

## *Original Paper*

# Designing Indicators for a Placement Test: Drawbacks and Affordances

Gabriel Brito Amorim<sup>1</sup>

<sup>1</sup> Languages Center, Federal University of Espírito Santo, Brazil

Received: September 9, 2021    Accepted: October 3, 2021    Online Published: December 29, 2021

doi:10.22158/fce.v2n2p67

URL: <http://dx.doi.org/10.22158/fce.v2n2p67>

### **Abstract**

*One of the greatest challenges when it comes to assessment is the lack of clarity in the objectives or, to make it worse, the absence of objectives whatsoever for what results the curriculum, course, or the test instrument wishes to achieve. According to Raupp and Reichle (2003), this kind of information usually exists as expectations pushing the examiner to transform expectations, general goals or poorly stated objectives into measurable objectives. Raupp and Reichle (2003) emphasize that very precise indicators need to be written or selected to measure whether or not the course/curriculum objectives proposed were achieved. However, this job becomes rather challenging when objectives are not clearly stated or are too vague. Indicators, then, ideally, resonate curricular goals, teaching methodologies, and testing. If one of these components is not well calibrated or is dissonant with the others, the whole assessment process is jeopardized (CEFRL, 2007). Indicators for placement testing, then, need to be carefully written or selected so that they can reflect an intended proficiency level as well as the school's curricular goals, teaching methodologies, and testing procedures (Brown, 2004; Richards, 2009). At the Federal University of Espírito Santo's Language Center (CL), the discontentment of instructors and coordination staff with the lack of reliability, validity and practicality with the CL's placement test motivated debates and work on the refinement of the indicators used for the referred exam. Through a literature review and with a practical example from a language center in Brazil, this paper addresses how teachers/practitioners can make informed decisions to write or select their own indicators for placement testing and, therefore, monitor the progress of their students and/or the success or pitfalls of their classes/programs/curricula and plan for possible welcome changes.*

### **Keywords**

*Placement Testing, Assessment, Indicators*

## Summary

Through a literature review and with a practical example from a language center in Brazil, this paper addresses how teachers/practitioners can make informed decisions to write or select their own indicators for placement testing and, therefore, monitor the progress of their students and/or the success or pitfalls of their classes/programs/curricula and plan for possible welcome changes.

## 1. Introduction

According to Raupp and Reichle (2003), in order for one to obtain efficient results in assessment, it is necessary to work with some guiding principles. The aforementioned authors recommend four principles: 1) assessment as a collaborative process of exploration, examining, questioning and search for other perspectives, 2) the availability and the use of valid and reliable information to make the process more efficient, 3) the use of the information obtained as the key point of the process; and, lastly, 4) assessment as a powerful and valuable catapult to increase the efficiency of projects and programs (p. 18).

Raupp and Kolb (1990) point out that, before designing an assessment system or instrument, one should consider the reason or reasons to assess. The authors list 10 possible reasons and claim that an assessment proposal may fall into one or more categories. The one we shall be focusing on this paper is number 2, “to make informed decisions”. Just as one ponders *why* to assess, they must also consider *what* to assess. Raupp and Kolb (1990) affirm that the information collected in the assessment corroborate a previous idea about the object being assessed. In other words, when placing an incoming student into a language program, the student might present certificates or diplomas from other schools, which indicate his/her overall proficiency level. A well-planned and well-executed placement test will collect valuable data that will confirm or refute that information and help the examiner place that student into the current language program.

When a patient goes to the doctor seeking a diagnosis and treatment to a health problem, he/she presents precious health information that help the doctor diagnose the problem and prescribe an effective treatment. The description of the symptoms, clinical exams, the history of the patient and the experience of the doctor will collaborate in the recovery of the patient. Currently, there is software that helps the doctor diagnose patients. This software provides a more precise diagnosis as it analyzes the data collected and contrasts with medical notes, medicine books, etc. After this thorough analysis, the software offers the doctor a list of possible health problems and prescriptions for the doctor to close the diagnosis. The doctor, then, decides on *how* to treat the patient. A placement test works similarly to the work of this software, but in a more analogical way.

