THE INDUSTRY OF TESTING AND MUTUAL RECOGNITION IN SOUTHERN EUROPE

Joaquín Manuel Cruz Trapero

Universidad de Jaén

Abstract

Following the publication of the Common European Framework of Reference, Spain’s central government has been unable to unify common policies for foreign language requirements in Spanish Higher Education. Language requirements are in Spain both partially unregulated and deemed compulsory to finish Higher education degrees. In the absence of national policies, different areas of the country have implemented different policies in this respect and, inevitably, Spanish universities have developed their own internal legislation, high-quality suites of language tests and mutual recognition systems. Spanish university language tests vary greatly from one to another, and are now competing in the industry of testing at both a national and a global level. This variegation has hampered mutual recognition and, consequently, academic and professional progress of stakeholders. The present paper focuses on the work developed by public universities in Andalusia, where we carry out our professional activity, to overcome the aforementioned problems. The universities in Andalusia, one of the 17 autonomous communities of Spain, have been able to agree on a set of unified criteria which favors mutual recognition of language certifications for +13K candidates on a yearly basis and have driven policy makers’ attention to their network, which is now considered as a benchmark.

Keywords: Industry of testing, mutual recognition, Spain, Andalusia, CEFR, reliability, validity, fairness.

Resumen

Tras la publicación del Marco Común Europeo de Referencia para las Lenguas, el gobierno central español no fue capaz de homogeneizar sus políticas en materia de requisitos lingüísticos para la Enseñanza Superior. Estos requisitos, que no están regulados en muchos casos, son al mismo tiempo obligatorios para concluir estudios superiores en la mayoría de las universidades de España. Ante la ausencia de políticas nacionales unificadas, diferentes comunidades autónomas han desarrollado distintas políticas en materia de acreditación lingüística y reconocimiento de títulos. Muchas instituciones han desarrollado, en paralelo o de forma conjunta, exámenes de acreditación lingüística de calidad que son reconocidos en diferentes ámbitos. Los exámenes de acreditación lingüística de las diferentes universidades españolas compiten ahora con otras pruebas tanto en el mercado nacional como en el internacional. Este artículo se centra en el trabajo desarrollado por las universidades públicas andaluzas en
materia de acreditación y reconocimiento lingüístico. Las universidades de Andalucía, una de las 17 comunidades autónomas de España, han sido capaces de acordar criterios comunes de confección y corrección de pruebas que favorecen el reconocimiento mutuo de los certificados expedidos por cada una de ellas. A su vez, dichos criterios favorecen a los más de 13.000 candidatos que anualmente realizan alguna prueba de acreditación lingüística en una universidad andaluza. El trabajo de las universidades andaluzas hacia la excelencia en las pruebas de idiomas y el reconocimiento mutuo no solo ha llamado la atención de los agentes políticos sino que, además, se ha convertido en un referente en España.

**Palabras clave:** industria de la evaluación, reconocimiento mutuo, España, Andalucía, MCER, fiabilidad, validez, equidad.

1. **Testing**

   Modern language testing is measured along two axes, psychometrics and social impact. These give raise to four main areas of interest, namely reliability, validity, fairness and results linkage. Reliability is concerned with the absence of measurement errors once repeated measurements have been carried out with the same instrument, validity focuses on the extent to which our tests actually measure what they are supposed to measure, while fairness guarantees equal conditions to all stakeholders regardless of their sex, age or cultural background. The last of the four big aspects, results linkage, is area-dependent, and is germane to the most widely used scales in each country or group of them, the CEFR being the main referent in Europe and, increasingly, in South America and Asia.

   Bachman points out that “testing almost never takes place in isolation. It is done for a particular purpose and in a specific context” (2). Paraphrasing Skinner’s opening lines to *Verbal Behavior* (1), we might say that tests act upon the world, and change it, and are changed in turn by the consequences of their impact. Like verbal behavior itself, tests are operants that modify the environment, the context in which they occur.

