

icCSBs 2016 : 4th Annual International Conference on Cognitive - Social, and Behavioural Sciences

The Big Four Skills: Teachers' Assumptions on Measurement of Cognition and Academic Skills for Non-Native Students

Sandra Figueiredo*, Margarida Alves Martins, Carlos Silva, O. Nunes

* Corresponding author: Sandra Figueiredo, sandradfigueiredo@ua.pt

^a Universidade Autónoma de Lisboa – DPS, Lisbon, Portugal, sandradfigueiredo@ua.pt

^b ISPA-Instituto Universitário, Lisbon, Portugal, mmartins@ispa.pt

^c Universidade de Aveiro, Aveiro, Portugal, csilva@ua.pt

^d Universidade Autónoma de Lisboa – DPS, Lisbon, Portugal, odete.nunes@ual.pt

Abstract

<http://dx.doi.org/10.15405/epsbs.2016.05.11>

The four-skills on tests for young native speakers commonly do not generate correlation incongruency concerning the cognitive strategies frequently reported. Considering the non-native speakers there are sparse evidence to determine which tasks are important to assess properly the cognitive and academic language proficiency (Cummins, 1980; 2012). Research questions: It is of high probability that young students with origin in immigration significantly differ on their communication strategies and skills in a second language processing context (1); attached to this first assumption, it is supposed that teachers significantly differ depending on their scientific area and previous training (2). Purpose: This study intends to examine whether school teachers (K-12) as having different origin in scientific domain of teaching and training perceive differently an adapted four-skills scale, in European Portuguese. Research methods: 77 teachers of five areas scientific areas, mean of teaching year service = 32 (SD= 2,7), 57 males and 46 females (from basic and high school levels). Main findings: ANOVA (Effect size and Post-hoc Tukey tests) and linear regression analysis (stepwise method) revealed statistically significant differences among teachers of different areas, mainly between language teachers and science teachers. Language teachers perceive more accurately tasks in a multiple manner to the broad skills that require to be measured in non-native students. Conclusion: If teachers perceive differently the importance of the big-four tasks, there would be incongruence on skills measurement that teachers select for immigrant pupils. Non-balanced tasks and the teachers' perceptions on evaluation and toward competence of students would likely determine limitations for academic and cognitive development of non-native students. Furthermore, results showed sufficient evidence to conclude that tasks are perceived differently by teachers toward importance of specific skills subareas. Reading skills are best considered compared to oral comprehension skills in non-native students.

© 2016 Published by Future Academy www.FutureAcademy.org.uk

Keywords: Scientific domain, teachers' training, four-skills, evaluation perceptions.

1. Introduction

Since the 1990s, studies have been analysing how teachers' instructional practices influence the learning skills of non-native students (Johnson, 1994). However, to date this analysis has focused mainly on higher education (Graham, 1984; Cho, Rijmen & Novak, 2013; Hazenberg & Hulstijn, 1996; Nation, 2001; Rosenfeld, Leung & Ottman, 2001). Foreign languages and arts teachers are expected to be more open to the multiple language skills method (Hinkel, 2012) with regard to immigrant school



population. The different tasks and skills tested are affected by the perception and choice of teachers according to their areas of expertise, since very recent studies show (Koizumi, 2015) that there is interaction between examiners and tests (in addition to the scores' variation explained by the proficiency differences among examinees), especially in the evaluation of task-specific language features (Koizumi, 2015, p.1).

As mentioned in the introductory section of this work, older studies in this field show that the expectations and representations of teachers about the knowledge and academic development of non-native students can affect the latter significantly (Creemers, 1994; Derwing, DeCorby, Ichikawa et al., 1999; Driessen & Whitagen, 1995; Jencks et al., 1979; Schneider & Yongsook, 1990) because, depending on the nationality and cultural perspective, non-native students, according to more recent research (Brok, Tartwijk, Wubbels et al, 2010; Horenczyk & Tatar, 2002) rely heavily on effective interpersonal relationships with their teachers, especially second generation immigrants (Brok et al., Callahan & Obenchain, 2013). On the other hand, authors like D'hondt, Eccles, Houtte et al. (2016) and Schaedel, Freung, Azaiza et al. (2015) have shown that the perceptions and expectations of teachers can affect just some minority groups because they rely heavily on other factors, such as parental investment. We believe that even parental investment depends on another variable, like nationality (Becker, 2010; Figueiredo, Alves Martins & Silva, 2015).

The L2 teacher has two profiles: the classical and the supportive (Tejada, Pino, Tatar et al, 2012.). And it is the second who is responsible for adjusting tests to the skills - reading, writing, speaking and listening - and respective sub-areas that require different cognitive strategies (Bialystok, 2002; Cummins, 2012; Hinkel, 2012; Koizumi, 2015). Hazenberg and Hulstijn (1996) and Nation (2001) had already detected the problem of the (complex) diversity of skills assessment and their inherent dimensions, which explains the multidimensionality of tests in L2 (Rosenfeld, Leung & Ottman, 2001). The multidimensionality is directly related to the concept of proficiency and academic competence presented earlier. Still, although recent studies show a tendency to organise tasks according to a comprehensive theory of evaluation (i.e. including a broad range of skills), other equally recent studies share the trend of the eighties and nineties regarding the notion of multiple language skills, which is confronted with the valuing of listening comprehension sub-skills, as in the recent studies by Sydorenko and Maynard (2014) and Yoon (2004), which evaluate the priorities of evaluating teachers regarding the discourses of non-native students, highlighting their evaluation preference for communication functions (i.e. recognition of intonation and cultural knowledge). The (academic) speaking skill test is also a priority in the evaluation of non-native teachers who are admitted to a teaching position in foreign universities (Choi, 2015).

In addition to the issue of which is the major skill evaluated in the tests, there is the question of the teachers' scientific teaching area and their formed perceptions. Teachers of scientific areas more related to the natural sciences are more negligent regarding the comprehensive teaching of the language, valuing the learning content of the syllabus of those subjects (Hinkel, 2012). This means that they prioritize tests based on specific skills without encompassing the four totals. Isaacs and Thomson (2013) draw attention to the need to clarify among younger teachers the content of an assessment in L2, as they have the greatest difficulty in distinguishing items within the test ranges. However, there are few studies that examine, by scientific area, the perceptions of teachers regarding the identification of the most important tasks for the satisfactory assessment of non-native students. Studies in this area examine above all the differentiation of assessment methods conducted by two groups of teachers: native and non-native. The study by Alemi and Tajeddin (2013) found that non-native teachers, in speaking tests of non-native students, are evaluators who differentiate fewer dimensions in the scale and have a more divergent perception. That is, native teachers pick up more strategies their non-native students use and differentiate them in terms of proficiency. However, few studies are known on the relationship between teacher groups divided according to academic areas (e.g. hard sciences, social sciences, MT teaching, FLs teaching and Visual Arts) and on their role as evaluators of non-native students.

