

## TOWARDS THE CRITERIA AND INDICATORS FOR ESSAY ASSESSMENT

Violeta JANUSHEVA, Milena PEJCHINOVSKA, Jove Dimitrija TALEVSKI  
University “St. Kliment Ohridski”, Faculty of Education,  
Republic of North Macedonia  
milena.pejcinovska@uklo.edu.mk

**Abstract:** *In the teaching practice, the essays assessment is quite challenging due to their complex nature. Thus, to increase the objectivity of their assessment, analytical lists (rubrics) are recommended. From here, this paper deals with an attempt to establish general criteria, and to examine the criteria and the indicators for essay assessment made by teachers regarding their measurement. This research is a qualitative one, and the sample consists of analytic lists made by teachers that refer to the introductory paragraph of a five paragraph essay which can be seen as an essay itself. Interpretative analysis, synthesis and comparison are the methods used for processing the data and reaching conclusions. The research points out that despite the efforts, some of the criteria and all of the indicators for assessment of the introductory paragraph cannot be made measurable and precise enough, primarily, due to the complex nature of the written text, which includes students' personal beliefs, attitudes, ideas and writing style which are subjected to the teachers' personal beliefs, attitudes, ideas and style.*

**Keywords:** *introductory paragraph; assessment; criteria; indicators; preciseness; analytical lists.*

### 1. Introduction

As it has already been pointed out (Black & Williams, 2001; Gojkov, 2003, p. 6; Nitko, 2004; Popovski, 2005, pp. 30-31; Janusheva & Pejchinovska, 2011; Brookhart, 2014; Talevski & Janusheva, 2015), the assessment of students' achievements is a very complex and multidimensional construct. Because of this, the assessment as a notion remains in the centre of the researchers' interest as well as of the direct and indirect participants in the teaching and learning process, thus having a meaningful value in the society. Many researchers point out the role and the significance of the formative assessment in the teaching and learning process as a continuous and ongoing process which essence is to collect and interpret various data for students' learning and achievements, to evaluate the achieved standards and goals and to determinate their value (Lubisi et al., 1997; Linsey, Webb, Panayiotidis, 1988; Black & William, 2001; Gojkov, 2003, p. 149; Nitko, 2004; Popovski, 2005; Janusheva & Pejchinovska, 2011; Wiggins, 2012; Brookhart, 2014; Talevski & Janusheva, 2015). Further, many researchers stress out the relevance of the criteria used to assess the students' achievement and their relation to the teaching goals (Popovski & Stojanovski, 1995, pp. 12-13; Huitt, 1996; Anastasi, 1998 in Bond, 1996; Gojkov, 2003, p. 162-165; Popovski, 2005, p. 199; Talevski, Janusheva & Pejchinovska, 2011). If the established criteria and especially the indicators are not precise and reliable and if they do not contain as more elements of students' achievements as possible, then the collected information as indicators of the achievements may be hard to interpret with the needed preciseness and certainty. Moreover, some researchers deal with the meaning of the learning outcomes which should precisely state what will students be able to do until the end of the teaching unit, i.e. what will they be able do with the content of that unit. They state that the learning outcomes should be well defined and stress the role of the Blooms' taxonomy in

facilitating the definition of the learning outcomes and teaching goals (Nitko & Brookhart, 2007; Clarck, 2014). In addition, the role of the feedback is largely discussed as an inseparable part of the teaching and learning process. It is comprehend as planned and well organized information that teachers give to students in order for them to independently plan their further learning, overcome the shortcomings and to take control over their own learning (Black & Williams, 2001; Pintrich & Schunk, 2002; Gojkov, 2003; Linn & Miller, 2005; Popovski, 2005; Brookhart, 2008; Janusheva & Pejchinovska, 2011; Chapuiss, 2012; Wiggins, 2012).

Though various aspects of the assessment have already been elaborated, the question of its objectivity and validity still provokes dilemmas and concerns. This especially refers to the essays, due to their complex and specific nature and, in particularly, to the criteria and indicators for their assessment, primarily regarding their preciseness as well as their measurement. Thus, this paper goes deeper in the nature of the essays, and examines the possibility of preciseness and measurement of the criteria and indicators for assessment of the introductory paragraph of the five paragraph essay.

