THE DEVELOPMENT OF CRITICAL THINKING SKILLS THROUGH SELF-EVALUATION IN A TERTIARY ESP COURSE

Violeta Janulevičienė, Galina Kavaliauskienė
Mykolas Romeris University, Institute of Humanities, Department of Foreign Languages
Ateities 20, LT-08303 Vilnius, Lithuania
Telephone (+370 5) 2714 613
E-mail: vjanul@mruni.eu; gkaval@mruni.eu
Received on 28 June, 2012; accepted on 5 September, 2012

Abstract. In academic settings the skills of critical thinking, closely linked with self- and peer-observation and self-evaluation, are of paramount importance in every subject. A university ESP (English for Specific Purposes) course requires application of these transferable skills, too. The study presents some findings of research which reveals how students at Mykolas Romeris university (MRU) evaluate their general language abilities, self-assess performance in various ESP class activities and how reflections on one’s performance lead to the development of critical thinking skills in ESP classes. The validity of the research findings is backed by a statistical processing of self-evaluation data by means of a SPSS (Software Package for Social Sciences) versus actual performance in formal tests. Authentic student reflections presented in the weblogs samples are believed to add further highly valuable insights into tendencies for strategies for learning ESP and ways to develop critical thinking skills.

Keywords: English for Specific Purposes, critical thinking skills, self-evaluation, formal testing.
Introduction

Critical thinking is thinking that assesses itself. There are many definitions of critical thinking, probably a most comprehensive of which was offered by M. Scriven and R. Paul: “Critical thinking is the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action.”¹ This means, first of all, the readiness of not taking for granted what one reads or hears, but consciously assessing the validity of something, e.g. statements made by people and media, arguments, performance of others, research, etc.

Critical thinking is a lifelong transferable skill of paramount importance in academic settings: students will not be able to improve on their own unless they can self-assess information, phenomena, data, presented or found via research.

However, these issues of critical thinking have not been extensively researched in an ESP context and very few of the existing reports employ objective statistical processing of the data. Therefore, it is considered important to look into the application of critical thinking skills in student self-assessment of different aspects while learning English for Specific Purposes at the university and to combine impartial statistical methods with authentic self-observation for highlighting the problematic issues and strategies to be followed. The additional novelty of this paper is believed to be the inclusion and presentation of the full authentic student blog texts, dedicated to the reflection on their learning process of ESP.

The objectives of the present research into critical thinking are: a) to find out how students, taking an English for Specific Purposes course at Mykolas Romeris university (MRU), evaluate their general language abilities; b) to let students self-assess performance in various ESP class activities; c) to process these research findings statistically in order to find the correlations between the students’ self-evaluation data and actual performance in formal tests, d) by presenting and analyzing authentic student blog samples, to demonstrate how reflections on one’s performance lead to development of critical thinking skills in ESP classes.

The respondents of this research were 57 full-time 1st year students of MRU, who studied ESP for Social Sciences. The design of the ESP course reflected the students’ needs in professional language, and the course was adjusted to the requirements for a Bachelor of Social Sciences degree. The majority of students were females in their early twenties. The level of proficiency was B2 or C1 according to the Common European Framework of Reference for Languages.

The methods of research include: 1) a questionnaire on self-evaluation of language skills (Appendix 1); 2) statistical processing of the research data by using a SPSS (Software Package for Social Sciences); 3) a graphic display of self-evaluation data versus actual performance in formal tests; and 4) an analysis of student reflections on their achievements or failures in learning ESP in their weblogs (Appendix 2).

1. Literature Overview

Each step in the process of thinking critically is tied to self-reflection. Because of the importance of self-assessment to critical thinking, it is important to bring it into the structural design of any academic course and not just leave it to episodic tactics. It is common knowledge that not only the teachers are expected to give feedback to students but also students themselves should give feedback to each other on the quality of their work. As students learn to assess others, they learn what is important about the learning, how their learning can be demonstrated, and how they shall identify and implement formative feedback. They develop the skill of evaluating their own work and the awareness of that there is always room for improvement.

