

Designing culturally sensitive assessment tools – How ready are student teachers for assessing multicultural classes?

Jarosław KRAJKA

Maria Curie-Skłodowska University

E-mail: jaroslaw.krajka@mail.umcs.pl, 

Abstract: While the significance of culture, the shift from linguistic to Intercultural Communicative Competence and translanguaging have become well-established principles of foreign language teaching, much less attention in language teacher training is devoted to the influence of culture on language assessment. The existence of multicultural classes, also caused by the appearance of migrants and refugees, brings about a challenge of designing assessment instruments not only in a valid and reliable, but also just and inclusive way. The question that arises is to what extent are student teachers of languages ready to encompass multicultural classes, including students from both majority and minority cultures, with their assessment instruments. To answer that question, the results of a two-stage study are presented: first, study curricula for English language teacher training from all Polish second-cycle university programmes are taken under scrutiny for aspects of cultural sensitivity in language assessment, second, a quasi-experimental research with two groups of student teachers of languages, getting ready to teach English, Spanish, French and Russian, is going to be presented. The analysis is based on student-made placement tests and their reflections in learning diaries.

Keywords: language assessment, cultural sensitivity, EFL, ESL, ELE

Introduction

While the interrelation of culture and language teaching has been established for a number of decades due to seminal research of a number of scholars (most importantly, Byram 1997, 2008; Byram et al. 2002; Kramsch 1993; Kramsch/ Sullivan 1996; Deardorff 2006, to mention just a few) and the firm grounding of intercultural communicative competence in the language proficiency model authorized by the Council of Europe as a part of the EU language policy in the Common European Framework of Reference for Languages (Council of Europe, 2001, 2018/2020), the interrelation between culture and language assessment has been much less often researched.

The purpose of the present paper is to report upon a two-stage research devoted to the exploration of the interrelation between cultural and linguistic research in the area of Language Assessment Literacy (LAL). In particular, we aimed at investigating the significance of Language Assessment Literacy in general and culturally sensitive testing in particular in English language teacher preparation programmes in Poland. Secondly, we intended to check the viability of developing cultural sensitivity in a LAL course devoted to culturally sensitive assessment.

1. Background to the study

The effect of culture(s) on the reality of foreign language teaching and learning has been extensively researched in the late 1990s, which even led to the emergence of the concept of “cultures of learning” (Cortazzi/ Jin/ Wang 2009; Jin/ Cortazzi 2011). These are taken-for-granted frameworks of expectations, attitudes, values and beliefs about successful learning and teaching, about learning and using different language skills in classrooms and about how interaction should be accomplished. According to Jin and Cortazzi (1997), assumptions about what should happen in the classroom, especially regarding the expectations of ‘good’ teachers, ‘good’ students, asking questions and interacting in different forms, add up to the culture of learning. Other aspects encompassed in this notion are roles and relationships in the classroom, learning and teaching styles, approaches and methods, classroom interaction and activities, forms and uses of textbooks and expectations of good work (Cortazzi/ Jin 1996; Cortazzi 1998).

An important area where the cultures of learning will lead to differentiation of foreign language teaching methodology is also assessment, as evaluating representatives of different cultures of learning together in a single classroom is a challenge that language teachers are not always ready to meet (Krajka 2010; Omeri 2014). Therefore, language teachers’ assessment literacy (LAL) (Bøhn/ Tsagari 2021; Coombe et al. 2012; Coombe et al. 2020; Fulcher 2012; Giraldo 2018; Rezai et al. 2021; Sohrabi et al. 2022; Vogt/ Tsagari 2014) should encompass not only the technical skills of choosing assessment approaches, designing measurements, conducting and scoring tests, analysing results and adapting teaching practices, but also accommodating multilingual learners with openness and respect while assessing them in a culturally sensitive way (Gottlieb 2024). Placing language assessment in a wider cross-cultural perspective is in line with the approach of ‘critical language testing’ (Shohamy 2022), which assumes teacher trainees’ reflection on the social and political consequences of the way assessment instruments are designed and administered as well as a thorough investigation of washback.