## **2. Previous research in the filed**

The assessment of essays refers to authentic assessment because it demands from students to show their competencies, to think critically and creative and to implement their knowledge. According to Palm (2008) and Wiggins (2012), this assessment takes into account the practical students' competencies. It has been already argued (Nitko, 2004) that the assignments which refer to the practical students' competencies give clearer picture for the significance of the learning outcomes and they assess the ability for performance. Many researchers (Nitko, 2004, pp. 185-186; Popovski, 2005; Linn & Miller, 2005; McMillan, 2008, pp. 83-84; Talevski & Janusheva, 2011; Becker, 2010/2011; Janusheva & Pejchinovska, 2016) define essays as questions that should be answered in a written way and connect them with knowledge that includes students' competencies to express their thoughts, ideas and attitudes in a coherent whole. They point out that the scope and the length of the essays depend on students because they have the complete freedom regarding the answer and decision on which information to include and how to organize them. Further, they indicate that essays assess students' abilities of production, organization and expressing ideas, as well as the abilities to critical and creative thinking, to develop logical arguments and to show original thoughts. However, they address many factors that can influence the assessment, such as two or more teachers assess the same essay with a different grade, primarily because they have different assessment criteria; teachers can be inconsistent regarding the criteria, i.e. teachers can interpret the same criteria differently over a period of time; the halo effect; they are not economic. Therefore, the analytic lists (rubrics) as a meaningful tool which can increase the objectivity of the essays assessment are recommended.

### ***2.1. Analytical lists (rubrics) for essays assessment***

Analytical lists (rubrics) are considered as a very useful tool for assessment of essays and it is stated that they could direct and facilitate teachers' work, ensuring them that their assessment is valid and objective. Many researchers suggest that these lists for assessment represent a sum of criteria that include indicators which are, in fact, description of students' performances quality (process or product) regarding the criteria, and that the results of the achievements are compared to these descriptions which serve as students' feedback. They agree that the criteria and the indicators for the used criteria in these lists should be as measurable and precise as possible (Moskal, 2000; Moskal & Leydens, 2000; Popovski, 2005; Shiel & Kitching, 2007; McMillan, 2008; Brookhart, 2013; Talevski, Janusheva & Pejchinovska, 2011; Talevski & Janusheva, 2011; 2015). According to Moskal & Leydens

(2000) if well made these lists provide a consistency of the assessment regardless of who does the assessment and when does he do it. Further, as Moskal (2000) claims, it is better to say: students' mathematical operations have mistakes because the number of the mistakes can be counted, thus making this indicator measurable instead of: students' mathematical operations are good because this asks the teacher to make a judgment for what is good and what is not, which, on the other side, cannot be measured. Talevski, Janusheva & Pejchinovska (2010) and Talevski & Janusheva (2011; 2015) prove Moskal's claim to be true by pointing examples where the criteria and the indicators are precise and measurable and examples where they are neither precise nor measurable. They indicate that the words: good, solid, bad etc. have relative meaning and depend on subjective factors, i.e. on teachers' judgment of what good or bad means.

### **3. Methodology of the research**

The aim of this paper is to establish general criteria, as well as to examine those criteria and indicators for assessment of the introductory paragraph of a five paragraph essay made by teachers regarding their measurement and preciseness. This research is a qualitative one and has a descriptive design. Qualitative methods used for processing the data and reaching conclusions in the research are the following: analysis of analytic lists for the introductory paragraph of a five paragraph essay, made from Macedonian language teachers and analysis of their preparation and content; synthesis; and comparison of the criteria and of the indicators for the used criteria in analytic lists in regard of their measurement and preciseness.

The sample units, that were deliberately chosen, consist of different classes (eighth and ninth grade of Primary School, and first grade of Secondary School – High School) where Macedonian language is thought, from which 20 Macedonian language teachers gave us the analytic lists from their everyday teaching practice, that refer to the parts that every introductory paragraph of a five paragraph essay should contain.

The received data - analytical lists (rubrics) were processed in a manner characteristic for the qualitative methods. In presenting the data, the verbal-narrative style and the description of the content of the introductory paragraph of a five paragraph essay relevant to the subject of research were applied.

### **4. Results and discussion**

When assessing an introductory paragraph of a five paragraph essay, many factors, besides those mentioned in p. 2, should be taken into account because they can also affect the assessment.

