VALIDITY OF ROLE PLAY IN SOCIOCULTURAL COMPETENCE ASSESSMENT IN YEAR 12 EXAMINATION IN LATVIA

VITA KALNBĒRZIŅA

University of Latvia, Latvia

Abstract. The article examines different facets of the role play task, its application in the Year 12 English language examination and the statistical analysis of its results with the purpose to validate the use of role play for testing the sociocultural competence in Latvia. The research method of the role play here combines the features of external validation (construct validation) with internal validation (contents and response validation). The statistical findings suggest that the role-play task ensures a reliable and valid method of assessment of the student performance; it also provides a means for reliable assessment of a large proportion of the curriculum, as the skills tested by role play differ from the skills tested by other examination tasks. The findings of the curricular validation suggest that removing role-play from the examination would undermine the curricular validity of the Year 12 examination. The contents analysis of the task suggests that the difference in the student performance level in Speaking is created by the impact of the assessment of the sociocultural competence as defined by Celce Murcia, Dörnyei, and Thurrell (1995).

Key words: assessment, role play, content analysis, sociocultural competence, intercultural competence, multicultural competence, transcultural competence, construct validity, curricular validity, content validity, face validity, reliability, curricular validation, content analysis

INTRODUCTION

Assessment of speaking is a complex and expensive endeavour, involving participation of item writers, test administrators, interlocutors and assessors, who have to treat the candidates one by one. This is not easy to organize within a classroom assessment situation, but on a state level it is an outstanding logistical undertaking involving a lot of resources; therefore, we need to be sure that we are using the resources in the most efficient way, see for example, May (2010) for a comprehensive overview of the recent studies focusing on the issues involved in assessment of interactional speech. In Latvia we can also hear complaints that structured role play is too cumbersome and depends on reading too much and should be removed from the examination, this is why this article is going to examine if the use of role play in the speaking test of the Year 12 examination is a valid choice. To do this, we will examine the construct, curricular and contents validity of the task and the sociocultural competence it is supposed to test, thus,

using both internal and external validation approaches (see Davies et al., 1999). All the test materials and the data analysis presented in this paper are based on the data of the Examination Centre and are available in Online 2.

CONSTRUCT VALIDITY

A construct validation exercise involves examining the theoretical basis of the task, in our case, the construct of sociocultural competence. The basis of this term is well grounded in sociocultural theory starting from Vygotsky's *Language and Thought* in 1934, where he first postulated that language is acquired as a social phenomenon, and only afterwards we do internalise the language processes. Sociocultural competence has been considered as a part of language learning since Canal and Swain (1980) introduced the term *communicative competence* in the 1980s.

Nowadays there are several approaches to defining sociocultural competence, e.g. van Ek and Trim's taxonomy developed within the Vantage level description, which would agree with the National Curriculum aiming at Common European Framework levels B2 and C1 (Online 4).

Van Ek and Trim's (2001: 97) construct of sociocultural competence contains knowledge about social practices: 'Sociocultural competence is the aspect of communicative ability which involves those specific features of society and its culture which are manifested in the communicative behaviour of the members of this society'. These features can be classified as universal experiences (everyday life procedures, living conditions, levels of formality, major values and attitudes), social rituals (body language, visiting rituals, eating and drinking rituals and linguistic rituals) and social or politeness conventions.

Van Ek and Trim's construct of sociocultural competence focuses on the knowledge about the society and its rituals, but the examples are mostly concerned with the practices involved in politeness conventions for British English:

1. Do not be dogmatic (use *I think, I believe*)
2. Be reluctant to say what may displease the partner (*I don't want to complain, but...*)
3. Do not force the partner to act (use *Please*, or *I wonder if you could help me*) (van Ek and Trim, 2001: 98).

The suggestions described in the Council of Europe Vantage level publication are the English language and English culture specific, and obviously are meant as an example, but they cannot be called universal as they will not be applicable for other socio-cultural situations in other languages.

