neofilolog

Czasopismo Polskiego Towarzystwa Neofilologicznego
ISSN 1429-2173, 2019, NR 53/2, 281-297
http://dx.doi.org/10.14746/n.2019.53.2.9
http://poltowneo.org/

Anna Czura
University of Wrocław
https://orcid.org/0000-0001-5234-6618
anna.czura@uwr.edu.pl

ADOLESCENT LEARNERS' PERCEPTIONS OF ASSESSING SPEAKING AND WRITING SKILLS – CRITICAL INCIDENT ANALYSIS

Spostrzeżenia nastolatków na temat oceniania ustnych i pisemnych kompetencji językowych – analiza zdarzeń krytycznych

Learners' perceptions of teacher assessment practices shape a unique classroom assessment environment that affects learners' willingness to engage in a task and motivation to learn (McMillan & Workman, 1998). Still, there has been limited research on learners' perceptions of classroom and high-stakes assessment. This article outlines a study exploring adolescent learners' perceptions of the speaking and writing assessment they experience in the English classroom. The Critical Incident Technique (CIT) was used to examine both positive and negative instances of oral and written assessment. During oral interviews, the learners reported and elaborated on the assessment experiences they personally perceived as particularly important, memorable and influential. The transcriptions of the interviews were content analysed in order to identify the properties of oral and written assessment that appeared salient to adolescent learners. The study indicates that the participants held predominantly positive views of oral and written assessment, emphasising such aspects as authenticity of assessment, freedom of expression and fair grading. It is worrying, however, that learners are not always provided with clearly articulated assessment criteria and that oral assessment is visibly neglected in the classroom.

Keywords: language assessment, oral assessment, writing assessment, learner perceptions, adolescent learners, critical incident technique (CIT)

Słowa kluczowe: ocenianie, ocenianie kompetencji ustnych, ocenianie kompetencji pisemnych, spostrzeżenia uczniów, uczniowie nastoletni, technika zdarzeń krytycznych

What would happen if we treated the student as someone whose opinion mattered?
(Fullan, 1991: 170)

1. Introduction

Assessment never takes place in a vacuum. Regardless of its form and content, it exerts a profound impact on individual learners and teachers, school systems and societies at large (Bachman & Palmer, 1996). Given the far-reaching consequences of assessment, it is essential that its design and administration are of high quality and produce valid and reliable results that accurately reflect learners' performance. Inspired by Freire's (1970) idea of democratic education and Messick's (1989) interest in the impact and consequences of assessment, many researchers (e.g. McNamara, 1998; Nevo, 1996; Shohamy, 2001) have set out to analyse the effects of assessment from a critical perspective, which emphasizes the role of ethics, fairness and power distribution. One of the principles of Critical Language Testing[1] (CLT) advocates that test-takers, whose voices are usually marginalised, play a more active role in the process of assessment and be given opportunities to offer a critical response (Shohamy, 2001); nevertheless, there has been limited academic research concerning learners' perceptions of both classroom and high-stakes assessment.

In light of the dearth of research centring specifically on learners' voices, this article outlines a study exploring adolescent learners' perceptions of the speaking and writing assessment they experience in English classrooms. The presentation of research findings is preceded by a theoretical analysis of the importance of learners' perceptions in ensuring high quality foreign language (FL) education and assessment. The present research study took a qualitative approach and used the Critical Incident Technique (CIT) (Flanagan, 1954) to examine both positive and negative instances of assessment of speaking and writing skills. During oral interviews, the respondents reported and elaborated on the assessment experiences they personally perceived as particularly important, memorable and influential. The transcriptions of the

[1] The terms 'testing' and 'assessment' are often used interchangeably in the subject literature. The principles of CLT refer to language assessment at large (cf. Shohamy, 2001)

interviews were content analysed in order to identify the properties of oral and written assessment that appeared salient to adolescent learners.

2. Theoretical background

In a teacher-centred classroom, it was the teacher who was responsible for creating the classroom environment, and selecting the content of teaching and designing assessment tasks, whereas the learners were reduced to passive recipients of knowledge (Brooks & Brooks, 1993). Nowadays, more learner-centred and person-centred approaches which promote learner autonomy, encourage meaningful learning and engage learners in the learning and teaching processes (cf. Jacobs & Renandya, 2016), are preferred in language pedagogy. Brooks and Brooks (1993: 60) note:

> Seeking to understand students' points of view is essential to constructivist education. (...) Students' points of view are windows into their reasoning. (...) Teachers who operate without awareness of their students' points of view often doom students to dull, irrelevant experiences, and even failure.