#### ***4.1. The number of words in the paragraph***

The number of the words in the introductory paragraph is quite often referred by teachers as one of the criteria for its assessment. It is obvious that the analytical lists (rubrics) with criteria and indicators that refer to the number of words can be easily created.

However, though it is important for students to respect the number of words, it can be said that this is not a very relevant criterion because it do not take in mind the quality of what is written with those words and how they are organized and connected to form a meaningful text. Surely, students can write a text within the required number of words, still the text may be not significant at all and also meaningless. Students can increase/decrease the number of words in the text. Still, the text may be a coherent unit which elaborates a significant issue.

Thus, the number of words in the introductory paragraph can be one of the general criteria that are precise and measurable. And precise and general indicators can be made. Yet, this

criterion and the indicators say nothing about the quality of the written text and about students' ideas and attitudes and about the way they are presented in the text.

#### **4.2. The formulation of the headline**

As pointed out by many (Bennett & Gorovitz, 1997, pp. 1-12; Crews, 1980; Whitaker, 2009; Mohrbacher, 2011; Velasquez, 2011; Uhlig, 2012; Moore, 2014; Pineteh, 2014, pp. 12-22; Ntereke, 2015; Gallagher, 2015; Inez, 2016; Can, 2017, pp. 267-271; Janusheva, 2017) the headline is a relevant element of every essay and its concretization and preciseness affect the flow of the thoughts and the writing quality. They agree that if the headline is not formulated in a clear and precise way, students can address the issue in the text in different manner and this can lead to shortcomings in the answers, which can harden their assessment. Therefore, the formulation of the headline can be one of the general criteria to assess an introductory paragraph.

In the analytical list below (a), the criterion is the formulation of the headline and five indicators are listed:

- a) Criterion > formulation of the headline
- Indicators:
- The headline has an excellent formulation (5)
  - The headline has very good formulation (4)
  - The headline has good formulation (3)
  - The headline has a satisfactory formulation (2)
  - The headline has no satisfactory formulation (1)

Though this criterion can be one of the general criteria when assessing an introductory paragraph, its measurement involves many aspects that need to be elaborated. In the five paragraph essay the headline should reflect students' stance (pro or con) on an issue (Janusheva & Pejchinovska, 2016; Janusheva, 2017). And the headline can be formulated in different ways depending on students' personal beliefs, thought, attitudes, style etc. Still, our teaching practice indicates that two different teachers may have different opinion on the formulation of the headline on the same topic, i.e. there are teachers who like better the headline *Euthanasia – yes or no?* And there are teachers who argue in favor of the headline *The euthanasia must be legalized*. This can lead to the conclusion that this criterion is not measurable and depends on teachers' personal attitudes on how the headline should be formulated. From a five paragraph essay view, both headlines promise that in the essay one stance will be considered, i.e. in the first headline student have to choose and to defend one alternative and the second headline contains the student's stance.

As for the indicators, it is obvious that they are described with words that have a relative meaning, such as: excellently, very good, good etc., thus they do not support the validity and objectivity of the assessment. Taking in account Moskal's& Leydens's (2000) claim that if well made these descriptions provide consistency of the assessment regardless of who does the assessment and when does he do it, it can be concluded that they are not made well. These results are in accordance with the research of Talevski, Janusheva & Pejchinovska (2011) and Talevski & Janusheva (2011; 2015) who claim that the indicators should not ask teachers to judge and to make decision of what is the difference among excellent, very good, good etc. Further, these indicators do not describe what is essential in the headline of a five paragraph essay, i.e. they contain no words which will be connected with the student's stance on the issue.

The same explanation could apply to the following indicators (b), regarding the headline:

b) Criterion > formulation of the headline

Indicators:

The headline reflects the main aspects of the essay completely (5)

The headline reflects the main aspects of the essay in a large extent (4) etc.

The headline reflects the main aspects of the essay in a good extent (3)

The headlines reflects the main aspects of the essay in a satisfactory extent (2) etc.