A crucial concept that comes into play when reflecting upon culturally sensitive language assessment is that of Cultural Intelligence (CQ – Earley/ Ang 2003; Earley/ Mosakowski 2004). As defined by Earley and Ang (2003), it is “an individual’s capability to function effectively in situations characterised by cultural diversity”, critical for expecting and addressing the unexpected during intercultural encounters. CQ is “a specific form of intelligence focused on capabilities to grasp, reason and behave effectively in situations characterised by cultural diversity” (Ang et al. 2007: 337). As a result of applying Cultural Intelligence in methodology training, language teachers will become better equipped to meet the challenges of the contemporary classroom, including assessing language proficiency in multilingual and multicultural contexts (Wysocka 2013; Krajka 2010). There is a relationship between the level of Cultural Intelligence and the organisational culture type – people with high Cultural Intelligence are able to meet the requirements of this organisational culture type and are also attracted to it, while students with low cultural intelligence prefer hierarchical organisations that value stability, predictability and control (Balogh/ Gaal/ Szabo

2011). Moreover, Cultural Intelligence proves to be a statistically significant predictor of adaptive performance in multicultural contexts (Şahin/ Gürbüz 2014). A higher level of Cultural Intelligence is directly related to more innovative work behaviour (Afsar et al. 2020). In the case of teachers, studies proved the relationship between Cultural Intelligence and interpersonal communication, psychological capital and organisational citizenship behaviour (Alifuddin/ Widodo, 2022).

As a part of a larger-scale investigation into the interrelations of culturality and language assessment (Krajka 2025), the present research aims to investigate Polish student teachers' of English readiness for culturally sensitive assessment and application of Cultural Intelligence in designing and adapting assessment instruments for multicultural classes.

2. Research context and participants

The first step in the research process was a comparative analysis of Language Assessment Literacy coverage in all English language second-cycle teacher training curricula from Poland. All tertiary education institutions (both public and private) offering English language teacher preparation (English studies or applied linguistics) were taken into consideration. The curricula for all ELT methodology courses related to language assessment were scrutinised for aspects of culturally sensitive language assessment in the following three areas: 1) learning outcomes, 2) study topics, 3) student-made products for assessment. For that purpose, the study curricula for FLT methodology-related subjects were retrieved from USOS university management systems or via contacts with teacher training department heads for 43 programmes offered by 21 universities.

The second part of the study was conducted in two teacher training programmes in a medium-sized university in Eastern Poland, embedded in the second-cycle study programmes of applied linguistics (English & German, French or Spanish, henceforth referred to as AL) and Spanish studies (henceforth referred to as ELE). The study was a part of a much broader investigation into the interrelations of culture, language assessment and Artificial Intelligence (Krajka 2025).

The two groups were subjected to the same treatment (the AL group in English, the ELE group in Polish) aimed at exposing student teachers to English/Spanish as a foreign/second language assessment materials from different countries, eliciting their opinions on their relevance to testing Polish students of English/Spanish, authoring placement tests and reflecting on the concept of cultural sensitivity in language assessment. Both groups had been sensitised to the effect of culture on foreign language teaching during other teacher training classes; however, they had never been formally instructed in intercultural language teaching and learning (Liddicoat/ Scarino 2013), nor had they had any formal training in interculturality in language assessment. The AL group was composed of 8 students (7 females and 1 male, 7 Poles and 1 Ukrainian), all volunteering to take the optional teacher training module during their second-cycle programme. The ELE group comprised 14 graduate students (12 females, 2 males, all Polish) who followed the Spanish as a foreign language programme during their M.A. Hispanic studies degree.

Even though the two groups were at the same moment of their education (second-cycle studies) and followed the same syllabus for the language assessment course, the competence and experience gap was visible enough to treat the AL group as advanced while the SFL group as beginning. This was because the student teachers in the Spanish as a Foreign Language (ELE) group only started their teacher preparation at the M.A. level, had completed only basic teacher training in the previous year and had served only one practicum, while the AL group completed much of their teacher training (in both foreign languages) together with a number of practicums during the first-cycle studies.

Both groups were intact, composed of all students volunteering to take the optional teacher training module and were exposed to the experimental treatment in their entirety with no option of withdrawal.