Celce Murcia, Dörnyei, and Thurrell (1995) provide a more general taxonomy for addressing the sociocultural issues in language teaching and assessment. They distinguish between social, cultural, stylistic and communicative factors. By social factors they understand the learner variables (age, gender, status and social

distance) and situational variables (time, place, social situation); the stylistic factors are politeness conventions, and specific register; the cultural factors are living conditions, social conventions, rituals, art, literature, values, beliefs and norms; the nonverbal communicative factors are kinesic, proxemic and haptic factors (Celce Murcia et al., 1995: 24).

Lately the sociocultural competence concept has been replaced by intercultural competence (see, for example, Byram, 1997, 2008, for the discussion of intercultural skills), multicultural competence (see, for example, Cartwright and Daniels, 2008, on assessing multicultural competence) and even transcultural competence (see, for example, Thorne, 2008, on mediating discourse online in transcultural context).
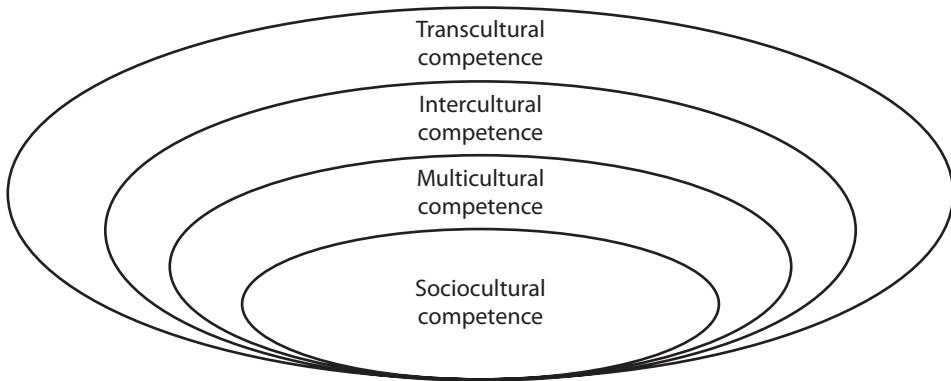

Figure 1 **Taxonomy of the cultural competence**

These terms are expanding the context of language acquisition from social to multicultural, intercultural, and finally transcultural contexts (see Figure 1), to include ever more cultural experiences, until we have come to realise that it is not scale that matters, but the ability of the individual to use language in different contexts; therefore, here we will stay with the original *term sociocultural competence* (used also in Latvian Foreign language curriculum) and use Celce Murcia et al's taxonomy (Celce Murcia et al., 1995) for a comprehensive language learning and assessment approach analysis in Latvia.

CONTENT VALIDITY

Now that we have examined sociocultural competence in theory, let us examine the contents of the Speaking test, to see if it is representative of the construct tested. The speaking test of the Year 12 examination in Latvia (Online 2) consists of 3 tasks: questions and answers, role play and a written text reproduction in speech. Each of the tasks is chosen by the learner separately on a different topic, for example, the topics of the questions and answers task (Task 1) in 2013 were: career, free time, sports, arts, television. The topic of the role play is not easily

described, as it mostly consists of a description of a situation, for example, Role play 1 in the year 2013 examination for the secondary school in English was:

> You are hosting an exchange student (played by your teacher). One morning he/she feels unwell. You have a minute to prepare. Then you will start. (Exam materials 2013)

The instruction gives a clear description of the social factors of the context: the role of the student as a host and the role of the teacher (the exchange student). As soon as the participants of the examination enter the role play, it changes the status of the participants of the conversation. If normally it is the teacher who is in charge of the situation, now it is the host (played by the student) who is in charge. The teacher has not only lost the grounding, being in someone else's country and someone else's home sick and in need of help from their host (played by the student).