Interestingly, this move from behaviouristic towards humanistic and constructivist approaches has not been fully reflected in research – although there is a relatively large body of studies on teachers' perceptions of and beliefs about different facets of teaching (Lee, Leong & Song, 2016; Pajares, 1992) and assessment (e.g. Leung, 2004; Sahinkarakas, 2012), our understanding of learners' perceptions of the foreign language (FL) learning and teaching, particularly in terms of assessment, is limited.

Learners' perceptions of teacher assessment practices, which involve the choice of elicitation tasks and assessment criteria, the form of feedback and grading policies, shape a unique classroom "assessment 'character' or environment" (Brookhart, 2004: 444). Determined by how learners understand and respond to the assessment practices they encounter, this group experience varies from classroom to classroom and affects how learners perceive the usefulness, value and significance of an assessment task (Entwistle & Entwistle, 1991). Classroom assessment environment also determines students' willingness to engage in a task and may contribute to an increase or a decrease of student motivation (McMillan & Workman, 1998). The impact of learners' perceptions on their willingness to engage in an assessment task was also confirmed in one of my earlier studies (Czura, 2013), in which a group of adolescent learners failed to engage in portfolio assessment due to an inconsistent grading policy and learners' insufficient awareness of the purpose and benefits of this form of assessment.

Whereas there exists a volume of research on learners' preferences as regards FL corrective feedback techniques (cf. Lee, 2005; Leki, 1991; McMartin-Miller, 2014), our knowledge about their perceptions of different facets of classroom-based assessment is rather fragmentary. On the whole, research suggests that learners perceive assessment and testing as an integral element of language learning (Kelly et al., 2010; Muñoz, Alvares, 2007); however, it appears they hold rather traditional views as regards their own role in the assessment process. Studies focused on corrective feedback to written work (Lee, 2004; Leki, 1991) and in-class assessment at large (Vavla & Gokaj, 2013) revealed that according to the majority of learners it was the teachers and teachers only that were responsible for feedback provision. Among the reasons the learners most frequently recalled were their own insufficient L2 competence, disinclination to engage in peer- and self-assessment and belief that assessment is the teacher's duty. In a similar vein, in a study conducted by Muñoz and Alvares (2007), most learners considered their teachers to be objective evaluators, whereas nearly half of them were distrustful of their own objectivity in the process of self-assessment. On the other hand, these conventional and teacher-centred perceptions may stem from the fact that some learners have never, or hardly ever, been afforded an opportunity to self-assess their own, or their peers' work (cf. Vavla & Gokaj, 2013).

Some researchers have aimed to harness learners' affective responses to assessment-related situations. For instance, a qualitative study based

on a draw-a-picture technique and interviews revealed that adolescent learners associated the assessment procedures they experienced at school with teachers' criticism and reprimand, low grades and the stress related to a test-taking situation. Significantly fewer learners referred to positive affective states related to assessment, and these mainly derived from having a sense of achievement, being praised by the teacher and receiving support from teachers and peers. Irrespective of their overall opinion about assessment, most learners viewed summative indicators such as grades and marks as an important element of self-image (Xiao & Carless, 2013).

The complexity and variability of secondary school learners' attitudes to high-stakes examinations is illustrated in a longitudinal study conducted in the Finnish context by Huhta, Kalaja and Pitkänen-Huhta (2006). Oral diaries kept over a time-span of 4 months enabled the researchers to gain insights into the evolution of participants' perceptions regarding the process of preparing for the final examination and their attitudes to test taking itself, as well as their reactions to their results. On the basis of the learners' perceived importance of the test, the time devoted to preparation and the self-perceived impact of this on test results, the researchers identified several roles the participants adopted to

talk about the final examination. The researchers termed the positive attributions *hard-working, skilled, lucky* and *cool*, and the negative attributions *lazy, poor, unlucky* and *nervous*. Additionally, it was reported that variability of perceptions was observed not only between learners, but also in the case of individual participants, whose perceptions changed over the course of the study (Huhta, Kalaja & Pitkänen-Huhta, 2006).

On the basis of a literature review, McMillan and Workman (1998: 22-23) formulated a list of assessment practices conducive to creating a favourable assessment environment, which involves (1) informing learners how learning will be assessed; (2) offering specific feedback that helps learners improve their performance; (3) using mistakes to improve future learning; (4) using assessment tasks of moderate difficulty; (5) collecting information about learners' performance by means of multiple and diverse assessment types; (6) employing authentic and meaningful assessment tasks; (7) using clearly articulated assessment criteria throughout the assessment; (8) familiarising the learners with the criteria before the assessment task is administered to facilitate systematic revision; and (9) offering individualised incremental feedback to increase learner self-efficacy.