The quasi-experimental treatment that both groups were subjected to was conducted between October 2024 and January 2025 in a 30-hour course spread over 15 weekly 90-minute meetings:

1. Pre-test: “Prepare a grammar and vocabulary test for a multilingual (Polish and Ukrainian) group of lyceum first-graders”;
2. Culturally rich instruction in testing and assessing grammar, vocabulary, pronunciation, listening comprehension, reading comprehension, speaking, writing and mediation: discussions of key aspects, analysis of tests from different countries, reflection on different modes of testing and peer assessment of self-produced test proposals;
3. Strategy training in Artificial Intelligence test generation and adaptation: detection of AI-appropriated tests, comparison of AI- and human-generated instruments, instruction on how to make more sophisticated prompts to make the most of AI tools in test generation and cultural appropriation;
4. Post-test final diary: evaluation of the pre-test adapted to the mixed Polish-Ukrainian group of students by a selected Artificial Intelligence tool, reflection on one’s experiences with AI tools in test generation and adaptation and reflection on one’s awareness of cultural sensitivity in language assessment.

Apart from the pre-test, qualitative data from student teachers were collected through diary posts, in which they were asked to reflect upon some of the topics raised in class, evaluate a self-assigned test from a different country (the ISCED 3 examinations), and evaluate an AI-generated placement test (the post-test final diary). The current study will explore only the initial part of the process, namely authoring placement tests and reflecting on their cultural sensitivity. The other components of the study were comprehensively reported in another publication (Krajka 2025).

3. Procedure and data analysis

In the first part of the study, the tertiary education institutions offering second-cycle English language teacher preparation programmes in either English studies or applied linguistics were located with the help of the Ministry of Science and Higher Education’s portal <http://studia.gov.pl>. Due to the different names of study programmes

used across the country, the search had to be wide enough to encompass all forms of English teacher training. Thus, such keywords as “anglistyka”, “filologia angielska”, “lingwistyka stosowana” as well as various forms of English studies with teacher preparation (“filologia angielska z pedagogiką/z przygotowaniem pedagogicznym/w zakresie nauczania języka/w edukacji/nauczycielska”) were entered.

The search yielded 38 second-cycle teacher training programmes from 19 public and 2 private universities in the following cities:

1. English studies:

- Lublin (*Maria Curie-Skłodowska University* – full-time English studies and applied linguistics, *Catholic University of Lublin* – full-time and extramural English studies);
- Wrocław (*University of Wrocław* – full-time and extramural English studies);
- Zielona Góra (*University of Zielona Góra* – full-time and extramural English studies);
- Gdańsk (*University of Gdańsk* – full-time and extramural English studies);
- Poznań (*Adam Mickiewicz University* – full-time and extramural English studies);
- Łódź (*University of Łódź* – full-time and extramural English studies);
- Szczecin (*University of Szczecin* – full-time and extramural English studies);
- Toruń (*Nicholas Copernicus University* – full-time English studies);
- Siedlce (*University of Siedlce* – full-time and extramural English studies);
- Rzeszów (*University of Rzeszów* – full-time and extramural English studies);
- Słupsk (*Pomeranian University in Słupsk* – full-time English studies);
- Nowy Targ (*University of Applied Sciences in Nowy Targ* – full-time English studies);
- Bielsko-Biała (*University of Bielsko-Biała* – full-time and extramural English studies);
- Katowice (*Silesian University in Katowice* – full-time and extramural English studies);
- Cracow (*University of the National Education Commission* – full-time and extramural English studies, *The Jesuit University Ignatianum* – full-time English studies);
- Kalisz (*Adam Mickiewicz University in Kalisz* – full-time English studies in education);
- Olsztyn (*University of Warmia and Mazury* – full-time English studies);
- Nowy Sącz (*University of Applied Sciences in Nowy Sącz* – full-time English studies);
- Warsaw (*University of Warsaw* – full-time and extramural English studies/linguistics).

2. Applied linguistics:

- Lublin (*Maria Curie-Skłodowska University* – full-time applied linguistics);
- Poznań (*Adam Mickiewicz University* – full-time applied linguistics);
- Warsaw (*University of Warsaw* – full-time applied linguistics).

To ensure comparability of results, only programs with these two core elements (English studies and applied linguistics) were taken into account. The curricula of FLT methodology subjects related to language assessment were analysed for references to cultural sensitivity in 1) learning outcomes, 2) study topics, 3) student-made products for assessment.

