

Chapter 2

Evidence Research Study: A Methodology for L2 Research

Abstract This chapter presents the performance differences of non-native students from different schools. Four hypotheses were addressed in this research to understand the school influence as variable for L2 learners' performance. First, the school differences were differentiated by quantity and quality of resources (including the language assessments provided by schools); second, it is expected that when the effect of the type of school is controlled, that there is an improvement of performance in the language and cognitive tasks in Portuguese as a second language. Those tasks will be each described here and presented the adaptation and procedures involved for the empirical study. Third, the performance differences and the school as a covariate will be considered against the students' mother tongue, fourth, against the students' nationality and the immigrants parents' socioeconomic status. It also expected that the probable performance difference is always observed in all the same tests within the set of tests used. Therefore, two samples will be used in the study to respond to the four hypotheses advanced: students and teachers, both from the same schools (i); and eight tests were determined with an Exploratory Factor Analysis in order to understand the clusters involved for the skills evaluation: verbal reasoning, vocabulary and writing, recall tests and phonological manipulation.

Keywords Urban Schools • Verbal Reasoning • Recall • Phonological Manipulation • Nationality • Mother Tongue • Socioeconomic Status

Students' sample: To determine the learners' sample were selected 108 immigrant students from Portuguese schools with Portuguese as L2. Participants were aged between 7 and 18 years of age ($M = 13$; $SD = 2.7$). 46 (43%) male and 59 (55%) female, from different countries: 25 (23%) from China, 6 (6%) from Latin American countries, 31 (29%) from Eastern Europe, 19 (18%) from Portuguese-Speaking African Countries (PALOPs), 12 (11%) from Western Europe, 14 (13%) from other Asian countries (other than China). Speakers of 28 home languages (or first language—L1) distributed according to six language groups (determined by language family): 33 speakers of Mandarin, 32 speakers of

Romance languages, 14 speakers of Slavic languages, 11 speakers of Portuguese-based creole, 10 speakers of Indo-Aryan languages, and 2 speakers of Afro-Asian languages. Concerning the arrival date (and age of onset), 14 students arrived to Portugal between 2001 and 2005, 21 between 2006 and 2009, 55 between 2010 and 2014. 61% immigrated more recently (the study was initiated in 2013/2014). Students come from lower to middle socioeconomic backgrounds (education and occupation of families was computed, income was not identified). Participants were right-handed (laterality was identified) and attended 11 state schools within the same geographical area: Lisbon.

From the 108 sample 36 participants were chosen from four schools to match the schools from the second sample: teachers. Subjects were classified by schools according to the proficiency levels defined by the Common European Framework of Reference for Languages—CEFR (European Commission, 2001) and our sample includes only the first three groups: A1 (beginner level), A2 (elementary level), and B1 (threshold level). Only 23 students were effectively assessed in school.

This criterion was based on the school match (teachers and students from the same schools to provide the comparison between students' achievement, school resources, and teachers' perceptions). ANOVA tests were carried out to compare results according to the participants' school and in relation to several variables considered in the study hypotheses. The results were: $F(3,39) = 9.956$, $p = 0.000$ for the students' nationality; $F(3,39) = 3.050$, $p = 0.040$ for the students' First Language; $F(3,39) = 10.737$, $p = 0.000$ for parents' nationality; and $F(1,14) = 18.951$, $p = 0.001$ for proficiency level assessed by the school (school resources). No significant difference was found in the different socioeconomic groups.

Schools' sample: To classify the schools in order to determine their resources for immigrant students, the answers of 77 teachers, aged between 32 and 62 years ($M = 47$ years, $SD = 7.4$) were examined, of whom 11 (14.3%) were male and 60 (77.9%) female, with an average of 22 years teaching experience ($SD = 6.7$). They represent nine schools/groupings in the district of Lisbon. Only four schools matching students' schools were selected for this study. 58 (75.3%) have experience of multicultural classes and 16 (20.8%) have never had non-native students in their classes. Respecting the Language Testing measures: only 46 (59.7%) reported to having administered them at the beginning of the year.