In developing these critical thinking peer- and self-assessment skills, a method of global analysis of the strengths and weaknesses of student performance in a class was applied. However, a successful analysis includes a number of prerequisites, students must be given 1) performance profiles (correlated with grades); 2) multiple opportunities to assess their own work and that of their peers using the performance profiles; 3) a thorough orientation on what is and is not expected in their self-assessment.

Assessment of critical thinking skills determines what students learn and how they learn. Measuring these skills informs teachers how students judge and analyze situations and how they make decisions. Applying several assessment techniques, such as peer reviews, portfolios and learning journals, gives teachers a broad sense of the skills learned during the critical thinking process.

Self-assessment provides an effective means of developing critical self-awareness. As a result, learners are able better to set realistic goals and direct their own learning. A second argument is that learners need to be involved in all the processes of learning,

---


4 Paul, R. W.; Elder, L., supra note 2.


including the process of evaluation. Skills-based self-assessment forms are designed to help students reflect on their approach and performance and provide teachers with a focus for learner training. Recent research by D. Bullock revealed a generally favourable attitude of the teachers towards students’ self-assessment, particularly highlighting that: a) when supported, learners benefit from assessing their own work; b) self-assessment raises learners’ awareness of their strengths and weaknesses; c) self-assessment stimulates motivation and involvement in the learning process.

2. Methods and Stages of the Research

The method of gathering data for this paper employed surveying learners’ self-evaluation of language abilities by a questionnaire (Appendix 1) and by analyzing their reflections on achievements or failures in performing various assignments. Self-reported data like questionnaires, interviews or diaries are the most frequent ways of identifying student attitudes. This information is available online in student weblogs. Student performance in formal testing of language skills was filed for analysis and compared with the relevant data of self-evaluation. Statistical processing of the data by means of a Software Package for Social Sciences (SPSS) included the computations of Cronbach’s Alpha coefficients of reliability and Spearman’s correlation coefficients, which indicate whether there is a correlation between the responses and their statistical significance. The findings are presented below.

3. Diagrammatic Representation of the Research Results

The obtained data is displayed in the four following charts. Each chart contains 5 or 6 double columns which show the percentage of students’ on Y axis versus the grades on X axis. The 1st columns present the percentage of students’ self-evaluated responses, while the 2nd columns show the percentage of students’ actual responses in testing. Grades on X axis are denoted by numbers 1 to 6, the lowest grade being 1 and the highest grade—6 (in Chart 1 the highest grade is 5, as none of the students self-evaluated themselves or got the highest possible score 10 “excellent”). Although in assessing student knowledge at MRU, a 10-point grading scale is used, we start with the lowest positive grade 5 “weak,” arbitrarily nominating it by the value 1, for practical reasons of clearer chart representation. In the other charts, the highest grade is 6, which means “excellent.”

Chart 1. Self- and actual evaluation of listening skills

Chart 1 demonstrates student listening skills. It shows that 60% of respondents evaluate their listening abilities by giving themselves average grades 3 or 4. Their actual performance in formal tests is nearly the same—65% of learners get the same grades.

However, an interesting tendency is highlighted by the Chart 1: students who do not perform very well, as well as the ones who are at the top of their class, underestimate their listening abilities, as 32% of the latter students give themselves low grades of 1 or 2, while in formal tests only 19% of students perform at that level. The same goes for the self-evaluation and performance at the top grade of 5.

This tendency for pronounced self-criticism came as an unforeseen insight which, in its own turn, contributes to the pedagogical results of the research.

Chart 2. Self- and actual evaluation of writing skills
Chart 2 reflects students’ writing skills assessment. The surprisingly vast majority—as many as 73%—of students self-evaluate their writing by giving themselves average grades of 3 and 4. In reality, however, only 54% of students were awarded such grades.

In other words, the results demonstrate a reverse tendency from the self-evaluation of the listening skills—that many students overestimate their writing skills.