The situational factors of the context are also clear from the instruction: it is at home in the morning. The introductory description of the situation is followed by the script for the student and the script for the teacher (see Table 1). An additional point of interest in the instruction is the remark *You will start* which puts the student in a further control position, as he/she decides when to start the conversation. These features were introduced in the Year 12 examination to promote a positive washback to the teaching process, hoping that the teachers will get used to giving their students more leverage during the learning process by allowing them initiation in the classroom discourse. Unfortunately, this was not followed up by washback research that would verify if this aim was achieved.

Table 1 **Sample role play from 2012/2013 examination materials**

Student	Teacher
1. Greet him/her.	1. Hi
2. Ask about his/her health.	2. Oh, I think I have a cold. Could you help me?
3. Answer the question and then suggest calling a doctor.	3. Oh, no, I'll just stay in bed. Will you inform the school?
4. Answer the question and then suggest preparing hot milk, Getting some medication.	4. Yes, please, I'd love some. No, thanks.
5. Offer to buy something tasty.	5. Thank you, I'd love some fruit. Could you buy some Oranges or apples?
6. Answer the question and then enquire about other wishes.	6. I'm fine, thanks. What do you do when you have a cold?
7. Recommend sleep as the best medicine.	7. OK, I'll try to sleep. Will you tell me the news when you return from school?
8. Answer the question and then say good-bye.	8. Bye-bye

As we can see the role play consists of 8 turns, the first is greeting in an appropriate style and inquiring of the wellbeing, then offering help, advice and recommendations as well as leave taking (stylistic factors). We can also see the role of cultural factors, in the instance of preparing a hot milk drink, which is a typical Latvian means of recovery from cold; thus, we can say that the situation is not assessing the student's ability to comply with British or American sociocultural conventions, but rather use the Latvian sociocultural competence in an intercultural context.

The next step takes us to the analysis of the marking scale (Table 2); as we can see the marking scale of the role play agrees with the task description (containing the descriptors of the situation management). The assessment scale is based on the degree of control of the given situation: students will get 4 points for the ability to maintain the conversation in the particular situation, 3 points for using only short phrases and sometimes hindering the situation, 2 points for handling the situation in spite of the misunderstandings, 1 point for frequent misinterpretation of what the teacher says, or reading from the script.

Table 2 **Marking scale Task 2: Role Play (Year 12 examination materials Online 2)**

4	Can maintain a conversation in the given situation but may sometimes have difficulties in saying what she/he would like to.
3	Can maintain a conversation in the given situation using mostly short phrases, which sometimes might hinder effective communication.
2	Can handle short social exchanges, but there are misunderstandings in communication.
1	Can handle very short and often inaccurate social exchanges; often misinterprets what is asked.
0	Not enough to evaluate.

CURRICULAR VALIDITY

Curricular validation is the examination of the task from the point of view of the curriculum, but before we move to examining the curricular validity, we will have a look if the construct agrees with the curriculum objectives.

The Curriculum of the Foreign languages for the secondary schools of Latvia (Online 1) provides that the compulsory content of the foreign language subject consists of two main types of competences, one is functional competence (language for learning, language for interaction, communication culture, language system and standard and language culture and its functional styles), the other being sociocultural competence (language for research and cooperation, language as a part of culture, language for integration, intercultural communication process in multilingual discourses).

If we compare the curriculum objectives (Online 1, translated by the author) to the theoretical constructs, we can see that the sociocultural competence addresses all the four factors mentioned by Celce Murcia et al. (1995: 24):

1. the social factors are represented in the following objectives: comprehends and adheres to the principles of communication and cooperation (point 8.9 in the Curriculum), perceives the similarities and differences in various cultures with understanding (8.8), listens to others and expresses his/her opinion applying the most convenient means of language;
2. the stylistic factors: adheres to the norms of communication in multicultural society (8.13);
3. the cultural factors: uses literature and art for the perception of other cultures and self-development (8.5), creates literary texts (8.6);
4. nonverbal communication: applies the knowledge of national verbal and non-verbal peculiarities of behaviour (8.14).