Most authors (Lee, Llosa, Jiang et al., 2016) point to the value that science teachers, for example, attribute to language-focused tasks and home language use, but related to the lexical content of science

subjects. There were no studies of these tasks with non-native students. However, Lee, Quinn and Valdés (2013) conducted a recent study at the National Research Council focusing on the need to separate the teaching of science from language teaching as they are two distinct and not interdependent learning, with implications on the reformulation that authors consider the European Framework of Reference (2001) should undergo. From a different perspective, other recent studies (Lee & Winke, 2013) point to the perceptions of learners regarding tests that they themselves may have to fill in to be evaluated. The authors conclude that the test developers should be aware of the multitude of optional items within a listening comprehension measurement test and that the varied choice of students regarding the items they consider more or less important is due to the plurality of sub-skills that tests measure, but which the raters are not clear about such multiplicity of factors.

Another dimension to analyse in the variability of teacher responses and assessment behaviours is their experience of teaching non-native students and use of materials in a L2 context.

Studies (Williams, Abraham & Negueruela-Azarola, 2013) point to the importance of the concept-based instruction (CBI) method that often teachers use in multicultural classrooms. This method was created in the US and has been rooted since the 1990s, focusing on linguistic techniques in the classroom (Shrum & Glisan, 2001; Stoller, 2008). Schools choose this methodology based on Stoller's (2008) assumptions, for whom the CBI method allows a natural approach when learning the four skills of a foreign language (FL). According to the author, CBI allows reading authentic materials, interpret and evaluate information on the subject under study and develop projects in which students must cooperate to develop oral and written skills. This study also features other advantages: 1) with CBI, students are exposed to the language, while learning its content; 2) the metalinguistic reflection is contextualized, only studying the part of grammar that is essential for the understanding of the message, avoiding work in isolation and artificial language fragments; 3) students can bring their knowledge of the world to class and increase their motivation to learn.

It is suggested that the themes to be worked must follow students' interests, and if necessary discussed with them (Allan & Stoller 2005). They must promote working methods able to develop cooperative learning, apprenticeship learning, experimental learning, and project-based learning. Finally, it should be noted that CBI is a student-centred model that follows their interests, thus allowing greater flexibility and adaptability in curriculum design. It is a teaching approach that is not completely in accordance with the textbooks (Williams et al.) and, therefore, it is important in order to analyse the perceptions of teachers with no experience and with experience (recent and experienced teachers) in relation to the CBI. They concluded that the foreign language teachers they analysed in their study sample showed preference for their traditional methods set in the textbooks, rather than switch to CBI even if it harms the development of L2 students.

Teaching non-native pupils actually turns teachers from all areas into language teachers. Although previous studies (Borg, 2003; Bree, Hird, Milton et al., 2001; Richards & Lockart, 1994) have examined the teaching methods, but not the evaluation methods of L2 teachers, with progressive focus on the bottom-up teaching method, yet current studies continue to draw mother tongue and foreign language teachers' attention to the need to assess and teach targeted on non-native students (Aldana, Rowley, Checkoway et al., 2012; Bailey & Butler, 2003; Horwitz, 2012).

As for teachers with experience of multicultural classes, it does not mean that they are prepared or familiar with the most effective methods of L2 teaching (Williams, Abraham & Negueruela-Azarola, 2013). But again, the studies we found focus mainly on the perceptions of students and not of the teachers in the context of L2 teaching and assessment (Barnes & Lock, 2013) and on the FLs teachers context (Horwitz, 2012), mainly English as FL. The first study indicates that students value the psychological attributes of their language teachers, such as empathy and patience, but also pedagogical aspects, such as clarity of the explanation and variety of questioning (Barnes & Lock, 2013). In the other studies mentioned earlier (Horwitz, 2012), FLs teachers stand out for teaching experience and not specifically for their experience with non-native students. The perspective does not matter in the L2 context described here, although the results are important because the more experienced teachers (in terms of teaching time) show they benefit more from teaching strategies related to knowledge of English spelling. We have not come across studies reporting differences between teachers according to

scientific field and experience with multicultural classes regarding the measurement and evaluation tests of non-native students. This is one of the objectives of this study.

The premise of our study based on literature review is that not all schools and not all teachers are prepared, or have the resources, to deal with such diverse contexts in the classroom (Forman 2014; Samson & Collins, 2012); neither are they prepared to receive three important minorities: immigrant students, students with low socioeconomic status and ethnic groups (Kaida, 2013; Keels, 2009; Loeb, Soland & Fox, 2014) and to differentiate methods for other group of students: foreign language learners (languages other than English). Our focus is the minority related to the immigrant population in Portuguese schools. The literature concludes that there will be strong and known differences among teachers with different levels of experience and between second language and foreign language contexts, but it did not find sound evidence of teachers differentiated according to scientific field and experience with L2 teaching methods. Therefore, in this study we expect to find differences in behaviour of Portuguese teachers regarding the tasks they select in order to make a satisfactory assessment test to evaluate the proficiency and academic skills of immigrant learners of Portuguese as L2. Teachers with larger teaching experience and in accordance with previous studies on foreign language teaching (William et al.), have more ability to implement measures, but there is no literature on the correlation between teachers who use specific techniques for L2 teaching and assessment and their perception of items throughout the range of the four skills. This study will also examine this aspect. Empirical evidence (Brok, Tartwijk, Wubbels et al., 2010; Callahan & Obenchain, 2013; Horenczyk & Tatar, 2002; Tejada, Pino, Tatar et al., 2012) suggests that better prepared teachers are the ones who will provide greater support and materials to students, which generates the interpersonal teacher-student relationship.