However, more proficient learners, who self-evaluated their skills in writing at 5 or 6, have again underestimated themselves: as many as 38% of respondents performed very well in comparison with the number of self-evaluated of 11%.

![Chart 3. Self- and actual evaluation of speaking skills](image)

Chart 3 reveals a similar tendency in students underestimating their speaking skills. It can be seen that 35% of students believe their speaking skills are poor and give themselves the lowest grades of 1 or 2. However, only 6% of students get such low grades in their speaking tests. 59% of students are awarded very good or good grades against self-evaluated of 34%.

![Chart 4. Self- and actual evaluation of vocabulary acquisition skills](image)
Chart 4 demonstrates self-evaluated and tested knowledge of the specialized vocabulary. It highlights a particularly interesting fact: again the majority of learners—67%—believe they have acquired good knowledge of professional vocabulary, although the tests reveal that only 41% performed at that level. There is a clear tendency to overrate the skills in the specialized vocabulary acquisition. It is further based by the data in the lowest grades section: 16% of learners estimate their knowledge by the lowest grades. Formal testing exposes that actually 43%—nearly half of all the students—perform at this level in mastering the ESP vocabulary. Moreover, just 16% of students get good grades in vocabulary tests.

The data analysis of the charts revealed some very important, from our point-of-view, tendencies which were not foreseen in the research hypotheses: students tend to underestimate their listening abilities and speaking skills, but there is a pronounced opposite tendency in self-evaluation of the writing and ESP vocabulary mastering spheres, the latter ones are clearly overrated by the students themselves. Such findings offer immediate insights for an ESP teacher, which skills in a tertiary ESP course should get extra attention. The revelation of the full scope of these findings would require further research, though.

However, the immediate objective of the current research was to determine if there are any correlations between self-evaluated and formal grading. The data analysis by chart presentation did not prove itself to be reliable and valid in those terms.

In order to solve the problem scientifically a statistical treatment was applied. Its results are described in the following section.

4. Statistical Processing of Data

The obtained responses have been processed statistically using a Statistical Package for the Social Sciences (SPSS) in order to determine how comparable and reliable the data are. Internal consistency reliability is usually estimated by computing Cronbach’s Alpha coefficient. Results are reliable if the value of Cronbach’s Alpha coefficient is at least 0.60 (Dornyei9).

Another procedure included the computation of the Spearman’s correlation coefficients which are often described as being “non-parametric” (Bachman10). First, a perfect Spearman’s correlation exists when variables are related by any monotonic function which is not necessarily known. The other sense of Spearman’s correlation being non-parametric means that it is not necessary to require the normal probability distribution of variables. However, as a matter of interest, the Kolmogorov-Smirnov tests have also been computed and in each case the distribution has been found normal.

Spearman’s correlation coefficient can range between minus 1 and plus 1. Positive coefficients indicate direct relationships, while negative coefficients indicate inverse relationships. The larger the coefficient is, positive or negative, the stronger the relationship is. Thus, if it is close to 1, either positive or negative, it indicates a very strong relationship, while if it is near zero it indicates very weak relationship (Bachman10).

The computations of non-parametric Spearman’s correlation coefficients and their statistical significance for each case have been conducted using the SPSS software and the data are displayed in Table 1.

**Table 1.** SPSS data computed for the data in Charts 1, 2, 3 and 4

<table>
<thead>
<tr>
<th>Language skill</th>
<th>Cronbach’s Alpha coefficient</th>
<th>Spearman’s correlation coefficient</th>
<th>Sig. level (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Listening</td>
<td>0.664</td>
<td>0.162</td>
<td>0.374 level (2-tailed)</td>
</tr>
<tr>
<td>Writing</td>
<td>0.765</td>
<td>0.065</td>
<td>0.722 level (2-tailed)</td>
</tr>
<tr>
<td>Speaking</td>
<td>0.839</td>
<td>0.807</td>
<td>0.01 level (2-tailed)</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>0.740</td>
<td>0.196</td>
<td>0.282 level (2-tailed)</td>
</tr>
</tbody>
</table>

It is seen that the values of Cronbach’s Alpha coefficient exceed 0.60: the lowest value is 0.664 for listening results and the highest value 0.839 for speaking results. According to the theory (Dorney11), even short scales of 3-4 items must have reliability in access of 0.60, so in our case the questionnaire items are reliable.

