Network Intelligence Studies
Volume V, Issue 9 (1/2017)

Zsuzsanna KISS,
Edit Barizsné HADHÁZI,
Domician MÁTÉ
Institute of Management and Organization Sciences, and Institute of Accounting and Finance,
Faculty of Economics and Business, University of Debrecen

ADVANCED LEVEL VOCATIONAL TRAINING STUDENTS' SELF ASSESSMENT

Case Study

Keywords
Self-assessment,
Self-evaluation,
Business education,
Higher education,
Students' academic performance

JEL Classification
A22, I23, M53

Abstract

This paper is intended to clarify the phenomenon that lower achieving students tend to evaluate their own academic performance less accurately than those who do better in their studies. Previous studies have found that lower performers generally overestimate while higher performers underestimate their performance. The current study analyses self-assessment behaviour and efficiency among Hungarian higher vocational education students. We found that the lowest level of higher education students typically overestimate their performance. Our results strengthen the empirical evidences from previous studies that showed that higher-achieving students evaluate their performance more accurately than their lower achieving fellows. Furthermore we found that higher-achieving students tend to over-assess their examination results to a lesser degree than low-achieving students. We also analysed the difference between the two genders. Compared to female students, males tend to overestimate their own performance.

INTRODUCTION

Self-assessment, or how we see ourselves (e.g. our characteristics, abilities, skills, and personality) plays an important role in our daily lives. For young people self-assessment plays an essential role in their decisions on further education or employment, because it determines how aware they are of their abilities, skills and knowledge (Keller, 2016). Accurate self-assessment and self-confidence have a positive effect when wage bargaining on the labour market, i.e. higher levels of self-confidence correlates positively with earnings (Keller, 2010). However, people are usually too optimistic about their social and intellectual activities, especially low-achievers, who tend to overestimate their own performance because of their incompetence they are unable to recognize their lack of skills (Ehrlinger et al. 2008). Many tutors experience the difficulties associated with exams (as predicted or suspected by students) and the expected results often differ from the actual performance. According to McDonald (2004), tutors have to confront the fact that the majority of students are not able to rationally assess their own readiness for the exam. Self-management of learning can be promoted by the development of self-assessment (Karnilowicz, 2012). This may occur if university teachers make greater effort to regularly account for different levels of knowledge and give feedback on results. Nicol, Macfarlane and Dick (2006) point out that students estimate their own abilities, and if these self-assessments are not accurate, they will make poor choices regarding their academic goals and efforts. If the modest abilities of students overestimate their performance, they invest less (or too little) power in learning the curriculum, so their goals and expectations will not be met. On the other hand, if students underestimate themselves, they will waste resources which could be invested in exploiting other opportunities. The present study aims to contribute to this topic by analysing data from higher vocational education students at the Faculty of Economics and Business, and the Faculty of Agricultural and Food Sciences and Environmental Management, both at the University of Debrecen. We analyse the accuracy of students' self-assessment and whether there is a general tendency to under- or overestimate. We try to find out whether higher-achieving students tend to overestimate their examination results less than the lower-achieving fellows. We also analyse whether there are any differences between the two sexes.

BRIEF LITERATURE REVIEW

According to Boud and Falchikov (1998) self-assessment is the involvement of students in shaping their view of themselves, especially as regards their results and other learning outputs. In a broader sense self-assessment is not only the evaluation of performance, but also the determination of standards, and so it is also linked to self-managed learning (see for example Karnilowicz, 2012). In the present study we understand the term in the former sense, i.e., when students assess their own performance (learning outputs).

The quality of self-assessment can be measured by two indicators, one being a review (estimation) of the accuracy of difference (the absolute value of the difference between the previously estimated and the actual achieved results), the other the direction of the difference (the signed difference).

Several studies provide empirical evidences that low-achiever students in higher education tend to predict and evaluate their own academic performance less accurately than those who perform better in their studies. Previously published papers have also supported the idea that low-performers generally over-evaluate while high-performers regularly underestimate their performance, or at least overestimate to a significantly lesser extent. These findings highlight the fact that poor skills and/or abilities are only one element of the low-achievers' handicap. Another serious problem is that they are unaware of these problems. This phenomenon is sometimes referred to as the 'Dunning–Kruger effect'.