2. Study

2.1. Sample

77 teachers aged between 32 and 62 years ($M=47$ years, $SD=7.4$), of whom 11 (14.3%) were male and 60 (77.9%) were female, with an average of 22 years teaching experience ($SD=6.7$). Teachers teach at nine schools/groupings in the district of Lisbon, with 9 being teachers of Portuguese language (11.7%), 12 of FLs (15.6%), 26 of Pre-school and Basic Education (33.8%), 8 of Hard Sciences (10.4%), 16 of History/Geography (20.8%) and 3 of Visual Arts (3.9%), distributed by the various levels of education (excluding higher education). 58 (75.3%) have experience of multicultural classes and 16 (20.8%) have never had non-native students in their classes. 46 (59.7%) used Non-Mother Tongue Portuguese Language (PLNM) measures and 19 (24.7%) admitted to never having used them.

ANOVA tests were carried out to compare results according to the participants' scientific domain and in relation to several variables: age, grade level, teaching experience, teaching experience with non-native students and experience with measures for second language learners. The results were: $F(5,66) = 3.518$, $p = .001$ for the age variable; $F(5,67) = 16.161$, $p = .000$ for the grade level; $F(5,68) = 3.198$, $p = .012$ for the teaching experience. No significant difference was found in the experience with non-native students (and measures used in their evaluation and learning).

2.2. Instrument

The *Inventory of undergraduate and graduate level – reading, writing, speaking and listening tasks* questionnaire by Rosenfeld, Leung and Ottman (2001) was used and adapted to the sample of Portuguese teachers. This questionnaire was originally developed by four scientific committees (framework teams) under the TOEFL and the Educational Testing Service in order to measure the importance, from the viewpoint of American university professors and students in education training courses, of the reading, writing, oral and listening tasks to be included in a test capable of assessing the academic competence and proficiency of non-native students. The original test has 42 items, of which we have adapted 40 distributed by the four scientific areas: reading (10 items), writing (10 items),

speaking (10 items) and listening (10 items). The original test in the English version has no information about its reliability properties, but the Portuguese version has a high Cronbach's coefficient (.94). The exploratory factor analysis was then conducted and all items were considered like in the original study that asserted that, although some items scored below 3.5, they were not excluded from the scale. We used the cut-off point established by the authors of the original version - 3.5. All items that scored below 3.5 revealed that teachers do not consider the task of integrating an assessment test for non-native students important.

The test was submitted to an exploratory factor analysis (EFA) to assess participants' answers and the factor structure of the test that was hypothesized as a four factor structure. The items with a factor loading of .40 or higher were used to define each factor. The Kaiser–Meyer–Olkin test showed that the sample size was adequate (.70) and the Bartlett test showed there was a good correlation index among the variables. As such, our acceptability rate allowed us to test our hypothesis ($p = .000$). By excluding no items, 11 factors were found and the first factor received the highest upload of almost all items from the “reading”, “writing”, “speaking” and “listening” scales. Eleven eigenvalues were higher than 1, explaining 75% of total variance. This was not expected considering that the EFA was hypothesized as a four-factor dimension structure. The original study did not produce an EFA. Fourteen items were selected.

The main component analysis showed that almost all 40 items loaded on the first factor, which was not expected considering that the first factor should correspond to the first scale (of four scales) - Reading. Items in this factor include items from the four academic skills. Three items (9. “Distinguish factual information from opinions”; 10. “Compare and contrast ideas in a single text and/or across texts”, 11. “Synthesize ideas in a single text and/or across texts”) from Reading scale loaded greatly only in the second factor.

2.3. Procedure

The data was collected in 2013 and 2015 in basic and high schools in the district of Lisbon. Contact was established with schools of the Lisbon district network to propose the study and disseminate the research aims. Communication with schools allowed identifying a vast group of teachers, which resulted in 77 teachers who fully completed the questionnaire. Following the informed consent and the demographic record of the selected school population, the questionnaire was answered and assessed (using points) according to the original test. Teachers responded to the questionnaire's forty questions on paper and returned it to the class Board and Department which, in turn, ensured it was returned to us. The procedure took place in the same way and in different academic periods in all schools.

The socio-demographic information was provided by the schools following informed consent after the beginning of each school year. The questionnaire was part of the empirical context of using linguistic and cognitive tests simultaneously with 108 immigrant students from different linguistic minorities who attended the same schools. Data were analysed using SPSS, version 21.

3. Results

This study aims to assess the knowledge and representations that these teachers have of evaluation tests, differentiating the groups according to scientific domain.

We made statistical analyses using SPSS and frequency tests were used (1) to compare means and standard deviations between groups in the different tasks (represented by the four specific scales) and to compare results with the original study; univariate analysis of variance to identify significant differences between groups and effect size (2); and regression analysis using the stepwise method (3).

3.1. Means, Standard Errors and Standard Deviations: comparison to the original study

The means, error deviation and standard deviation of all items of the scale of Part I of this study were first analysed, in accordance with the statistical steps carried out in the original study by Rosenfeld, Leung and Ottman (2001), in order to examine and distinguish the differentiated importance of each task set out in the Scale's factors. The cut-off point was set by the authors at 3.5 to determine the tasks considered to be most important in the composition of a FL evaluation test. Considering the total sample, the items of this study showed means ranging between 3.33 (mean in the writing skill items - "awareness of audience needs and write to a particular audience or reader" - and speaking skill - "developing and structuring hypotheses") and 4.70 (mean calculated for the listening comprehension item - understanding the teacher's instructions in class). The results are very similar to the original study, considering the total sample, as follows: the writing skill item with a mean of 3.33 ("moderately important" or "important") is also one of the least scored (less perceived as important by teachers) in the original study; also, the reading skill item (understands written instructions) with 4,57 and one with the highest score ("extremely important") also has a mean in the original study. On the other hand, this study in Portugal has higher scores regarding the importance of the tasks in the perception of teachers. And no task, as in the previous study, was considered irrelevant ($M < 3.0$) to measure the academic skills and proficiency of non-native students.

Comparing the mean response between the groups and considering the sample as determined by the variable "scientific area", in the answers of the Portuguese teachers (Mother Tongue) the mean below the cut-off point was only found in the item mentioned above for the speaking skill "developing and structuring hypotheses" ($M=3.22$) and for a writing item "awareness of audience needs and write to a particular audience or reader" ($M=3.44$). But the most scored item was also the item of listening comprehension related to understanding teachers' instructions ($M=5.00$).

In the answers of FLs teachers, the lowest mean ($M=3.22$) was also in speaking items (one already identified above, and the other related to the competence of orally comparing ideas and arguments: "making comparisons/contrasts", but also in two writing skills items (i.e. "produce sufficient quantity of written text appropriate to the assignment and the time constraints"). This same item was identified in the original study and also had a mean below the cut-off point. FLs teachers value most the listening comprehension item related to understanding details and facts ("understand factual information and details"), as seen in higher means in the previous study.