In the second part of the study, during a large-scale research described in another publication (Krajka 2025), the student teachers were asked to design a 45-minute grammar-and-vocabulary placement test for a multilingual (Polish and Ukrainian) group of lyceum first-graders. The pre-test was preceded by two introductory classes devoted to types of assessment, kinds of tests, purposes for testing, writing test specifications, recommended procedure for test design and principles of properly constructed tests. Moreover, approaches to placement, achievement and diagnostic testing (discrete-point vs. integrative) were elaborated upon with examples of instruments provided by major ELT/ELE publishers.

Following instructor explanations, home readings, in-class discussions and group analysis of presented testing instruments, the student teachers were supposed to prepare their own placement test for a multilingual group of lyceum first-graders. The multilingual context was required due to the reality of the Polish language education – starting September 2024, all Ukrainian refugees residing in Poland wishing to receive the monthly governmental child benefit of 800 PLN (so-called “800+”) needed to take part in regular education in Polish schools, with assessment conducted on the same terms as Polish students.

The student teachers were given a two-week deadline to submit their tests, together with assessment criteria and a grading key, via the course Moodle assignment. The task was accomplished by eight student teachers from the AL group and 15 from the ELE group. The analysis of submitted tests and reflections in student diaries was grouped around three major topics – test format, accommodations for a multicultural class and cultural content.

4. Results and findings

4.1. Comparative analysis of teacher training curricula

The analysis of the curricula for all the institutions that made the research possible (either through making course curricula available publicly or by request) showed a highly insufficient coverage of Language Assessment Literacy (LAL) in teacher training second-cycle English studies/applied linguistics programmes. In most cases, LAL is not granted an independent subject (with the exception of 4 programmes discussed in more detail below) and is relegated only to 1 or 2 90-minute meetings as part of a didactics course.

However, it is striking that there are some programmes in which LAL does not receive even a single mention, which is not in line with the ministerial standards for teacher development. In the case of three universities, the contacted department heads admitted that there used to be a separate subject devoted to building assessment literacy, however, it is no longer offered as it was replaced by another subject (for instance, teaching English to young learners).

Only three universities in five programmes (for both the University of Łódź and University of the National Education Commission, a full-time and an extramural English studies programme, for University of Applied Sciences in Nowy Targ, a full-time English studies programme) offered a separate course devoted to Language Assessment Literacy (named, respectively, “Testing and Language Assessment”, “Assessment in foreign language teaching” and “Workshops in English language didactics: designing integrated skills tests”). However, a closer analysis of learning outcomes, class topics and assessment products showed that in only one case (“Assessment in foreign language teaching”) the problem of culturally sensitive testing may be alluded to when dealing with the washback effect of assessment instruments (the topic named “Influence of examinations on different stakeholders”).

Taking all of the above into account, it can be safely concluded that the Language Assessment Literacy of English language teachers is not adequately developed in pre-service teacher training in Poland. The largely peripheral role it is assigned is reflected in the absence of separate courses devoted to language assessment, inclusion of LAL only in a very tiny timeframe during general didactics courses and in the absence of learning outcomes devoted to assessment skills and attitudes. Unfortunately, culturally sensitive testing is absent in virtually all analysed curricula, even those very few devoted exclusively to LAL. The analysis, thus, calls for a greater interest of course developers and teacher trainers in the cultural dimension of language testing and assessment.

4.2. Student teachers and culturally sensitive placement testing

In the second part of the research, during a quasi-experimental mini-study conducted in the author’s institution in the applied linguistics and Spanish studies programmes, the tests were prepared by student teachers informed by different approaches to placement testing (in particular, the discrete-items and the integrated skills approach). However, no guidance on the cultural dimension of testing had been provided. The resultant tests contained both multiple-choice tasks (in various formats, ranging from isolated words for synonyms/antonyms, through single isolated sentences to mini-dialogues) and more integrative tasks, such as gapped text or completing mini-dialogues. Still, the tests were predominantly focused on selected grammar and vocabulary items: “cross the odd word out”, “transform sentences with keywords or sentence beginnings and endings given”, “choose sentences with a similar meaning”, “complete sentences with the correct form of the verbs in brackets”, “complete sentences with the words formed from the words given in brackets”, “match left and right items to form phrases or find synonyms/antonyms”, “group the words given into categories” and “choose one tense or the other to complete the sentences”. Some open-ended tasks appeared as well, such as “write a sentence to illustrate the meaning of the phrasal verbs given”, “answer questions using the words given” and “describe a person using the prompts given”. A few student teachers also added a receptive skills component to their placement instruments, giving short texts with True/False, multiple-choice or open-ended questions to answer. Overall, all used discrete-item tasks to account for at least half of the points of the whole test, while only around one-fourth added more

integrative and communicative activities such as open-ended question-answer, person description, writing sentences to respond to situations or reading given texts for details.