Again, here (b) teacher is in position to judge or to think what the differences between “completely” and “in a large extent” are, and what does this mean. Moreover, he has to explain further to student 1 and to student 2 what does this difference mean, because these two students earn five or four points for the teacher’s judgment. It is naturally and it is expected that various teachers will have various opinions on whether the topic reflects the main aspects of the essay “completely” or “in a large extent”. Our daily practice confirms this. Asked to make a distinction between completely and in a large extent, for example, teachers state that they know this difference intuitively and that these words need an additional interpretation. From here, our point that these indicators cannot be made measurable is justified. Further, the noun group “main aspects” is debatable too. What does it mean when saying that the headline reflects the main aspects? “Main aspects” can be interpreted in a different way by teachers. Therefore, where an interpretation is needed there cannot be objectivity in terms of preciseness or measurement.

The same counts for the following indicators (c):

(c) Criterion > formulation of the headline

Indicators:

The headline reflects the student’s stance completely (5)

The headline reflects the student’s stance in a large extent (4) etc.

The headline reflects the student’ stance s in a good extent (3)

The headlines reflects the student’s stance in a satisfactory extent (2) etc.

### ***4.3. The structure of the introductory paragraph***

According to many researchers (Crews, 1980; Bailey, 2003; Harvey, 2009; Whitaker, 2009; Uhlig, 2012; Janusheva & Pejchovska, 2016; Janusheva, 2017; 2018) the introductory paragraph of a five paragraph essay contains a) topic sentence; b) 3 to 5 sentence which carry information that should direct and develop the topic sentence, and c) closing sentence which should contain the topic and the thesis which are to be presented in the essay, bearing in mind that the topic and the thesis are not referring to the same, for example: Euthanasia should be legalized – topic and because the patient’s right to choose should be protected + 2 reasons (Janusheva, Kostadinovska-Stojchevska, 2017).

#### ***4.3.1. The topic sentence***

The topic sentence is a significant part of the introductory paragraph. There are many steps on how to write an appropriate topic sentence and many researchers indicate them in their research. Further, the topic sentence is the first sentence in the introductory paragraph and can quite easily be distinguished among other sentences. Thus, as a main part of an essay, the topic sentence can be one of the general criteria when assessing the introductory paragraph.

The list below (d), concerns the topic sentence from the introductory paragraph of a five paragraph essay with the following headline: “Reading a book is much better than watching the film version of that book”.



(d) Criterion > topic sentence is a declarative sentence

Indicators:

The topic sentence is a declarative sentence that is meaningfully connected with the headline and contains all the key terms from the headline and/or their synonyms (5)

The topic sentence is a declarative sentence that is meaningfully connected with the headline, but it does not contain all the key terms from the headline and/or theory synonyms.

One key term is missing (4)

The topic sentence is a question and/or it does not contain all the key terms from the headlines and/or their synonyms. One or more key terms are missing (3)

The topic sentence is not related with the headline and the topic. It does not contain the key terms from the headline and/or their synonyms. It involves only one of the key terms (2)

However, notifying that the introductory paragraph has a topic sentence, though it can be one of the general criteria, said nothing about the nature and the character of this sentence and is a poor indicator that cannot contribute to the validity and objectivity of the assessment. Teacher can get 30 introductory paragraphs, for example, and every of them may have different topic sentence, some of them good and some of them not so good according to the teachers' opinion. What will be the difference in regard of the points between two topic sentences from two different paragraphs? So far, it is clearly noticed that defining the indicators with words having relative meaning is not satisfactory.

Further, the sentence: "Many people love watching the film version of the book over reading the book" is a declarative sentence, and it is connected with the headline and promises that the essay will go on as expected. However, the topic sentence can be presented in other way, for example: "Many people love reading the book over the film version of the book". This sentence is also a declarative one, it is connected with the headline and has the key words from the headline. This sentence also promises that the essay will express the main idea. It seems that nothing can be said about these two topic sentences and that they both fulfill the given criteria. Yet, one teacher can like the first better than the second. This is confirmed by our everyday teaching practice. Moreover the topic sentence can be written as: "Are people prone to love watching the film versions of the book over the book itself"? Now, it is clear that the sentence is a question sentence, yet still connected with the title of the essay and still has the key words from the title. Does this sentence seem not so good only because it is not a declarative sentence? No one can argue in favor of that. Finally, the topic sentence can be: "Many people love watching the film version of the book over the book itself"! It is now an exclamation sentence, which expresses the surprise of the student which is considered as an excellent way to catch the attention of the audience. Does this sentence seem not good? No one can argue that too. The sentence is connected with the headline and has the key words from the headline. Though it is not a declarative sentence it still promises that the essay will describe the main idea.