3. Research methodology

3.1. Research objectives

This qualitative study based on the critical incident technique aims to investigate adolescent learners' perceptions of lived experiences of assessment in the school context. The following research questions were addressed:

1. What instances of oral and written assessment do adolescent learners consider as particularly positive and negative?
2. What properties of oral and written assessment appear salient to adolescent learners?

The analysis of both positive and negative experiences will also offer valuable insight into classroom assessment practices, in particular as regards grading policies, the use of criteria, forms of feedback provision and variety of task types used.

3.2. Participants

The study was conducted in a lower secondary school in a large city in Poland. Convenience sampling was used to select the participants. To avoid any bias in further analysis, the researcher was not affiliated with the school in any way and had not known the participants prior to data collection. Twenty Polish adolescent learners (11 girls and 9 boys) aged 14-16 took part in the

research. The learners represented A1+/A2 levels of proficiency, as estimated by the school teachers. Even though they were taught by different English teachers, the language instruction they received was otherwise comparable: there were 3 English lessons per week, and the teaching was based on the same syllabus and textbook. The language assessment procedures were delineated by an internal set of assessment regulations that had been developed and approved by all the language teachers in the school. According to these regulations, FL assessment should involve written tests and revision quizzes as well as more formative techniques (e.g. group/individual presentations, projects, in-class and homework assignments) that would aim to elicit the four language skills and aspects of language (grammar, vocabulary and pronunciation). Conducting the study in one school only aimed to ensure comparable teaching content and assessment regulations. It also helped to explore the assessment environment in a specific context.

3.3. Instrument and procedures

The Critical Incident Technique (CIT) was adopted in the present study to examine learners' perceptions of the oral and writing assessment they have experienced during lessons of English as a foreign language. The CIT can be defined as

> A qualitative interview procedure which facilitates the investigation of significant occurrences (events, incidents, processes, or issues) identified by the respondent, the way they are managed, and the outcomes in terms of perceived effects. The objective is to gain understanding of the incident from the perspective of the individual, taking into account cognitive, affective, and behavioural elements. (Chell, 1998: 56)

Since the CIT offers both a description and "an interpretation of the significance of the event" (Tripp, 1993: 3), this technique enables researchers to gain insights into participants' behaviour, perceptions and attitudes. Gremler (2004) highlights that in order to use a critical incident for further analysis, the researcher needs to (1) define explicitly how a critical incident is understood in the investigated context and (2) provide criteria for including or excluding the incident from the study. Introduced by Flanagan (1954) as a set of clearly defined procedures for the purposes of retrospective qualitative research, CIT was first used mainly in the social sciences. Recently it has also been applied in studies of foreign language learning and teaching (e.g. Czura, 2017; Finch, 2010; Gabryś-Barker, 2012; McClure, 2007).

Written parental permissions were obtained at the onset of the study. The data were collected during individual oral interviews that took place in

a separate room. During the interviews, the learners were asked to elaborate on and interpret one particularly positive and one particularly negative instance of FL speaking and writing assessment they had experienced. The incidents could refer to both in-class and homework assignments. When needed, the researcher asked facilitative questions to encourage the participants to interpret the recalled situation, examine their own reactions and critically evaluate the meaning of the incident and its consequences. Additionally, further questions were asked as regards the administration of the assessment task, grading and the criteria used. The interviews were audio-recorded, transcribed and then content analysed by means of NVivo 9. Inductive content analysis was used to identify main categories that emerged from the data. As many of the responses expressed multiple meanings, the number of codes exceeds the number of critical incidents reported by the participants (as illustrated in Figure 1 and 2 in the results section).

3.4. Results

Following Gremler's (2004) specification presented above, for the purposes of the study, a critical incident referred to a distinctive and clearly identifiable situation that was directly connected with classroom-based assessment of either writing or speaking. Additionally, all the critical incidents had to refer to assessment-related situations that took place in the school in which the data was collected. Any responses that failed to meet these criteria were excluded from further analysis.

3.4.1. Critical incidents of writing assessment

Fourteen accounts of positive incidents underwent further analysis – 4 reports were excluded as they offered a general description of the assessment process rather than an account of a single incident, and 2 participants failed to recall a singular positive situation. Of the 20 responses, 12 were classified as negative incidents. While one person considered the assessment of written skills as predominately repetitive and was thus unable to point to one particularly negative situation, as many as 7 participants did not indicate any negative incidents, claiming that they either could not recall or they had not experienced any negative assessment situations. All the writing assessment tasks recalled by the learners were graded, except for letters addressed to exchange partners, which, although not graded directly, were taken into account in end-of-semester grading.