   Departing from an analysis of the current state of the art in the industry of testing, the present paper analyses the main drawbacks which nine universities from southern Europe (the context) have faced over the course of their ongoing race towards a common standard for foreign language tests and mutual recognition, together with the solutions that they have been able to find for their stakeholders (the consequences).

1.1. **The industry of testing**

   Some tests are certainly high-stake. They may be a requirement for obtaining employment, or provide the key for immigrants wishing to gain entry to a country. Besides those who simply seek to measure their linguistic abilities
for personal achievement, tests are likely to interest millions of people hoping for a place at college, for a better job or for a better future. In the USA, for example, TOEFL tests are a must-have for foreign students who seek to access universities. In 2014, according to internal documentation of their Spanish branch, ETS implemented 50M tests in more than 180 countries. Cambridge tests, which celebrated their first hundred years in 2013, are taken by 4M candidates in more than 130 countries. Out of these candidates, 250K are Spanish. Such figures are a remarkable success in marketing and adaptation for Cambridge, when we consider that the first modern Cambridge test, held in 1913, lasted 12 hours and was taken by only three candidates, all of whom failed (Hawkey and Milanovic 22). Cambridge and TOEFL are just two examples out of the many existing test boards (ACTFL, TOEIC, EIKEN, IELTS, Apts, Trinity, DELE, eLade, etc.) that can give an idea of the impact that tests have for candidates after policy makers, at all sorts of institutions, decide to turn them into a requirement, into keys that open doors to immigration or to academic and professional progress.

Political moves towards a particular test may define, in a way that no other theoretical approach can, test takers’ preferences for one particular brand. The industry of testing moves millions of euros worldwide and has even borrowed many practices typical of pharmaceutical marketing. The expectation generated amongst platinum centers, distribution centers that conform to Cambridge’s most ambitious business development programs, is just one measurement of the current industry of testing. Trinity tests, for example, have gathered momentum in the Spanish market following their recognition by a number of important institutions. In Spain, Trinity tests have surpassed in popularity ETS and even the most popular suite of exams in the past twenty years in Spain, Cambridge. Leaving aside reliability and construct validity concerns, Trinity has been quick at meeting the requirements of a region, Andalusia, with a poor tradition in foreign language recognition, and has been able to build very powerful face validity through aggressive marketing campaigns. Once again, Trinity and its market penetration in Andalusia are just one example of the many tests which are nowadays ubiquitous around the world. With different degrees of penetration, we are now living the heyday of language testing.

In this context, test developers have to balance two unrelated and inharmonious fields, linguistics and psychometrics, at the same time as coping with institutional, economic, social and political demands (Spolsky 4) and, of course without losing sight of the competitive market of which they are a part. The different big test brands which operate worldwide are constantly competing for the market while they struggle to maintain reliability, validity, fairness and linkage standards. This fight is shaping the market and the industry of testing. In this context, smaller initiatives, less ambitious but equally necessary tests, also struggle to operate, as in the case of Andalusia.
2. The context of Europe

The idea of Europe has not always been the same. The modern process of European construction was launched back in 1949 when the Council of Europe was founded. After two world wars, the Council of Europe aimed at protecting democracy and human rights and at promoting European unity by fostering cooperation on different matters. The European Cultural Convention of the Council of Europe of 1953 was the first of a series of acts and treaties oriented to set forth the fundamental rights and freedoms which were the core concern of post-war Europe. Culture and eventually language were among these concerns as well.