From the 77 teachers' sample, 36 were selected from the same four students' school to compare school resources as the main variable for this study. In total there are 77 participants, teachers and students in equal numbers, and the four schools were divided by different municipalities in the district of Lisbon: two schools in the center and two others in the periphery. We assigned a number to each school based on its characteristics, which we obtained from univariate analysis of variance in order to compare schools regarding the students' age, school year according to the students' country of origin and the host country, host period, L1 instruction and proficiency assessment through school tests. The results were significant for students' age ($F(3,40) = 15.027$, $p = 0.000$), for the grade level ($F(3,40) = 21.176$, $p = 0.000$), for school grade obtained in the country of origin ($F(3,25) = 6.685$, $p = 0.002$), and for assessed proficiency ($F(1,14) = 4.846$, $p = 0.046$). No

significant difference was found in the SES, and the L1 parallel instruction was not applicable considering that these schools do not offer language support for L1-based instruction. As for the homogeneity of variance, the Levene's test revealed that the variances differ only in groups of assessed and non-assessed students (only two schools meet the requirement to conduct diagnostic assessment). Considering this result, in order to define the characteristics of the schools regarding the assessment variable (school resources), we used the Kruskal–Wallis test for independent samples since the ANOVA cannot be used in this context in which the parameters criteria are not respected (homogeneity of variances) for this type of statistical tests. Thus, we found that school 1, followed by school 3 and compared to the others, has older students (over 10 and under 17 years of age); students of schools 1 and 3 are in more advanced education levels, as opposed to school 4 that has the youngest immigrant students studying 1st cycle education. Naturally, students of schools 1 and 3 have more schooling completed in their countries of origin, unlike those in schools 2 and 4. However, school 1 has the lowest host period (students arrived in Portugal and to school 1 later) as opposed to school 3. None of the schools provides L1 support to students and only schools 1 and 2 conducted assessment tests. We did the independent sample Kruskal–Wallis test (nonparametric) on these two schools and found that there were differences between the two schools regarding the tests' results and the median comparison analysis shows that school 1 has higher proficiency students than students from school 2 ($p = 0.025$).

2.1 Instruments

2.1.1 *Students*

A 15-task test (with multiple items) was used to assess L2 learners' skills in Portuguese, using, for example, verbal reasoning, vocabulary, writing, word recall, and retelling tests. The tasks were developed and adapted based on literature review of available and recent tests in international repositories in the field of performance assessment in a foreign language and in second language. The recognized validity of the original tests from which the tests to test the Portuguese population were adapted was one of the criteria to ensure their suitability to groups of students according to low proficiency levels. On the other hand, we were interested in using the tests with several groups classified by the CEFR as low proficiency to check how their performance showed in our tests and comparing them to the levels (A1, A2, and B1) awarded to students by schools. It should be noted that all schools were duly informed about the proficiency levels when there was testing and classification to be given. One of the gaps noted was the absence of tests or speaking diagnostic assessment in schools for minority newcomers. These tests (of the present research study) were carefully assessed for internal consistency and correlation (between tests and related items), revealing appropriateness and validity for almost all tests. Some of the tests also indicated non-native students' level of lexical

knowledge in Portuguese, and we crossed the data collected in all tests and noticed the performance influence on vocabulary tests on the remaining tests, i.e., the more correct the vocabulary or word recall lists, the better the performance observed in different tasks.

The vocabulary chosen for each test followed a criterion found to be the most appropriate: vocabulary difficulty index according to the frequency and difficulty profiles already validated in Portugal for a lexical index—the Corlex (Bacelar, 2001). It should be noted that the proper use of this instrument relates to the need to detect inconsistencies in the tests' scores for each student and it should be able to

Table 2.1 Exploratory factorial analysis: all the tests of the battery administered to the L2 Portuguese learners

	Factor I	Factor II	Factor III	Factor IV	Factor V
	Verbal reasoning, writing and vocabulary	Recall	Oral comprehension	Phonological manipulation	Unfamiliar sounds
Measures					
Naming task	0.546				
Semantic associations	0.761				
Verbal analogy	0.477				
Extraction	0.729				
Vocabulary match	0.748				
Writing	0.688				
Cognates	0.758				
Metaphor language	0.724				
Syllable awareness				0.648	
Writing comprehension	0.776				
Reading recall		0.516			
Accent detection					0.802
Non-words					0.664
Conversion non-words	0.447				
Blending				0.693	
Retelling		0.840			
Words recall		0.804			
Oral comprehension			0.711		
Comprehensibility			0.679		
% of explained variance	0.36	0.87	0.8	0.7	0.65

identify more correct properties for future tests to be used in schools (Gándara, 2015). According to recent data collected by Edele, Seuring and Kristen (2015), we also followed the assumption that one should use real performance tests more than self-assessment often found in L2 research, which in the past decades has proven to be of limited validity (Brantmeier & Vanderplank, 2012). Moreover, the measures used are not always previously validated (Carter & Dunning, 2015), which we consider to be one of the greatest weaknesses of the existing Core Common Standards in Europe and elsewhere. These guiding documents favor mostly qualitative self-assessment tests that are far from informing about the actual performance of immigrant pupils. On the other hand, as Edele et al. indicate, samples of studies with self-assessment measures are too small (Ross, 1998) or the tests have a few number of points. One of the striking results of the fragility of the validity of the tests used is the huge correlation variability values (although positive) between the tests and/or items. This is one aspect that we noted in this study, which reports positive consistency far from the variability found in other studies. Another problem identified in the area of instruments also by Edele et al. results from the use of Foreign Language tests with L2 students who are different scenarios, because they are distinct populations. The variability of correlations will probably be larger and less valid. Yet another limitation pointed out by the above authors is the little heterogeneity of the samples regarding age. This is another aspect we have tried to keep in the students' sample as it encompasses a higher age range.