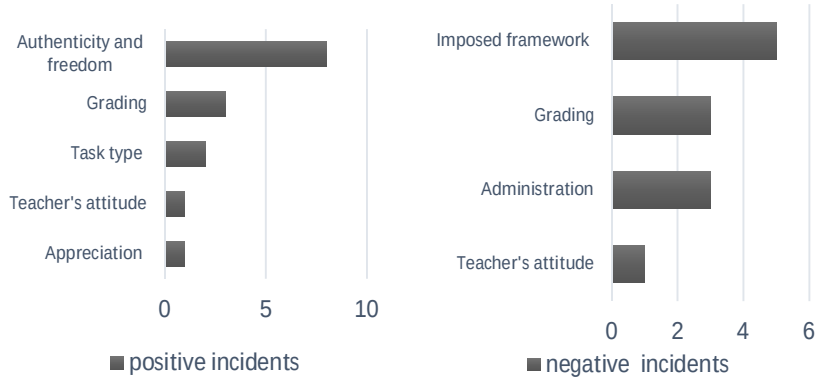

Figure 1: The categories of learners' perceptions of writing assessment.

On the basis of content analysis, five main categories of positive incidents were found. The frequencies of codes within each category are provided in Figure 1. The results reveal that eight learners attached considerable importance to the freedom of expression and authenticity of written tasks. As regards positive examples of text types, the learners mentioned a biography of a famous person, description of a selected painting from the National Museum, personalised texts about their animals and other aspects of life, an entry about their city for a guide book, or a letter to an exchange partner abroad. What they particularly appreciated about these tasks was the possibility of using their knowledge and expressing opinions. As one of the learners noted, "While [describing a city] I could use my extra knowledge, present something other students might not know. I could also express my own opinion – write what I like and what I dislike about my city" (P15).

The student involved in an international exchange programme underlined the benefits of personalised content and communicative language use in an authentic context: "Firstly, I liked being able to communicate with a different country, but it was not the most important element. We could simply write what we wanted, in any form, on any topic. We wrote what we really wanted. Then we answered their questions and so on" (P20). The fact that the teacher evaluated these letters did not seem to be of much importance to the learner. Finally, one learner viewed writing a story as a means of stimulating his/her creativity: "I'm a humanistic and very creative person. I often forget about the form and focus too much on the content I want to convey" (P7). Interestingly, a written task about volunteer work had a more substantial impact on one of the participants:

P18: We had to imagine we work as volunteers in a charity organisation and to write what we do there. I liked it a lot – it was great!
R (researcher): What did you like about it?
P18: Why? It was simply interesting (...) and encouraged me a bit to take up voluntary work.

Moving onto the next category, it was observed that for three people it was a good grade that made the assessment situation particularly memorable. The reasons varied – for one person a four[2] was enough, whereas other learners particularly appreciated being awarded the maximum number of points, or a five. Participant P1 recalled a situation in which his/her efforts were additionally appreciated by the teacher in front of the whole class: "I wrote [a biography] really well as I had revised it at home. So, it was relatively easy and I got a five. Also, the teacher praised me in front my peers. It was so great".

In the two incidents coded as "task type", the learners pointed out writing a short story (P3) and being engaged in collaborative writing (P4), which was seen as a solution to the heavy time-load associated with writing tasks: "I think it's better to work together in class, do some activities in groups (...). Otherwise, instead of boasting to my parents about the five I got at school, I would need to work on another written assignment at home". Finally, one participant underlined the importance of the teacher's attitude and the quality of corrective feedback: "[The teacher] really made an effort to check our work. [The teacher] did it honestly and with great care" (P2).

One of the learners made an interesting observation about writing assessments in the classroom that might shed light on the relatively high number of positive incidents and learners' predominantly positive attitudes to writing assignments and teacher feedback in the present study. It appeared that writing tasks were graded more leniently in comparison to other assessment tools: "When it comes to writing tasks, we hardly ever get a bad grade. We usually get worse grades for homework and tests" (P15).

As regards the negative incidents (see Fig. 1), the learners mainly complained about written tasks in which the form, the word limit or the topic were imposed by the teacher. As one of the learners noted, "I dislike the monotony – those questions in emails or letters are so repetitive. One can learn them by heart as the content of the letters is the same over and over again" (P17). As one participant added, a description of a day that needs to imitate an imposed pattern and include a set of expressions cannot be regarded as an authentic

[2] The Polish grading system is based on the 1 to 6 scale, where 1 is the worst mark (fail) and 6 is the best (exemplary).

text: "It is not MY day!" (P13). The learner wanted to be able to identify with the text he/she authors and then to "stand in front of the class and read a text about what really happened" (P13). Another critical incident within this category referred to an imposed topic that proved too difficult for the learners: "I wouldn't have known what to write in Polish on this topic" (P18). One learner complained about an imposed word limit that made it impossible to express his/her opinion and use a wide range of vocabulary and grammatical structures.