After some initial attempts, it took European higher education almost 40 years to unify cultural goals. It was not until 1988 when the rectors of 388 universities signed the Magna Charta Universitatum (1988) (MCU henceforward) on the 900th anniversary of the University of Bologna. The MCU was a two-page document designed to lead Europe into a culture-based new millennium which contained principles of academic freedom and autonomy as a guideline for good governance and mutual recognition of universities. Nine years later, in 1997, the Council of Europe and the UNESCO drafted the Lisbon Convention (1999), which was designed to streamline the legal framework at a European level and to replace six previous conventions in matters of higher education. In 1998, four education ministers (from France, Germany, the United Kingdom and Italy) participating in the celebration of the 800th anniversary of the University of Paris shared the view that the segmentation of the European higher education sector was outdated and harmful. As a consequence, they agreed to sign the Sorbonne Declaration (1998). The document put forward a number of ideas about the European Credit Transfer and Accumulation System (ECTS) as well as the two main cycles of the system, undergraduate and graduate, and, for the first time, officially called for a recognition system able “to remove barriers and to develop a framework for teaching and learning, which would enhance mobility and an ever closer cooperation” (1998:1). One year later, in 1999, the European Higher Education Area (EHEA) was finally shaped, through the signing of the Bologna Declaration (1999). Thus the idea of mutual recognition was officially born in Europe in 1999 with the Bologna Declaration, three-quarters of a millennium after the first universities came into being in the Old Continent. In retrospect, the Bologna Declaration has become one of the most influential documents in the modern history of European higher education. After 1999, other communiqués have been issued (Prague, 2001; Berlin, 2003; Bergen, 2005; London, 2007; Leuven/Louvain-la-Neuve, 2009; Bucharest, 2012 and Yerevan, 2015), articulating the previous agreements, including some partners from beyond Europe, and weaving a brand new European network. European EHEA policies have also been updated at the Budapest-Vienna Ministerial Conference (2010) and the Bucharest Ministerial Conference (2012) as well as through further Declarations (Salamanca, 2001; Graz, 2003; Glasgow, 2005; Lisbon, 2007 and Budapest-Vienna, 2010) and through three Bologna...
Policy Forums (2009, 2010 and 2012). All these are ordered chronologically in the following timeline:

- 1949: The Council of Europe is founded.
- 1953: European Cultural Convention of the Council of Europe
- 1988: Magna Charta Universitatum
- 1997: Lisbon Convention
- 1998: Sorbonne Declaration
- 1999: Bologna Declaration
- 2001: Prague Communiqué
- 2001: Salamanca Convention
- 2003: Berlin Communiqué
- 2003: Graz Declaration
- 2005: Bergen Communiqué
- 2005: Glasgow Declaration
- 2007: London Communiqué
- 2007: Lisbon Declaration
- 2009: Leuven/Louvain-la-Neuve Communiqué
- 2009: First Bologna Policy Forum (Leuven/Louvain-la-Neuve)
- 2010: Budapest-Vienna Declaration
- 2010: Second Bologna Policy Forum (Vienna)
- 2012: Bucharest Communiqué
- 2013: Third Bologna Policy Forum (Bucharest)
- 2015: Yerevan Communiqué

The Council of Europe “provides a pan-European forum for sharing expertise and experience based on common values and respect for the diversity of contexts” (Extra and Yağmur 7). It acknowledges the particularities of the state members and lessens privileges among them. All in all, the most influential of the achievements of the Council of Europe has been, perhaps, the confection of the CEFR.

### 3. The context of Spain

In Spain, where the boom in language testing is unprecedented, marketing arguments seem to be leading the choices of test takers in the first part of the 21st century. Some tests are held in massive venues such as hotels or trade fair parks which host thousands of candidates. Sometimes, as massive numbers of test takers make their way towards test venues, it is hard to tell whether they are actually going to sit a test or to watch the local football team in the final match of a major competition. Such is the amount of people that language tests are able to bring together.

The washback from these tests has also been important in a country that relied heavily on traditional methods of language teaching, certification and accreditation. The language teaching methods used in Spain until very recently
were inherited from the Grammar-Translation principles which were used extensively in the teaching of dead languages, but are decidedly unsuited to teaching modern, living ones. Communicative approaches were only introduced in Spain in the last decade of the 20th century. Foreign languages were taught at schools with little or no emphasis on communication and this produced several generations of Spanish students who were able to read or write, but unable to speak in the foreign language they had been studying for years. Catching up with the rest of Europe has necessitated profound changes in the mindsets of professionals and, even nowadays, at times, Spain seems to be stuck in second gear while the rest of Europe is working at full speed.

