III METHODOLOGY

This chapter informs the reader about the methodology used in the study. It introduces first the institution in which the study took place, gives a short explanation about the different programs they use and which of them were relevant for the study, and then let the reader know how the participants were selected. A next subsection explains in details the procedure of the testing: which tests have been chosen, where they come from and how the participants were assessed. The following subsection gives information about the coding of the tests (how they were checked and how the points were given) and the last one informs the reader about the pilot that has been conducted before the actual study started: it gives a short overview of the results obtained and if the materials needed an adaptation for the actual study.

3.1 Place of the investigation and description of the institution' programs

The part of the research concerning collecting data took place in a bilingual school, located in the area of the city of Puebla, in Mexico. For the investigation, only children in the primary school were considered. The school offers different programs according to the German level of the pupil and in each primary grade, there are five groups with different levels of German. Table 1 below indicates the division of the groups within one grade. The division of the groups is the same from the first to the sixth grade. In each grade, there are three different programs: the first one for the native Spanish speakers, the second one for the bilingual pupils and the last one for the native German speakers.

Name of the group	Number of groups within one level	Particularity of the group	Students native language in the group
DaF-group (*Deutsch als Fremdsprache:* German as a foreign language)	3	Spanish native speakers with German as a foreign language (7 hours/ week)	Spanish
DFU-group (*Deutscher Fachunterricht:* Content classes in German)	1	Bilingual students with content classes in German (18 hours/week)	• Spanish and German (Spanish as the home language and German as the educational language)
DM-group (*Deutsch als Muttersprache:* German as the native language)	1	German native speakers with content classes in German only and Spanish as a second language	• German

Table 1: Description of the groups within one level

3.2 Participants

The students who participated in the research are children from the DaF-group (monolingual) being Spanish native speakers learning German as a foreign language and pupils from the DFU-group (bilingual) who speak Spanish at home and German at school. For the investigation, the importance of the children of the bilingual group is that they are bilingual, irrespective of whether they acquired both languages at the same time or one after the other. The participants are 20 pupils of each group: monolingual and bilingual. In total, 40 children contributed to the research. The two groups were as shown in table 2:

	Group	Number of participants	Number of girls within the participants	Number of boys within the participants	Spoken language(s)
1. Monolingual group	DaF	20	10	10	Spanish
2. Bilingual group	DFU	20	10	10	Spanish (home language) German (education language)

Table 2: Participants

The participants are between 10 and 12 years old. Both groups have more than 20 pupils, so the pupils have been randomly selected according to the Systematic Random Sampling procedure (Trochim, 2006) designed at Cornell University in the Center for Social Research Methods. All possible participants were first divided into the four initial groups (N = entire population within one group between 10 and 12 years old at the bilingual institution): monolingual girls, monolingual boys, bilingual girls and bilingual boys. Then, on each of the four lists, the pupils received a number, starting with 1 and going on chronologically until each student had one number. Ten participants of each list were needed (n = 10). For example, there were 29 bilingual girls between 10 and 12 years old at the bilingual institution and only 10 bilingual girls were needed for the study. The question is now how to proceed to select them randomly? The interval size, k, is N/n and is needed to be able to count the intervals between each participant, starting on a randomly selected number between 1 and k. Every k^{th} pupil has been chosen. The same procedure was repeated for each of the four groups. N was in each group a different number (according to the entire population of each group

between 10 and 12 years old available at the institution) whereas n = 10 has been used for each group (10 participants were needed in each group). The total of the participants reached 40 pupils.

The example below shows how the bilingual girls had been chosen, following the steps of the Systematic Random Sampling (Trochim, 2006). The pupils that already participated in the pilot were taken out of the population.

Selection of the bilingual girls for the study:

N = 29 (number of bilingual girls between 10 and 12 years old at the bilingual institution)

n = 10 (number of bilingual girls needed for the study)

k = N/n = 29/10 = 2.9 → 2.9 has been rounded up to 3.0

Randomly selected starting number between 1 and k: 3

Each third sample has been selected starting with the third one. Once the complete list has been used, the count goes back to the first sample again until *n* is full.

Chosen samples: 3, 6, 9, 12, 15, 18, 21, 24, 27, 1.

Each participant kept his anonymity in the study.