The value of statistical significance of Spearman’s correlation coefficients is important for the interpretation of the relationship between two samples. In other words, if it is equal at least to 0.05, it means that the relationship is not likely to be due to chance. Larger than 0.05 values of the significance level, even if there is a correlation coefficient close to +1.00 or -1.00, mean that the probability of the significant relationship between two items is smaller than 95% and, therefore, the relationship is likely to be due to chance.

In Table 1 the computed values of Spearman’s correlation coefficient for listening, writing and vocabulary are rather small: 0.162, 0.065 and 0.196, respectively. Moreover, the levels of significance in these cases are too large. It means that there is no correlation between self-evaluated and formally graded data. The only exception is the skill of speaking. The significance level here is 0.01, i.e. probability is 99%, and the value of the correlation coefficient is 0.807. It means that in speaking there is a significant correlation between the students’ self-evaluated grade and the one received from testing.

---

11 Dornyei, Z., supra note 9.
This is a very interesting finding, because speaking is less susceptible for objective self-evaluation. However, the computed correlation indicates that students can assess speaking skills in ESP quite realistically and objectively.

5. Student Reflections

Student reflections are very important in the process of learning, as each person thinks about his/her own performance and analyzes achievements and failures. Such analysis helps students decide what strategies should be taken to improve language skills. At the end of the semester students are asked to write a self-assessment of their performance in the weblogs. Some passages from student weblogs (Appendix 2) demonstrate the ongoing changes in learning. The presented excerpts cover progress in writing, listening and speaking skills and vocabulary learning. The language of student entries has not been corrected.

Conclusions

The findings of the research disclosed a fair ability of the students to self-assess their speaking skills in ESP. The statistical treatment of experimental data has shown the good correlation between self-evaluated speaking grade and performance in formal tests.

However, correlations have been obtained for writing, listening and vocabulary data. Students seem to be unable to evaluate their writing and listening skills and knowledge of professional vocabulary impartially and objectively. This fact suggests the necessity of developing critical skills in learning by encouraging students to analyze their own ongoing performance within different fields of ESP usage. Students’ reflections on achievements and failures are useful techniques that make learners analyze their work and resort to critical thinking on how best to master language skills in listening, ESP vocabulary retention, and cohesive and coherent writing.

References


---

**Appendix 1. Questionnaire. Self-evaluation of language skills.**

1. I believe that my grade in listening is ____.
2. I believe that my grade in writing is ____.
3. I believe that my grade in speaking is ____.
4. I believe that my grade in ESP vocabulary is ____.

**Appendix 2. Some excerpts from students’ weblogs.**

**http://mburauskaite.blogspot.com/**

In terms of summary writing, I learned some new things. Firstly, I don’t write examples, but generalize the main ideas. Secondly, I don’t use personalization and reported speech. Also, I don’t start new sentences with wrong linking words. However, sometimes I don’t know or forget grammar and vocabulary rules and in my summaries are some mistakes. Listening online: I think that frequent listening practices were very beneficial because today I can do these exercises easily and understand the main idea and more sentences. Listening to cassettes has also been a good practice to improve language skills. ESP Vocabulary. 1) *Formal tests in class*. I learned a lot of new words and definitions. Actually to write definition is easier for me than filling in the gaps, because sometimes it’s difficult to choose and remember one word, which is appropriate in that sentence. 2) *Class tests online*. These exercises can help to prepare for formal tests in class and solidify my knowledge. 3) *Moodle tests* help to prepare for formal tests in class, because I can find the answer in a book and learn unknown definitions.