The problem we faced at the Languages Center (CL) of the Federal University of Espírito Santo (UFES, Brazil) was very much like that. The discontentment of instructors and coordination staff with the lack of reliability, validity and practicality with the CL’s placement test motivated debates and work on the refinement of the indicators used for the referred exam. Also, the staff saw the need to refine the test

instrument in order to align with the refined indicators.

The CL was established in 1997 as a university extension project to offer language courses to the academic community at UFES and the community at large at a reasonable cost. Additionally, the CL provides an intensive language teacher education program for the Letters students at UFES. The CL holds an average of 5,000 students and 60 teachers (permanent and trainees). Since its conception the CL offers regular language courses that range from 8 semesters to 10 semesters of duration. The largest program is the English program, with the highest number of students, teachers and coordinators. For the purposes of this paper, we will be addressing the challenges faced by the English department only.

In the early 2000's, the CL hired an assessment specialist to assist the teachers and coordination staff transition from a paper-and-pencil assessment to a performance-based assessment system in the four abilities, namely, reading, writing, listening and speaking. The urge to change grew out of a dissatisfaction of both teachers and students with the mismatch between teaching and testing in the school. At that time, a brand-new performance-based assessment system was developed with the participation of permanent teachers and trainees. The system was and is used across language departments at the center. A principal document was elaborated with learning objectives for, mainly, three levels: beginner, intermediate and advanced. The document reflected the school's views on language and language learning and had the communitive approach as its central theoretical basis. An assessment matrix, ranging from grade A to grade D, was designed with indicators that reflected the expected performance of the test-takers in the four abilities. The instruments were also redesigned and included tasks for reading comprehension, listening comprehension, writing expression and oral expression. As for the placement tests for incoming students of the English department (focus of this paper), a multiple-choice exam with 100 questions was created and encompassed a listening comprehension section, a grammar and vocabulary section and a reading comprehension section. There was also a second type of placement test which consisted of an essay exam and an oral interview.

In 2013, due to the university's internationalization process, the CL felt the need to adjust its assessment documents to international documents like the Common European Framework of Reference for Languages (CEFRL). So, the original documents that divided the benchmark levels were adapted to A1—corresponding to the beginner levels which encompassed levels 1, 2 and 3—, A2—corresponding to the pre-intermediate levels which encompassed levels 4, 5 and 6—, B1—or the intermediate levels, englobing levels 7 and 8—, and, lastly, B2—which stands for the upper-intermediate levels 9 and 10. The language courses that were only 8 semesters long adopted a different configuration. Since this time, the CL teachers and coordination staff have debated on the refinement of the assessment matrix and, consequently of the instruments they have used, including the placement tests.

The motivation to rethink and, possibly, redesign the assessment documents and placement testing instruments of the English department came up as teachers and students reported that students were being misplaced over the course of years. In many occasions, the teachers, along with the coordination staff, had to adjust students' levels after the placement testing procedures. When this situation became too

frequent, teachers and coordinators decided it was time to change (Note 1).

The laborious task of refining the assessment documents started back then when an assessment committee was formed and when it was decided that work should begin with an immersion in the literature of the area in order to make informed decisions. A series of teacher education sessions on assessment was carried out with different specialists and members of the staff. According to Raupp and Reichle (2003), for the daunting task of (re)designing assessment documents to be successful, it is essential that the designers know what they are aiming for. In order for one to know what a (language) program goal is, the objectives or indicators need to be carefully and clearly written. That was the beginning of our work.

The objective of this paper is to present the refinement process of the indicators of the placement tests used in the CL at UFES. After an in-depth literature review and the analysis of assessment documents both from the institution and also of public domain (like the CEFRL), the designers began working on the assessment documents, which culminated in a revisited and calibrated assessment matrix and new placement test instruments. The research and data collection process was a collaborative task amongst the members of the CL staff, including myself, reporting the results of this endeavor.

## **2. Methodology**

Given the history of the CL and the problems or difficulties the staff was facing with placement testing, teachers and coordination staff decided to delve into the literature of the area, which included theories, public documents like the CEFRL and internal assessment documents. As the team advanced in their readings and discussion prior to the beginning of the heavy work, the methodology to carry out this task became clearer to all. The team decided for a qualitative approach (Creswell, 2007) for its participatory and self-reflective nature (Rossman & Rallis, 1998).

