

Perceptions of teacher assessment practices in ESL

Predodžbe o nastavničkom vrednovanju u kontekstu engleskog kao stranog jezika

¹Tania Žuvela Blažević, ²Mara Ruža Blažević

¹Faculty of Law, University of Split, Ulica Domovinskog rata 8, Split

²Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture,

University of Split, Ulica Ruđera Boškovića 32, Split

e-mail: ¹tblazevic@pravst.hr, ²mrblazev@fesb.hr

Abstract: *This paper is based on research investigating how assessment practices used by teachers in secondary schools in Croatia affected student perception of fairness. In particular, this research focussed on the extent to which teachers consistently implemented assessment practices and made judgements of student achievement, and how these judgements were likely to affect these student perceptions. Consequently, studying the application of teacher assessment led to an evaluation of the effects of teacher assessment practices on student perceptions of fairness. The need to assess influences the creation and implementation of assessment, which in turn influences the end result of assessment. Three secondary school subjects, including English taught as a second language in Croatia, involving 6 teachers from two Croatian secondary schools formed the research sample. Ensuing results in terms of student perception of teacher performance is focused on with results showing students rated a teacher of English the most favourably in terms of student perception of teacher fairness in implementing fair assessment practices.*

Key words: *teacher assessment, student perceptions, ESL, fairness*

Sažetak: *Ovaj se rad bazira na provedenom istraživanju o tome kako praksa vrednovanja od strane srednjoškolskih nastavnika u Hrvatskoj utječe na učeničke predodžbe o pravednosti*

iste. Ovo se istraživanje posebice usredotočuje na to koliko nastavnici primjenjuju dosljednu praksu vrednovanja i donose sudove o učeničkim uspjesima te na to koliko je vjerovatno da ovi sudovi utječu na učeničke predodžbe. Posljedično tome, proučavanje primjene nastavničkog vrednovanja iziskuje evaluaciju utjecaja nastavničke prakse vrednovanja na učeničke predodžbe pravednosti. Potreba za ocjenjivanjem utječe na kriterije i primjenu vrednovanja, što zauzvrat utječe na krajnje rezultate samoga vrednovanja. Tri srednjoškolska predmeta, uključujući engleski kao strani jezik, i 6 nastavnika iz dvije hrvatske srednje škole predstavljaju uzorak ovoga istraživanja. Istraživanje se također usredotočuje na proizlazeće rezultate u smislu učeničkih predožbi o nastavnikovoj pravednosti u radu te rezultate koji pokazuju da su nastavnici engleskoga jezika procijenjeni u najboljem svjetlu, odnoseći se na učeničke predodžbe o nastavnikovoj primjeni pravedne prakse vrednovanja.

Ključne riječi: *nastavničko vrednovanje, učeničke predodžbe, engleski kao strani jezik, pravednost*

1. Introduction

How positively students perceive their teachers can have the potential to influence short- and long-term consequences. Short term, it may impact on motivation and how much effort students put into the subject. Long term it can impact students' attitudes towards education and institutional policy. Students' well being can be affected by their perception of teachers and what students believe can affect the way they perceive the world (Truman, 1976).

2. Processes

Assessment involves a number of processes including determining what is going to be assessed, what form the assessment will take, creating the task, implementing the task, judging the evidence resulting from the task, and reporting the information gained from evidence of task implementation. Conclusions are often made on the basis of this information and these decisions assist in formulating educational decisions having wide reaching effects.

Processes are generally organised and facilitated by teachers (Brookhart, 2015; Strahinić, 2011); often to comply with requirements set by education authorities or school leadership. Teachers and students accept this because they are both parties that are involved in the formal education process in the roles of teacher and student respectively, which automatically makes them the assessor and assessee (Domovic, 2004). This means that teachers implement assessment tasks regardless of the possible effects this can have on students. Furman (2009)

talks of the negative effects formal assessment can have on students yet despite this, teachers feel forced to implement formal assessment processes on students and students comply with teachers' demand from teachers. Even if students react negatively, both parents and school leaders accept the place of formal assessment, which continues to take place regardless (Furman, 2009). This is not unusual in Croatia where many view the teacher as the sole assessor of knowledge, and not primarily as a facilitator for the students to create knowledge and self-assess that knowledge (Pivac, 2009).

