

Self-Assessment Ability of Pre-Service Teachers

DOI: 10.15804/tner.2017.48.2.17

Abstract

Knowing one's own level of knowledge is an important characteristic of an individual. It enables individuals to objectively evaluate their abilities and properly adapt to their advantages and disadvantages. In this paper, we present the results of the empirical research, where pre-service teacher students had to perform self-assessment after their seminars and mathematics classroom performance. We compared their self-assessments to their teachers' assessments. Results show that the students' self-assessments on average deviate from their teachers' assessments. We also noticed that the Dunning-Kruger effect is present both for seminars and mathematics classroom performance. The students that received low assessment scores from their teacher provided too high self-assessment scores.

Keywords: mathematics, self-assessment, seminar paper, mathematics classroom performance

Introduction

Every human being has to constantly make decisions. Some of those are more important than others, but all of them influence a person's everyday life. The impact of an individual's own perception of their skills, talents, personality and expectations is determined by their own decisions (Dunning, Health & Suls, 2004).

Self-assessment refers to the involvement of learners in making judgements about their own learning, particularly about their achievements and the outcomes of their learning (Boud & Falchikov, 1989). Fitzpatrick (2006) defines self-assess-

ment as a form of assessment in which the student reflects their strengths and weaknesses in order to improve their performance.

The competence of self-assessment also develops through formal education. One of the goals of higher education is, therefore, also to develop students' self-assessment abilities (Boud, Lawson & Thompson, 2013). Slovenia's education institutions' curriculums and goals for all subjects are oriented towards learning-to-learn competence, where the capability of self-assessment is emphasized. The European key competences define the learning-to-learn competence as "the ability to pursue and persist in learning, to organise one's own learning, including through effective management of time and information, awareness of one's learning process and needs, identification of available opportunities, and the ability to overcome obstacles in order to learn successfully" (Lifelong learning – key competences, 2006, p. 8).

The Slovenian White Book (2011) emphasizes, in addition to the essential tasks of school, the ability of learning-to-learn. The realization of this ability can take place via the integration of self-assessment topics in curriculum. Unfortunately, some studies conducted in Slovenia point out that this is not carried out at the level of education of future primary school teachers (Hus & Koprivnik, 2016).

Many studies have been conducted in the area of student self-assessment skills (e.g., Falchikov & Boud, 1989). Some have attempted to find out if students' self-assessment grades are equal to the teachers' grades (e.g., Dunning, Johnson, Ehrlinger, & Kruger, 2003). Others attempted to find out if students' self-assessment can eventually become more objective (e.g., Pallier, 2003).

Several studies focused on the impact of students' individual abilities on their objective self-assessment. When researching this question, we found parallels with the Dunning-Kruger effect. Kruger and Dunning (1999) tested the assessment ability of psychology students. In the test, students had to assess their own ability and predict their own achievement on a test. The results show that the students with high abilities tended to underestimate themselves. On the other hand, the students with lower abilities tended to overestimate themselves. Other studies have also confirmed these findings (Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008; Dunning, 2011). The findings are important because some people do not realize that their own judgement is not always objective; i.e., less competent individuals in general have higher confidence about their knowledge (Dunning, 2014). Additionally, it seems that feedback does not help to objectify their opinions (Ehrlinger et al., 2008). The Dunning-Kruger effect does not occur in every culture (DeAngelis, 2003). The reason for this could lie in different approaches to education (Falk, Heine, Yuki, & Takemura, 2009). It seems that the capability of

self-assessment can be taught, especially if we start to develop it early on (Brown & Harris, 2014).

Some studies reported some problems which emerged during the evaluation of students' self-assessment. Falchikov (2004) reported that in students' opinion, self-assessment is insignificant because it is the teacher's job to evaluate students. Liang (2006, as cited in El Koumy, 2010) pointed out that students felt humiliated when they were asked to report their weaknesses. Errors in self-assessment could occur because of the inconsistency between students' and teachers' goals (Oscarson, 1997; as cited in El Koumy, 2010).