Spanish policy makers have passed seven different major educational laws over the past 33 years, each of them intended to substitute the previous one: LGE (1970), LOECE (1980), LODE (1985), LOGSE (1990), LOCE (2002), LOE (2006) and LOMCE (2013). To complicate matters further, there are 17 autonomous communities in Spain, all of which have their own laws on education (see section 4). The resulting variegation has hampered mutual recognition of foreign language levels. As a result, depending on the community chosen, the linguistic proficiency of two different students may vary by up to two CEFR levels in the same academic year. It is because of this lack of intra-regional standardization that it has become necessary to establish external language tests whose results can clearly be linked to the CEFR.

The recent history of this situation dates back to the year 2007, when the Spanish Ministry of Education and Science passed a Royal Decree (Real Decreto 1393/2007) which regulated official higher education in Spain. This decree set the future for a series of new university degrees by recognizing the importance of the European educational policies generated following the Bologna Declaration, referred to in the first paragraph of the decree. In this decree, the references to foreign languages are vague and yet, through the acceptance of the European policies derived from the Bologna Declaration, it implicitly agrees upon the importance of foreign languages in higher education for transnational mobility of students and for their employability. In terms of mutual recognition, the decree proposes using ECTS credits.

At a different level, on November 18th 2010, following European regulations and recommendations, the CRUE (Conferencia de Rectores de las Universidades Españolas, Spanish University Rectors’ Conference) commissioned a report on language teaching and accreditation which was drafted in February 23rd 2010 and finally passed at the General Meeting held by the CRUE on September 8th, 2011 in Santander, Spain. This report, a type of unintended follow-up of Real Decreto 1393/2007, pointed out that there existed considerable diversity in procedures and requirements, and that this lack of homogeneity was leading to confusion. The report, entitled Propuesta sobre la acreditación de idiomas (PAI henceforward), also highlighted the fact that training and certification were not always differentiated in Spain, and wished that all universities integrated in the CRUE should issue language certificates which would be mutually recognized at both a national and an international level. To
reach these conclusions, the CRUE took into account the experience of 50 Spanish universities and other educational institutions. At the same time, they agreed to work towards mutual recognition of language levels to access higher studies and to converge on accreditation mechanisms. For this purpose they established ACLES (Asociación de Centros de Lenguas en la Enseñanza Superior, Association of Language Centers in Higher Education) as the standard of quality for language tests, and agreed to recognize other certifying institutions such as Cambridge, the Alliance Française and the Goethe Institut.

From this moment on, many regions in Spain have passed laws to meet the standards previously mentioned. The report La Acreditación del Nivel de Lengua Inglesa en las Universidades Españolas (Accreditation of English Language Level at Spanish Universities) (Halback and Lázaro) first published by the British Council in Spain in 2010 and updated in 2015 is, perhaps, the most up-to-date and comprehensive analysis of the impact of Spanish regional policies on higher education. This report gathers data from 50 Spanish universities and confirms that the coordination and homogenization of certification processes and mutual recognition has improved between 2010 and 2015. Likewise, the report looks forward to further clarification of standardization processes, clear national linguistic policies, the implementation of quality standards and a more pro-active role by the Spanish central government, among other possible improvements. All these are laws and reports are ordered chronologically in the following timeline:

- 1970: LGE (Ley General de Educación, General Education Law)
- 1980: LOECE (Ley Orgánica por la que se regula el Estatuto de Centros Escolares, Organic Law for the Regulation of the Statutes of Schools)
- 1985: LODE (Ley Orgánica del Derecho a la Educación, Organic Law for the Right to Education)
- 1990: LOGSE (LEY de Ordenación General del Sistema Educativo, Law on the General Organization of the Education System)
- 2002: LOCE (Ley Orgánica de Calidad de la Educación, Organic Law on the Quality of Education)
- 2006: LOE (Ley Orgánica de Educación, Organic Law on Education)
- 2011: CRUE Report on Language Teaching and Accreditation
- 2013: LOMCE (Ley Orgánica para la Mejora de la Calidad Educativa, Organic Law for the Improvement of Educational Quality)