The team did not have a research question in mind, but a hypothesis that the assessment matrix being used to evaluate the incoming students was not really aligned with the learning objectives of the course nor with the test instruments being used. If this hypothesis could be translated into research questions, something along these lines would come up: To what extent are the assessment documents (matrix and instruments) influencing the misplacement of incoming students? How can the calibration of these documents improve results of placement testing?

The characteristics of the qualitative research fit well into the scenario that the CL teachers and staff were inserted. They were inside the problem and could have access to information first-hand. As Creswell (2007) points out, qualitative research utilizes multiple methods of data collection and are more interactive and humanistic (p. 188). According to Creswell (2007), the analysis of documents, interviews and observations are some of the data collection methods applied in this kind of research. Another trait of the qualitative research is the flexibility of the process of data collection, which may change as the researcher encounters better opportunities to understand better the phenomenon in question.

The CL research was of interpretative nature. That means that the investigator makes an interpretation of the phenomenon (or the data) through his/her own lenses (Creswell, 2007). In the case of the CL endeavor, the researchers (considering that all members of the staff contributed to the construction of this research) were very experienced teachers and coordinators, who have participated in the construction of the assessment documents of the CL and also had a lot of background knowledge on assessment. So, they were very aware of the context in which they were inserted to make the conclusions they needed to make and the alterations or calibrations they need to make.

As per the phases of the research, first, the team observed the inconsistency of placement test results for some time (recurring topic in staff meetings and in the assessment committee meetings) and decided to invest time and effort in digging the problem to its roots. We call it the *Observation Phase*. Next, in phase 2, the team decided to delve into the literature of the area and into public and internal assessment documents in order to build on their expertise and find a direction to the problem posed. Staff meetings to discuss an appropriate manner to handle the problem were scheduled regularly. We call it the *Theoretical Embedding phase*. Then, in phase 3, using the data they had collected, the team started calibrating the internal assessment documents and designed one specifically for placement testing. We call it the *Calibration Phase*. Finally, in phase 4, the team began designing new test instruments for placement tests and piloting them in subsequent semesters. This was the *Design and Piloting Phase*. The four phases of the project were completed between 2018-2019.

In the next section, a literature review is presented aiming to show the reader the theories underlying the changes made by the CL team. This body of work shed some light into the doubts the team had about how the assessment documents were interfering in the placement test results and served as guiding principles for the calibration of the internal documents.

### 3. Literature Review

As mentioned elsewhere, the objective of this literature review is to provide the theoretical foundations for the assumptions made by the CL team in regards to the institution's assessment documents and their direct or indirect influence in the placement test results. As part of the research endeavor carried out by the team, they went back to the theory, from placement testing (Brown, 2004) to reference documents like the CEFRL (2001) to find the answers they needed.

Coombe, Folse and Hubley (2007) start off by listing eight theoretical concepts, namely: 1) *utility*: supported by Bachman and Palmer's (1996) test instrument utility model, Coombe, Folse and Hubley (2007) attest that this is the most important principle to give legitimacy to the evaluation process, that is, an evaluation instrument must be created for a specific objective and with a target audience in mind; 2) *validity*: refers to its purpose, that is, whether it evaluates what it intends to evaluate, one of its ramifications being construct validity, that is, whether the methodology chosen to evaluate is consistent with the instrument used; 3) *reliability*: refers to the consistency of the instrument, that is, if it presents the same results (or similar results) if applied on different occasions and by different examiners; 4)

*practicality*: when thinking about the evaluation instrument, one must also consider the costs for development and maintenance, the administration time, the necessary resources, among other aspects, and for a large-scale evaluation it is necessary to determine practical factors for the collection of information about the object to be evaluated; 5) *washback*: it is related to the effect that the evaluation results produce, which can be positive or negative, and one should not conduct the evaluation in a way to mask the results for both positives and negatives; 6) *authenticity*: refers to the fact that the instrument is linked to a real situation, a real objective such as pointing out solutions or directions for a current problem; 7) *transparency*: it relates to the clarity of the information to be collected, the way it will be collected and why it will be collected, which should be made available to the parties involved in the evaluation process and 8) *security*: it relates to the preparation and administration of the evaluation instrument and in the case of large-scale assessments, for example, it is common for an instrument to be designed and administered by a group of people, which can compromise both the confidentiality of the instrument and the provision of information to the target audience, which may mask the results obtained (p. XXII).