Previous studies in the literature have frequently examined the role of sex among self-assessment influencing factors. Most studies did not find significant differences between the two sexes (see e.g. Boud and Falchikov, 1989; Kruger and Dunning, 1999; O'Neill et al., 2006; Basnet et al, 2012; Hobohm, 2012; Kun, 2016a; Máté et al., 2016), while some studies found a tendency to overestimation among men (Edwards et al., 2003; Macdonald, 2004;). These results fit well with the findings of Grijalva et al (2015) who showed that the propensity to narcissism is higher in men.

The type of questions in exams plays an important role in the accuracy of evaluation (Csehné, 2013; Kun, 2016b).

Several studies have observed a general trend towards overestimation (see for example Kruger and Dunning, 1999; Basnet et al, 2012; Tejeiro et al., 2012; Kun, 2016a), although Mehrdad, Bigdeli and Ebrahim, 2012 did not confirm this phenomenon.

Most studies produced similar results in terms of better-performing students estimating more accurately (Boud and Falchikov, 1989; Kruger and Dunning, 1999; Dunning et al., 2003; Karnilowicz, 2012; Tejeiro et al., 2012; Kun, 2015a; Máté et al., 2016).

The papers consulted by the authors found that those students who have achieved higher results

tend to overestimate themselves less (Boud and Falchikov, 1989; Kruger and Dunning, 1999; Hodges, Regehr and Martin, 2001; Edwards et al., 2003; Karnilowicz, 2012; Kun, 2016a; Máté et al., 2016).

Based on the findings of the literature reviewed above, the current study forms three hypotheses:

- ***H1:*** Higher-achieving students evaluate their examination results more accurately than their lower achieving fellows.
- ***H2:*** Higher-achieving students tend to over-assess their examination results less than low-achieving students.
- ***H3:*** Compared to female students, male tend to overestimate their own performance more.

SAMPLE AND METHOD

The data collection took place at the Faculty of Economics and Business, and the Faculty of Agricultural and Food Sciences and Environmental Management at the University of Debrecen among higher vocational education students. They were in the first semester of the academic year 2016/2017 and on the labour economics element of the course, and taking an examination. The exams were publicised between 21th December 2016 and 6th February 2017, 8 times in total. The students had to answer 18 multiple choice and 14 true or false questions, and give 8 definitions. We asked them to guess their performance, i.e. how many points they would achieve (maximum 40).

The students wrote 508 examinations altogether, some of them sitting it more than once, if they had failed or wanted to get a better mark (it was possible).

The structure of the sample by major, sex and full- or part time students is presented in Table 1.

The mean of students' scores is 24.373 (the minimum is 5, the maximum is 38), the standard deviation is 5.664. The mean of the tutors' scores is 21.537 (the minimum is 4, the maximum is 38), and the standard deviation is 5.681.

We analyse the accuracy and direction of students' self-assessment. Accuracy is defined as the absolute value of the difference between the student-assessed and the tutor-assessed test score, while direction is the signed (positive or negative) difference. We test our Hypothesis 1 with a linear regression model, Hypothesis 2 with a binomial logistic regression model, and use a dummy variable (sex) for testing Hypothesis 3.

The dependent and independent variables are listed below:

- ADIFFSC: the absolute value of the difference between the student's and the tutor-assigned estimation (dependent variable of the first model),
- OVEREST: 1, if the student overestimated his/her total test scores, 0, if not,
- FINALS: tutor-assigned final scores,
- SEX: 1, if the student is female, 0, if male,
- TIME: 1, if the student is part-time, 0, if full-time.
- FAILED: 1, if the student failed, 0, if not,
- MAJOR: as dummy variables: SFM: 1, if the student is on the Stud Farm course, 0, if not; PPPT: 1, if the student is on Plant Production and Protection Technology, 0, if not; AM: 1, if the student is on Agricultural Management, 0, if not; EM: 1, if the student is on Ecological Management, 0, if not; MRD: 1, if the student is on Management and Rural Development, 0, if not; BM: 1, if the student is on Business and Management, 0, if not; IB: 1, if the student is on International Business, 0, if not; FA: 1, if the student is on Finance and Accounting, 0, if not; TH: 1, if the student is on Tourism and Hospitality, 0, if not. The Commerce and Marketing course was excluded.