The accommodation strategies adopted in the authored placement tests were as follows (ordered from the most to the least frequent):

- No linguistic or cultural adaptations, the whole test in Spanish/English – 6 persons;
- Linguistic adaptation of instructions, parallel instructions in Polish and Ukrainian – 5 persons;
- Linguistic adaptation of instructions, parallel instructions in Spanish/English and Polish – 4 persons;
- Full linguistic adaptation of instructions, instructions only in Polish – 3 persons;
- Linguistic adaptation of instructions, parallel instructions in Spanish/English, Polish and Ukrainian – 2 persons;
- Linguistic adaptation of instructions, parallel instructions in English/Spanish and Ukrainian – 1 person;
- Test differentiation, with separate tasks (first and last) clearly marked “for Ukrainian students”, with instructions in Ukrainian, with the remaining tasks with instructions in Polish – 1 person;
- Separate language versions of the same test (Polish and Ukrainian) – 1 person.

The analysis of authored placement tests, prepared without any guidance on culturally sensitive accommodations, showed a lack of cultural sensitivity in the Language Assessment Literacy of the student teachers and a need for awareness-raising instruction in this respect, as they largely either ignored the multicultural nature of the test audience or did not go beyond the simplest measure of parallel instructions. Out of the more creative solutions, the separation of the test’s content into two language groups within the same test does not seem to be plausible and inclusive.

After the authored placement tests were submitted, the student teachers were asked to reflect upon the cultural dimension of language assessment in a diary responding to the following questions:

1. What was the first thing you thought when you read that the test was going to be for a mixed-culture group of students? Do you think such a mixed-culture audience makes test design harder? Why?
2. When preparing your tasks, did you try to use some that might be regarded as more universal, no matter what culture students come from, or did you try to have a similar number of tasks relevant to each cultural group?
3. When you found out that, apart from Poles, your audience was also comprised of students of Ukrainian origin, did you try to do any research on how Ukrainian students learn English at primary school and how they are tested in Ukrainian primary schools? If yes, what did you find out?

The feelings about test design for multicultural contexts displayed in student teachers' diaries varied from fear and anxiety ("I was a little nervous I won't be able to do it") through concern ("The first thing I thought was that it is a complicated task. Preparing tests for multicultural groups is more difficult, especially for teachers who have not had contact with students from other cultures so far.") to curiosity or even satisfaction with the challenge ("I was happy when I saw that the test would be for children of different nationalities. I think it is worth preparing now to teach foreigners, especially when they constitute a high percentage of the population of our country."). Overall, out of the whole range of attitudes, the more positive ones prevailed (with only 2 voices expressing fear), most student teachers treated test design for multicultural groups as another ordinary part of the teacher's job, and only 2 expressed greater enthusiasm about it.

As evidenced by the diaries, the multicultural context for language assessment in itself did not trigger increased attitudes of cultural sensitivity and the resulting actions of individual research and inclusive test task design. Only less than half of the student teachers would go beyond the surface level of manipulating the language of instructions. Even here, the ways that the multilinguality of instructions was conceptualised differed a lot, from English/Spanish-only, English/Spanish-Polish to Polish-Ukrainian, all justified as ways of accommodating the minority group and making the test more universal. The analysis of the diary data demonstrated a need for awareness-raising sessions, alerting student teachers to the effect that the multiculturality of the tested groups should have on test format, situation/topic choice and task design.

While preparing their tests, the student teachers only partially tried to gather knowledge of the minority students' educational and assessment systems to inform their choices. Actually, in the Spanish as a Foreign Language group, only three student teachers out of 14 made any attempt to gather knowledge of the Ukrainian language teaching and assessment context; however, they admitted to being rather unsuccessful. The remaining ones either claimed they had not thought about doing such research or ignored the question altogether. The English as a Foreign Language group exhibited a similar pattern – out of 8 student teachers, 1 exhibited a full awareness of the Ukrainian language assessment system due to his being Ukrainian, one other student reported searching for materials with a satisfactory result, 3 reported unsuccessful or minimal attempts and made their assumptions about the cultural specificity of the minority group, while the remaining half of the group (4 out of 8) either did not respond to the question or openly admitted that the assessment tool needs to be universal without making any special concessions towards the minority group.