We conducted a series of univariate analyses of variance for all identified tasks, except for the tests excluded after the exploratory factorial analysis (Table 2.1), in order to identify performance differences among groups established according to mother tongue (L1; named also as first language or home language), nationality and the SES, but considering the influence of the school covariate. Differences were found between groups for the tests listed in this section:

2.1.2 *Picture Naming*

The picture naming test of the Diagnostic Test of Portuguese as Second or Foreign Language (Mateus, 2009) includes 36 pictures and has high internal consistency (0.94). Students are asked to name, in writing rather than orally, as in the original test, the pictures as they are sequentially shown over 5 pages. The test originally aims to identify the vocabulary skills of immigrant students in second or foreign language and in a noncomplex frequency level. This criterion was established according to CORLEX (Bacelar, 2001), the Portuguese index of vocabulary frequency and difficulty. The total score for this task is 12 points.

2.1.3 Semantic Associations

The 6-item semantic association test was adapted from the Woodcock-Munoz Language Survey-Revised (WMLS-R, 2005). The task adapted to Portuguese showed high internal consistency (0.86). Regarding the size of vocabulary, this test aims to assess students' ability to identify semantic relationships between words at random by completing the six items with the respective synonyms and antonyms of each word in Portuguese. The task is assessed as follows: 2 points for each correct answer (total score: 12 points). Example: Word: "rich _____ (synonym) _____ (antonym)."

2.1.4 Verbal Analogies

The 6-item verbal analogy test was adapted from *Verbal Analogies* (Test no. 2) by Woodcock-Munoz Language Survey-Revised (WMLS-R, 2005) and the Portuguese adapted version has a Cronbach's alpha of 0.60. The test aims, within verbal reasoning, to assess how the student, whose mother tongue is not Portuguese, completes six sentences based on vocabulary and the proposed analogies. The verbal analogy task aims to measure understanding of logical association within random phrase contexts. The test scores as follows: 1 point for each correct answer (total score: 6 points). Example 1: "Fill in the missing word by logic association: Star is to sky as fish is to _____."

2.1.5 Morphological Extraction

The morphological extraction test (or morphological change) has 4 items and was adapted from the *Morphological Extraction Test* by August, Kenyon, Malabonga et al. (2001). In this study, it has a low Cronbach's alpha of 0.53. The test aims to assess individuals' ability to make modifications and extractions from a derived word in a given sentence context. In the case of L2 learners, it is intended to further check, in terms of vocabulary and verbal reasoning, the influence of transfer between mother tongue and Second Language during the morphological change process. The change is analyzed in terms of morpheme conversion into new words in the dominant language. The task is calculated as follows: 2 points for each correct extraction, 1 point for partially correct extraction (total score: 8 points). Example: "Word: Friendship/Sentence: The classmates are my _____."

2.1.6 Vocabulary Match

The test aims to assess the lexical competence of individuals by choosing the correct answer from five groups of three options, and by completing a series of exercises whose complexity increases gradually per item. The choice of vocabulary followed the order of complexity and frequency content of the CORLEX (Bacelar, 2001). The CORLEX has the frequency and difficulty levels of the lexicon in European Portuguese. For this task, we had words of low and average levels of difficulty. The test has a Cronbach's alpha of 0.85. Among the options, several distractors using very similar words but with different semantics were used. The score was estimated from 0 to 3, and 3 represents over 9 correct answers in a total of 15 items. Example: "Match the appropriate synonym to the word "understand." " Correct answer: "realize."

2.1.7 Text Recall

The recall test was created to evaluate the attention and memory skills of non-native students after their reading of three short texts. They should remember in writing as many words as they can, as well as the events according to that order that they appear in the original texts previously read. The selected texts are unrelated and they are excerpts from authentic Portuguese books that appear in the contents Program of K-12 levels. Students were instructed on the test procedure: to read a sequence of three short texts in Portuguese; then to remember as many words they can and report in writing. For the scoring only the words and facts correctly recalled were positively considered (2 points distributed for word recalled and event correctly ordered).