For basic education teachers, none of the items' mean was below the cut-off point and they are the most positive in valuing the tasks in an assessment test for non-native students. The highest mean ($M=4.77$) was found in the reading field ("determining the basic theme (main idea) of a passage").

Regarding teachers of other groups not related to language(s) teaching, the results show a smaller selection of important items to evaluate students' skills. As for teachers of hard sciences, we found different results because there are several items with means below the cut-off point with values between 2.25 and 3.4 in items above the area of reading and speaking skills.

Sharing some of the above identified items with lower means, items with higher mean levels (and curiously always with the same mean = 4.63) are related to two reading tasks ("read written instructions/directions concerning classroom assignments" and "read text material (...) to remember major ideas") and to a listening comprehension task ("understand main ideas and their supporting information").

Concerning teachers of History and/or Geography, there are also several items with means below the cut-off point, ranging from 2.94 to 3.44 and covering mainly items devalued in writing and speaking skills. Two of the items, in two respective skills, and considering the computation of the overall sample, indicated to be of less importance "awareness of audience needs and write to a particular audience or reader" - and speaking skill - "developing and structuring hypotheses". Also in this overall computation it was found that one of the most valued items was the listening comprehension related to teachers' instructions. Here, for the specific group of History and Geography teachers, the same was also found ($M=4.88$).

Finally, as regards the group of Visual Arts teachers, the items considered the less to assess the skills and proficiency of non-native students are not, like in the previous group, represented in the writing skills. Some of the items are the same in both groups regarding the 'moderate' importance (below the cut-off point). However, the highest mean is also found in the item listening comprehension of teachers' instructions (M=4.67).

In conclusion, it should be noted that teachers with more positive perceptions are language teachers and basic education teachers, who understand more tasks, focusing on all the skills to satisfactorily assess immigrant pupils. Social sciences teachers (History and Geography) and hard sciences (and visual arts) select fewer items, thereby minimizing the number of tasks and focusing more on reading and listening skills. The means and standard deviations of the comparison between the groups of teachers according to scientific area and to the various items and test factors are described in Table 1.

Table 1. Means and standard deviations of the comparison among groups according to scientific domain and training

	Scientific domain	Mean	P.Deviation
Reading (writing instructions)	Portuguese	4,89	,333
	Foreign Languages	4,36	,674
	Kindergarten and 1-4 grades	4,58	,584
	Sciences	4,57	,787
	History	4,60	,507
	Arts	4,33	,577
	Reading (distinguish facts from opinions)	Portuguese	4,00
Foreign Languages		3,45	,688
Kindergarten and 1-4 grades		3,96	,690
Sciences		2,71	,951
History		3,67	1,234
Arts		3,33	,577
Writing (awareness of audience needs)		Portuguese	3,44
	Foreign Languages	3,45	,688
	Kindergarten and 1-4 grades	3,79	,779
	Sciences	2,14	,690
	History	3,00	,845
	Arts	3,33	,577
	Speaking (clarity toward teacher)	Portuguese	4,44
Foreign Languages		4,64	,505
Kindergarten and 1-4 grades		4,38	,711
Sciences		3,57	1,272
History		3,60	,986
Arts		3,33	,577
Speaking (arguing)		Portuguese	4,22
	Foreign Languages	3,82	,751
	Kindergarten and 1-4 grades	4,08	,584
	Sciences	3,86	,900
	History	3,47	,640
	Arts	3,67	,577

Speaking (to compare and contrast facts and opinions)	Portuguese	3,78	,667
	Foreign Languages	3,45	,820
	Kindergarten and 1-4 grades	3,71	,806
	Sciences	3,29	,488
	History	3,53	,743
	Arts	3,67	,577
Speaking (structuring hypotheses)	Portuguese	3,22	,667
	Foreign Languages	3,09	1,221
	Kindergarten and 1-4 grades	3,67	,917
	Sciences	3,43	,787
	History	2,93	,884
	Arts	4,00	1,000

3.2. Univariate analysis of variance (ANOVA): effect size and post-hoc analyses

Through several ANOVA's the effect size was determined on the groups mean differences, using the "scientific domain" variable. For all the items there was substantial effect size (η^2 ranging between .178 and .358). Cohen's benchmarks for statistical value of η^2 were established as norm (Cohen, 1988). Mean results, effect sizes and Tukey test (post-hoc) of the univariate analysis for all tasks (reading, writing, speaking and listening comprehension) are shown in Table 2.

Reading: ANOVA results evidenced significant differences among the groups determined by "scientific domain". For the "theme comprehension" ability, the groups differed significantly ($p=.010$; $\eta^2 =.210$). The group with the highest mean in this task was the group of History and Geography teachers ($M=4.87$), followed by the group of basic education teachers ($M=4.75$), as opposed to the group of Hard Sciences teachers ($M=4.00$) who valued this item less compared to previous groups.

For the item "distinction of facts/opinions", the groups also showed significant different behaviour ($p=.027$; $\eta^2=.178$) between them as regards the valuation of the item. Teachers of Portuguese and basic education scored higher ($M=3.96$), as opposed to science teachers ($M = 2.71$). As for the task "distinction and comparison of texts", the groups showed significant differences as indicated by the effect size ($p=.009$; $\eta^2=.211$). The basic education teachers are the ones who valued this task more ($M=3.71$) with a significant difference compared to science teachers ($M=2.43$).

A post-hoc (Tukey) test revealed significant differences among the groups of teachers from different scientific domains ($p<.05$), considering differences for the following specific Reading items: theme comprehension ($F(5.63)=4.705$; $p=.001$), facts and opinions ($F(5.63)=4.319$; $p=.002$), distinction and comparison of texts ($F(5.63)=.931$; $p=.009$). The differences were between teachers of Sciences and teachers of History/Geography, basic education teachers and Portuguese Language teachers and the means were presented above; no significant differences were found among the other teachers' groups (for example, Foreign Language teachers and Arts teachers).

Writing: ANOVA results evidenced significant differences among the groups according to "scientific domain". The groups differed significantly ($p=.000$; $\eta^2 =.317$) in the "writing for an audience" ability. The group with the highest mean in this task was the group of basic education teachers ($M=3.79$), followed by the group of Portuguese teachers ($M=3.44$) and History/Geography teachers ($M=3.00$), with Science teachers scoring this task lower ($M=.2.14$) as being important for the evaluation of non-native students. For the "time constraints" item, groups also showed significant different behaviour ($p=.007$; $\eta^2=.219$) in relation to the valuing of the item. Teachers of Portuguese value it most ($M=4.33$), unlike FLs teachers ($M=3.27$), whose mean is below the cut-off point.