3.3 Procedure for the realization of the tests

Each group performed four cognitive tests. The explanation of the tests was given in German for the bilinguals and in Spanish for the monolinguals. The questions of the test 3 were originally in Spanish and had been translated into German by a professional translator and the questions of the test 4 were originally in English and had been translated into German and Spanish by the same professional translator. Tests 1 and 2 were taken on the computer and could not be copied and thus are

not in the appendix of this thesis and the tests 3 and 4 were printed on paper (see appendix 1 – 4).

The choice of the languages for the written tests had been made according to the language of education of the groups (language in which the students take the mathematic class). The monolingual group was taught mathematics in Spanish whereas the bilingual group was taught in German. The languages of education needed to be kept in order for the results to be compared. The pupils took the tests in the same language used in class. Table 3 indicates the cognitive skills and the test used to evaluate these skills of the participants.

	Cognitive skill	**test**
1.	velocity	spatial speed test
2.	memory	visual memory → pattern recognition
3.	logic	logical thinking test
4.	abstraction	abstract reasoning test

Table 3: Cognitive skills and their tests

I had four reasons to have selected these skills for my study and not others. The first reason is that they involve the most important cognitive skills and second, because they are all being used in other subjects than mathematics. For example, velocity is needed for languages or music, memory skills for history or geography, logical thinking for sciences such as chemistry and abstraction for arts. A further reason was to choose cognitive skills for which adequate tests have been designed to assess the skill. There are plenty of tests that can be found on the internet, but for this study, it was important to select tests that had been designed by linguists,

professors or neuroscientists. The last reason was the limitation of time. I needed to take into account the time I had to conduct this investigation and according to it, I limited the choice for these four cognitive skills.

The participants performed only one test per day in a classroom at the bilingual institution. The first group (monolingual) was asked to enter, one child at the time, and sat in front of the computer (for tests 1 and 2) where the participant received the instructions. Only after the child had completed the task, did the next child come into the room. The participants were not able to see each other during the procedure. They were waiting in a separate room before they took the test. The same procedure was repeated with the second group (bilingual) on the same day for the same test. On the next day, they all completed test 2, repeating the very same procedure. For tests 3 and 4, the participants of each group were in the same room, one child per desk. Each group (monolingual and bilingual) took the test separately and the same instructions have been given to each group, respecting their language of instruction. Both groups took the same test on the same day.

3.3.1 Test 1: Spatial speed test

For the Spatial Speed Test, the student saw one geometric figure, had to remember it (see figure 3) and then saw another figure and had to choose between two options: if the figure is the same as the one before or if it is different. He had to perform the task as fast as possible. This test was created by Luminosity Lumos Labs (2009), a laboratory in California specialized in brain functions. Lumos Labs is

a cognitive neuroscience research and development company that builds software tools for improving brain health and performance. The software was created by a team of nine neuroscientists from Stanford University[1]. This test investigates the speed of identifying geometrical figures in spatial setting according to the neuroscientists of Luminosity Lumos Labs. Figure 3 shows an example of one question of this test.

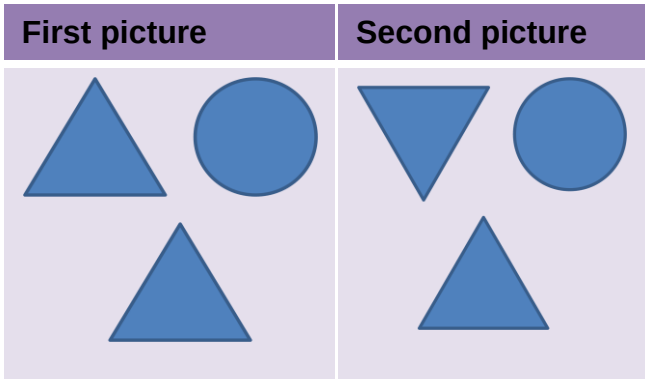

Figure 3: Sample of test 1

3.3.2 Test 2: Visual memory test

For the Visual Memory Test, the student had to memorize 12 figures, one at a time, as shown in figure 4 under memorization phase. After the 12 figures had been shown, the student had to recognize them by choosing between pairs (one correct and one wrong) until the 12 figures had been identified (see recognition phase in figure 4). The maximum amount of points for this test was 12. This test was also

[1] Moriah Thomason, Ph.D., Cris Niell, Ph.D., Russell Fernald, Ph.D., Michael Walker, Ph.D., Elizabeth Race, Ph.D., Jennifer Tsui, Ph.D., Elizabeth Buchen, M.S., Ph.D., Raag Airan, MD, Ph.D., Wesley C. Clapp, Ph.D.

created by Luminosity Lumos Labs (2009) and according to the neuroscientists who designed the test, it investigates the short-term visual memory.