2.1.8 Cognates Awareness Test

The Cognates Test integrates 5 items and was adapted from the *Cognates Awareness Test* (August et al., 2001). The Portuguese adapted version showed a Cronbach's alpha of 0.70 (consistent with the original test). In the verbal reasoning domain, this task examines the decoding ability based on transfer strategies of L2 learners considering a words list containing cognates (similar words in visual-phonological features across L1 and L2). Attending to the heterogeneity of language speakers, we expect better performances for romance language speakers (L1: Spanish, French, Catalan) considering that linguistically the languages with origin in the same family share more similarities in vocabulary and semantics. On the other hand the cognates also are along with the 'false friends' (words despite visually similar are distinct in meaning) across approximate languages. The "false

friends” are the main obstacle for the cognates identification. Score: 1 point for each correct answer (total score: 5 points), 0 for incorrect answers.

2.1.9 Metaphor Comprehension

This test aims to measure the understanding that non-native students have of different idiomatic expressions (so particular to L2) and how they can write their reasoning in relation to each of four metaphors. It has a Cronbach’s alpha of 0.60. Students are requested to explain each item literally. The difficulty index of this test is high because it is figurative language, decoded in a language other than the mother tongue. Each metaphor is assigned 1 point if the answer is correct, 0 if it is incorrect (total score: 4 points).

2.1.10 Questionnaire to Identify Schools’ Resources

For teachers of the schools studied in this work, we used a questionnaire based on the Alberta Education indicators (2012, p. 2) to know the level of evidence of certain school practices and resources (*nonexistent, emerging or evident resources/practices*) in the specific area of support provided to L2 learners of Portuguese (and their families). Following the Alberta Education indicators, we adapted 10 items distributed by the following dimensions and in this order: differentiated practices by groups of individuals (L1), by students at different levels of proficiency, planning of the school’s staff in terms of support and definition of practices, support offered to groups with background educational deficit, physical or digital measures that stimulate immigrant students’ parental investment, physical materials for L2, and targeted support to L1 students (see Table 2.2). All items were answered on a scale of 1–3 with 1 being equivalent to “missing resource” and 3 “evident resource.” Regarding the teachers and school resources, we considered the schools that have implemented, or not, proficiency tests according to the Common European Framework Reference for Languages (European Commission, 2001).

Teachers and students were assessed between 2013 and 2016 in the same primary and high schools in the district of Lisbon. The 108 students were first examined, then the 77 teacher completed the questionnaire, further reduced to 36 to match the four schools contemplated in this study. After obtaining the informed consent and the demographic record of the selected school population, the 10-item questionnaire for the teachers was answered and assessed (using points) according to the original test. Only four schools were selected for the statistical analyses conducted in this study in order to match students and teachers’ schools.

Table 2.2 Existent resources (perceived by teachers) in each school concerning L2 learners' education

School	Different practices for learners	Explicit practices in classroom	Support for lower proficiency learners	Staff	Support for learners with educational deficit	Support to families	Didactic materials	Support for learners' L1	Software/technology resources
1	<i>M</i>	2.00	2.22	2.89	2.44	2.67	1.33	1.78	1.50
	<i>N</i>	9	9	9	9	9	9	8	8
	<i>SD</i>	0.707	0.972	0.333	0.726	0.707	0.707	0.667	0.707
	Median	2.00	3.00	3.00	3.00	3.00	1.00	2.00	2.00
2	<i>M</i>	1.50	2.00	2.00	2.00	2.00	1.50	1.50	1.50
	<i>N</i>	2	2	2	2	2	2	2	2
	<i>SD</i>	0.707	1.414	1.414	1.414	1.414	0.707	0.707	0.707
	Median	1.50	2.00	2.00	2.00	2.00	1.50	1.50	1.50
3	<i>M</i>	2.40	2.43	2.33	1.87	1.80	1.53	1.67	1.79
	<i>N</i>	15	14	15	15	15	15	15	14
	<i>SD</i>	0.632	0.756	0.617	0.834	0.862	0.743	0.724	0.802
	Median	2.00	3.00	2.00	2.00	2.00	1.00	2.00	2.00
4	<i>M</i>	2.14	2.13	2.00	1.50	2.14	1.29	1.75	1.38
	<i>N</i>	7	8	8	8	7	7	8	8
	<i>SD</i>	0.900	0.835	0.756	0.926	0.690	0.488	0.707	0.518
	Median	2.00	2.00	2.00	1.00	2.00	1.00	2.00	1.00