As for the "linguistic rules" (or grammar) item, the groups differ significantly with large effect size ($p=.000$; $\eta^2=.358$) regarding the value of the item. The means show the following decreasing order: Portuguese language teachers ($M=4.33$), basic education teachers ($M=4.04$), History/Geography

teachers (M=3.67), Science teachers (M=3.29) and FL teachers (M=3.27). The groups of language teachers differ substantially from each other in the importance they attach to this writing item.

A post-hoc (Tukey) test revealed significant differences in the “writing for a specific audience” (F(5.63)=.166;p=.000), “time constraints” (F(5.63)=1.768;p=.007), and “grammar” (F(5.63)=1.311;p=.000) items between Science teachers and History/Geography teachers, Basic Education and Portuguese Language teachers. As the means suggest (Table 1), differences were also noticed between Portuguese Language teachers and Foreign Languages teachers on ability to write facing time constraints; significant multiple differences ($p < .05$) were observed among teachers’ groups for the item related to writing according to grammar knowledge in L2: between Science teachers and History/Geography teachers, Basic Education teachers, and Portuguese language teachers. And, in the same task, other groups behaved differently *inter se*: this was the case between History/Geography teachers and Portuguese Language teachers, Foreign Language teachers and Basic Education Teachers (Table 2).

Speaking: ANOVA results evidenced significant differences among the groups according to “scientific domain”: in terms of ability of “questioning”, the groups differed significantly ($p = .003$; $\eta^2 = .241$). The group with the highest mean in this task was the group of FLs teachers (M=4.64), and the teachers of History/Geography had the lowest mean (M= 3.60).

Regarding the task “clarity during participation in classroom context”, the groups differed significantly ($p = .005$; $\eta^2 = .227$). The group of basic education teachers had the best score (M=4.25), followed by History/Geography teachers (M=3.40) and by Science teachers (M=3.29).

As for item “clarity during presentation in classroom context”, the groups differed significantly ($p = .002$; $\eta^2 = .249$). The Portuguese language group of teachers had the highest mean (M=4.11), followed by FLs teachers (M=4.09) and by basic education teachers (M=4.00). Science teachers paid less attention to the task (M=2.86).

A post-hoc (Tukey) test revealed significant differences among the groups of teachers from different scientific domains ($p < .05$), with regard to specific speaking task items: “questioning” (F(5.63)=3.167;p=.013), “clarity during participation in classroom context” (F(5.63)=3.432;p=.008), and “clarity during presentation in classroom context” (F(5.63)=.697;p=.002). As previous means values suggested, the differences were observed only between History/Geography teachers and Foreign Language teachers; regarding speaking with clarity during participation in class: the differences ($p < .05$) were between Basic Education teachers and History/Geography teachers; regarding speaking with clarity for presentation in class: there were differences between Science teachers and the other teachers’ groups (Portuguese, Foreign Languages and Basic Education, see Table 2).

Listening: For all the listening items/tasks, ANOVA results revealed no significant differences concerning effect size. The post-hoc tests did not show significant differences among the teachers’ groups.

Table 2. Comparison among groups (means, pearson and effect sizes): teachers perceptions according to teaching service, experience with multicultural classes and L2 measures application.

Scientific domain	Mean	<i>p</i>	η^2
Reading	Theme comprehension	.010	.210
	Distinction facts/opinions	.027	.178
	Comprehension of texts	.009	.211
Writing	Audience needs	.000	.317
	Time constraints’	.007	.219
	Linguistic rules	.000	.358
Speaking	Questioning	.003	.241
	Clarity during participation in classroom	.005	.227

3.3. Linear Multiple Regression Analysis

Considering the sequence of past results, and having particularly noted how the groups of subjects determined according to different independent variables behaved in the valuation of tasks of the four scientific areas, it was decided to resort to linear regression analysis to ascertain the main predictors among the group of independent variables under study using the stepwise method, and how the model is used to establish the importance of specific tasks to be given to non-native students by teachers in Portuguese schools (Lisbon district). Only the tasks that showed significant differences in the groups and significant effect sizes were considered, according to the results of the previous tests.

Reading

For the task “compare and contrast ideas in a single text and/or across texts” regression results also showed that the teachers’ scientific domain variable has predictive value ($b=-.333, p=.009$) but also the “experience with measures applied to foreign students” ($b=-.249, p=.041$), as opposed to the other 3 factors – teaching experience, age, experience with foreign students in classes - where no significant predictive power was shown. The importance of that specific reading items is affected by the perceptions of the different teachers (by scientific domain). In order to clarify this result, and having examined the overall result (through a frequencies previous test) of the answers to all the items of the scale (14 factors), it was found that Portuguese language and FLs teachers are the ones who had more positive perceptions regarding all the tasks listed in the questionnaire. Teachers of other scientific fields value different items and attach less importance to items. The teaching areas are important factors in predicting tasks and enforcing them among linguistic minorities in schools. As for the other variable that the model shows as being the second and final model predictor, it appears that the negative experience of teachers (absence of experience) predicts lower application of measures for students inside the classroom. There were 2 reading tasks (theme identification and facts/opinions distinction) that regression analysis revealed as having no significant predictive value for any independent variables. The results are summarized in table 3.

Writing

Regarding both tasks “awareness of audience needs and write to a particular audience or reader” (1) and “time constraints” (2), results showed that only the teachers’ scientific domain variable has predictive value (task 1: $b=-.261, p=.044$; task 2: $b=-.613, p=.000$) as opposed to the other 4 factors – teaching experience, age, experience with foreign students in classes, measures applied to foreign students in classroom - where no significant predictive power was found. Importance for that specific writing item is affected by the scientific domain of teachers, meaning that there are perceptions of teachers, according to the teaching area, that produce differences on the importance attributed to that task, for foreign/immigrant students. The results are summarized in table 3.