4. The context of Andalusia

Andalusia is one of the 17 autonomous communities that exist in Spain. Similar to the German federal Länder system, Andalusian autonomous communities hold devolved powers over education. In practical terms this means that each community has exclusive competences in educational affairs, which leaves room for a great deal of heterogeneity in accreditation and certification.
In the face of such heterogeneity, in 2011, nine Andalusian universities signed an agreement, the Convenio de Colaboración (2011) (CC henceforward) to define the standards which would regulate the recognition of foreign language levels in their autonomous community. The minimum level of competence had already been set at B1 in 2010 by the General Direction of Universities of Andalusia (Dirección General de Universidades Andaluzas). The CC was drafted by the AGAE (Agencia Andaluza para la Acreditación, Andalusian Agency for Accreditation), nowadays DEVA (Dirección de Evaluación y Acreditación, Board of Evaluation and Accreditation). The nine signing universities are now represented in a working group with the most widespread representation in Andalusia in terms of recognition.

The CC was particularly important because, for the first time ever, the same regional model of test was defined in Andalusia, with the idea of mutual recognition in mind. The CC also included a list of other international certifications which would be recognized by the signing institutions. Although the CC left many areas open to interpretation, thanks to the debate on how these aspects should be interpreted, it was possible to start implementing and fine-tuning the original model of test. In 2013, two years after the CC was signed, a follow-up meeting was organized in Málaga, in which, for the first time, test developers were allowed to participate. The meeting proved to be a great opportunity to identify the vulnerabilities of the common specifications after two years of implementation. From the beginning, it was clear that the nine signing universities shared certain problems, the foremost being the lack of homogeneity in the design of the tests. Since each university had been designing their own tests, work was repeated in some areas and languages (English), while others (German, Russian, French, etc.) were almost unattended since the demand for tests in such languages was much more reduced. There was no centralized source of information and test developers received different messages in different ways, and this made it very difficult to identify the appropriate path to follow. The autonomy of each university, which was recognized in the CC, originated differences in the frequency with which tests had to be designed, as well as in policies regarding the temporal validity of external certificates, exemption criteria for the handicapped, and differences in the criteria regarding the great variety of requests to recognize certificates which were not originally included in the agreement.

This meeting of representatives and test developers in Málaga triggered the constitution of a board of experts, a working group, with one representative per university, who would ensure compliance with the agreement signed in 2011 through yearly follow-up meetings. The first of such meetings was held in Cádiz (October, 2013); others have followed in Jaén (July, 2014), Málaga (January, 2015) and Sevilla (May, 2015). The frequency of these meetings is a clear example of the commitment of all the members of the working group. Each one of the meetings has enabled a follow-up of the implementation of the agreement and has also updated important questions such as the recognition and
certification of Andalusian processes by ACLES and other independent quality-control bodies.

The consensus of the nine signing institutions has facilitated the certification and mutual recognition of language certificates for the +25K students which these universities host in their language programs, as well as for the countless other stakeholders who have had their external certifications validated by any one of the universities involved. In practical terms this means that a certificate issued by any Andalusian university is automatically recognized by the other universities. Table 1 exemplifies the impact and the scope of Andalusian recognition for the academic year 2013/14, the first one for which records were compiled.