3. Practices

Assessment practices are the methods used for assessing student performance. Practices are different to processes in that processes are sets of activities whereas practices are what is specifically done. Scott, Webber, Lupart, Aitken & Scott (2014) believe that these methods should be accurate and should be aligned with the purpose of the assessment, the context of the assessment, scoring procedures should be appropriate, the interpretation of the assessment results should be accurate and communication of these results should be clear. Scott et al stated what they found to be fair and equitable in student assessment as a result of study involving questionnaires and interviews. It was outlined that improvement in teachers' professional abilities in student assessment and knowledge of student diversity are needed in order to achieve optimal fairness and equity. This process needs to commence right from teacher training and continue into the classroom.

That assessment practices must be fair and equitable is a value that must be promoted (Scott et al., 2014). Scott et al. believe assessment must be fair and non-discriminatory and this should be reflected in education policy and practice, which incurred various changes due to the study undertaken in Scott and al.'s study. They emphasise that all participants in the school community must strive for fairness and equity in assessment. Fairness is an important concept in education, which includes teacher assessment (Klenowski, 2013). Klenowski states that the concept of fairness in assessment is giving students equal opportunities to demonstrate acquisition of knowledge and skills while taking into consideration various social contexts. Granting the same conditions and chances for all students to demonstrate abilities provides the base for achieving fairness in assessment in that results achieved under the same conditions are as fair as possible and are perceived by students to be as fair as possible (Alm & Colnerud, 2015; Brookhart, 2015; Scott et al., 2014). It is important to consider what

precedes assessment as in access and resources as well as the consequences of assessment and the interpretation of the results of assessment.

Assessment practices tend to be a private concern and there is a lot of variation among teachers in the practices they adopt (McMillan & Workman, 1988, cited by Alkharusi, 2015). Furthermore, there is the research of McMillan and Workman, which examined student perceptions of teacher assessment.

4. Perceptions

Assessment practices result in student perception of teacher assessment. Here, understanding student perceptions of teacher assessment becomes an issue. There seems to be a variation in perception among students of teacher assessment. The results from the research done by Alkharusi, (2015) confirm that there is a shared common experience about assessment and a common student perception in the same class about their teacher's assessment practices. Alkharusi expands that, not only should individual student perceptions of teacher assessment be considered, but also the average perceptions of all the students in that class environment. Alkharusi then classifies that particular class environment as the objective environment. This is the environment, which is believed to be real. This means that the perceived classroom assessment environments were taped at a collective classroom level rather than at an individual level. Then, ensuing from this environment there were differential class results arising from teacher assessment practices (Alkharusi, 2015). That is, different assessment practices were implemented resulting in different results. For example, there were different class perceptions based on the gender weighting of the class. In conclusion, Alkharusi believes that the perceived classroom assessment environment requires more precise measurement.

However, where an individual student's perception of teacher assessment practices is different to that of the whole class, this student's perception might be diminished or less relevant. This reflects the concept of *majority rules* where the majority's view of the teacher's assessment practices becomes the dominant one. The student experience becomes a group one and varies between classes and teacher assessment practices affect this group experience (Alkharusi, 2008). The student experience is better interpreted as a collective rather than as an individual response. Alkharusi (2008) mentions that students will have a more positive perception of the classroom environment if there is better student – teacher communication about classroom assessment. This study revealed that the perceptions of student research participants on

assessment tasks all resulted in positive influences on student self-efficiency and motivation levels. The conclusion was that teacher assessment practices related with student perception of the assessment. Even though better classroom communication between teacher and student had no real influence on student achievement, it resulted in a more positive student perception of the classroom practices.

Dayal and Lingham (2015) discuss the differing conceptual understandings and perceptions held by teachers about assessment. These can be assessment for improving teaching and learning, accountability both for students and institution, and finally that assessment should be rejected because it is irrelevant and invalid. In other words, that assessment produces negative outcomes and therefore should not be included. Teacher perception is very important as it will influence teacher practice (Dayal & Lingham, 2015). Dayal and Lingham's study included investigating Fijian teachers' conceptions of assessment through interviewing over seventy teachers. This revealed that most teachers believed assessment was for formative purposes while these teachers generally held a learning view of assessment. Assessment for them had a learning purpose.