Speaking

For the tasks “questioning teacher” (1), “participation toward other students” (2) and “presentation toward other students” (3), results showed that only the teachers’ scientific domain variable has maintained predictive value (task 1: $b=-.421, p=.001$; task 2: $b=-.335, p=.009$; task 2: $b=-.365, p=.004$). For the task “giving instructions/directions” (4) results showed that experience with multicultural classes is the only predictor ($b=.270, p=.037$). For the task “structuring hypotheses” (5), results showed that the teachers’ scientific domain variable has maintained the predictive value ($b=-.389, p=.002$), but also other variable emerged from the model as a predictor: age variable ($b=-.287, p=.017$). In the ANOVAs the age variable did not display significant differences among the groups for any items.

Listening

For the task “recognize the speaker’s attitudinal signals”, results showed that the experience with measures for foreign students is the only predictor ($b=-.272, p=.034$). The importance of that item is affected by teachers’ experience in dealing with pedagogical measures (considering the descriptive statistics of ANOVA, we determined that the teachers with experience of pedagogical measures

attribute greater importance to this task than the group with no experience or knowledge of those measures). The results are summarized in table 3.

Table 3. Linear regression analysis of tasks relevance (*dependent variables appeared in the prediction model).

Model	Standardized coefficients	
	Beta	Sig.
Reading		
compare and contrast ideas in a single text and/or across texts (1)		
Scientific domain*	.333	.009
Experience with L2 measures (2)		
Scientific domain*	.249	.041
Writing		
awareness of audience needs and write to a particular audience or reader (1)		
time constraints (1)	.261	.044
Scientific domain*	.613	.000
Speaking		
questioning teacher (1)	.421	.001
participation toward other students (2) presentation toward other students (3)	.335	.009
Scientific domain*	.365	.004
giving instructions/directions (4)	.270	.037
Experience with multicultural students* structuring hypotheses (5)		
Scientific domain*	.389	.002
Age*	.287	.017
Listening		
recognize the speaker's attitudinal signals (1)		
L2 measures administered*	.272	.034

4. Discussion

Concerning the question of the study, the knowledge and representations that different teachers have regarding evaluation tests vary mainly according to the predominance of the scientific area compared to other factors that have been confirmed as not being strong predictors in the model: age, length of service and experience with multicultural groups. Regarding the scientific area, in the early part of the paper we present the differences in perceptions that teachers, by scientific area, denote and explain their valuing of reading and listening comprehension tasks to the detriment of the writing and speaking tasks. As noted, previous studies suggest that teachers have inadequate knowledge and representations of instruction and assessment of priority tasks in L2 (Graham & Peri, 2007; Littlewood, 2007; Veenman, 1984).

This study concludes that the teachers with more positive perceptions are language teachers and basic education teachers, who understand more tasks, focusing on all the skills to satisfactorily assess immigrant pupils. These data contradict the study by García-Nevarez, Stafford & Arias (2010), which examined an American sample (Arizona State) of basic education teachers regarding the importance of adjusting teaching to non-native students (English as L2), and detected a huge variability of answers,

which depended on the type of training these teachers had had, with those qualified to teach bilingual education being more favourable and more supportive of L2 learners, unlike older teachers (bilingual and monolingual) who, compared with the younger ones, had negative attitudes towards non-native MT students.

The same variability of answers, albeit dependent on other variables, was also found in teachers (Michigan) examined by Karabenick and Noda (2004), supporting the teachers' perception regarding the students' bilingualism as an advantage compared to non-bilingual L2 pupils. In this study, in more advanced levels of basic education and high school, social sciences teachers (History and Geography) and sciences (and visual arts) teachers select fewer items, thereby minimizing the number of tasks and focusing more on reading and listening comprehension skills. These results are consistent with data advanced by Hansen-Thomas and Cavagnetto (2010), who found that teachers in the area of mathematics do not distinguish between tasks that these students should complete to be able to develop the language of the subject. Also, a study by Reeves (2006) showed that high school teachers had insignificant positive attitudes regarding including immigrant students in regular classes, due to the specific teaching for L2 and to the change of plans and programmes to adapt to these students.

Hansen-Thomas and Cavagnetto (2010) and Rubinstein-Avila and Lee (2014) drew attention to the lack of preparation of high school teachers and changing approaches of L2 teaching to a language-based approach that sensitizes teachers who are not language teachers to realize that to teach hard sciences, for example, the approach has to be that one and not merely the content of the subject without the linguistic competence. This type of results is close to those obtained in our study, in that high school teachers, unlike basic education teachers, are the ones who value items as being less important for non-native students' tasks. The study on the cognition of teachers teaching non-native students (in the context of L2) has been the subject of several international research (Borg, 2003; Tsui, 2003; Venman, 1984; Wright, Eisenhardt & Mainzer, 2010). Teachers are examined with regard to inconsistencies between their knowledge of L2 teaching and practice in the classroom. Studies have also focused on the difference between younger and more experienced teachers, with the inconsistencies having a negative impact on students (Reagan, 2006; Tsui, 2003). The main problem of teachers in relation to the multiplicity of tasks is the decision and planning of those tasks for different contexts (students), prominently in high school where the skills of academic language focus on the content of the subjects and simultaneously cut across them (Schleppegrell, & O'Hallaron, 2011).

In other empirical studies (Hasan, 2010; Siegel, 2013), teachers, especially FLs ones, prioritise listening and reading in tasks that share the same demands (understanding teacher's instructions, for example), seriously devaluing writing skills aspects (listening and grammar) and speaking (Brown, 2009; Khuwaileh & Shoumali, 2000). This teachers' priority was also confirmed in this study and is consistent with the research that confirms the positive relationship between reading and listening tests, in that they also share specific aspects in the field of cognitive processing (Zeeland & Schmitt, 2012; Harding, Alderson & Brunfaut, 2015). This is the advantage of the interdependence between skills, and teachers seem to be aware of this positive resource for assessing and teaching students in a L2 context (Chien, Palau and Sun (2014).

The results presented in this study are an important contribution especially in two aspects: the analysis of teachers' perception of relevant tasks in L2 is pioneer in Portugal. On the other hand, it presents a corpus of results that corroborate and contrast those of previous international studies, with implications for education and concepts of practices that teachers from various scientific fields reveal about L2 teaching and the type of tasks to consider in tests and in the classroom.

The data suggest that teachers may be developing inadequate practices and concepts, especially considering the differences according to scientific field and high school level; that they undervalue the grammar component of all skills to be developed by the students; that they overemphasize listening comprehension and its relationship with reading; that they follow closely a L2 teaching model (originally of American design, Horwitz, 1985) but only basic education teachers (for students aged 4-11 years); and that they have poor notions regarding L2 tasks and evaluation tests, in general.