Our multivariable regression model:

$$Y_{ADIFFSC} = \beta_o + \beta_1 X_{FINALSC} + \beta_2 X_{TIME,} + \beta_3 X_{FAILED} + \beta_4 X_{SEX} + \beta_5 X_{MAJOR} + \varepsilon$$

Our binomial-logistic regression model:

$$Y_{OVEREST} = \beta_o + \beta_1 X_{FINALSC} + \beta_2 X_{TIME,} + \beta_3 X_{FAILED} + \beta_4 X_{SEX} + +\beta_5 X_{MAJOR} + \varepsilon$$

where

$$Y_{OVEREST} = ln \frac{P\{OVEREST = 1|u\}}{P\{OVEREST = 0|u\}}$$

RESULTS

The histograms and distributions can show the difference between the tutor's evaluation and the students' self-assessment (see Figure 1). Students' estimated scores slipped to the right compared to the tutor-assessed scores. This result suggests that students typically overestimate their abilities.

According to the H1 hypotheses, multivariate linear regression models should be tested, where the dependent variable is the accuracy of the students' estimations (ADIFFSC) measured by the absolute difference value of the student-estimated test scores and the tutor-assigned test scores. The FINALSC is substituted by the tutor-assigned test scores one independent variable among others. In our regression models the dummies of SEX, TIME, FAILED and MAJORS (see above) are selected to maximize the 'goodness of fit' (R^2, as the percentage of the response variable variation) of the linear regression models. The first (Model 1) contains all the available independent variables and the other (Model 2) is restricted to those that are significant at least at the 10% p-level.

Table 2 contains the statistics of our linear regression models. In Model 1 and Model 2 we found a significant linear connection between the accuracy of the students' evaluation and the tutor's assessment. We found a negative correlation between these two variables. This means that higher achieving students (who achieve a higher final scores) miscalculate less than lower achieving fellows. Essentially, the effect of tutor-assigned final scores on the absolute value of the differences of self and tutor assessment does not seems to be large, but in both models the student results correlated negatively with accuracy. Consequently, we can accept the H1 hypotheses; the higher achieving students seem to be able to evaluate their examination results more accurately than their lower achieving fellows. Our results conform to previous empirical studies (see above).

In order to identify the relationship between the students' achievement and the accuracy with which they overestimate their own performance, a binary logistic regression method might be an appropriate tool for our analysis. Table 3 contains the statistics of our binary-logistic regression models. The dependent variable indicates the likelihood of students' over-assessment. Those cases where the students evaluate their own performances accurately are estimated without an error and left out of the sample. The proportion of variance explained by the predictors (measured by Cox and Shell's, and Nagelkerke's pseudo R^2) of the binary logistic regression models are relatively high – indeed high enough – to agree with our results. As a result, for every one-unit increase in the tutor-assigned test scores (i.e. for every additional point, and holding all other independent variables constant), we found a decrease in the pre- and post-examined self-assessment differences. Consequently, higher achieving students tend to overestimate their own examination performance less, so we can similarly accept the H2 hypotheses, as well. Our results in the terms of overestimation also conform to previous empirical studies (see above).

In addition, we found that students, who failed the exam, or who are on part time courses, tend to overestimate their examination results more than students who have not failed, or who are on full-time courses.

However, there are several studies that could not identify any gender related effects of over-estimation, such as Kruger and Dunning (1999); O'Neill et al. (2006) and Hobohm et al. (2012) etc. We have also paid particular attention to variations in gender. According to our results, in Model 1 and Model 2 of overestimation, gender (SEX) has a negative and significant effect on accuracy. We found that the female students' probability of overestimation is significantly lower than males. To strengthen our empirical evidence we apply an independent sample *t*-test for equality of means. The average absolute difference between students and tutor scores among women is 5.000, and 5.906 among men. We can assume equal variances (F=3.700, sig=0.055). According to the results of the independent samples *t*-test, the difference between the two gender groups is significant at 5% p-level (t=2.218, sig=0.027). Hence, H3 hypotheses can be accepted because female students seem to overestimate their examination results less than their male colleagues.