How assessment is perceived is related to whether it is viewed as being valid and reliable. Both teachers and students can regret assessment and both can believe that in some cases the assessment implemented was unjust (Alm & Colnerud, 2015). This study looked at teachers' view of unfair grading. Many student questionnaires contain inquiries into how students perceive assessment and the degree to which they perceive assessment to be a fair measure of their work (Dhindsa & Omar & Waldrip, 2007; Dorman & Knightley, 2006; Dorman, Waldrip & Fisher, 2008). Discussion on validity and reliability ensue as entailing perceiving the assessment as measuring what it is supposed to measure and as being an accurate measure.

5. Research

This research dealt with unearthing teacher assessment practices in terms of fairness and consistency and student perception of this in terms of fairness and consistency. The data was collected in Croatia, which is a transitional country transferring from communism to democracy where problems in fairness and consistency in teacher assessment and student perception of this exist.

In the reliable scales of Congruence and Authenticity, Teacher of English (Teacher 1), scored the highest means of 3.62, and 3.45 respectively. This was scored the highest mean above by at least 0.20 from all the other teachers. For the third reliable scale of Transparency, Teacher 1

scored second highest at 2.40 higher only achieved by Teacher 6 (Biology) who scored 2.44. In the unreliable scale of Consultation teachers 1, 2 (Croatian) and 3 (Biology from school 2) scored a mean response above 3 often with Teacher 1 having the highest mean level for consultation at 3.42. In the second unreliable scale of Diversity, Teacher 1 scored third highest at 2.13 and Teacher 4 (another teacher of English) second highest at slightly higher 2.14 with Teacher 2 scoring the highest at 2.23. Individual case studies for each teacher now follow.

In this section, analysis of each teacher will be presented. For each of the six teachers involved, firstly, the results of analysis of the quantitative survey data will be given, followed by an analysis of the qualitative data from interviews and open-ended questions from the survey.

6. Research results

6.1. Case study 1

Table 1 represents a summary of sample targets. Teacher 1 who was female was from School 1 and taught the subject of English as a second language in classes 1 and 6. Bearing in mind the numbers of students who actually completed the survey, these classes comprised 19 and 28 students respectively. The former class was comprised of 14 females and 5 males, and 22 females and 6 males in the latter, making a total of 47 students in these two classes.

Research involved examining selected teachers and their students regarding the fairness of, or perceptions of assessment practices. Three secondary school subjects were targeted, namely Biology, Croatian (as an official language) and English (as a foreign language). Therefore, our research included teachers and students in three different subject areas from two secondary schools. The two secondary schools would both be *gimnazija* (secondary schools with focus on gaining entry to university). All of the targeted students and teachers were involved in the three selected subjects in their respective teaching programmes. The selection of teachers were 2 teachers who taught each of those three chosen subjects in that particular school.

Table 1

Summary of sample targets.

Subjects	3 subjects -mixture of humanities/science: Biology, Croatian, English
Schools	2 High Schools with same timetabled programme of these 3 subjects
Classes	Entire classes of final year students in these 3 subjects
Students	Approximately 30 students per class making a total of 330 students
Teachers	2 teachers from each subject making a total of 12 teachers

Source of literature: Authors' own research

The data collection methods here included **structured interviews** of each of the teachers selected for the study and **self-completion questionnaires** for the students of the subjects being researched. In addition, a structured set of assessment documents was collected from each teacher to conduct a **document analysis**. The questionnaire was based on two existing instruments. Firstly, the *Student Perceptions of Assessment Questionnaire (SPAQ)* developed by Dorman, Waldrip and Fisher (2008) and secondly the student questionnaire developed by the Centre for Schooling and Learning Technologies (CSaLT) for the Digital Forms of Assessment project (Newhouse, 2012). The SPAQ has five scales labelled *Congruence with planned learning, Authenticity, Student consultation, Transparency and Diversity*. The CSaLT instrument provided some open-ended questions and some questions related to the experience of students in completing assessments.

6.2. Results in terms of teachers

The results of analysis of all the data (students' questionnaire and teacher interviews) for each of the six teachers are now presented as case studies. Each teacher had two classes of the same subject Biology, Croatian or English. Both classes of each teacher were homogenous in that they were only grouped according to alphabetical order of surname, not ability. Students entering that school had to achieve the minimum entry requirement, which was an A grade average from their results in Grade 7 and 8 in primary school. Therefore, the two classes were taught the same syllabus and curriculum as stipulated by the Croatian Education Department. To assist in comparing between these teachers in the case studies, the means for the scales from the student questionnaire for the six teachers are presented in Table 2. and Figure 1 in tabular and graphical form respectively.