The inconsistency in perception and practices across different groups of teachers and respective experiences results from the variability in their responses and from the statistically significant

differences in the specific tasks they have chosen as being relevant and irrelevant. However, when compared to the group of teachers from one of the samples of the original study (Rosenfeld, Leung & Ottman, 2001), Portuguese teachers are more positive when differentiating tasks. And despite the fact that the variables related to experience with multicultural classes and application (knowledge) of the PLNM measure had little predicting value in the analysis of the results, we consider these factors to be important to improve teachers' perceptions of tasks to be done in classrooms with different minorities, especially considering that over the last few months Portuguese schools have been receiving refugee students.

Acknowledgements

This work was supported by the Foundation for Science and Technology (FCT) under the Grant n.º SFRH/BPD/86618/2012; and Center of Psychology Research of Universidade Autónoma de Lisboa, Lisbon Portugal.

References

- Aldana, A., Rowley, S., Checkoway, B., & Richards-Schuster, K. (2012). Raising ethnic-racial consciousness: The relationship between intergroup dialogues and adolescents' ethnic-racial identity and racism awareness. *Equity & Excellence in Education*, 45(1), 120-137.
- Alderson, C., Brunfaut, T., & Harding, L. (2015). Towards a Theory of Diagnosis in Second and Foreign Language Assessment: Insights from Professional Practice Across Diverse Fields. *Applied Psycholinguistics*. <http://apllj.oxfordjournals.org/content/early/2014/01/03/applin.amt046.abstract>
- Alemi, M., & Tajeddin, Z. (2013). Pragmatic rating of L2 refusal: Criteria of native and non-native English teachers. *TESL Canada Journal* 30(7), 63.
- Alan, B., & Stoller, F. (2005). *Maximizing the benefits of project work in foreign language classroom*. English Teaching Forum: eca.state.gov
- Bailey, A., & Butler, F. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document*. Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies. LA: University of California.
- Barnes, B., & Lock, G. (2013). Student Perceptions of Effective Foreign Language Teachers: A Quantitative Investigation from a Korean University. *Australian Journal of Teacher Education* 38(2).
- Bialystok, E. (2002). Acquisition of literacy in bilingual children: A framework for research. *Language learning* 52(1), 159-199.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language teaching* 36(2), 81-109.
- Bree, M., Hird, B., Milton, M., Oliver, R., & Thwaite, A. (2001). Making sense of language teaching: teachers' principles and classroom practices. *Applied Linguistics*, 22(4), 470-501. doi: 10.1093/applin/22.4.470
- Brown, A. (2009). Students' and Teachers' Perceptions of Effective Foreign Language Teaching: A Comparison of Ideals. *The Modern Language Journal*, 93, 46-60. doi: 10.1111/j.1540-4781.2009.00827.x
- den Brok, P., van Tartwijk, J., Wubbels, T., & Veldman, I. (2010). The differential effect of the teacher-student interpersonal relationship on student outcomes for students with different ethnic backgrounds. *British Journal of Educational Psychology*, 80, 199-221. doi: 10.1348/000709909X465632
- Callahan, R., & Obenchain, K. (2013). Bridging Worlds in the Social Studies Classroom: Teachers' Practices and Latino Immigrant Youths' Civic and Political Development. *Sociol Stud Child Youth*. 16, 97-123. doi: 10.1108/S1537-4661(2013)0000016009
- Chien, K., Palau, V., & Sun, M. (2014). *Reading and Writing Achievement Test for Advanced 1 Students in the Community English Program*. Teachers College, Columbia University.
- Cho, Y., Rijmen, F., & Novák, J. (2013). Investigating the effects of prompt characteristics on the comparability of TOEFL iBT™ integrated writing tasks. *Language Testing*, 30(4), 513-534.
- Choi, T. (2015). The impact of the 'Teaching English through English' policy on teachers and teaching in South Korea. *Current Issues in Language Planning* 16(3), 201-220.
- Creemers, B. (1994). *The Effect Classroom*. London: Cassell.
- Cummins, J. (2012). The intersection of cognitive and sociocultural factors in the development of reading comprehension among immigrant students. *Reading and Writing*, 25 (8), 1973-1990.
- Derwing, T., DeCorby, E., Ichikawa, J., & Jamieson, K. (1999). Some Factors That Affect the Success of ESL High School Students. *The Canadian Modern Language Review*. DOI: <http://dx.doi.org/10.3138/cmlr.55.4.532>
- Driessen, G., & Withagen, V. (1999). Language varieties and educational achievement of indigenous primary school pupils. *Language Culture and Curriculum*, 12(1), 1-22.
- European Commission. (2001). *European Common Framework for Languages*. Porto: Edições Asa.
- Figueiredo, S., Alves Martins, M., Silva, C. & Simões, C. (2015). Second language education context and home language effect: language dissimilarities and variation differences in immigrant student's outcomes. *International Journal of Multilingualism*. DOI:10.1080/14790718.2015.1079204.