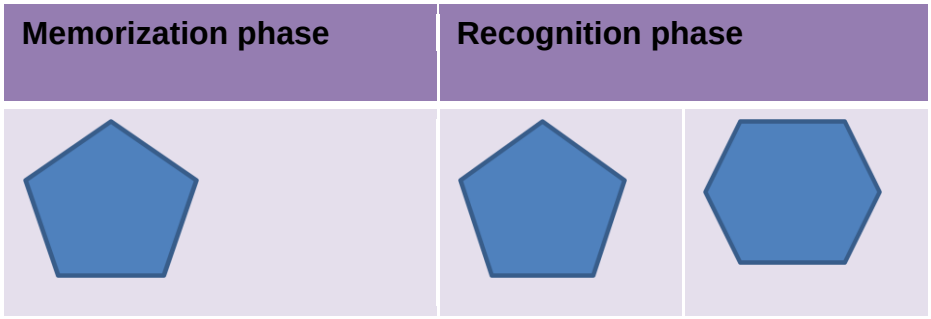

Figure 4: Sample of test 2

3.3.3 Test 3: Logical thinking test

In the Logical Thinking Test, the student answered 10 questions about logical thinking. The test was an adaptation of the TOLT test invented by Tobin & Capie (1980, cited in Trifone, 1987). After having chosen the answer, the student had to find an explanation from five multiple choices options. The maximum amount of points was 10. The original TOLT test could not be used because it is meant for adults but the idea of the TOLT test was respected where a question was asked and five possible answers were proposed. The content of the questions asked was kept, but the level of difficulty needed to be adapted to the mathematical knowledge of children between 10 and 12 years old, therefore the five questions were taken from the *ENLACE exam 2008* for fifth graders from the mathematical section which is a national written exam in Spanish performed in every primary school in Mexico. The purpose of the exam is to compare the academic level of the schools. The mathematical section tests logical thinking ability: the student needs to use common sense and logic to find out the answers; it is not an exam where

the content can be learned by heart previously. Yet none of the participants had taken this test previously because at the time it was applied, they were in the third and fourth grades. This is why the questions were taken from the 2008 exam, that is the exam that was applied to fifth graders two years ago. (Test 3 is in the appendix 1 in Spanish and 2 in German.)

3.3.4 Test 4: Abstract reasoning test

The Abstract Reasoning Test required the student to look at a series of symbols and find out which one completes the sequence. An example can be seen in figure 5. The student had to identify the relationship between the figures in order to find out which one was correct and fit to the schema. This test was created by Psychometric Success (2009) an organization that creates tests to improve students' cognitive skills, for example before applying for a job. According to the designers of Psychometric Success, the test has been created to find out the ability to analyze symbols on an abstract level. To perform the test, students need to picture themselves these symbols that follow a sequence. Without an abstract reasoning ability, the student would not be able to find the correct answer. (Test 4 can be seen in the appendix 3 in Spanish and 4 in German.)

Example:

1. Which symbol in the Answer Figure completes the sequence in the Problem Figure?

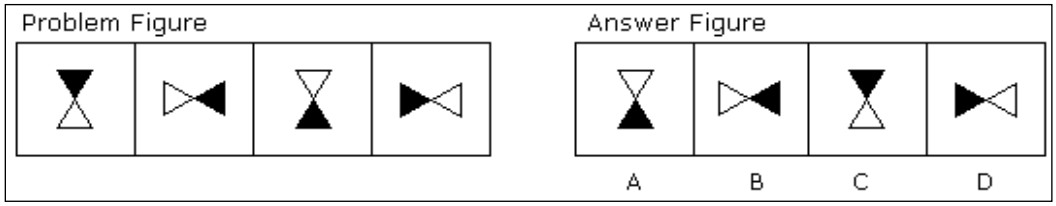

Figure 5: Sample of test 4

3.4 Coding of the tests

Each test was coded separately by the researcher. The results are generally organized and visualized in tables, according to the two main groups – monolingual and bilingual learners – and by gender. For each test, there are two main tables, one for the results in general with the percentages of correct answers (see table 4) and another one for the significance of the results (see table 5). The results are shown and explained in chapter 4.