Thus, we can conclude that all the factors represented in the theoretical framework of socio-cultural competence are mirrored in the Language Curriculum and, therefore, should also be included in the language assessment.

After the 2013 examination the Ministry of Education carried out an extensive validation exercise involving all the foreign language item writers as well as the author of this article (see the results of the study in Online 3). The item writers were asked to go through all the curriculum objectives writing next to each which task assessed which curriculum objective. Once this had been done, all the language examination data were put on one scale, and the item writers came to the conclusion that the sociocultural competence was mostly assessed by the role play tasks not only in the English language examination, but also in the examinations of German, French and Russian. It was also discovered that students had certain difficulties in performing the role play 'connected with the students failing to use the given script to maintain a natural conversation. The main problems were:

1. skipping or rearranging parts of the script;
2. inability to use the cues due to the lack of vocabulary;
3. misunderstanding the required function (e.g. explaining instead of suggesting);
4. reading out the lines without changing them into more appropriate exchanges;
5. little participation in the dialogue;
6. not responding or commenting on the information received from the interlocutor;
7. grammar mistakes in either forming direct questions or reported questions' (Online3).

As we can see from the analysis of the student recordings, out of the 7 difficulties mentioned by the experts, one of the issues is connected with

following the procedure (1), two are comments on the linguistic competence (2 and 7), but the other four are the experts' comments on the appropriacy of the students' participation in the situation; thus, we can conclude that the qualitative analysis of student response suggests that the role play task is assessing the sociocultural competence.

RESULTS AND DISCUSSION

1 QUANTITATIVE ANALYSIS OF THE TEST TAKER PERFORMANCE

Apart from qualitative analysis, the Examination centre also carries out quantitative analysis, assessing the means, the correlations between the different parts and different assessors' performance as well as item statistics of the receptive skills tests.

The quantitative analysis of the overall performance of the students in the foreign language examinations in 2013 shows that all the foreign language examinations are statistically compatible; they all test the same curriculum objectives and follow the same test specifications.

Table 3 **Mean for all the tests in all the foreign languages**

	English	**Russian**	**French**	**German**	**Mean**
Reading	48.59	56.51	67.80	69.72	62.04
Listening	52.80	66.92	59.18	64.43	60.83
Language Use	50.64	68.65	60.14	62.31	60.44
Writing	55.83	52.18	65.70	61.83	58.89
Speaking	65.73	76.54	77.82	71.35	72.86
Mean	54.72	64.16	66.13	65.93	62.73

As we can see from Table 3, the overall mean of the examinations ranges from 66.13 out of 100 points for the French examination, to 54.72 for the English language examination. Here we need to note that the English language examination is taken by the whole population, while French is a choice examination, see Table 4 for the examination population statistics breakdown by language.

Table 4 **Overall statistics in foreign language examinations 2013**

	Total	**B1**	**B2**	**C1**	**Not qualified**
English	16971	7780	4572	62	4557
Russian	2610	1363	1056	0	191
German	227	66	109	12	40
French	53	26	25	0	2
Total	19861	9235	5762	74	4790

The statistical analysis of the student performance in the 4 foreign language examinations shows that the speaking test was the easiest in all languages if compared to other skills tests (mean 72.86), but the most difficult was the writing test (mean 58.89), except in the English and French examination.