- Forman, R. (2014). How local teachers respond to the culture and language of a global English as a Foreign Language textbook. *Language, Culture and Curriculum*, 27 (1): 77-88.
- García-Nevarez, A., Stafford, E., & Arias, B. (2005). Arizona Elementary Teachers' Attitudes Toward English Language Learners and the Use of Spanish in Classroom Instruction. *Bilingual Research Journal*, 29(2), 295-317. DOI: 10.1080/15235882.2005.10162837
- Shrum, L., & Glisan, E. (2001). *Teacher's Handbook*. Heinle Cengage Learning.
- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3), 445-476.
- Graham, S. (1984). Communicating sympathy and anger to Black and White children: The cognitive (attributional) consequences of affective cues. *Journal of Personality and Social Psychology* 47(1), 40.
- Hansen-Thomas, H., & Cavagnetto, A. (2010). What Do Mainstream Middle School Teachers Think About Their English Language Learners? A Tri-State Case Study. *Bilingual Research Journal: The Journal of the National Association for Bilingual Education*, 33(2), 249-266. DOI:10.1080/15235882.2010.502803
- Hasan, A. (2010). Learners' Perceptions of Listening Comprehension Problems, *Language, Culture and Curriculum*, 13(2), 137-153, DOI: 10.1080/07908310008666595
- Hazenbergh, S., & Hulstijn, J. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, 17, 145-163. doi:10.1093/applin/17.2.145.
- Hinkel, E. (2012). *Cognitive-code learning*. Encyclopedia of the Sciences of Learning. USA: Springer.
- Horenczyk, G., & Tatar, M. (2002). Teachers' attitudes toward multiculturalism and their perceptions of the school organizational culture. *Teaching and Teacher Education*, 18(4), 435-445.
- Horwitz, E. (1985). Using Student Beliefs About Language Learning and Teaching in the Foreign Language Methods Course. *Foreign Language Annals*, 18, 333-340. doi: 10.1111/j.1944-9720.1985.tb01811.x
- Horwitz, E. (2014). Becoming a Language Teacher: A Practical Guide to Second Language Learning and Teaching. *The CATESOL Journal*, 25(1), 163-192.
- Isaacs, T., & Thomson, R. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly* 10(2), 135-159.
- Johnson, K. (1994). The emerging beliefs and instructional practices of preservice English as a second language teachers. *Teaching and Teacher Education*, 10(4), 439-452.
- Kaida, L. (2013). Do host country education and language training help recent immigrants exit poverty? *Social Science Research*, 42 (3), 726-741.
- Karabenick, S., & Noda, P. (2004). Professional Development Implications of Teachers' Beliefs and Attitudes Toward English Language Learners. *Bilingual Research Journal*, 28(1), 55-75, DOI: 10.1080/15235882.2004.10162612
- Keels, M. (2009). Ethnic group differences in early head start parents' parenting beliefs and practices and links to children's early cognitive development. *Early Childhood Research Quarterly*, 24 (4), 381-397.
- Khuwaileh, A., & Shoumali, A. (2000). Writing errors: a study of the writing ability of Arab learners of academic English and Arabic at university. *Language, Culture and curriculum*, 13(2), 174-183.
- Koizumi, R. (2015). Task and rater effects in L2 speaking and writing: A synthesis of generalizability studies. *Language Testing*, 0265532215587390.
- Lee, O., Llosa, L., Jiang, F., Haas, A., O'Connor, C., & Van Booven, C. (2016). Elementary teachers' science knowledge and instructional practices: Impact of an intervention focused on english language learners. *Journal of Research in Science Teaching*.
- Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. *Educational Researcher*, 0013189X13480524.
- Lee, H., & Winke, P. (2013). The differences among three-, four-, and five-option-item formats in the context of a high-stakes English-language listening test. *Language Testing* 30(1), 99-123.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language teaching*, 40(3), 243-249.
- Loeb, S., Soland, J., & Fox, L. (2014). Is a Good Teacher a Good Teacher for All? Comparing Value-Added of Teachers With Their English Learners and Non-English Learners. *Educational Evaluation and Policy Analysis*, 36 (4), 457-475.
- Nation, I. (2001). *Learning vocabulary in another language*. Ernst Klett Sprachen.
- Reeves, J. (2006). Secondary Teacher Attitudes Toward Including English-Language Learners in Mainstream Classrooms. *The Journal of Educational Research*, 99(3), 131-143.
- Richards, J., & Lockhart, C. (1994). *Reflective teaching in second language classrooms*. Cambridge University Press.
- Rosenfeld, M., Leung, S., & Oltman, P. (2001). *TOEFL monograph series: The reading, writing, speaking, and listening tasks important for academic success at the undergraduate and graduate levels*. Toefl Monograph Series MS-21. Princeton, NJ: Educational Testing Service.
- Rubinstein-Ávila, E., & Lee, E. (2014). Secondary Teachers and English Language Learners (ELLs): Attitudes, Preparation and Implications. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 87(5), 187191. DOI:10.1080/00098655.2014.910162
- Samimy, K., & Lee, Y. (1997). Beliefs about language learning: Perspectives of first-year Chinese learners and their instructors. *Journal-Chinese Language Teachers Association*, 32, 40-60.

- Samson, J., & Collins, B. (2012). *Preparing All Teachers to Meet the Needs of English Language Learners: Applying Research to Policy and Practice for Teacher Effectiveness*. Center for American Progress.
- Schaedel, B., Freung, A., Azaiza, F., Hertz-Lazarowitz, R., Boeum, A., & Eshet, Y. (2015). School Climate and Teachers' Perceptions of Parental Involvement in Jewish and Arab Primary Schools in Israel. *International Journal About Parents in Education* 9(1), 77-92.
- Schleppegrell, M., & O'Hallaron, C. (2011). Teaching academic language in L2 secondary settings." *Annual Review of Applied Linguistics*, 31, 3-18.
- Schneider, B., & Yongsook, L. (1990). A Model for Academic Success: The School and Home Environment of East Asian Student. *Anthropology and Education Quarterly*, 21(4), 358-377.
- Siegel, J. (2013). Second language learners' perceptions of listening strategy instruction. *Innovation in Language Learning and Teaching*, 7(1), 1-18.
- Stoller, F. (2008). "Content-Based Instruction." In N. Hornberger (Ed.), *Encyclopedia of Language and Education* (pp. 1163-1174). US: Springer.
- Tejada, A., Pino, R., Tatar, M. & M. Sayáns, P. (2012). Spanish as a foreign language" teachers' profiles: inclusive beliefs, teachers' perceptions of student outcomes in the TCLA program, burnout, and experience. *European Journal of Psychology of Education*, 27(3), 285-298.
- Tsui, A. (2003). *Understanding expertise in teaching: Case studies of second language teachers*. Ernst Klett Sprachen.
- Veenman, S. (1984). Perceived Problems of Beginning Teachers. *Review of Educational Research*, 54(2), 143-178.
- Williams, L., & Lee B. Abraham, L., & Negueruela-Azarola, E. (2013). Using concept-based instruction in the L2 classroom: Perspectives from current and future language teachers. *Language Teaching Research*, 1362168813482950.
- Wright, E., Eisenhardt, P., Mainzer, A., Ressler, M., Cutri, R., Jarrett, T., & Stanford, S. (2010). The Wide-field Infrared Survey Explorer (WISE): mission description and initial on-orbit performance. *The Astronomical Journal*, 140(6), 1868.
- Yoon, B. (2004). Uninvited Guests: The Influence of Teachers' Roles and Pedagogies on the Positioning of English Language Learners in the Regular Classroom. *American Educational Research Journal*, 5, 495-522.
- Zeeland, H., & Schmitt, N. (2013). Lexical Coverage in L1 and L2 Listening Comprehension: The Same or Different from Reading Comprehension? *Applied Linguistics*, 34 (4), 457-479. doi: 10.1093/applin/ams074.