Thus, these indicators are also in opposition with the research of Nitko (2004, pp. 185-186), Popovski (2005) Linn and Miller (2005); McMillan, 2008, pp. 83-84; Talevski & Janusheva, 2011; Becker, 2010/2011) who argue that essays measure the following learning outcomes: production, organization and expressing ideas; creation of original stands; the ability for critical analysis, creativity and resuming etc. If the teacher states that the topic sentence should be a declarative sentence, then he neglects the critical and creative approaches that students can have regarding the topic sentence.

What is said above, counts for the following indicators too (e):

(e) Criterion > structure of the topic sentence and its connection with opposite attitudes

Indicators:

The topic sentence is excellently structured and it is related with one of the opposite attitudes from the headline (5)

The topic sentence is very well organized and it is related with one of the opposite attitudes from the headline (4)

There is a lack of skill in organizing the topic sentence though it is clear that it is related to one of the opposite stance (3)

There is a lack of skills in organizing the topic sentence and the argument is barely précised though it is connected with one of the attitudes (2)

Here (e), we see that the structure and the connection of the topic sentence with the student's stance are taken into account. As for the structure, it is clear that words with relative meaning are exploited. As for the connection of the topic sentence with the student's stance, it can be said that this connection does not provide the objectivity of the assessment automatically. The indicator 4 says: "it is barely precise", thus again refer to the word with relative meaning. What does it mean when one says that the argument is barely précised? What would be the exact words (which would have no relative meaning) which would be a precise, concrete and measurable at the same time? It appears that there are no such words and that the description will depend on teachers' subjective beliefs, attitudes and opinions.

#### 4.3.2. *Three to five sentence*

As it has been argued, after the topic sentence, the next step is to write three to five sentences which aim is to develop the topic sentence, i.e. student's thought on the issue and direct it to the closing sentence. Thus, these sentences can also serve as one of the general criteria when assessing the introductory paragraph. They can be easily counted in the text, but this does not ensure that their nature and quality are satisfactory for the teacher and can lead to wrong teacher's impression that his assessment is valid and objective. It is well known that these sentences can be written in a different manner.

In the list below (f), the indicators are subjected to interpretation and explanation, though they cannot be measurable:

#### (f) Criterion > three to five sentences

Indicators:

The sentences explain why the problem is discussable and contain information that represents two different views for the topic which make it debatable (5)

These sentences are similar to the topic sentence and they begin to explain the problem, yet they do not include information connected with the opposite views that make the topic sentence discussable (4)

The sentences which follow the topic sentence are repeating the topic sentence in some extent with minimal extension. Though they offer new information for the topic, they do not represent the opposite attitudes (3)

The sentences which follow the topic sentence include rhetorical questions and/or retell personal experiences and do not contain any information on opposite attitudes for the topic (2)

Besides the words with relative meaning in the description, the criterion 1 and 2, introduce a new dilemma. What will the differences be between: "the student explains" and "the students begin to explain"? Is explanation of some problem something that is considered as exact and uniform? What kind of information, i.e. what kind of explanation teachers are expecting? What kind of explanation would satisfy two different teachers and what kind of explanation would be considered regarding the five or four points? It is obvious that teachers should now explain the difference in explanation to the students that gain five and four

points. Moreover, the teacher gives 4 points to a paragraph regarding the three to five sentences, though they do not contain information about the different views on the topic, which is incorrect and therefore the teacher's decision to give 4 points can be questioned.

The same could be said for the following descriptions, i.e. indicators:

(g) Criterion > three to five sentences

Indicators:

These sentences give five data connected with the topic of the essay (5)

These sentences give four data connected with the topic of the essay (4)

These sentences give three data connected with the topic of the essay (3)

These sentences give two data connected with the topic of the essay (2)

(h) Criterion > three to five sentences

Indicators:

The given sentences explain and support the topic sentence excellently (5)

The given sentences explain and support in a very well manner (4) etc.

(i) Criterion > three to five sentences

Indicators:

The given sentences are nicely chosen, they explain the headline in detail (5)

The given sentences are well chosen, though they explain the headline there is space for additional information from the student (4)

There is a space for the student to give clearer and better information (3) etc.