CONCLUSIONS

In this study the first objective was to analyse the self-assessment behaviour and efficiency among advanced level vocational training students which is the lowest level of higher education. This analysis provides new empirical results for the literature from a sample of Hungarian advanced level vocational training students. Using various statistical methods, the results confirm the hypothesis that high-achieving students are more accurate in their examination self-assessment. This result is in accordance with the conclusion of Boud and Falchikov, 1989; Kruger and Dunning, 1999; Dunning et al., 2003; Karnilowicz, 2012; Tejeiro et al., 2012; Kun, 2015a; Máté et al., 2016.

A further conclusion is that higher-achieving students are less likely to overestimate their performance, which are supporting the results of Boud and Falchikov, 1989; Kruger and Dunning, 1999; Hodges et al. 2001; Edwards et al., 2003; Karnilowicz, 2012 ; Kun, 2016a; Máté et al., 2016.

This study found a difference between the sexes, with female students overestimating significantly less. Edwards et al. (2003) and McDonald (2004) identified a higher tendency to self-overassessment in the case of male students, but Kun (2016a) and Máté et al. (2016) do not support these findings.

An overall tendency among students to over-rate their own examination performance is also explored in papers by Kruger&Dunning, 1999; Basnet et al, 2012; Tejeiro et al., 2012; Kun, 2016a etc.

Since the analysed sample of students is from one year of a given university, sitting an examination in a given undergraduate subject, any generalization of the results should be approached with caution. However, most of the findings are supported by some element of the previous studies, thus taken in context, the results of this paper can contribute to better understanding of the wider picture of students' self-assessment. Moreover, we expect to implement further analyses in the coming years to explore and expand the extent to which other determinants may explain and make comparable the self-assessment of students, including for example learning time, frequency of exams,

previous experiences, demographic variables, academic area, and ethnicity.

REFERENCES

[1] Basnet, B., Basson, M., Hobohm, C., & Cochrane, S. (2012). *Student's self asssessment of assignments, is it worth it?* Proceedings of the 2012 AAEE Conferecne, Melbourne, Victoria.

[2] Boud, D., & Falchikow, N. (1989). Quantitative studies of students' sefl-assessment in higher education: a critical analysis of findings. *Higher Education 18*(5), 529-549.

[3] Csehné Papp, I. (2013). A társadalomtudományok oktatásának módszertani jelentőségei [The methodology of teaching social sciences]. in. Karnilowicz, T.J. (Ed) . *Társadalomtudományi gondolatok a harmadik évezred elején [Toughts of Social Sciences at the beginning of the third millennium]* (312-317). Komarno: International Reseearch Institute.

[4] Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directionsin Psychological Science 12*(3), 83-87.

[5] Edwards, R. K., Kellner, K. R., Sistrom, C. L., & Magyari, E. J. (2003). Medical student self-assessment of performance on an obstetrics and gynaecology clerkship. *American Journal of Obstetrycs and Gynaecology 188(4*), 1078-1082.

[6] Ehrlinger, J., Johnson, K., Baner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behaviour and Human Decision Process 105*(1), 98-121.

[7] Grijalva, E., Newman, D. A., Tay, L., Donellan, M. B., Harms, P. D., Robins, R. W., & Yan, T. (2015).Gender differences in narcissm: A meta-analytic review. *Psychological Bulletin (141),* 261-310.

[8] Hobohm, C, Cochrane, S., Basson, M., & Basnet, M. (2012). Students' Self-assessment of Assignments – Is It worth It? Proceedings of the 2012 AAEE Conference. Melbourne, Victoria.

[8] Hodges, B., Regehr, G. & Martin, D. (2001). Difficulties in Recognizing One's Own Incompetence: Nowice Physicians Who are Unskilled and Unaware of it. *Academic Medicine: Journal of The Association of American Medical College 76*(10), 87-89.

[10] Karnilowicz, W. (2012). A Comparison of Self-Assessment and Tutor Assessment of Undergraduate Psychology Students. *Social Behacviour and personality 40*(4), 591-604.

[11] Keller, T. (2010). Az önbizalom kereseti hatása [The impact of self-confidence on earnings]. *Közgazdasági Szemle 62*(3), 241-260.