Despite the fact that some problems with self-assessment were detected, the positive effects of using self-assessment are well documented. Self-assessment increases students' awareness of their own thinking processes (Zohar, 2004). Self-assessment also helps students make their own guidelines for future learning (Chamot & O'Malley, 1994; as cited in El Koumy, 2010). Self-assessment also helps students to achieve better school grades because they are aware of their strengths and weaknesses (Harris, 1997). Self-assessment is also an appropriate tool for teachers because it makes them aware of students' needs (Blanche, 1988).

Research Problem

In our study, we wanted to find out if students are able to objectively and realistically self-assess their seminar papers and mathematics classroom performance. We also wanted to find out if there are any differences between teachers' grades and students' self-assessments of their seminar papers and classroom performance. In addition, we focused on the connections between students' ability to self-assess and the Dunning-Krueger effect. This kind of verification of students' ability to self-assess was chosen mainly because our goal is to educate students in a way that allows them to objectively evaluate their work.

Research Focus

The aim of our study was to verify the self-assessment skills of pre-service teachers. We focused on the area of mathematics education. We wanted to find answers to the following research questions:

1. How do students perform at realistically self-assessing their seminar papers, which was introduced to other students?

2. How do students perform at realistically self-assessing their mathematics classroom performance, which was prepared with the help of their faculty teacher and tested in an authentic situation in the classroom?
3. Is there any difference between students' self-assessment of the seminar paper and their mathematics classroom performance?

Research Methodology

Research General Background

At the Faculty of Education, University of Maribor, future primary teachers are educated in the first and second Bologna cycle according to the 4 + 1 Model. The first authentic experience in teaching a particular subject takes place in the third year. Students prepare teaching scenarios under the surveillance of an educator who has specialized in the specific subject. Scenarios are developed at three individual meetings, lasting half an hour on average. After the teaching performance, a guided reflection takes place, pointing out the shortcomings and strengths of the lecture. Criteria of performance evaluation are established during the preparation and students are closely involved in the process. Students write seminar papers from the very beginning of their education, so they are relatively well-informed about evaluation criteria.

Research Sample

There were two groups of participants in our research sample. There were 76 third year students and 76 fourth year students of the Primary Education Programme. The first group of participants (third year students) was included in both parts of the research study, while the second group (fourth year students) was included only in the second part. The research was conducted during the winter semester of the 2015/16 academic year in the subjects Mathematics Education I and Mathematics Education II.

Instrument and Procedures

After the presentation of the seminar paper and their mathematical classroom performance, each student had to evaluate their work using a numerical grade. A teacher also evaluated the work of each student using a numerical grade. The students were familiar with the assessment criteria before their self-assessment. The teacher evaluated the presented seminars right after each presentation. Moreover, for the mathematical classroom performance, the students were numerically

assessed at the end of the semester, when their performance was finished. In both cases, the teacher's grades were numerical with an added descriptive explanation. We compared the teachers' and students' grades for the same activity.

The study was based on a descriptive and causal non-experimental methodology in pedagogical research. The inferential statistic χ^2 was used to analyse differences among the teachers' grades and the students' self-assessment grades.

Data Analysis

The IBM SPSS Statistics 23 program was used. We collected the data concerning the teachers' grades and the student's self-assessments in a table. Next, we calculated the difference between the teacher's grade and student's grade for each student. We subtracted the student's grade from the teacher's grade in order to quantify the gap. The gap was considered to be a measure of the students' self-assessment abilities; the smaller the gap, the better their abilities to self-assess. In the next phase, the data were checked at two stages; first, we used descriptive statistics and, second, we used inferential statistics when searching for differences between the teacher's grade and the student's self-assessment

Research Results

The results are presented in Tables 1–3. The first table shows basic descriptive statistical values for the data relating to the students' self-assessment and the teachers' grades of the seminar paper and mathematics classroom performance.

Table 1. Teachers' grades and students' self-assessments of the seminar paper and classroom performance

		N	M	σ	SKEW	KURT
Seminar paper	Teachers' grades	76	8.13	1.427	-0.577	-0.243
	Students' self-assessment	76	8.53	0.945	-0.467	0.058
Classroom performance	Teachers' grades	136	8.82	1.032	-1.116	1.069
	Students' self-assessment	136	8.83	0.756	-0.962	1.639

The first table shows that the averages of the teachers' grades and students' self-assessments are relatively high. The average numerical values of the teachers' grades are lower than the average numerical values of the students' self-assessments.