Three learners believed their performance had been graded unfairly. Three cases stemmed from the lack of clearly stated criteria and insufficient formative feedback on learners' written performance. As one of them noted, "Although I had fewer mistakes than my friend, who is more or less on the same level, I got a lower grade" (P19). Another learner was dissatisfied with the four he/she got from an optional task: "I think that in the case of optional tasks, it is the learner who should decide whether or not to accept the grade". The three incidents that were assigned to the "administrative aspects" category referred to situations in which the learners failed to complete the tasks successfully due, in their opinion, to receiving an ambiguous or imprecise instruction for the task. Finally, one learner was disheartened with the fact that the teacher failed to correct or to provide any feedback on his/her written work.

3.4.2. Critical incidents of oral assessment

The participants provided accounts of 10 positive and 8 negative incidents of assessing oral skills. The small number of incidents in these two categories results from the fact that, in the participants' opinion, oral skills were not assessed at all (7 learners), or because the participants failed to recall any particular situation, either positive or negative.

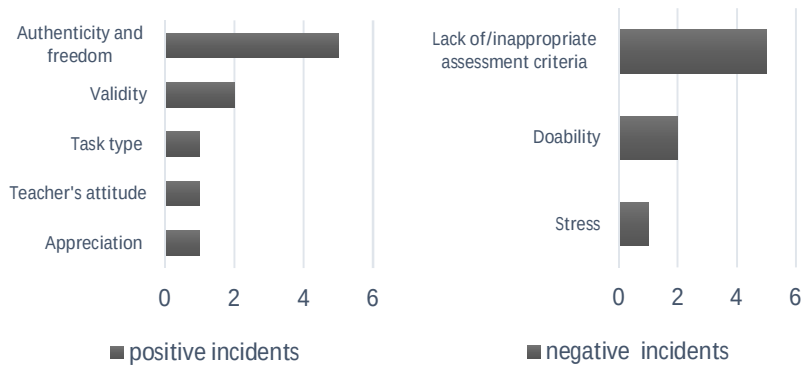

Figure 2: The categories of learners' perceptions of oral assessment.

Similar to the results obtained in the previous section, oral assessments that involved authentic, purposeful and meaningful content were most frequently singled out as examples of positive incidents (see Fig. 2). Within this category, the participants appreciated tasks that involved meaningful contexts ["it was just a casual conversation about shopping – it prepared us for real-life communication" (P15)], a free choice of topic and content, creativity and fluent speech [e.g. "The teacher assessed our fluency and whether we say what we really want to say, I mean, our opinions. It was the most important thing" (P9)].

Validity, understood here as the alignment of assessment tasks with the objectives and content of teaching, appeared to be the second most important aspect of assessment. The learners underlined that the assessment tasks had been practised in class or at home before the assessment took place: "the good thing was that the teacher asked questions and we had two or three days to prepare the responses. Then we answered the questions and got a grade" (P13). Finally, the least frequently mentioned positive aspects of oral assessment involved the task type ["I liked when I was able to answer all the questions in a reading task" (P2)], a sense of appreciation ["I was very active that day and I got a five. I was so glad the teacher had appreciated my work" (P7)] and the teacher's helpful hints during oral assessment.

The categories of negative critical incidents of oral assessment appeared to be markedly different in comparison with the data presented previously. The first broad category referred to assessment criteria. Two learners reported that they had not been informed about assessment criteria before and/or after the oral task was assigned – as one respondent noted, "I didn't like that the teacher hadn`t told us how the task had been graded" (P17). Three other participants whose responses were coded under this category considered the assessment criteria applied as unfair or inappropriate. In one case, with the support of the rest of the class, the learner eventually managed to convince the teacher to award him/her a higher grade for an oral presentation.

In two further negative critical incidents, the task proved undoable due to a restrictive time limit and the difficulty of the task. Despite a generally positive view of an oral presentation that afforded a choice of topic, one learner complained about an overly rigid time limit: "On the negative side, you were given 5 minutes and your presentation can be neither slightly longer nor slightly shorter. It was problematic for me as I speak quite quickly when I talk to the rest of the class. I remember that although my presentation was four pages long, it took less than 5 minutes" (P16). One participant raised the issue of the stress associated with talking in front of the class: "Making a mistake in front of the whole group is quite stressful (...). The teacher asked me about

a word, and to translate it into English, and I gave a wrong one, and the whole class knew the word and everyone stared at me. It was unpleasant" (P20).