Table 2

Mean scores for the student questionnaire scales for each teacher

Scale	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6
Congruence	3.62	3.48	3.41	3.38	3.37	3.45
Authenticity	3.45	3.12	3.11	3.28	3.29	3.31
Consultation	3.42	3.09	3.08	2.99	2.96	2.89
Transparency	2.40	2.37	2.18	2.41	2.29	2.44
Diversity	2.13	2.23	2.05	2.14	1.82	2.06

Source of literature: Authors' own research

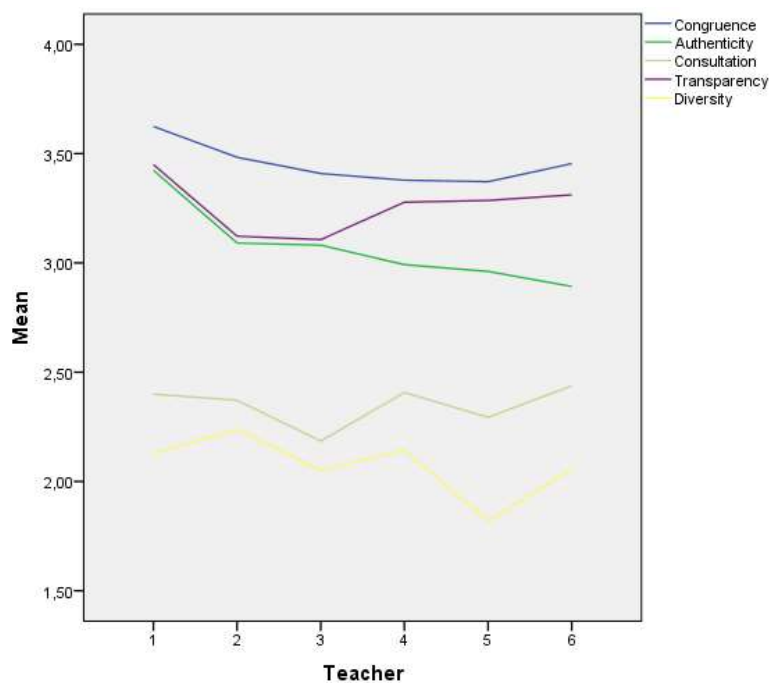


Figure 1 Mean scores for student questionnaire scales

Figure 1 shows mean scores for the student questionnaire scales. In the reliable scales of Congruence, and Authenticity, Teacher 1 scored the highest means of 3.62, and 3.45 respectively. This was scored the highest mean above by at least 0.20 from all the other teachers. For the third reliable scale of Transparency, Teacher 1 scored second highest at 2.40 higher only achieved by Teacher 6 who scored 2.44. In the unreliable scale of Consultation teachers 1, 2 and 3 scored a mean response above 3 often with Teacher 1 having the highest mean level for consultation at 3.42. In the second unreliable scale of Diversity, Teacher 1

scored third highest at 2.13 and Teacher 4 second highest at slightly higher 2.14 with Teacher 2 scoring the highest at 2.23. Individual case studies for each teacher now follow.

In this section, analysis of each teacher will be presented. For each of the six teachers involved, firstly, the results of analysis of the quantitative survey data will be given, followed by an analysis of the qualitative data from interviews and open-ended questions from the survey.

6.2.1 Quantitative data results

As discussed earlier, the mean scores of the student questionnaire responses were highest for Teacher 1 (see Table 2) and significantly higher than the other teachers for the reliable scales of Congruence, Authenticity and Transparency. On average, this indicates that students perceived Teacher 1's assessments to be a more commensurate measure of their abilities than the others. The only reliable scale for which Teacher 1 did not achieve the highest mean score was Transparency, suggesting this teacher's assessments were not viewed by students as the most coherent of all. Scores for all scales were then analysed by means of an ANOVA test to detect differences and similarities between classes 1 and 6 taught by Teacher 1.

A significant difference ($p < 0.0001$) in mean scores can be seen for the unreliable Diversity scale. For the three reliable scales, Congruence, Authenticity and Transparency, no significant differences emerged between the teacher's two classes. It can therefore be reasonably concluded that students' perceptions of assessments were similar for these measures in both of Teacher 1's classes.