2.2 Data Analysis

Repeated analyses of univariate variance tests (effect size included) and covariance test (ANCOVA) were used (SPSS version 23) to identify whether there were significant differences among the four selected schools as regards the existence of resources and non-native student' performance, and how this difference would be reflected in the groups' performance established according to mother tongue, nationality and socioeconomic status. We also used the Structural Equation Model (SEM), using AMOS, to proceed with a Confirmatory Factor Analysis (CFA) to corroborate the univariate analyses' information: tests' scores are significantly impacted by the school variable and its resources for immigrant children. The hypotheses of Part I of this study were:

1. It is expected that the resources of the four schools, perceived by teachers, determine the school's influence on the results of the performances, by test. The conceptual domain "resources" include the proficiency tests used, or not, by each school and respective teachers. From this hypothesis we can specify the following conditions:
 - There are differences between schools that do not use the proficiency tests and the low scores of immigrant children in the tests;
 - There are differences between schools with more verifiable support programs (including physical and digital materials) and better performance of non-native students;
2. It is expected that when the effect of the type of school is controlled, the mother tongue significantly influences the difference in performance in the tests between groups of participants (students);
3. It is expected that when the school effect is controlled, the nationality of the students significantly influences the difference in performance in the tests between groups of participants (students);
4. It is expected that when the school type effect is controlled, the socioeconomic status significantly influences the difference in performance in the tests between groups of participants (students).

Analysis of Variance (ANOVA's and Kruskal–Wallis) and One-Way Analysis of Covariance (ANCOVA)

In all the hypotheses, repeated measures of univariate analysis of variance tests were used to identify significant differences among groups and effect sizes, as well as One-Way Analysis of Covariance to establish if the "school" covariate influences and/or increases the effect of other variables with regard to performance in each test. In order to use ANOVA tests and then ANCOVA, the assumptions of the sample to use the above parametric tests were first reviewed, considering the independent variables (school, L1, nationality, SES and measurement levels) and the dependent variables (tests and school resources). The sample normality criteria were verified (using the Shapiro–Wilk test) as well as the homogeneity variance (through

Levene's test) in the relationship between the independent variables (L1 and school type) and dependent ones (the tests).

It was concluded that for the L1 independent variable, the retelling (recall), naming, semantic relationships and vocabulary tests showed lack of homogeneity (Levene values: $p > 0.05$). However, in most of the tests the sample remained homogeneous and parametric tests were conducted to verify the confirmation/rejection of our hypotheses, having ensured the other assumptions (normality). In the case of ANCOVA, it was decided to keep the test considering its statistical feasibility (nonparametric ANCOVA) for, even in cases with nonparametric tests, the test is valid as the studies show due to the robustness of ANCOVA for conditions where normal assumptions are violated (Akritas & Arnold, 2000; Akritas, 2004; Wang & Akritas, 2006).

For the nationality variable, the sample showed lack of homogeneity in only two tests (Levene values: picture naming: $p = 0.038$ and extraction: $p = 0.001$). The Shapiro–Wilk test revealed normality problems ($p < 0.05$) only in groups with fewer individuals (<10) identified in the Latin America, Eastern Europe, and Asia (Southern) groups. However, in most of the tests the sample remained homogeneous and parametric tests were carried out to verify the confirmation/rejection of our hypotheses, the other assumptions having been ensured and the effect sizes examined (according to Cohen's statistical patterns). For the SES variable, the Levene's test revealed variance homogeneity in all tests except in reading comprehension. For the normality test, only one of the socioeconomic groups had limitations (the group of families with academic qualifications demanding professions) in general in the various tests. Thus, the tests for all the independent variables with homogeneity and normality problems were taken into account as to their statistical limitation for covariate analysis.

References

- Akritas, M., & Arnold, S. (2000). Asymptotics for analysis of variance when the number of levels is large. *Journal of the American Statistical Association*, 95(449), 212–226.
- Akritas, M. G. (2004). Nonparametric survival analysis. *Statistical Science*, 615–623.
- Alberta Education. (2012). Language proficiency assessment. Retrieved February 12, 2013 from <http://www.learnalberta.ca/content/eslapb/languageproficiencyassessment.html>
- August, D., Kenyon, D., Malabonga, V., Louguit, M., Caglarcan, S., & Carlo, M. (2001). *Extract the base test—English*. Washington, DC: Center for Applied Linguistics.
- Bacelar do Nascimento, M. F. (2001). Um novo léxico de frequências do português. *Biblos*.
- Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me? Individual self-assessment by skill and level of language instruction. *System*, 40(1), 144–160.

Learning Portuguese as a Second Language

Figueiredo, S.

2017, IX, 58 p. 13 illus., Softcover

ISBN: 978-3-319-55818-9