4. Discussion

The higher number of positive critical incidents as well as some respondents' inability to recall examples of negative incidents may be indicative of relatively positive perceptions of speaking and writing assessment. The learners in the sample most valued speaking and writing assessment for providing them with opportunities for authentic and meaningful FL use, self-expression and a creative outlet. This finding is in agreement with McMillan and Workman's (1998) principle of the good assessment practices that underscore the use of authentic elicitation tasks. Additionally, the learners appreciated the alignment of assessment with teaching content, and task variety, which ranged from answering a list of questions, to more creative and collaborative exercises. This finding highlights the importance of assessing speaking and writing skills on a number of occasions and by means of diverse task types in order to cater for the needs of learners with different skills, learning styles and levels of FL competence. Interestingly, although assessment is typically perceived as a source of anxiety, assessment-related stress was mentioned by only one person.

Despite the predominantly positive views of speaking and writing assessment, some worrying observations should be addressed here. First of all, as reported by the learners, the assessment of speaking skills was visibly neglected, which has been observed in a number of other educational contexts (Luoma, 2004; Vavla & Gokaj, 2013). The planning and administration of the assessment of speaking skills may raise a number of challenges to teachers, such as the need to define the construct, establish assessment criteria, select appropriate speaking tasks, manage time and provide reliable and informative feedback. Regardless of these practicality issues, speaking assessment should not be downplayed in the classroom as it is the productive skills that stand at the heart of communication. The position of speaking skills in Polish schools is alarming, especially in the light of the "Learning and teaching foreign languages in lower secondary school" study (Gajewska-Dyszkiewicz et al., 2015), which indicated that excessive focus on practising receptive skills and the teaching of grammar and vocabulary at the cost of productive skills is commonplace (see also Ellis, 2015). It is worth noting that although both adolescent learners and teachers in that study viewed speaking as the most important skill, there is a marked discrepancy between the perceptions of the frequency of speaking assessment in the school context – whereas as many as 51% of teachers claimed oral skills are assessed on a regular basis, only 23%

of learners shared this opinion (Ellis, 2018). The neglect as regards assessment of speaking skills may also result from the negative washback effect of the end-of-school examination in the foreign language that comprises only written tasks. As a result, writing task types that are likely to appear in the examination tend to be over-practised. A similar negative impact of final assessment on learning and teaching can be observed across different educational contexts (Ellis, 2018; Ene & Hryniuk, 2016; Kremmel, Eberharter & Maurer, 2018; Zawadowska-Kittel, 2013).

Based on the learners' accounts in the present study, it can be stated that the teachers' approaches to grading writing and oral assignments and the use of assessment criteria were occasionally inconsistent, which, coupled with setting unrealistic time limits and minor doability problems, indicates that the results of some assessments may not be reliable. The violation of the reliability criterion, one of the key principles of good assessment (Bachman, 1990), is even more pronounced in the incidents that pointed to situations in which the criteria were not disclosed to the learners. Insufficient familiarity with assessment criteria, also observed in other studies (Aitken, 2011; Czura, 2016), may decrease the formative value of feedback on learners' performance and, in the long term, inhibit learner motivation (Huhta, Kalaja, & Pitkänen-Huhta, 2006; McMillan & Workman, 1998). Insufficient reliability of assessment additionally raises the issue of ethical assessment – only the use of clear and consistent criteria can ensure fair, unbiased feedback that can help learners improve their future performance.

Although the interviews touched upon the sensitive topic of evaluation and assessment, the learners seemed to have been relatively unbiased in their evaluations and eagerly shared their stories of positive and negative assessment experiences. They tended to evaluate a specific situation objectively, without resorting to attacking teachers personally. Naturally, some of the problems listed by the learners, for instance, the complaints regarding word/time limits, suggest that the learners misunderstood the objective of the assessment task, which calls for a need for tasks and criteria to be presented in greater detail in the future.

Finally, it is worth analysing the use of CIT in this study on learners' perceptions. On the positive side, the CIT did not restrict learners' choices in any way – they were free to choose any type of writing or oral assessment they viewed particularly positive or negative. Throughout the study, the learners provided well-balanced and objective responses, and were, in most cases, able to provide well-developed justifications of their choices. Learners' perceptions, expressed through open-ended and unrestricted responses, revealed some variability as regards the preference for assessment task types, the perceived importance of grading and the expectations of the teacher's role in speaking and writing assessment. The CIT, therefore, offered a valuable

insight into the properties of oral and written assessment which adolescent learners consider as particularly important. However, the picture of assessment practices in the investigated context is not comprehensive as, by definition, examples of critical incidents presented by learners focused only on selected aspects of language assessment. Moreover, the results of the study refer to a small sample of learners and are constrained to the unique assessment environment of one specific school, and therefore should not be generalised to other settings. Further studies are required to gain more in-depth understanding of learners' perceptions of language assessment. Additionally, future researchers could compare learners` and teachers` views of assessment by means of both qualitative and quantitative methods.