3.4.1 Spatial speed test

The spatial speed test, as mentioned previously, tested the speed of dealing with geometric figures. The participant took the test on the computer, where the time was recorded in seconds. When the child finished the task, the final time was given. The time of each participant was filed and an average of each group was calculated. Each group (monolingual, bilingual, boys and girls) also received its own average. Table 4 shows how the averages are compared in chapter 4.

	boys	**girls**	**average**
monolinguals			
bilinguals			
average			

Table 4: Table for the results of the tests

A second table (see table 5) shows then if the results are significant or not for test 1. The significance has been calculated with the t-test in an Excel program. The t-test has the purpose to compare the means of two groups and shows if the results are significant or not, according to the critical value. If the t-score is greater

than the critical value, the results are significant. In the case that the t-score is smaller than the critical value, the results are not significant. In order to find out the t- score, the following formula was needed:

$$t = \frac{M1 - M2}{\sqrt{\frac{(SD1)^2}{N1} + \frac{(SD2)^2}{N2}}}$$

M1 = Mean of group 1
M2 = Mean of group 2
SD1 = Standard deviation of group 1
SD2 = Standard deviation of group 2
N1 = Number of subjects in group 1
N2 = Number of subjects in group 2

The same formula was used for each test. N1 and N2 are the same for all tests whereas M1, M2, SD1 and SD2 are different numbers in each test.

	Monolinguals	**Bilinguals**
Mean		
Standard deviation		
Variance		
t - score		
$\rho < .05$ (confidence level)	*df* = 39 (population)	1.697 (critical value)
Interpretation	The results are significant if the t-score is greater than the critical value.	

Table 5: Table for the significance of the tests

3.4.2 Visual memory test

The visual memory test was computerized and the score of the participant automatically appeared at the end of the test. The scores of the participants were calculated in an average so that they can be compared. In total, the test had four averages: one for monolingual girls, one for monolingual boys, one for bilingual girls and one for bilingual boys, as shown in table 4. The same table used for the results of the test 1 is kept as such for the results of the test 2 (see table 4). Also, a second table (see table 5) shows if the results of test 2 are significant or not. There was no time limit for this test.

3.4.3 Logical thinking test

The logical thinking test was a written test which was given in a classroom. Every participant of the same group took the test at the same time and when every student finished, the tests were recollected and corrected by the researcher. For each correct answer, the participant got one point and for each correct reason given, the learner also got one point. In total, the participant could get up to 10 points. The results of each participant of the same group (monolingual boys, monolingual girls, bilingual boys and bilingual girls) were put in an average. The results are represented in the same table, as shown in table 4 followed by the indication of the significance of the results of test 3 illustrated in a table identical to table 5. There was no time limit for this test: the participants handed in their test, once they were finished.

3.4.4 Abstract reasoning test

The abstract reasoning test was a written test. The same coding procedure for test 3 was repeated for test 4. Each participant got a maximum of 4 points, one for each correct answer. The tests were corrected by the researcher. The averages are represented in a table (see table 4) where the scores of each group can be compared and a second table (see table 5) indicates afterwards if the results of test 4 are significant or not. There was no time limit for test 3: the participants handed in their test once they finished answering it.

3.5 Pilot of the study

The materials (the four tests) were piloted in order to find out if they were adequate. The pilot study took place at the same private institution where the actual study was done.

3.5.1 Participants of the pilot study

The tests were piloted with 20 participants, half of the number of the total participants of the final study. They were randomly selected from the groups that participated in the study, described previously in table 1. The participants of the pilot were all between 10 and 12 years old. The participants of the pilot were eliminated from the lists so that they did not appear again as participants of the actual study. Table 6 shows how the participants selected for the pilot study were divided into four groups.