(j) Criterion > three to five question

Indicators:

The given sentences contain nicely chosen quotes and they deepened and extend the topic sentence (5)

The given sentences contain quotes that are not quite good and they deepened and extend the topic sentence in a large extent (4)

The quotes are not well connected with the rest of the paragraph, however they do not deepen and extend the topic (3)

The given sentences present poor choice of quotes which lack explanation and deepening the topic (2) etc.

Besides that these indicators contain words with relative meaning, it is good to point out the following: In (i) we read that there is a space...: what does it mean that there is a space, what information are considered to be better and clearer? If the teacher states the information that could be regarded as additional, then the teacher is neglecting the student's choice, his ability to think creative and critically etc.

#### 4.3.3. *The closing sentence*

This sentence as said above, should include the topic and the thesis, i.e. it has to be obvious whether the student will present the negative or the positive aspect of the topic. Thus, this sentence can be one of the general criteria when assessing an introductory paragraph from a five paragraph essay or the essay itself. However, noting that the paragraph contains such a sentence is not contributing to the assessment validity.

(k) Criterion > the closing sentence

Indicators:



The paragraph ends with a sentence that indisputably lists the arguments (for or against the topic) and/or the attitude of the author's is presented (5)

The closure sentence mentions the author's stance and/or contains arguments, but they are not precise enough (4)

The closure sentence does not present all the arguments and/or does not point out to the authors' attitude (3)

The closure sentence list only one argument and/or the writer's attitude is not presented. The closure sentence is a (rhetorical) question or with general information on the problem (2)

What was discussed above can be said about these indicators. Many of them contain words with relative meaning, such as: not precise enough, general information etc. Besides that, they involved parts that should be further explained. In addition, the indicator three says that this sentence does not list all the arguments (three) and that it does not point out the student's stance. Thus, the number of point given to this essay is questionable. What has student write in the closure sentence that is worth to be assessed with three points, but does not contain the arguments and the student's stance? Moreover, what does it mean when saying "the closing sentence indisputably lists all the arguments and the student's stance is presented" and "the closing sentence mentions the student's stance and contains arguments which are not precise enough"? What would be considered as a precise argument?

Various aspects of the introductory paragraph can be taken into consideration as criteria and indicators for assessment. However, the descriptions below show the same that has already been mentioned:

(l) Criterion > paragraph coherence

Indicators:

The paragraph is readable and flow of thoughts can be easily followed (5)

The paragraph is readable. However, flow of thoughts can be followed with minor difficulties (4)

The flow of thoughts in the paragraph is quite difficult to be followed, thoughts are unorganized, and connecting words are not present (3) etc.

(m) Criterion > paragraph coherence

Sentences are sequenced, i.e. one follows the other. They are supported by facts that are proven with previous research results (5)

Sentences are sequenced, i.e. one follows the other. They are not supported by facts that are proven with previous research results (4)

Sentences are not sequenced and students do not support information with facts (3) etc.

The indicator (5) says "reader can follow the flow of the thoughts". What does this mean? How can one concretize this indicator? It is well known that flow of the thoughts means that the one idea is a result of the previous one, then that many transitional words have to be used between ideas etc. Still, no one can measure the flow of the thoughts, i.e. no one can truly write down and make this sentence more concrete.

## 5. Conclusion

Learning outcomes related with an introductory paragraph of a five paragraph essay which is comprehended as an essay itself (students' competencies to express themselves in a written form in which they have to think in a critically and creative way, to analyze and to

evaluate, to show their creative skills etc.) are hard to be measured, primarily because unlike objective tests, these written answers contain more than one correct answers and their assessment is not consistent regard of who does the assessment and when the assessment occurs. Unfortunately, as the analysis suggests, though general criteria can be created, criteria and indicators for assessment of these essays are almost impossible to be made measurable and precise enough, due to the complex nature of these questions which contain personal beliefs, attitudes, stances, ideas, opinions, explanations and interpretations that, on the other hand, are inevitably subjected to teachers' individual beliefs, attitudes, stances etc. And, where teachers' explanation and interpretation of criteria and indicators are involved there is not a valid and objective assessment. To sum up, it seems that is almost impossible to write down criteria and indicators that will enable consistency regardless of who assesses the text. The only criteria and indicators that can be precise enough are those referring to what can be count, i.e. number of words, notifying presence of paragraphs' component, although they are not so relevant in regard to objectivity and validity. The only valid criteria and indicators referring to a written text are those connected with the grammar because the mistakes can be easily counted. Thus, teachers should be very careful when grade the students leaning only on essay questions and be aware of the advantages and disadvantages regarding their assessment.