Brown (2004) suggests the foundations of designing and revising tests by providing 5 critical questions: 1) What is the purpose of the test? 2) What are the objectives of the test? 3) How will the test specifications reflect both the purpose and the objectives of the test? 4) How will the test tasks be selected and the separate items arranged? What kind of scoring, grading, and/or feedback is expected? (p. 42). Knowing the purpose of a specific type of test helps to identify the foundational principles of assessment, namely, utility, validity, reliability, practicality, authenticity, transparency, security and wash back and, therefore, helps to determine whether a test is effective or not.

The purpose of a placement test, for example,

[...] is to place a student into a particular level or section of a language curriculum or school. A placement test usually, but not always, includes a sampling of the material to be covered in the various courses in a curriculum; a student's performance on the test should indicate the point at which the student will find material neither too easy nor too difficult but appropriately challenging (Brown, 2004, p. 45).

So, the placement test must fit the program or curriculum for which the incoming student is applying. There are some standardized tests available (some even online) which may be very practical for the institution administering the exam, but they might not really reflect the goals of the curriculum the student is going to encounter in that specific school or program.

Richards (2009) explains that most schools and language programs describe their curricular goals in terms of learning objectives and aims. The author claims that “[T]he nature of aims and objectives, however, is not necessarily straightforward because they refer to knowledge, skills, and values that educational planners *believe* learners need to develop” (p. 112, emphasis in the original). This means that the role of teaching and learning of a language is reflected in the curricular goals of a language program. Richards (2009) helps us understand the complex task of describing curricular goals by

giving us three common assumptions: 1) People are generally motivated to pursue specific goals; 2) The use of goals in teaching improves the effectiveness of teaching and learning; and 3) A program will be effective to the extent that its goals are sound and clearly described.

As mentioned elsewhere in this manuscript, stemming from a dissatisfaction of both teacher and students about their assessment system and teaching practices, the CL at UFES started rethinking its curricular goals back in the early 2000's, which resulted in the design of a brand-new performance-based assessment system. Part of the tasks back then was to define curricular goals for the language program and design institutional assessment documents to guide the staff. After the adjustment and refinement of the documents around 2013, the staff from all departments felt the need to revisit the assessment documents. This refers to Phase 1 and 2 described in the methodology section. Alongside this movement of change, the teachers and coordination staff of the English department began to work on the placement testing documents and procedures (Phase 3) (Note 2).

During Phase 3, it became very clear to the English department staff that considerable time and effort should be spent on the indicators or descriptors (We will use indicators henceforth.) that were the direct link to the program curricular goals. According to Almeida (2018), important and valuable data can be collected through indicators, which can be assessed both quantitatively and qualitatively. Indicators present evidence of whether a goal was achieved or not.

A document that served as reference and model for the English department staff when redesigning their own indicators was the CEFRL. The CEFRL was intentionally created with an action-oriented approach to aid language practitioners to elaborate their own frameworks, according to their specific needs and contexts across Europe. The manuscript became largely diffused and has served as theoretical basis for language programs around the world. Here is the opening paragraph of the document:

The Common European Framework provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe. It describes in a comprehensive way what language learners have to learn to do in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively. The description also covers the cultural context in which language is set. The Framework also defines levels of proficiency which allow learners' progress to be measured at each stage of learning and on a life-long basis (CEFRL, 2001, p. 1).

Barrett *et al.* (2018) put forth some criteria for the formulation of indicators, namely: 1) brevity, that is, having a maximum of 25 words; 2) positivity, expressing a skill or goal in affirmative sentences; 3) clarity, that is, no jargon and simple structure; 4) independence, in other words, being independent of another indicator; and 5) accuracy, that is, describing whether a certain behavior or a certain goal was achieved or not.