[12] Keller, T. (2016). Ha a jegyek nem elég jók... Az önértékelés szerepe a felsőoktatásba való jelentkezésben [If the marks are not good enough... The role of self-assessment in applicatoin to higher education]. *Közgazdasági Szemle 63*(1) 62-78.

[13] Kun, A. I. (2016a). A comparison of self- vs. tutor assessment among Hungarian undergraduate business students. *Assessment and Evaluation in Higher Education 41*(3), 350-367.

[14] Kun, A. I. (2016b). Önértékelés és teljesítmény az üzleti felsőoktatásban [Self-assessment and performance in business higher education]. *Taylor: Gazdálkodás- és Szervezéstudományi Folyóirat: A Virtuális Intézet Közép-Európa Kutatására Közleményei 8*(3), 112-119.

[15] Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessment. *Journal of Personality and Social Psychology 77*(6), 1121-1134.

[16] Macdonald, A. (2004). Student sefl-evaluation of coursework assignments: a route to better perception of quality. *Learning and Teaching in Higher Education 1*(1), 102-107.

[17] Máté, D., Kiss, Zs., Takács, V. L., & Molnár, V. (2016) Measuring Financial literacy: A case study of self-assessment among undergraduate business students in Hungary. *Annals of The University of Oradea Economic Science 25*(1), 690-697.

[18] Mehrdad, N, Bigdeli, S., & Ebrahimi, H. (2012). Institutional Comparisons of Educational Attainment.A comparative study of self, peer and teacher evaluation to evaluate clinical skills of nursing students. *Procedia, Social and Behavioral Science 47*(1), 1847-1852.

[19] Nicol, D. & Macfarlane, D. (2006). Formative assessment and self-regulated learning: a model and seven priciples of good feedback practice. *Studies in Higher Education 31*(2), 199-218.

[20] O'Neill, P., Holzer, C., & Lynn, D. J. (2006). "Relationships between Self-assessment Skills, Test Performance, and Demographic Variables in Psychiatry Residents. *Advances in Health Sciences Education 11*(1), 51–60.

[21] Tejeiro, R. A., Gomez-Vallecillo, J., L., Romero, A., F., Pelegrina, M., Wallace, A., & Emberley, E. (2012).Summative self-

assessment in higher education: implication of its counting towards the final mark. *Electronic Journal of Research in Educational Psychology 10*(2), 789-812.