Conclusion

Given the demographic changes in societies on the one hand and increased mobility and employability patterns of university graduates on the other, it is more than likely that graduates of English language teacher training programmes are going to teach in bi-, tri- or multicultural settings. While intercultural language teaching and intercultural communication do get at least some coverage in ELT training curricula, this is not the case with language testing and assessment. Language Assessment Literacy is

not given the place it deserves, usually being relegated to brief mentions or lecture-style coverage. For LAL to be addressed properly, it is recommended that

- language assessment needs to be given its proper place as a separate course, with a systematic coverage of principles and techniques of assessing language subsystems, skills, communicative competence, intercultural competence, both in the traditional and formative/alternative model;
- student teachers should be exposed to as wide a range of assessment instruments (both publisher-made and student-made), from different countries and testing different languages, to ensure a wider picture and more in-depth reflection;
- test adaptation to fit learners with diverse educational needs, including multicultural backgrounds, should be one of the tasks to explore during practicums, with student teachers encouraged to interview (multiple) school teachers to better explore the problem.

The mini-research conducted in two language teacher training programmes (applied linguistics and Spanish studies), oriented at building cultural sensitivity in language testing and assessment, proved a need for such targeted training. Crossing the boundaries of disciplines to bring cultural and linguistic research together seems more than proper to enable future language teachers to both teach and assess in a fairer, more equitable and inclusive way.

While the study findings show an interesting picture, the current research did have its limitations that need to be taken into account. Most of all, the participants were in intact groups with no choice of reconfiguration or withdrawal due to the organisation of the teacher training modules. Another limitation was the size of the groups – due to the decreasing popularity of graduate studies and teacher training in general, larger participating groups are rather difficult to obtain. Finally, while the curriculum analysis attempted to be comprehensive in terms of all English teacher training institutions, it is possible that test adaptation issues might be embedded in more general training of teaching language skills and subsystems, which was not indicated clearly in the syllabi under analysis. Interviewing representatives of each institution would help to gain a wider picture of the investigated issue.

References

Afsar, B./ B.M. Al-Ghazali/ S. Cheema/ F. Javed (2020), *Cultural intelligence and innovative work behaviour: The role of work engagement and interpersonal trust*, (in:) “European Journal of Innovation Management” 24(4), 1082–1109.

Alifuddin, M./ W. Widodo (2022), *How is Cultural Intelligence related to human behaviour?*, (in:) “Journal of Intelligence” 10(1), 3.

Ang, S./ L. Van Dyne/ C. Koh/ K.Y. Ng/ K.J. Templer/ C. Tay/ N.A. Chandrasekar (2007), *Cultural intelligence: Its measurement and effects on cultural judgment and decision making, cultural adaptation and task performance*, (in:) “Management and Organization Review” 3(3), 335–371.

Balogh, A./ Z. Gaal/ L. Szabo (2011), *Relationship between organizational culture and Cultural Intelligence*, (in:) "Management & Marketing Challenges for the Knowledge Society" 6(1), 95–110.

Bøhn, H./ D. Tsagari (2021), *Teacher educators' conceptions of language assessment literacy in Norway*, (in:) "Journal of Language Teaching and Research" 12(2), 222–233.

Byram, M. (1997), *Teaching and Assessing Intercultural Communicative Competence*. Clevedon.

Byram, M. (2008), *From Foreign Language Education to Education for Intercultural Citizenship*. Clevedon.

Byram, M./ B. Gribkova/ H. Starkey (2002), *Developing the Intercultural Dimension in Language Teaching. A Practical Introduction for Teachers*. Strasbourg.

Coombe, C./ S. Troudi/ M. Al-Hamly (2012), *Foreign and second language teacher assessment literacy: Issues, challenges, and recommendations*, (in:) C. Coombe/ P. Davidson/ B. O'Sullivan/ S. Stoynoff (eds.), *The Cambridge Guide to Second Language Assessment*. Cambridge, 20–29.

Coombe, C./ H. Vafadar/ H. Mohebbi (2020), *Language assessment literacy: What do we need to learn, unlearn, and relearn?*, (in:) "Language Testing in Asia" 10(3), 1–16.