Taking the literature from the area and the analysis of the internal and public assessment documents into account, the CL staff initiated Phase 4 which encompassed the design of new assessment instruments aligned with the redesigned assessment documents. In the following section, we present one of the results of this research endeavor, specifically taken from the English department staff. This department engaged in producing an assessment matrix for placement testing aligned with the other institutional assessment documents which, in turn, were aligned with the revised program curricular goals. As we present the assessment matrix for placement testing, we will make the connections with the literature.

#### **4. Results and Discussion**

As mentioned before in this manuscript, the main outcomes of this project were the redesign of the institutional assessment documents and, subsequently, the assessment instruments. The English Department at the CL, in particular, underwent a surgical process of (re)construction of the assessment matrix for placement testing and its instruments, which will be used here as illustration of how the staff used the theory and knowledge of the program curricular goals to attenuate the issue they were having with misplaced students.

The process of rethinking the assessment documents for placement testing started with the redesign of indicators for that purpose, which had to be carefully and meticulously aligned with the course curriculum. Before we analyze the process of creating the indicators, we shall describe the redesign of the placement testing instruments. There are two types of placement tests at the CL's English Department: 1) one multiple-choice exam with a reading section, a grammar and vocabulary section and a listening section; and 2) one with a written test (essay) and an oral interview.

As part of a screening process prior to the exam, the incoming students fill out an application form in which they describe their language learning experience. Depending on the information they fill in the form, the coordination staff will assign test 1 (multiple-choice) or test 2 (essay and interview) for the candidate. Test 1 is geared towards candidates who have had less numbers of hours of instruction and test 2 is assigned to candidates who have had an extensive number of hours of instruction (around 300h).

Test 1 was completely redesigned question by question going from a 100-item test to an 80-item test (with a listening section, a reading section and a grammar and vocabulary section), which contained questions from all across the curriculum (Brown, 2004) and were aligned with curricular goals of teaching and learning (Richards, 2009). Decreasing the number of questions made the test more practical to administer (the CL administers up to 250-300 tests each turn—usually happening twice in the academic semester), which increased security, validity, reliability and provoked positive washback (Coombe, Folse & Humbley, 2007).



Test 2 was completely redesigned as well. The test is still divided between an essay part and an oral interview. This time, candidates are presented with 2 images and guiding questions and have 40 minutes to write a 150-word-minimum essay about 1 one them. Upon finishing their writing exam, candidates are directed to the interview in which they are supposed to talk about the same image they have chosen in the writing part. The role of the interviewer is more of a facilitator to welcome the candidate, to unwind by talking about personal information and then approaching the topic with some discussion questions (the same as in the writing part). The interviewer is only supposed to interfere if he/she perceives that the candidate needs, and will benefit from, a supportive or provocative question. The two parts of the test add validity and reliability to the test as it provides more data for a better-informed decision about the level of the candidate. The examiner has the chance to verify and explore both in written and also in oral form different structures of the English language, from simple to more sophisticated, vocabulary range, pronunciation and, ultimately, the candidate's command of the language (CEFRL, 2001; Coombe, Folse & Humbley, 2007).

The meticulous process of designing the indicators for placement test 2 (written test and oral interview) was informed by the literature presented here and resulted in an assessment matrix for written expression and another one for oral expression. These two documents have helped the staff attenuate the problem of misplaced students in the sense that: 1) the indicators have been aligned with the course or program curricular goals (Richards, 2009); 2) with this alignment of the indicators and curricular goals, validity and reliability (Coombe, Folse & Humbley, 2007) were increased; 3) the specifications or the indicators of the test reflect its purpose (Brown, 2004); 4) the indicators echo the specifications of the CEFRL (2001), making it easily relatable to standardized proficiency levels; 5) the indicators are carefully and accurately written expressing the skill or goal the learner is expected to achieve (Barrett *et al.*, 2018); and finally, 6) because this project was undertaken as a collaborative task, the staff of the English Department was able to have a more comprehensive understanding of the complexity of why, what and how to assess in order to get better results (Raupp & Kolb, 1990; Raupp & Reichle, 2003).