**http://sdarguzis.blogspot.com/**

Summary writing is stressful because I find it hard to write something in other words. Since the time I was at school I disliked online materials, but online listening is easier than to cassettes. For me recordings are hard to follow because speakers talk too fast.
I really enjoyed writing definitions in formal ESP vocabulary tests, but I disliked gap filling tests: I have never been able to perform well. Moodle tests are my favourite: it is really a good way to remember what I studied before and get a good mark.

http://rcepulyte.blogspot.com/
Comparing the results of my last term’s summaries with summaries of this term, I can say that I improved my writing skills. My personal evaluation average grade increased from 6 to 7.

Listening task was always quite difficult for me: it depends on how fast speaker is talking and on some other facts. I didn’t notice my improvement at this, I still need to listen to the task several times if I want to understand it clearly.

Formal testing is good to check my knowledge. For me it was easier to write definitions for given terms, so when task format was changed it was too difficult to get a good mark. That was the reason why my grade went down. I was quite good at class tests online not only because of my knowledge but also of intuition. Moodle tests were good and easy to get a good mark.

http://rstanisauskas.blogspot.com/
Writing is my strong point. I didn’t enjoy writing summaries. Usually I just write a summary on the draft paper without even checking it. The idea of rewriting frustrated me.

Listening tasks were quite enjoyable because I can understand different accents without putting any efforts. I can’t tell the difference between cassette and online listening.

http://rstankevic.blogspot.com/
Writing a summary isn’t hard. My main problem is grammar. I make a lot of mistakes, it is the same for me in other languages. I tried a few strategies to deal with this problem, but they did not help.

Listening is the field I am the best at. Although it is a bit harder with cassettes due to background noises and sometimes accent of the speaker.

The biggest problem I faced was ESP vocabulary. The main reason I haven’t studied well because of my laziness.

http://vstankute.blogspot.com/
I have written a number of summaries and believe that I have improved my writing skills significantly.

The most difficult part of the course has been ESP vocabulary. Usually I did quite well with definition tests, but gap filling tests have been challenging and hard.

http://monikakrikscikaite.blogspot.com/
It was not easy to write summaries at the beginning of the first semester and even during the whole second semester. I was making the same mistakes and it was also hard to distinguish the main points of a text and not to write anything else that was not
so important. But at the end of the second term I noticed that I get even better marks, at least, it was not very high – from 5 to 6, but anyway I was hoping to improve my summary writing skills. Perhaps I can say that I succeeded, because my last summary was graded 10. I do believe that the same success will go along with me on the exam day.

Writing ESP tests always was a great opportunity to get a good mark. To be honest, I have to say that I was learning definitions by heart. Not because it was difficult to understand or translate them, but because usually they were too scientific to write in your own words and not change the real meaning in this way. I was surprised, when the format of tests changed in the second semester. Writing a correct definition became not the main task, it was much more important to remember the material of the topic. Firstly, it looked impossible to do, but after a few failed tests I found the way to improve my mark. It was not only learning definitions, but also doing Moodle tests online as well as online tasks on each topic of Feldman’s online course book straight after reading the Module and revising everything before writing the ESP test. Unfortunately, I had never got 100% again after such turning, but I believe that it was more useful for me.

http://ikuniejeute.blogspot.com/

Summary writing seemed one of the easiest tasks for me. At the end of the second term I managed to select the most important information from the articles and write summary fast. But even though I knew all the rules how to write a good summary, sometimes there were some mistakes in my summaries. Maybe it was due to lack of attention when I was writing them.

I see vocabulary as my weakness so that is why ESP vocabulary tests were the hardest task of all. Even though online tests which we did during our classes weren’t very hard but formal tests in classes were really difficult. It took a lot of time for me to prepare for them and my performances weren’t as good as I wanted them to be. There were a lot of definitions so I always confused them. The easiest part of vocabulary tests was Moodle tests. It was useful because I could check what I remember from reading the modules in the course book, and if I didn’t know anything I could check it in the book. But I sometimes made some mistakes because of lack of concentration. Anyway, I’m satisfied with my Moodle performance.