## References

- Bailey, S. (2003). *Academic writing - a practical guide for students*. UK: Routledge.
- Becker, A. (2010/2011). Examining rubrics used to measure writing performance in U.S. Intensive English programs. *The CATESOL Journal*, 22 (1), 113-130. <http://bit.ly/3ar441C>, 16.10.2014.
- Bennett, J. , Gorovitz, S. (1997). Improving academic writing. *Teaching philosophy*, 20 (2), 1-12. <https://bit.ly/3bkODfs>, 20.2. 2017.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education: Principles, policy and practice*, 5 (1). <http://bit.ly/2uQ70sK>, 15.8.2014.
- Black, P. & Wiliam, D. (2001). *Inside the black Box. Raising standards through classroom assessment*. London: King's College.<http://bit.ly/3aoMoH9>, 15.8. 2014.
- Brookhart, M. S. (2008). *How to give effective feedback, Chapter 1. Feedback: An overview*.ASCD: Alexandria, Virginia. <http://bit.ly/38jCOE3>, 2.10.2014.
- Brookhart, M. S. (2013). *How to create and use rubrics for formative assessment and grading*. Chapter 1. ASCD: Alexandria, Virginia. <http://bit.ly/39miUtw>, 12.10.2014.
- Can, R. (2017). Analysis of written expression revision skills of the students in faculty of education. *Educational research and reviews*, 12 (5), 267-271. <https://bit.ly/2VBUYN8>, 10.8.2017
- Clark, D. (2014). *Bloom's taxonomy of learning domains*, updated July 2014. <http://bit.ly/3aoT9st>, 13.10.2014.
- Chappuis, J. (2012). How am I doing? *Educational leadership*, 90 (1), 36-41. <http://bit.ly/39hk8WQ>, 2.10.2014.
- Crews, F. (1980). *Distinguishing between subject area, topic and thesis*. Center for teaching and learning: The Random house handbook.<https://bit.ly/2KdME0E>, 10.11. 2016.
- Gallagher, E. (2015). *Oral and writen expression. What shoud you expect?* Uno International. <https://bit.ly/2VEmTfd>, 15.5.2016.
- Gojkov, G. (2003). *Dokimologija, priručnik (drugo izmenjeno izdanje)* [Docimology, second altered edition]. Vršac: Viša škola za obrazovanje vaspitača.
- Huitt, W. (1996). Measurement and evaluation: Criterion- versus norm-referenced testing. *Educational psychology interactive*. Valdosta, GA: Valdosta State University. <http://bit.ly/2uPnMbq>, 16.10.2014.