Table 1 shows the Oral Assessment Matrix for Placement Testing, one of the products of the project, which is now in its Phase 4, the piloting stage. Some adjustments have been made on the punctuation for the levels at the CL and in the instruments of Test 2. Candidates are now required to take the multiple-choice test, then the essay exam and, lastly, do the oral interview. There was consensus that the candidates screened for test 2 required that more information was collected in order for the examiner to make a more informed decision about his/her level (Raupp & Kolb, 1990; Raupp & Reichle, 2003). The document was designed in Portuguese, candidates' and teachers' native language, for the sake of transparency and also for it to be a model for the other departments at the CL.

**Table 1. Placement Testing Matrix for Oral Assessment**

Níveis/Faixas de Pontuação/Pontuação	VALORES	FLUÊNCIA	VOCABULÁRIO	GRAMÁTICA	PRONÚNCIA
B2 13-16 pontos ING 9 e ING 10	4	Discurso com certa espontaneidade, coerente, com raras pausas para aprimorar e/ou reformular ideias, sem prejuízo ao fluxo da fala.	Expõe opinião sobre temas globais e ideias abstratas com variedade lexical.	De modo geral, faz uso de estruturas complexas com raras inadequações que não comprometem a inteligibilidade da fala.	Pronúncia clara com entonação próxima da língua alvo. Raros lapsos que não comprometem a inteligibilidade da fala.
B1 9-12 pontos ING 7 e ING 8	3	Discurso coerente, porém, simples, com pausas curtas para resolver algum problema de construção linguística, exigindo algum esforço para expressar-se.	Expõe posicionamento sobre temas cotidianos e ideias em geral com léxico simples.	Faz uso de estruturas gramaticais complexas com algumas inadequações decorrentes de interferência de L1.	Pronúncia, em geral, clara, com entonação regular e lapsos pontuais.
Menor ou igual a ING 4 e ING 3	2	Discurso compreensível com pausas evidentes para resolver algum problema de construção linguística, que causam interrupções frequentes ao fluxo da fala.	Discorre sobre temas do cotidiano e expressa ideias com léxico simples e interferências ocasionais da L1.	Faz uso de estruturas simples com inadequações decorrentes de interferência da L1 e que, por vezes, comprometem a inteligibilidade da fala.	Pronúncia com forte interferência da L1 e/ou de outras línguas que, por vezes, compromete a inteligibilidade da fala.
A decisão final se dará a partir do alinhamento entre Entrevistador e Corretor.	1	Discurso com constantes pausas e hesitações, que causam interrupções ao fluxo da fala.	Expressa ideias simples com frases curtas, com interferências constantes da L1.	Faz uso de estruturas sem flexão de tempos verbais e/ou concordância, com ordem invertida de palavras.	Pronúncia com forte interferência da L1 e/ou de outras línguas que frequentemente afeta a inteligibilidade da fala.

Source: Núcleo de Línguas, Programa Cursos de Línguas (UFES).

## 5. Conclusion: Drawbacks & Affordances

In this paper, we described the process of redesign of indicators of the assessment documents of the Languages Center at the Federal University of Espíto Santo, Brazil. The staff of the English Department, in special, took on an extra project, which was to redesign the placement testing documents. The desire to implement such project grew out of the observation from the teachers and coordination staff of misplaced students after an adjustment to the assessment system around 2013. After observing the phenomenon, the English Department staff took action and set off on a qualitative research approach which encompassed a theoretical embedding phase, a calibration phase and a piloting phase. The outcome was a new assessment matrix for placement test type 2 (see above) for oral and written expression and also new placement testing instruments. As per the hypothesis and research questions we posed, as we observe the consistency of results of placements for the past year, this suggests that the assessment documents, indeed, to some extent influenced the misplacements happening before the changes. However, a more detailed study should be carried out to confirm.