6.2.2 Qualitative data results

Teacher 1 was asked the standard interview questions. A summary follows, together with explanations about each question in turn. Student responses from both classes to the open-ended questions supplemented the discussion. This procedure was repeated for each teacher and her two classes. Since they taught the same subject to both classes, teachers were interviewed once – their responses are paraphrased and discussed in further detail below.

Question 1: What assessment tools do you use?

Teacher 1 used both oral and written tasks for tests and assessments. The specific ratio of oral to written tasks was not stated, but it is common practice in Croatia to use both in equal measure. She also used essay and grammar tasks, and endeavoured to assess all four foreign-

language skills: reading, writing, listening and speaking. Assessment tasks included learning texts off by heart from textbooks and reproducing them orally, as well as grammar tests requiring rewriting of sentences, filling in gaps with correct verbs or prepositions. Other tasks could be found in matriculation examination papers, such as cloze (gapped text for students to provide missing words), word formation, multiple-choice, reading, comprehension and composition.

Question 2: How often do you use these assessment tools?

This teacher orally assessed each student three times a year and conducted written tests (English writing tasks given to the whole class under test conditions) four to five times a year, representing the obligatory number of assessments as stipulated by the state education department. Between two and four assessments took place per semester. Students were required to write an essay twice a year, in addition to two grammar and two vocabulary tests. Oral examinations were undertaken twice a year.

While written assessments were carried out with the whole class under test conditions, oral assessments did not involve all the students at the same time. This form of assessment is prevalent in Croatia, but unlike written tests, students are not all examined at the same time, nor asked the same questions under the same conditions. In oral assessments, teachers typically ask questions of students one at a time while the others listen, so the process is lengthy, drawn out and lacking in rigour and consistency.

Question 3: Do you give the students assessment criteria?

Teacher 1 clearly explained what she expected from her students without providing specific information on how assessments would be scored. She firmly believed that students were clear about assessment requirements.

Question 4: If so, in what form?

Teacher 1 stated her students had to write down the assessment criteria after she clearly explained them and copied the county grids provided by the Croatian Education Department. She believed it was important for students to be aware of the assessment criteria because they explained the marks, how they were allocated and their meanings. Teacher 1 informed the researcher that it was the responsibility of the subject teachers in each county to decide the criteria for assessments, write them up and make them available in the school. It should be noted that all schools, by decree of the Croatian Education Department, were required to draw

up assessment criteria for all their subjects at a county level. This was usually done by the subject teachers who got together and formulated the assessment criteria.

Question 5: How consistently do you apply them?

This teacher claimed emphatically that she applied the criteria consistently, but later admitted that she did not always do so fully. To some extent her consistency depended on the task, and she did not always apply all criteria to the same task, possibly deeming some criteria not as important as others. She explained that students' grades for written tasks were sometimes erratic, and claimed not to place students into categories after just one test, a frequent practice in Croatia whereby teachers continue to award some students the same mark awarded for their first assessment task.

Question 6: What do you base your assessment most on?

In Teacher 1's opinion, thoroughness, hard work, diligence, and depth were the most important features of students' work. She also considered grammatical accuracy, vocabulary, word choice, and sentence construction important aspects for assessing a foreign language.

Question 7: What do you think students think about how fair your assessment is?

Teacher 1 thought many students' first impression of her was that she was too strict, but as they got used to her and the way she assessed them, their perceptions changed and they became more positive. She believed students were satisfied with her assessments because they knew what she expected.

6.2.3 Commentary on student and teacher responses

In her interview, Teacher 1 said she applied the assessment criteria very consistently, but acknowledged there were sometimes inconsistencies, because in her view, not all assessment criteria applied to every task in the same way. Her seemingly conflicting statements could be construed as evidence of inconsistency, yet the students' comments appear to indicate high levels of satisfaction with this teacher's assessment practices and procedures. Students in both classes 1 and 6 indicated in their open-ended comments section of the questionnaire that assessments were fair. One student in class 6 corroborated Teacher 1's remark about some students who undeservedly continued to receive consistently high grades for assessment tasks. The most frequently mentioned negative issues associated with assessments were concerned with stress, and in some cases, inappropriate assessment tasks for evaluating students' capability. For example, some students perceived literature as unimportant for learning a

language. Nevertheless, students in this particular English class were evidently satisfied overall, especially in comparison with the others.