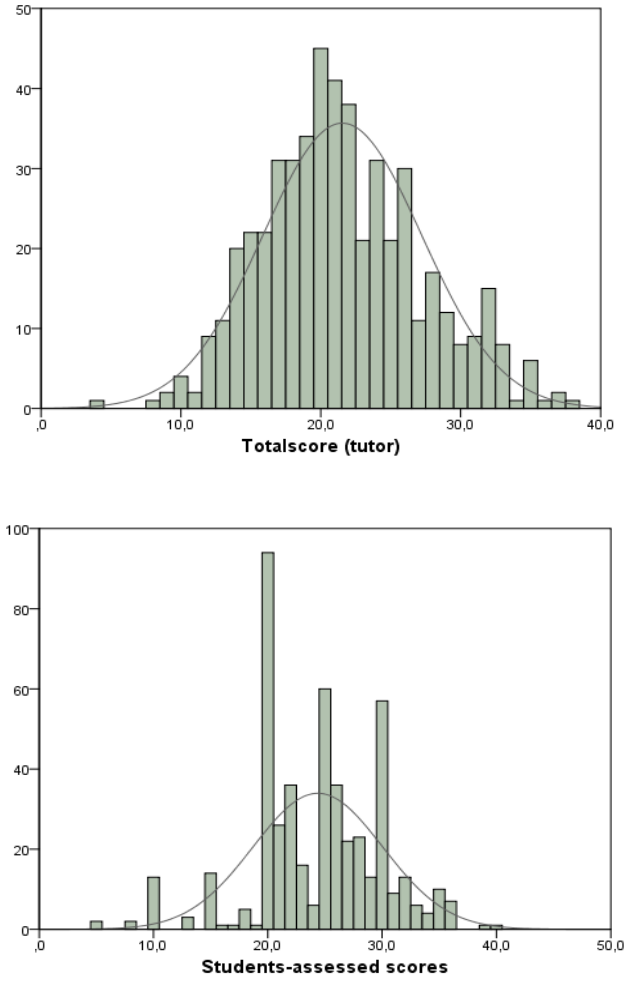
ANNEXES

Table 1. Sample structure by major, sex and full- or part-time students

		Faculty										
		FAFSEM				FEB						
Structure		Major										
status	sex	SFM	PPPT	AM	EM	MRD	BM	CM	IB	FA	TH	Total
full-time	male	5	11	38	11	18	17	40	24	17	22	203
full-time	female	4	0	13	13	15	32	34	36	49	53	249
part-time	male	0	0	8	2	1	4	5	0	2	0	22
part-time	female	0	0	3	1	0	3	4	9	10	4	34
Total		9	11	62	27	34	56	83	69	78	79	508

Source:primary data
Note: FAFSEM = Faculty of Agricultural and Food Sciences and Environmental Management, FEB = Faculty of Economics and Business, *SFM = Stud Farm, PPPT = Plant Production and Protection Technology, AM = Agricultural Management, EM = Ecological Management, MRD = Management and Rural Development, BM = Business and Management, CM = Commerce and Marketing, IB = International Business, FA = Finance and Accounting, TH = Tourism and Hospitality*

Figure 1. The tutor's evaluation (upper) and students' self-assessment (lower)



Source: authors' own data
Note: vertical axis: number of students, horizontal axis: examination scores

Table 2. Results of linear regression models for the self-assessment

Dependent variable:	**ADIFFSC**			
	Model 1		**Model 2**	
	β	***t***	***β***	***t***
CONSTANT	**10.827*****	11.590	**10.691*****	14.087
FINALSC	**-0.245*****	-6.663	**-0.246****	7.207
SEX	-0.612	-1.443	-	
TIME	0.221	0.336	-	
FAILED	0.279	0.629	-	
MAJOR-SFM	0.354	0.229	-	
MAJOR-PPPT	-0.183	-0.204	-	
MAJOR-AM	0.049	0.050	-	
MAJOR-EM	0.239	0.312	-	
MAJOR-MRD	-0.379	-0.504	-	
MAJOR-BM	-0.290	-0.400	-	
MAJOR-IB	0.032	0.022	-	
MAJOR-FA	0.790	1.116	-	
MAJOR-TH	0.234	0.333	-	
N	**507**		**508**	
R^2	**0.102**		**0.093**	
Adjusted R^2	**0.078**		**0.091**	
Durbin Watson	**1.905**		**1.884**	

Source:authors' own data

*Note: Letters in the upper index refer to significance: ***: significance at 1 per cent, **: 5 per cent, *: 10 per cent. P-values without an index mean that the coefficient is not significant even at the 10 per cent level.*

Table 3. Results of the linear regression models for the self-assessment

Dependent variable:	**OVEREST**			
	Model 1		**Model 2**	
	β	***W***	***β***	***W***
CONSTANT	**5.451*****	0.632	**5.315*****	91.622
FINALSC	**-0.218*****	75.854	**-0.214****	78.955
SEX	**-0.435***	3.246	**-0.537****	4.611
TIME	**0.719***	3.781	**0.781****	12.660
FAILED	**0.923*****	11.386	**0.960*****	5.622
MAJOR-SFM	-0.396	0.217	-	
MAJOR-PPPT	0.056	0.003	-	
MAJOR-AM	0.314	0.440	-	
MAJOR-EM	-0.575	1.218	-	
MAJOR-MRD	0.436	0.649	-	
MAJOR-BM	-0.130	0.092	-	
MAJOR-IB	-0.234	0.324	-	
MAJOR-FA	0.093	0.053	-	
MAJOR-TH	-0.501	1.535	-	
N	**482**		**482**	
Cox&Snell R^2	**0.259**		**0.248**	
Nagelkerke R^2	**0.354**		**0.34**	
Omnibus χ^2 test	**144.236*****		**137.504**	
HL χ^2 test	**9.089**		**3.349**	

Source:authors' own data

*Note: Letters in the upper index refer to significance: ***: significance at 1 per cent. **: 5 per cent. *: 10 per cent. P-values without an index mean that the coefficient is not significant even at the 10 per cent level. HL: Hosmer and Lemeshow χ2test.*