 80% of the maximum score. The average scores for all five positive subscales of the Intrinsic Motivation Inventory (IMI) were above the middle point of the scale. The negative subscale had a very low score. The highest level of the post experimental non-linguistic goals motivation was 5.4 on a 6 point scale, or 90% of the maximum score. These high levels were for "Attitude toward learning Spanish", "Interest in foreign languages", "Desire to learn Spanish," and "Integrative orientation." The "Spanish course anxiety" and "Spanish use anxiety" scores were relatively low. Low level of anxiety is a very desirable feature of language learning software.

Two of the AMTB subscales were used to evaluate the RS software and the process of learning Spanish with RS. Users evaluated the software very favorably with average score of 5.3 on a 6 point scale (or 88.3% of the maximum score). The one month language course was also highly valued by the participants. The average score of this subscale was 5.2 on a 6 point scale or 86.7% of the maximum score.

About 90% of the users considered the software to be easy, helpful, enjoyable, and satisfying. The majority of them (90.2%) stated that they would recommend the software to others.

The overall conclusion based on this study is that Rosetta Stone® users were able to maintain and increase all aspects of their foreign language intrinsic and extrinsic motivation. In addition, their post experimental levels of linguistic goals motivation and non-linguistic goals motivation were very high. The participants had very favorable views of the software and the learning process.

References

Cochran, W. G. (1977), Sampling Techniques, Third Edition, New York: John Wiley & Sons.

Cronbach L.J. (1951). Coefficient alpha and the internal structure of tests, Psychometrika 1951;16:297-333.

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. (1994). Facilitating internalization: The self-determination theory perspective. Journal of Personality, 62, 119-142.

Fairchild, A.J., S.J. Horst, S.J. Finney, K.E. Barron, Evaluating New and Existing Validity Evidence for the Academic Motivation Scale, James Madison University.

Gardner, R. C. & Tennant, J. (1998, May). Attitudes, motivation and anxiety in learning French: Their relation to classroom experiences and achievement. Paper presented at the Trends/Tendances 98 conference, Carleton University, Ottawa, Canada.

Gardner, R. C. (1985). Social psychology and second language learning: The role of attitudes and motivation. London: Edward Arnold.

Gardner, R. C. (1996). Motivation and second language acquisition: Perspectives. Revue de l'ACLA/Journal of the CAAL, 18, 19-42.

Gardner, R. C. (2000). Correlation, causation, motivation and second language acquisition. Canadian Psychology, 41, 10-24.

Gardner, R. C., & MacIntyre, P. D. (1993). On the measurement of affective variables in second language learning. Language Learning, 43, 157-194.

Gardner, R. C., & Tremblay, P. F. (1994). On motivation, research agendas, and theoretical frameworks. The Modern Language Journal, 78, 359-368.

Gardner, R. C., & Tremblay, P. F. (1995). On motivation: Measurement and conceptual considerations. The Modern Language Journal, 78, 524-527.

Gardner, R. C., Lalonde, R. N., & Moorcroft, R. (1985). The role of attitudes and motivation in second language learning: Correlational and experimental considerations. Language Learning, 35, 207-227.

Gardner, R.C., & Tremblay, P. F. (1998). Specificity of affective variables and the trait/state conceptualization of motivation in second language acquisition. In R. Agnihotri, A. L. Khanna, & I. Sachdev (Eds.), Social psychological perspectives on second language learning (pp. 31-52). New Delhi, India: Sage.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. Research Quarterly for Exercise and Sport, 60, 48-58.

Plant, R. W., & Ryan, R. M. (1985). Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: An investigation of internally-controlling styles. Journal of Personality, 53, 435-449.

Ratelle, C.F., F. Guay, R. J. Vallerand, Autonomous, Controlled, and Amotivated Types of Academic Motivation: A Person-Oriented Analysis, (2007), Journal of Educational Psychology, 2007, Vol. 99, No. 4, 734–746

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. Journal of Personality and Social Psychology, 43, 450-461.

Ryan, R. M., Connell, J. P., & Plant, R. W. (1990). Emotions in non-directed text learning. Learning and Individual Differences, 2, 1-17.

Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Varied forms of persistence: When free-choice behavior is not intrinsically motivated. Motivation and Emotion, 15, 185-205.

Ryan, R. M., Mims, V., & Koestner, R. (1983). Relation of reward contingency and interpersonal context to intrinsic motivation: A review and test using cognitive evaluation theory. Journal of Personality and Social Psychology, 45, 736-750.

SAS Help and Documentation, SAS Institute Inc., Cary, NC: SAS Institute Inc., 2002-2004.

SAS OnlineDoc, SAS Institute Inc., , Cary, NC: SAS Institute Inc., 2002-2005.

Tsigilis, N., & Theodosiou, A. (2003). Temporal stability of the Intrinsic Motivation Inventory. Perceptual and Motor Skills, 97, 271-280.

Vallerand, R.J., Pelletier, L.G., Blais, M.R, Brière, N.M., Senécal, C., & Vallières, E.F. (1992). The academic motivation scale: a measure of intrinsic, extrinsic, and amotivation in education. Educational and Psychological Measurement, 52, 1003-1017.

Vallerand, R.J., Pelletier, L.G., Blais, M.R, Brière, N.M., Senécal, C., & Vallières, E.F. (1993). On the assessment of intrinsic, extrinsic and amotivation in education: Evidence on the concurrent and construct validity of the academic motivation scale. Educational and Psychological Measurement, 53, 159-172.