KRITINIO MĄSTYMO ĮGŪDŽIŲ PLĖTOTĖ PER ĮSIVERTINIMĄ UNIVERŠITETINĖSE DALYKINĖS KALBOS STUDIJOSE

Violeta Janulevičienė, Galina Kavaliauskienė

Mykolo Romerio universitetas, Lietuva

Santrauka. Akademineje aplinkoje kritinio mąstymo įgūdžiai, glaudžiai susiję su savistaba ir savęs vertinimu, yra svarbus kiekvieno dalykų studijoms. Šie bendrųjų kompetencijų gebėjimai yra būtini ir profesinės anglų kalbos studijose.
Tyrimu siekta išsiaiškinti, kaip Mykolo Romerio universiteto (MRU) studentai vertina savo bendrosios ir dalykinės anglų kalbos gebėjimus ir kaip šios refleksijos pėtoja kritinio mąstymo įgūdžius anglų kalbos studijose. Tyrine dalyvavo 57 nuolatiniių socialinių mokslų studijų pirmo kurso studentai, kurių bendrosios anglų kalbos žinios atitiko B2 ar C1 lygį pagal Bendrųjų Europos kalbų metmenų skalę.

Pirminiame tyrimo etape studentai atsakė į klausimyną (1 priedas), prašantį įsivertinti bendrinės ir dalykinės anglų kalbos klausymo, skaitymo, rašymo bei profesinės kalbos žodyno įvaldymo įgūdžius. Atsakymų apibendravimai, juos palygintus su objektyviu formaliuoju minėtų kalbos įgūdžių vertinimu, pateikiami straipsnio diagramose. Tačiau buvo įsitikinta, kad grafinis rezultatų pavaizdavimas neišryškina koreliacijų tarp studentų įsivertinimo ir faktinio objektyvaus jų kalbinių įgūdžių vertinimo. Todėl tyrimo duomenys apdoroti ir statistinių metodų rezultatais, pateiktais lentelėje, naudojant kompiuterinių statistinių socialinių mokslų programų paketą SPSS (Software Package for Social Sciences).


Tyrimo duomenys statistinis apdorojimas išryškinėjo papildomus, tyrimo hipotezėje neįmanytus, studentų dalykinės bendrosios anglų kalbos įgūdžių plėtotės aspektus – studentai pervertinti savo anglų dalykinės kalbos žodyno įsisavinimo bei rašymo įgūdžius ir akivaizdžiai nepakankamai vertina savo supratimo įsivertimus ir kalbėjimo sukalbinės kalbos įgūdžiai. Šios tyrinėtojų atskleistos įžvalgos yra labai svarbios sudėtingai akcentus universitetiniame dalykinės kalbos kurse ir praktinio mokymo kaip rekomendaciją, atitinkanti laiką ir dėmesį kalbinių įgūdžių lavinimui dalykinės kalbos studijose.

Tyrimo metu studentų rašomos refleksijos taip pat paskatino besimokančiuosius analizuoti savo kalbinės veiklos rodiklius ir pasirinkti efektyviausius metodus jos tobulinimui. Neabejotina, kad dalykinės kalbos studijų metu lavinti kritinio mąstymo įgūdžiai bus plėtojami ir taikomi akademinėje ir profesinės veiklos sferose.

**Reikšminiai žodžiai:** dalykinė/profesinė anglų kalba, kritinio mąstymo įgūdžiai, įsivertinimas, formalusis vertinimas.

Violeta Janulevičienė, Mykolo Romerio universiteto Humanitarinių mokslų instituto Užsienio kalbų katedros docentė. Mokslinių tyrimų kryptys: specialybės anglų kalbos mokymas(is), kognityvinė lingvistika ir leksikologija.

Violeta Janulevičienė, Mykolas Romeris University, Institute of Humanities, Department of Foreign Languages, Associate Professor. Research interests: ESP teaching/learning, cognitive linguistics and lexicology.
Galina Kavaliauskienė, Mykolas Romeris University, Institute of Humanities, Department of Foreign Languages, Associate Professor. Research interests: ESP teaching/learning, lifelong learning, teaching/learning quality.