Amongst the very few negative responses were claims from three students in one class that the topics/textbooks were boring, one student even questioned the relevance of the contents. The remainder of the student responses (15) were overwhelmingly positive regarding assessments and content. Responses about teacher involvement and notifications of upcoming assessments were positive, and students spoke affirmatively about their participation with the teacher. This was consistent with the results of the survey scales, showing high positivity for Teacher 1 compared to other teachers, and the highest mean score of all teacher participants in this research (see Table 2). It is possible there was a particular connection between this group of students and the teacher that may have influenced their generally positive perception of the subject. A typical comment from students in classes 1 and 6 was “*the teacher is understanding and involved*”. The majority of students (7) found that assessments reflected real life and were useful.

The only significant difference between Teacher 1’s two classes was noted in the unreliable scale of Diversity. In all other respects, the students in her two classes had similar views in relation to the other four scales, so it can reasonably be concluded that the students’ perceptions of this teacher’s assessments were similar for classes 1 and 6 in the subject of English.

Teacher 1 scored the highest for the Transparency, Authenticity and Consultation scales, consistent with students’ mostly positive comments. Other comments on what students liked and disliked about assessments in this subject were all positive and dispute any belief that assessment was a negative experience for students. Some comments about lack of objectivity were clarified by students to mean the task’s lack of objectivity (e.g. literature not relevant to foreign language learning), not the teacher’s lack of clarity in explaining the task.

6. Conclusion

Teacher 1 was viewed most positively. However, the limitations of this research must be emphasised. An important one are the limits of the student, teacher, subject and school research sample. It would need to be carried out on a much greater varied sample to increase validity. Student perception of teachers is multi-scaffolded area which can be viewed from many aspects. Given perspective’s many possible consequences, it most certainly warrants

further research and discussion. The time of educational reform in Croatia as a result of Croatia's entry into the European Union is extremely appropriate for undertaking this task.

Works cited

- Alkharusi, H. (2008). Effects of Classroom Assessment Practices on Students' Achievement Goals. *Educational Assessment*, 13(4), 243-266.
- Alkharusi, H. (2015). An evaluation of the Measurement of Perceived Classroom Assessment Environment. *International Journal of Instruction*, 8(2), 45-54
- Alm, F., & Colnerud, G. (2015). Teachers' Experiences of Unfair Grading. *Educational Assessment*, 20:2 132-150, DOI: 10.1080/10627197.2015.1028620
- Brookhart, S. (2004). *Local Districts Assessments: one indicator for the validation of Nebraska's standards, assessment, and accountability system*. Final report prepared for the Nebraska Department of Education, December.
- Brookhart, S. (2013). The use of teacher judgement for summative assessment in the USA. *Assessment in Education: principles, Policy and Practice*, 20(1), 69-90. doi: 10.1080/0969594X.2012.703170.
- Brookhart, S. (2015). Graded achievement, Tested Achievement, and Validity. *Educational Assessment*, 20:268-296 DOI: 10.1080/10627197.
- Brookhart, S. et al (2016). A Century of Grading Research: Meaning and Value in the most common educational measure *Review of Educational research* 86 (4) 803-848.
- Brookhart, S., (1997) A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education* 10,161-180.
- Dayal, H. C., & Lingam, G. I. (2015). Fijian Teachers' Conceptions of Assessment. *Australian Journal of Teacher Education*, 40(8).
- Domovic, V. (2004) *Školsko ozračje i učinkovite škole (School environment and school effectiveness)* Naklada Slap.
- Dhindsa, H., & Omar, K., & Waldrip, B. (2007). Upper Secondary Bruneian Science Students' Perceptions of Assessment. *International Journal of Science Education*, 29(10), 1261-1280. doi:10.1080/09500690600991149.
- Dorman, J. P., Waldrip, B. G., & Fisher, D.F. (2008). Using the Student Perceptions of Assessment Questionnaire (SPAQ) to develop an assessment typology for science classes. *Journal of Science Education*, 1(9), 13-17.

Furman, C. (2009). The Dangerous Necessity of Assessment: A Teacher's dilemma. *The Coalition of Essential Schools*

Pivac, (2009). *Nacionalni okvirni kurikulum* (National Curriculum Frameworks), 2011).

Strahinić, C. (2012). *Škola bez ocjena* (School without grades). Osijek: Grafika doo.

Truman Whitfield (1976). How students perceive their teachers, *Theory Into Practice*, 15:5, 347-351, DOI: 10.1080/00405847609542657