The distribution curve for the whole examination in the English is normal, but the distribution curve for the speaking test is negatively skewed (see Figures 2 and 3). A similar pattern can be observed in the distribution curves of other foreign language examinations (see Online 3).
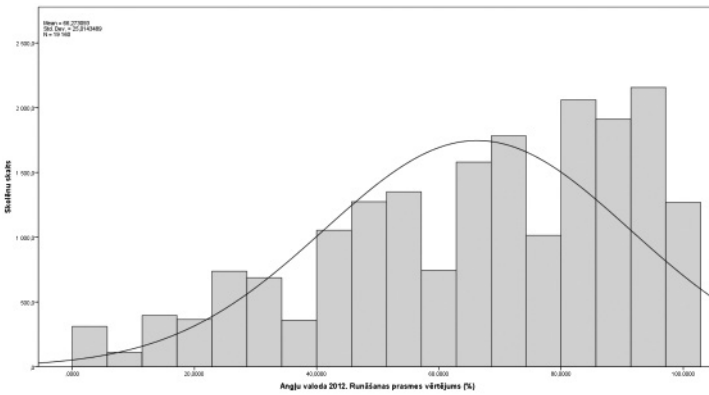

Figure 2 **Speaking test distribution curve (Exam Centre data)**

This would agree with the finding that the speaking test tests a different competence (sociocultural competence) from the other skills tests, where the weighting of linguistic competence is more pronounced.
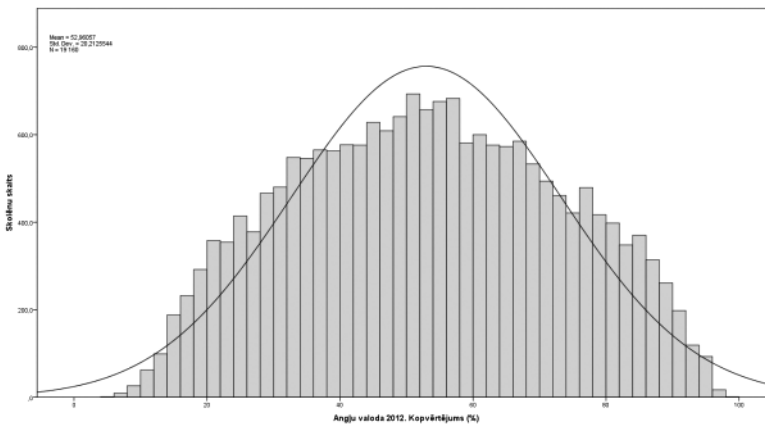

Figure 3 **Comparison of the total test distribution curves (Examination Centre data)**

Thus the statistical analysis and the qualitative analysis allow us to conclude that it is the sociocultural competence that allows the students to gain the highest score in all the foreign language examinations if we can provide evidence that the speaking test assessors have assessed the student performance reliably. This is why we will turn to the correlation indices.

2 RELIABILITY OF THE ASSESSMENT

To examine the reliability of the marking of the year 2013 examination, we used linear correlation coefficients to examine the reliability of the marking quality of the whole population. As we can see in Table 5, the results suggest that the interrater marking reliability of the Speaking tests is acceptable as the correlations are significant and strong, although slightly weaker than the writing test results.

Table 5 **Double marking correlation coefficients**

	English	**Russian**	**French**	**German**
Speaking	0.722	0.887	0.815	0.896
Writing	0.799	0.825	0.963	0.912

The second reason for the use of the correlation coefficients is to examine the contents of the examination, that is, to see if all the skills are examining the same kind of competence. Table 6 shows the correlation coefficients for the English language examination within each separate skill test. As we can see, the correlation coefficients are significant and within the normal range for all the skills (0.6–0.8). Interestingly enough, the speaking does not correlate too highly with the Reading test (0.676), which renders invalid the oft-repeated complaints that the role play task performance depends on the ability of reading of the instructions. Instead, the highest correlation for speaking is with the Writing test which would be logical as both these skills depend on the use of productive strategies.

Table 6 **Correlation for English language tests**

	Listening	**Reading**	**Speaking**	**Language Use**
Reading	0.817			
Speaking	0.706	0.676		
Language Use	0.837	0.850	0.707	
Writing	0.700	0.691	0.748	0.729

If we look closer into the student performance of role play in the English language examination, in Table 7 we can see that the easiest of the speaking tasks was the questions and answers task. The role play was slightly more difficult (mean 4.02), while speaking on the text was the most difficult (mean 3.73).