### Table 1. Candidates in Andalusian tests for the academic year 2013/14

<table>
<thead>
<tr>
<th>University</th>
<th>Languages</th>
<th>CEFR Levels</th>
<th>Candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Universidad de Almeria</td>
<td>EN, FR, IT</td>
<td>B1</td>
<td>475</td>
</tr>
<tr>
<td>Universidad de Cádiz</td>
<td>EN, FR, GER, IT</td>
<td>B1-B2</td>
<td>1128</td>
</tr>
<tr>
<td>Universidad de Córdoba</td>
<td>EN, FR, GER, IT</td>
<td>A1-B2</td>
<td>561</td>
</tr>
<tr>
<td>Universidad de Granada</td>
<td>EN, FR, GER, IT</td>
<td>B1-B2</td>
<td>1334</td>
</tr>
<tr>
<td>Universidad de Huelva</td>
<td>EN</td>
<td>B1-B2</td>
<td>195</td>
</tr>
<tr>
<td>Universidad de Jaén</td>
<td>EN, FR</td>
<td>B1-B2</td>
<td>707</td>
</tr>
<tr>
<td>Universidad de Málaga</td>
<td>EN, FR, GER, IT</td>
<td>B1-B2</td>
<td>1,584</td>
</tr>
<tr>
<td>Universidad Pablo Olavide</td>
<td>EN, FR, GER</td>
<td>A1-C1</td>
<td>1,679</td>
</tr>
<tr>
<td>Universidad de Sevilla</td>
<td>EN, FR, GER, IT, POR</td>
<td>B1-B2</td>
<td>5,810</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>13,473</strong></td>
</tr>
</tbody>
</table>

Six of the above universities had their language tests certified by ACLES in the academic year 2013/14 (the languages in bold in Cádiz, Córdoba, Granada, Jaén, Pablo Olavide and Sevilla). The University of Málaga joined this list in year 2014/15, as did the German, French, Italian and Portuguese tests of the University of Sevilla. The candidates at the University of Huelva did not sit tests designed by their university, but by external institutions. Four of these

---

1For a full and up-to-date list of the tests included, check ACLES’s official list at www.acles.es.

universities also host tests of international brands (Córdoba, Granada, Huelva and Jaén), which are not included in the table.

Convergence has recently been enhanced through the design of a course on language testing for test developers of the different universities. The bi-monthly course, designed by international experts, brought together 30 of the test designers of the nine universities to be trained in good practices for test development. The contents of the course ranged from specifications design to validation through Classical and Modern Test Theory. This leap forward in training provides ample opportunities to re-evaluate the weaknesses and strengths of all the tests designed in Andalusia. As a consequence, mutual recognition of test results and certifications is being reinforced and constantly re-assessed at two different levels: at an institutional level (Board of Directors) and at a practical level (test developers). Thus, the board acts as a link between the linguistic demands of the stakeholders at universities and the different rectorates, balancing the impact of language testing policies in the Autonomous Community, while test developers ensure the practicality of initiatives and maintain a high quality standard in the tests which each Andalusian university designs. The most important milestones in Andalusia are ordered chronologically in the following timeline:

- 2011: Convenio de Colaboración
- 2013: Meeting of Directors and test designers in Málaga
- 2013: Meeting of CACLU’s Board of Directors in Cádiz
- 2014: Meeting of CACLU’s Board of Directors in Jaén
- 2015: Meeting of CACLU’s Board of Directors in Málaga
- 2015: Meeting of CACLU’s Board of Directors in Sevilla
- 2015: Course on Language Testing

5. Conclusions

Up until recently there has been heterogeneity in the way in which accreditation and recognition are being tackled in Spain. Due to the lack of clearly defined national policies, the different Spanish autonomous communities have passed laws which, when implemented at universities across Spain, have led to significant disparity in the requirements for students at different universities. To avoid such differences and to enhance mutual recognition, institutions should foster common criteria and perform an active role in unifying them. Andalusia is a clear example of how such goals are attainable on a large scale.

Acknowledgements

This research was possible thanks to the contribution of the colleagues of the universities of Almería, Cádiz, Córdoba, Granada, Huelva, Jaén, Málaga, Pablo Olavide and Sevilla, who provided insight, expertise and data regarding
their students, protocols and tests. I would also like to thank Antonio Bueno and Marisa Pérez, from the University of Jaén, as well as Theresa O’Shea, from the University of Málaga, for proofreading the paper.

WORKS CITED


Received: 19 May 2016
Accepted: 7 September 2016