	Group	Number of participants	Number of girls within the participants	Number of boys within the participants	Spoken language(s)
1. Monolingual	DaF	10	5	5	Spanish

group					
2. Bilingual group	DFU	10	5	5	Spanish (home language) German (education language)

Table 6: Participants of the pilot

3.5.2 Results of the pilot study

Table 7 shows an overview of the most important aspects of the results of each test. The first column indicates if the bilingual group scored higher or not and, according to the literature discussed in chapter 2, it was expected that the bilingual students would score higher. The second column shows which group performed the highest on each test. The third column confirms if the test needs an adaptation according to the results obtained in the pilot. An adaptation is needed if for example the participants had difficulties understanding the instruction or the questions, or if the level of the content was not adequate for children between 10 and 12 years old. The pilot had the purpose to find out this kind of information. According to the results of the pilot, I personally took the decision of adapting the test or not. Finally the last column shows if the test needs to be piloted again.

test	**Bilinguals scored higher**	**Best group**	**Needed to be adapted**	**Needed to be piloted again**
velocity	yes	Bilingual boys	no	no
memory	no	Monolingual boys and bilingual boys	no	no
Logical thinking	yes	Bilingual boys	no	no

Abstract reasoning	yes	Bilingual boys	yes	yes

Table 7: Summary of the results

As illustrated in table 7, the bilingual group performed better in three out of four tests, so this means that in general they have better skills than monolingual learners and that they can apply them in mathematics. The fact that monolingual students have better memory skills might be the consequence of having too few participants since the average of the bilingual group lies just under the average of the monolingual group with a difference of 0.1.

Every group in general scored poorly on test 4. The cause could be that the test was too difficult for the age of the participants. Based on the fact that each group had a low score, this test was adapted for the thesis study. The first and the last tasks were kept the same as in the pilot but the second and the third ones were changed to have less difficult tasks. The adapted questions were taken from an IQ test written by McConochie[2] (1999), from the section *abstract reasoning*. The two questions taken from this test were, as well, testing the pattern recognition ability in abstract reasoning. A series of figures were given to the child who needed to find the correct figures that followed the first ones. He had to choose one out of four or five. The test was piloted again after the changes in order to see if the adaptation was adequate or not. The population was smaller as for the other pilot: 6 bilingual participants (3 boys and 3 girls) and 6 monolingual participants (as well

[2] Dr. William McConochie has a B.A. from Carleton College, a Master's in School Psychology and Ph.D. in Counseling/Clinical Psychology from IIT in Chicago, and a post-doc in Clinical Psychology from Northwestern University School of Medicine.

3 girls and 3 boys) which is a total of 12 participants for the second pilot of test 4. The participants of the first pilot could not be chosen again for the second pilot and the participants of the second pilot of test 4 could not take part of the actual study. The first reason to select only 12 participants was that I did not need a large population to see if test 4 was properly adapted and adequate for the age of the participants, meaning that the level of difficulty had decreased. The second reason was that I would not have enough bilingual participants between 10 and 12 years old for the actual study if I selected too many for the pilot study.

An interesting aspect that can be seen is the one in the second column of table 7 where bilingual boys always had the best scores. Once they performed as well as the monolingual boys (see table 7), but the fact that they always were the best leads to the conclusion that boys are better at mathematics than girls and that they are able to apply their cognitive skills more efficiently than girls. The gender variable, as mentioned earlier, was taken into consideration for the thesis study.

Tables 8 and 9 below indicate the results (the numbers indicate the averages of correct answers in the group) of the test 4, abstract reasoning test, before and after the changes.

Abstract reasoning test before the modifications:

	Boys	**Girls**	**average**
Monolinguals	2.0	1.6	**1.8**
bilinguals	2.2	2.0	**2.1**
average	**2.1**	**1.8**	

Table 8: Results of test 4 of the pilot before modifications

Abstract reasoning test after the modifications:

	Boys	Girls	average
Monolinguals	3.6	3.6	**3.6**
bilinguals	4.0	3.8	**3.9**
average	**3.8**	**3.7**	

Table 9: Results of test 4 of the pilot after modifications

All the participants scored higher after the modifications had been made. The test was kept as such for the study. Table 10 shows a summary of the results of test 4 after the changes.

test	Bilinguals scored higher	Best group	Needed to be adapted	Needed to be piloted again
Abstract reasoning	yes	Bilingual boys	no	no

Table 10: Summary of the results of test 4 after modifications

The next chapter shows the results of the four tests and gives an explanation for the results obtained. A statistical analysis calculated with the Excel program will define if the results were significant or not.