Cortazzi, M. (1998), *Curricula across cultures. Contexts and connections*, (in:) J. Moyles/ L. Hargreaves (eds.), *The Primary Curriculum. Learning from International Perspectives*. London and New York, 205–216.

Cortazzi, M./ L. Jin (1996), *Cultures of learning: Language classroom in China*, (in:) H. Coleman (ed.), *Society and the Language Classroom*. Cambridge, 169–206.

Cortazzi, M./ L. Jin/ Z. Wang (2009), *Cultivators, cows and computers: Chinese learners' metaphors of teachers*, (in:) T. Coverdale-Jones/ P. Rastall (eds.), *Internationalizing the University. The Chinese Context*. Basingstoke, 107–129.

Council of Europe (2001), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*.

Council of Europe (2018/2020), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume*.

Deardorff, D.K. (2006), *The identification and assessment of intercultural competence as a student outcome of internationalization at institutions of higher education in the United States*, (in:) "Journal of Studies in International Education" Fall 2006, 10, 241–266.

Earley, P.C./ S. Ang (2003), *Cultural Intelligence: Individual Interactions Across Cultures*. Stanford.

Earley, P.C./ E. Mosakowski (2004), *Cultural Intelligence*, (in:) "Harvard Business Review", October 2004.

Fulcher, G. (2012), *Assessment literacy for the language classroom*, (in:) "Language Assessment Quarterly" 9(2), 113–132.

Giraldo, F. (2018), *Language assessment literacy: Implications for language teachers*, (in:) "Profile: Issues in Teachers' Professional Development" 20(1), 179–195.

Gottlieb, M. (2024), *Assessing Multilingual Learners. Bridges to Empowerment* (Third Edition). Thousand Oaks, Ca.

Jin, L./ M. Cortazzi (1997), *Expectations and questions in intercultural classrooms*, (in:) "Intercultural Communication Studies" 7(2), 37–68.

Jin, L./ M. Cortazzi (2011), *Re-evaluating approaches to second language teaching and learning*, (in:) E. Hinkel (ed.), *Handbook of Research in Second Language Teaching and Learning* vol. II. London, 558–575.

Krajka, J. (2010), *Preparing language teachers for multinational classrooms – guidelines for foreign language methodology training*, (in:) H. Komorowska/ L. Aleksandrowicz-Pędzich (eds.), *Coping with Diversity*. Warsaw, 242–262.

Krajka, J. (2025), *Language Assessment Across Cultures – From Intercultural Communicative Competence to Artificial Intelligence*. Göttingen.

Kramsch, C. (1993), *Context and Culture in Language Teaching*. Oxford.

Kramsch, C./ P. Sullivan (1996), *Appropriate pedagogy*, (in:) "ELT Journal" 46(4), 340–349.

Liddicoat, A.J./ A. Scarino (2013), *Intercultural Language Teaching and Learning*. Oxford.

Omeri, A. (2014), *Teaching a foreign language in a multicultural context*, (in:) "Journal of Educational and Social Research" 4(2), 254–256.

Rezai, A./ G. Alibakhshi/ S. Farokhipour/ M. Miri (2021), *A phenomenographic study on language assessment literacy: Hearing from Iranian university teachers*, (in:) "Language Testing in Asia" 11, 26.

Şahin, F./ S. Gürbüz (2014), *Cultural intelligence as a predictor of individuals' adaptive performance: A study in a multicultural environment*, (in:) "International Area Studies Review" 17(4), 394–413

Shohamy, E. (2022), *Critical language testing, multilingualism and social justice*, (in:) "TESOL Quarterly" 56(4), 1445–1457.

Sohrabi, Z./ N. Ghanbari/ A. Abbasi (2022), *Iranian EFL teachers' perceptions of writing assessment literacy: A countrywide study*, (in:) "Language Testing in Asia" 12, 23.

Vogt, K./ D. Tsagari (2014), *Assessment literacy of foreign language teachers: Findings of a European study*, (in:) "Language Assessment Quarterly" 11(4), 374–402.

Wysocka, M. (2013), *Teaching English as an International Language: Problems and research questions*, (in:) K. Droźdżał-Szelest/ M. Pawlak (eds.), *Psycholinguistic and Sociolinguistic Perspectives on Second Language Learning and Teaching*. Heidelberg, 41–55.