5. Conclusions

Assessment plays a powerful role in education; therefore, every attempt should be made to provide high quality assessment tasks that are valid, reliable, practical, authentic and aligned with the learning objectives. Many incidents of assessments recalled by the participants meet these criteria, which is indicative of a generally positive assessment environment in this school. Although the final exam undoubtedly exerts a mild impact on the choice of writing tasks, the learners are often invited to write creative, personalised texts that encourage freedom of expression and authentic language use. However, there is some room for improvement, particularly as regards setting assessment criteria and informing learners about them, providing coherent feedback and following a clearly-defined grading policy. Given the subjective nature of providing feedback on open-ended task types, these aspects have a profound impact on the reliability, consequential validity and ethics of assessment. Another alarming phenomenon is the visible neglect of speaking assessment, which despite its reputation for being time-consuming and difficult to conduct, should be treated as an inherent element of classroom-based assessment.

Finally, despite a growing interest in critical language testing, in the investigated context, the oral and written assessment tend to be teacher-centred – it is the teacher who decides about the types of tasks, assessment criteria and administrative issues. Many of the misunderstandings reported by the learners in this study could be easily avoided provided teachers engage in dialogue with learners, which, as Lambert (2003: 34) stresses, may have "an almost magical effect on what we say: issues and problems are held at arms' length and examined from all sides instead of being subjected to quick opinions and ready solutions".

BIBLIOGRAPHY

Aitken E. N. (2011), *Student voice in fair assessment practice* (in) Webber C. F., Lupart J. L. (eds.), *Leading student assessment. Studies in educational leadership.* Dordrecht: Springer Netherlands, pp. 175-200.

Bachman L. (1990), *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman L., Palmer A. (1996), *Language testing and practice.* Hong Kong: OUP

Brookhart S. M. (2004), *Classroom assessment: Tensions and intersections in theory and practice* (in) "Teachers College Record", No 106, pp. 429-458.

Brooks J. G., Brooks M. (1993), *In search of understanding: The case for constructivist classrooms*. Alexandria, VA: ASCD.

Chell E. (1998), *Critical incident technique* (in) Symon G., Cassell C. (eds.), *Qualitative methods and analysis in organizational research.* London: Sage, pp. 51-72.

Czura A. (2013), *Implementing portfolio assessment in lower-secondary school* (in) "ELOPE: English Language Overseas Perspectives and Enquiries", No 10 (1), pp. 83-94.

Czura A. (2016), *Wpływ oceny koleżeńskiej na lęk przed pisaniem u studentów* (in) "Teraźniejszość – Człowiek – Edukacja", No 19 (4), pp. 97-114.

Czura A. (2017), *Adolescent learner perceptions of foreign language assessment: Critical incident analysis* (in) "Glottodidactica. An International Journal of Applied Linguistics", No 44(2), pp. 25-39.

Ellis M. (2015), *Learning and teaching foreign languages in lower secondary school (BUNJO 2012): Implications for the education of teachers of English* (in) "Neofilolog", No 44 (*2*), pp. 171-182.

Ellis M. (2018), *The impact of a national examination in foreign language on teaching, learning and the work of schools.* Kraków: Wydawnictwo Naukowe Uniwersytetu Pedagogicznego w Krakowie.

Ene E., Hryniuk K. (2016), *Worlds apart, but in the same boat: How macro-level policy influences ELF writing pedagogy in China, Mexico, and Poland* (in) Ruecker T., Crusan D. (eds.), *The Politics of English Second Language Writing Assessment in Global Contexts.* New York: Routledge, pp. 15-28

Entwistle N. J., Entwistle A. (1991), *Contrasting forms of understanding for degree examinations: the student experience and its implications* (in) "Higher Education", No 22, pp. 205-227.

Finch A. (2010), *Critical incidents and language learning: Sensitivity to initial conditions* (in) "System", No 38 (3), pp. 422-431.

Flanagan J. C. (1954), *The critical incident technique* (in) "Psychological Bulletin", No 51*(4),* pp. 327-358.

Freire P. (1970), *Pedagogy of the oppressed*. New York: Continuum.

Fullan M. (1991), *The new meaning of educational change*. New York: Teachers College Press.

Gabryś-Barker D. (2012), *Reflectivity in pre-service teacher education. A survey of theory and practice*. Katowice: University of Silesia Press.

Gajewska-Dyszkiewicz A., Marczak M., Paczuska K., Pitura J., Kutyłowska K. (2015). *Efekty nauczania języka angielskiego na III etapie edukacyjnym. Raport końcowy Badania uczenia się i nauczania języków obcych w gimnazjum i Badania umiejętności mówienia*. Warszawa: Instytut Badań Edukacyjnych.