In the seminars, the teacher’s grades and also the students’ self-assessments were lower than the assessments for the mathematics classroom performances. The highest dispersion of grades is shown in the teachers’ grades of the seminars. This means that the diversity of the teachers’ grades is higher than the diversity of the students’ self-assessments. The same phenomenon is observed in the assessments of mathematical classroom performance, but the diversity of grades is lower in that case. All of the distributions are skewed towards the left and also the majority of distributions are leptokurtic.

In Table 2, we present the data which show the relationship between the teachers’ grades and the students’ self-assessments of the seminar. In order to facilitate the transparency of the results, we classified the teachers’ numerical grades into three categories: below average (numerical grades 5, 6, and 7), average (numerical grades 8 and 9), and above average (numerical grade 10). The data are shown in Table 2.

Table 2. The relation of teachers’ grades to students’ self-assessment of seminar papers

Teachers’ grade	Self-assessment of seminar paper										Total	
	-1		0		1		2		3		f	f %
	f	f %	f	f %	f	f %	f	f %	f	f %		
Below average	2	8.3	4	16.7	6	25.0	10	41.7	2	8.3	24	100
Average	6	15.8	20	52.6	12	31.6	0	0	0	0	38	100
Above average	6	42.9	8	57.1	0	0	0	0	0	0	14	100
Total	14	18.4	32	42.1	18	23.7	10	13.2	2	2.6	76	100

$\chi^2 (1r) = 46.725, P = 0.000$

The data in Table 2 show that slightly more than 40% of the students’ self-assessments of seminars match the teachers’ grades. Additionally, we can observe that the majority of the students (75%) who had received a below-average grade in the seminar self-assessed their seminar as deserving a higher grade. Half of the participants who received an average grade from their teacher made the same self-assessment as the teacher. The highest share of the students’ self-assessments which were lower than the teachers’ grades can be seen among the students who got the teacher’s above-average grade.

The differences between the teachers’ grades and the students’ self-assessments of mathematics classroom performance are presented in Table 3.

Table 3. The relation of teachers' grades, students' self-assessments of classroom performance.

Teachers' grade	Self-assessment of mathematics classroom performance										Total	
	-2		-1		0		1		2		f	f %
	f	f %	f	f %	f	f %	f	f %	f	f %		
Below average	0	0	4	11.4	9	25.7	15	42.9	7	20.0	35	100
Average	0	0	7	10.3	52	76.5	9	13.2	0	0	68	100
Above average	2	6.1	22	66.7	9	27.3	0	0	0	0	33	100
Total	2	1.5	33	24.3	70	51.5	24	17.6	7	5.1	136	100
$\chi^2 (1r) = 98.800, P = 0.000$												

In this case, the differences are much more visible than the differences in Table 2. We can clearly see that the students who received the teachers' below-average grades have much higher self-assessments than others do. On the contrary, the students with the teachers' above-average grades have lower self-assessment.

Discussion

Our study shows that the teachers' grades and the students' self-assessments were both quite high (on the grading scale, equivalent to the grade 'very good'). The results also show that the students' self-assessments of seminars are on average lower than their self-assessments of their mathematical classroom performance. The same can also be seen in the teachers' grades. In Slovenian primary schools, we notice that students' grades are relatively high. The same phenomenon can be observed in internal parts of the school final exam; on average, students obtain 90% of all available points in those parts (Zupanc & Bren, 2010). Some authors also point out the danger of this trend, which starts in primary school, but which has nowadays spread throughout the whole education system. A similar phenomenon was observed in Canadian schools, where Anglin and Meng (2000) noticed that average grades were increasing. Some authors reported students' aggressive behaviour when the teacher's grade was lower than they had expected (Vaillancourt, 2013). Isely and Singh (2005) reported a detected connection between students' grades and the results of a student questionnaire, which every student has to complete at the end of semester.