Table 7 **Speaking task correlations and Means**

Speaking Task	Correlation with Speaking total	Correlation with total	Mean
Questions and answers	0.862	0.717	4.30
Role-play	0.862	0.670	4.02
Text analysis	0.872	0.622	3.73

The last task has also the lowest correlation with the overall speaking performance, which could be caused by the importance of the reading ability for the performance there.

CONCLUSIONS AND SUGGESTIONS

Thus, having examined the contents and the results of the foreign language examinations in 2013, we can conclude that the role play task possesses the features of construct, curricular and content validity and can be used in the future Year 12 examinations because:

1. It agrees with the construct of the sociocultural competence description in the theoretical sources.
2. It complies with the demands of the secondary school curriculum.
3. It provides a reliable and valid task format for the assessment of the sociocultural competence across the foreign languages.
4. It provides a reliable and valid framework for the development of assessment instruments (tasks and marking scales) across the foreign languages.

The issue that remains unexamined here is the face validity (test popularity among the test developers and users) of the role play task, which maybe the real cause of discontent among the examination developers and users; therefore, it should be studied separately.

REFERENCES

Cartwright, B. Y. and Shuqiang Zhang, J. D. (2008) Assessing multicultural competence: perceived versus demonstrated performance. *Journal of Counselling and Development,* 86 (3): 318–322.

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

Byram, M. (2008) *From Foreign Language Education to Education for Intercultural Citizenship. Essays and Reflections.* Clevedon: Multilingual Matters.

Canale, M. and Swain, M. (1980) Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics* 1(1): 1–47.

Celce Murcia, M., Dörnyei, Z. and Thurrell, S. (1995) Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics,* 6 (2): 5–35.

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. and McNamara, T. (1999) *Dictionary of Language Testing.* Cambridge University Press.

May, L. A. (2010) Developing speaking assessment tasks to reflect the 'social turn' in language testing. *University of Sydney Papers in TESOL*, 5: 1–30.

Tamcke, M., De Jong, J. and Klein, L. (eds.) (2013) *Europe – Space for Transcultural Existence.* Gottinger: Universitäts Verlag.

Thorne, S. L. (2008) Transcultural communication in open Internet environments and massively multiplayer online games (pp. 305–327). In S. Magnan (ed.) *Mediating Discourse Online*. John Benjamins: Amsterdam.

Van Ek, J. A. and Trim, J. L. (2001) *Vantage.* Cambridge University Press.

Vygotsky, L. (1986) *Thought and Language*. Cambridge. MA: The MIT Press.

ONLINE SOURCES

1) *Noteikumi par valsts vispārējās vidējās izglītības standartu, mācību priekšmetu standartiem un izglītības programmu paraugiem* http://likumi.lv/doc.php?id=257229 [Accessed on 13 October 2014].

2) *Skolotāja materiāls Mutvārdu daļas uzdevumu paraugi angļu valodā, 2011. gads* http://visc.gov.lv/vispizglitiba/eksameni/dokumenti/ce_paraugi/20120427_sv_val/an_val/skolot_materials.pdf [Accessed on 13 October 2014].

3) *2012./2013.mācību gada centralizētā eksāmena svešvalodās rezultāti, secinājumi un ieteikumi* http://visc.gov.lv/vispizglitiba/eksameni/dokumenti/metmat/2013_ce_svesvalodas_analize.pdf [Accessed on 13 October 2014].

4) *Common European Framework of Reference for languages: learning, teaching, assessment* http://www.coe.int/t/dg4/linguistic/source/framework_en.pdf [Accessed on 13 October 2014].

Vita Kalnbērziņa (Dr. Philol., Assoc. Prof.) is currently working at the University of Latvia. Her research interests include language testing and language acquisition. Email: vita@lanet.lv