- Harvey, G. (2009). *A brief guide to the elements of the academic essay*. Harvard college writing program. <http://bit.ly/1RZFlaB>, 25.12.2014.
- Inez, M. S. (2016). *How to narrow a topic and write a focused paper*. Kibin. <https://bit.ly/2XGbEWm>, 1.7. 2017.
- Janusheva, V., Pejchinovska, M. (2011). Formative assessment in the teaching practice through the prism of the teachers. *International conference: „Educational Technologies“. Announcements of Union of Scientists – Sliven*, 19, 71-76.
- Janusheva, V., Pejchinovska, M. (2016). The role of the five paragraph essay in promoting the students' written expression in the Republic of Macedonia. *International journal „Teacher“*, 12, 112-119. Bitola: University „St. Kliment Ohridski“, Faculty of Education.
- Janusheva, V. & Kostadinovska-Stojchevska, B. (2017). *Akademsko pishuvanje 1* [Academic writing 1]. Bitola: University “St. Kliment Ohridski”.
- Janusheva, V. (2018). The differentiation of the topic and thesis in the five-paragraph essay – requirement for the quality of the students' written expression. *International journal of education „Teacher“*, 15 (1), 28-34.
- Linn, R. L. & Miller, M. D. (2005). *Measurement and assessment in teaching* (9th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- McMillan, H., James. (2008). *Assessment essentials for standards-based education*, 2<sup>nd</sup> edition, Chapter 5: Assessment, methods, items and techniques, 83-84, Corwin press. <http://bit.ly/2PMpbqm>, 13.11.2014.
- Mohrbacher, C. (2011). *Choosing and narrowing an essay topic*. <https://bit.ly/2XMRJ7Z>, 20.12. 2016.
- Moore, C. (2014). *Writing process tips: How to narrow your research paper ideas*. <https://bit.ly/3exuIfq>, 20.12. 2016.
- Moskal, M., B. & Leydens, A., J. (2000). Scoring rubric development: validity and reliability. *Practical assessment, research and evaluation*, 7 (10). <http://bit.ly/2Twp0k4>, 12.10.2014.
- Ntereke, B. B. (2015) Effectiveness of Academic writing activities and instruction in an academic literacy writing course at the university of Botswana. *Journal of pedagogy development*, 5(3). <https://bit.ly/2yrXptr>, 15.2. 2016.
- Nitko, A., J. (2004). *Essay, assessment, task*, chapter 9. <http://bit.ly/32QEzHA>, 13.11.2014.
- Palm, T. (2008). Performance assessment and authentic assessment: A conceptual analysis of the literature. *Practical assessment, research & evaluation*, 13 (4). <http://bit.ly/2TmS7rd>, 14.11.2014.
- Pineteh, A. E. (2014). The academic writing challenges of undergraduate students: A South African case study. *International journal on higher education*, 3 (1), pp. 12-22. <https://bit.ly/2xT1Z44>, 22.12. 2016.
- Pintrich, P. & Schunk, D. (1996). *The Role of expectancy and self-efficacy beliefs. Motivation in education: Theory, research & applications*, Ch. 3, Englewood Cliffs, NJ: Prentice-Hall. <http://bit.ly/39ovfx3>, 4.10.2014.
- Popovski, K. (2005). *Uchilishna dokimologija. Sledenje, proveruvanje i ocenuvawe na postiganjata na uchenicite* [School docimology. Following, checking and assessment of the students' achievements]. Skopje: Kitano.
- Popovski, K. & Stojanovski, Z. (1995). *Testovi na znaenje* [Tests of knowledge]. Skopje: prosveten rabotnik.
- Shiel, G., Kitching, K. 2007. *Improvement of school-based assessment*. Skopje: USAID, Macedonia. <https://bit.ly/3blfMyP>, 11.11.2014.
- Talevski, D. J. & Janusheva, V. (2011). Assessment of the essay questions with extended answer (EQEA). Book of proceedings – VI International Balkan congress for education

- and science: The modern society and education*, 409-415. Skopje: „Ss. Cyril and Methodius University“, Faculty of pedagogy „St. Kliment Ohridski“.
- Talevski, D. J., Janusheva, V., Pejchinovska, M. (2011). Teacher's development in relation to evaluation – Learning teams as a possibility for more effective assessment process of student's achievement. *Journal plus education*, 7 (1), 44 -57.
- Talevski, D. J., Janusheva, V. & Pejchinovska, M. (2014). Formative assessment and its effects in the teaching practice. *Modern social and educational challenges and phenomena – Polish and Macedonian perspective*, 212-221. Warszawa – Bitola: WyzszaSzkołaPeadogiczna w Warszawie, University „St. Kliment Ohridski“.
- Talevski, D. J. & Janusheva, V. (2015). Aspekti an ocenuvanjeto [Aspects of the assessment]. Bitola: Univeristy “St. Kliment Ohirdski”, Faculty of Education.
- Ternes, R. (2011). *Overview and introduction*. <https://bit.ly/2yBYQps>, 18.2. 2017.
- Uhlig, H. L. (2012). *Choosing a successful paper topic*. The writing center at GULC. <https://bit.ly/3aD1J6G>, 12.12. 2016.
- Velasquez, B. T. (2011). Investigating difficulties in elementary school students' writing. *Zona proxima*, 14. <https://bit.ly/2KvfpWP>, 20.1. 2016.
- Wiggins, G. (1990). The case for authentic assessment. *Practical assessment, research & evaluation*, 2 (2). <http://bit.ly/2TEbawb>, 14.11.2014.
- Wiggins, G. (2004). *Assessment as feedback*. <http://bit.ly/2TEvcqh>, 28.9.2014.
- Whitaker, A. (2009). *Academic writing guide. Step-by-step guide to writing academic papers*. Slovakia: City University of Seattle. <http://bit.ly/2pu2UiT>, 20.12. 2016.