In comparison to the teacher's grade, the student's self-assessment of mathematical classroom performance was more accurate than the student's self-assessment of seminars. The student's self-assessment and the teacher's grade of mathematical classroom performance were the same for more than half of the participants. On the other hand, the student's self-assessment and the teacher's grade of seminars were the same less often. We can conclude that the students are more competent in their self-assessment of the mathematical classroom performance activity, where they know for sure what constitutes good performance and what does not. The same cannot be said about seminars. Obviously, students have a lot of knowledge about the properties of 'good' classroom performance from their previous pedagogical and didactic knowledge.

The most obvious difference between these two assessment activities is the different types of preparation for these activities. Before mathematics classroom performance, students have three individual tutorials with their teacher. Before seminars, students are not provided with any individual help from the teacher. It follows that students consequently receive higher grades for mathematical classroom performance than they do for seminars. Additionally, individual tutorials help students to clarify the criteria which students use to improve their self-assessments (Falchikov & Boud, 1989).

On the other hand, these findings are inconsistent with the hypothesis that precise criteria help students to make more accurate self-assessments. Unfortunately, we can conclude that the seminar criteria were not presented carefully to the students. We think that teacher-student cooperation in deciding criteria might help. Other authors have reported positive effects of collaboration between teachers and students in creating criteria (Orsmond, Merry, & Reiling, 2002).

In our study, the Dunning-Kruger effect was detected. It was noticed in the students' self-assessments of seminars and also of the mathematical classroom performance. We noticed that the students with the teachers' lower grades made much higher self-assessments. However, the students with the teachers' higher grades made much lower self-assessments. The students with below-average grades self-assessed their performance with a higher grade (up to three numerical grades higher). The students who received the teachers' above-average grades self-assessed their performance one numerical grade lower. This was observed especially in the self-assessments of mathematical classroom performance. Boud and Falchikov (1989) also detected the same trend; the best students think they do not deserve the best grades. Additionally, students with teachers' average and above-average grades are more competent at self-assessment than students with teachers' below-average grades.

The students' reflections about their mathematical classroom performance reveal that the students with the teachers' below-average grades and high self-assessments reported on detected problems in the classroom. They know that they made some mistakes during their teaching, but this fact does not influence their self-assessment. Here, we quote a reflection from a student with an average self-assessment (numerical grade 8) and the teacher's below-average grade (numerical grade 6):

In general, I'm satisfied with my mathematical classroom performance, but I'm aware that I could do better if our pupils behaved better. Because the pupils did not behave well, my whole teaching plan was ruined. I lost too much energy during the teaching hour and made a lot of interventions about discipline. Because of that, I was completely confused and the whole situation in the classroom was not as I had planned and expected.

We can see from this testimony that the student saw problems caused by pupils and not in their own performance. They also detected more didactic problems than mathematical problems. A similar phenomenon was observed by Lipovec, Podgoršek and Antolin Drešar (2013), when they reported on the locus of control of pre-service primary school teachers.

Conclusions

Our study shows that future primary teachers' self-assessment abilities are average. In more than half of the cases, the participants' self-assessments were the same as the teachers' grades. In this study, we confirmed the connection between students' self-assessment abilities and teachers' grades, which is also known as the Dunning-Kruger effect. Even though we only checked the abilities of university students' self-assessment, it is necessary for teachers at all stages of education to be aware of the described phenomenon because it could be observed for the different levels of education. Incorrect perceptions of a person's own advantages and disadvantages could lead to unsuitable choices (e.g., when making decisions about their profession). On the other hand, they could also have an effect on the individual's satisfaction. Our results are consistent with Brown and Harris's (2014) assumption that realistic self-assessments are more likely when students are involved in the process of establishing criteria for evaluating work outcomes.

In future research, one could check the ability of students' self-assessments using the same participants after a given time period. One could explore whether

their ability of self-assessment has improved over time. Further research could also check the connection between the Dunning-Kruger effect on students' perceptions and students' knowledge about a certain subject. Additionally, differences between students' perceptions of their knowledge and their performance in the mathematics classroom or in seminars could be investigated. We think that this approach could lead to important insights which could help both students and teachers.