Gremler D. D. (2004). *The critical incident technique in service research* (in) "Journal of Service Research", No 7(1), pp. 65-89. https://doi.org/10.1177/1094670504266138

Huhta A., Kalaja P., Pitkänen-Huhta A. (2006). *Discursive construction of a high-stakes test: The many faces of a test-taker* (in) "Language Testing", No 23(3), pp. 326-350. https://doi.org/10.1191/0265532206lt331oa

Jacobs G. M., Renandya W. A. (2016), *Student-centred learning in ELT* (in) Renandya W. A., Widodo H. (eds.), English language teaching today: Linking theory and practice. Switzerland: Springer, pp. 13-24.

Kelly S. P., Weiner S. G., Anderson P. D., Irish J., Ciottone G., Pini R., Griffoni S., Rosen P., Ban K. M. (2010). *Learner perception of oral and written examinations in an international medical training program* (in) "International Journal of Emergency Medicine," No 3(1), pp. 21-26, doi:10.1007/s1224 5-009-0147-2.

Kremmel B., Eberharter K., Maurer M. (2018), *Righting Writing Practices? An exam reform's impact on L2 writing teaching and assessment* (in) Ruecker T., Crusan D. (eds.), *The Politics of English Second Language Writing Assessment in Global Contexts.* New York: Routledge, pp. 122-137.

Lambert L. (2003), *Leadership capacity for lasting school improvement*. Alexandria, VA: Association for Supervision and Curriculum Development.

Lee I. (2005), *Error correction in the L2 writing classroom: What do students think?* (in) "TESL Canada Journal," No 22(2), pp. 1-16.

Lee H. H., Leong A. P., Song G. (2016), *Investigating teacher perceptions of feedback* (in) "ELT Journal", No 71(1), pp. 60-68. doi:10.1093/elt/ccw047

Leki I. (1991), *The preferences of ESL students for error correction in college -level writing classes* (in) "Foreign Language Annals", No 24(3), pp. 203-218.

Leung C. (2004), *Developing formative teacher assessment: Knowledge, practice and change* (in) "Language Assessment Quarterly", No 1, pp. 19-41.

Luoma S. (2004), *Assessing Speaking*. New York: Cambridge University Press.

McClure J. W. (2007), *International graduates' cross-cultural adjustment: experiences, coping strategies, and suggested programmatic responses*

(in) "Teaching in Higher Education", 12 (2), pp. 199-217. doi:10.1080/13562510701191976

McMartin-Miller C. (2014), *How much feedback is enough? Instructor practices and student attitudes toward error treatment in second language writing* (in) "Assessing Writing," No 19, pp. 24-35.

McMillan J. H., Workman D. J. (1998), *Classroom assessment and grading practices: A review of the literature*, Richmond, VA: Metropolitan Educational Research Consortium. (ERIC Document Reproduction Service No. ED453263)

McNamara T. (1998), *Policy and social considerations in language assessment* (in) "Annual Review of Applied Linguistics", No 18, pp. 304-319.

Messick S. (1981), *Evidence and ethics in the evaluation of tests* (in) "Educational Researcher", No 10, pp. 9-20.

Messick S. (1989), *Validity* (in) Linn R. (ed.), *Educational measurement*. New York: ACE/Macmillan, pp. 447-474.

Muñoz, A., Alvares, M.E. (2007), *Students' objectivity and perception of self-assessment in an EFL classroom* (in) "The Journal of Asia TEFL", No 4(2), pp. 1-25.

Nevo D. (1996), *School based evaluation*. Oxford: Pergamon University Press

Pajares F. (1992), *Teachers' belief and educational research: Cleaning up a messy construct.* (in) "Review of Educational Research", No 62(3), pp. 307-322.

Sahinkarakas S. (2012), *The role of teaching experience on teachers' perceptions of language assessment* (in) "Procedia - Social and Behavioral Sciences", No 47, pp. 1787-1792. doi:10.1016/j.sbspro.2012.06.901

Shohamy E. (2001), *The power of tests: A critical perspective on the uses of language test* London: Longman.

Tripp D. (1993), *Critical incidents in teaching: Developing professional judgment*. New York: Routledge.

Vavla, L., Gokaj R. (2013), *Learner's perceptions of assessment and testing in EFL classrooms in Albania* (in) "Mediterranean Journal of Social Sciences", No 4 (11), pp. 509-515. doi:10.5901/mjss.2013.v4n11p509

Xiao, Y., Carless D. R. (2013), *Illustrating students' perceptions of English language assessment: Voices from China* (in) "RELC Journal," No 44(3), pp. 319-340. https://doi.org/10.1177/0033688213500595

Zawadowska-Kittel E. (2013), *Nowa Matura z języków obcych: szanse i zagrożenia*. Piaseczno: Studio LEXEM.