As for the drawbacks encountered on the way, speaking specifically of the English Department and their attempt to redesign the placement testing documents, because this is the largest department at the CL, with the greatest concentration of novice teachers, it was (still is!) difficult for novice teachers to conduct the placement test (type 2) using these documents. The novice teachers may not have a very comprehensive understanding of the program curricular goals and how they branch out into the benchmark levels. Not only do they have little grasp of the curricular goals for different proficiency levels, but also, they had little experience teaching the levels of the course to see it in practice. This fact concentrated the work of placement tests on more experienced teachers, who represent a smaller contingent at the CL. Ideally, all teachers should be able to conduct the placement test. This issue has been recognized and training sessions are in order to tackle the problem. A strategy used when designing the indicators was to attribute numerical values for the areas and indicators so that the sum of the indicators represents the level placed. Another difficulty faced by the staff when designing the indicators was to make them independent from one another (Barrett *et al.*, 2018). Because we understand the skill or linguistic knowledge being assessed in holistic way, it was hard not to make one indicator dependent on the other. There were countless and long debates on one single word or line, which made the process quite exhausting.

On the brighter side of things, we can count some affordances from this undertaking. The staff has noticed less complaints about misplacements both from teachers and from students. There have been, of course, some small level adjustments, which are part of the placement process. As mentioned elsewhere in this paper, the fact that this was a teacher-led project, as opposed to a top-down decision made by the coordination staff, the redesign of the indicators was beneficial to all as we had to go back to the foundations of assessment and reassure our language teaching and learning philosophies. This exercise helped us look at both the internal (e.g., the CL's assessment documents) and the public documents (e.g., the CEFRL) and make the necessary links between them to have a clearer idea of what

our program curricular goals are and also to have a more global view on language proficiency. In times of the internationalization of education worldwide, this understanding is crucial to make the institution move forward. Finally, the experience of designing indicators for a placement test in dialogue with internal and external documents gave us the experience to work with assessment grids, which are used worldwide to assess students and help us minimize the subjectivity of any assessment endeavor.

### Acknowledgements

The author would like to acknowledge the work of the UFES' Languages Center English Department staff during this project. Special thanks to: Sandra Prates, Leni Puppini, Wendell Máximo, Gustavo Reges, Danilo Marinho, Manoel Sampaio Júnior, Danuza Fonseca e Sandlei Moraes.

### References

- Almeida, J. (2018). *Approaches and Tools for Internationalisation at Home*. Bolzano, Italy. 08 set. Conference.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Barrett, M. et al. (2018). Reference Framework of Competences for Democratic Culture, Descriptors for democratic culture. *Council of Europe, 2*.
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices: United States of America*. A Pearson Education.
- Coombe, C. A., Folse, K. S., & Huble, N. J. (2007). *A practical guide to assessing English language learners*. University of Michigan Press.
- Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- Creswell, J. W. (2007). Projeto de pesquisa métodos qualitativo, quantitativo e misto. In *Projeto de pesquisa métodos qualitativo, quantitativo e misto*. Artmed.
- Raupp, M., & Kolb, F. A. (1990). *Evaluation Management Handbook*. Andover, MA: Network, Inc.
- Raupp, M., & Reichle, A. (2003). *Avaliação: Ferramenta para melhores projetos* (p. 251). Santa Cruz do Sul: EDUNISC.
- Richards, J. C. (2001). *Curriculum development in language teaching*. Ernst Klett Sprachen. <https://doi.org/10.1017/CBO9780511667220>
- Rossmann, G. B., & Rallis, S. (1998). Analyzing and interpreting data. In G. B. Rossmann, & S. Rallis (Eds.), *Learning in the field: An introduction to qualitative research*. Sage Publications.

**Notes**

Note 1. The CL uses a unified performance-based assessment system and the staff from all departments felt the need to change/redesign/refine the assessment documents. The staff of the English department went a step further and decided to invest time and energy on assessment documents for its placement testing procedures. In this paper, we are referring to the whole process of refinement of the documents, but with a special focus on the placement testing procedures and documents of the English department.

Note 2. There were two tasks going on at the same time: 1) the redesign of the institutional assessment documents being carried out by the CL's whole staff, and 2) the redesign of a specific assessment document/matrix for placement testing purposes being carried out by the English department staff only.

We are focusing on the latter.