We would like to point out that even though self-assessment is highly rated as a competence, it is not clear how to develop it. Both teachers and students will need training before they can engage with self-assessment as a competence.

References

- Anglin, P.M., & Meng, R. (2000). Evidence on grades and grade inflation at Ontario's universities. *Canadian Public Policy/Analyse de Politiques*, 26 (3), 361–368.
- Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal*, 19 (1), 75–93.
- Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher education: A critical analysis of findings. *Higher Education*, 18 (5), 529–549.
- Boud, D., Lawson, R., & Thompson, D.G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, 38 (8), 941–956.
- Brown, G.T.L., & Harris, L.R. (2014). The future of self-assessment in classroom practice: Reframing self-assessment as a core competency. *Frontline Learning Research*, 2(1), 22–30.
- DeAngelis, T. (2003). Why we overestimate our competence. *Monitor on Psychology*, 34 (2), 1–60.
- Dunning, D. (2011). The Dunning-Kruger Effect: On Being Ignorant of One's Own Ignorance. *Advances in experimental social psychology*, 44, 247.
- Dunning, D. (2014). We are all confident idiots. Retrieved 6/6/2016, from <https://psmag.com/we-are-all-confident-idiots-56a60eb7feb7#edwe8bxea>
- Dunning, D., Heath, C., & Suls, J.M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5 (3), 69–106.
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12 (3), 83–87.
- Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, 105 (1), 98–121.
- El-Koumy, A.S.A. (2010). Student Self-Assessment in Higher Education: Alone or Plus? *Education Resources Information Center (ERIC)*. 1–28.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59 (4), 395–430.

- Falchikov, N. (2004). Involving students in assessment. *Psychology Learning & Teaching*, 3 (2), 102–108.
- Falk, C.F., Heine, S.J., Yuki, M., & Takemura, K. (2009). Why do Westerners self-enhance more than East Asians. *European Journal of Personality*, 23, 183–203.
- Fitzpatrick, J. (2006). Self-assessment as a strategy to provoke integrative learning within a professional degree programme. *Learning in Health and Social Care*, 5 (1), 23–34.
- Harris, M. (1997). Self-assessment of language learning in formal settings. *English Language Teaching Journal*, 51 (1), 12–20.
- Hus, V. & Koprivnik, M. (2016). Development of some notions of learning to learn competence of students of primary education in Slovenia. *The New Educational Review*, 43 (1), 17–27.
- Isely, P., & Singh, H. (2005). Do higher grades lead to favorable student evaluations? *The Journal of Economic Education*, 36 (1), 29–42.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77 (6), 11–21.
- Lifelong learning - key competences (2006). Retrieved 31/12/2016, from <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3Ac11090>
- Lipovec, A., Podgoršek, M., & Antolin Drešar, D. (2013). Točka kontrole in osredotočenost študentov razrednega pouka. [Locus of control and focus of preservice primary teachers] *Pedagoška obzorja*, 28 (3/4), 157–170.
- Orsmond, P., Merry, S., & Reiling, K. (2002). The use of exemplars and formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, 27 (4), 309–323.
- Pallier, G. (2003). Gender differences in the self-assessment of accuracy on cognitive tasks. *Sex Roles*, 48 (5–6), 265–276.
- The White Book (2011). Retrieved 24/5/2016, from http://pefprints.pef.uni-lj.si/1195/1/bela_knjiga_2011.pdf
- Vaillancourt, T. (2013). Students aggress against professors in reaction to receiving poor grades: An effect moderated by student narcissism and self-esteem. *Aggressive Behavior*, 39 (1), 71–84.
- Zohar, A. (2004). *Higher order thinking in science classrooms: Students' learning and teachers' professional development*. Dordrecht: Kluwer Academic Publishers.
- Zupanc, D., & Bren, M. (2010). Inflacija pri internem ocenjevanju v Sloveniji. [Inflation in the internal part of assessment in Slovenia] *Sodobna pedagogika*, 3, 208–228.