TRAMES, 2013, **17**(**67**/**62**), 4, 417–432

WHO IS WILLING TO TAKE LOW-STAKES ASSIGNMENTS?

Mari-Liis Mägi, Liina Adov, Karin Täht, and Olev Must

University of Tartu

Abstract. The main purpose of this study is to explore which students of Estonian higher education institutions are willing to take low-stakes tests which have no direct consequences for them. Altogether 603 first-year undergraduates from different institutions of higher education participated in the study – 46.3 per cent of them took the low-stakes cognitive test. Female students were more willing to participate in the survey. Test-takers, compared to the students who did not take the low-stakes test measuring mental ability, had lower levels of self-evaluation and higher results in national examinations taken at the end of high school. Substantial differences between genders emerged. For male students, previous performance predicted test-taking activity, whereas no variables predicted test-taking activity in female students. When predicting test results, paradoxical relationships with motivation appeared – female students who had higher levels of motivation had lower results in the low-stakes test. It is important to take into account that when interpreting low-stakes tests significant differences could be overlooked when genders are considered together.

Keywords: higher education, low-stake assignements, academic motivation, self-evaluations, gender differences

DOI: 10.3176/tr.2013.4.07

1. Introduction

The process of acquiring a higher education in Estonia has been changing over the last two decades. Estonian researchers have written about the possible devaluation of higher education – as higher education has become more obtainable, the number of students in Estonia has almost tripled in the last twenty years (Unt, Täht, Saar, and Helemäe 2013). The researchers found that, for many students, a degree in itself has become more important than the quality of the education they acquire. As the main purpose for many students is to pass and get a degree, not to achieve mastery of the subject matter, many education-related tasks are low-stakes

for students. Low-stakes tasks are, for example, tests that are rather trivial for students – failure or a bad grade does not result in serious consequences. So, we argue that, for many students in Estonia, the process of acquiring a higher education could be a sequence of taking low-stakes tests.

From another point of view, these kinds of tests are the quickest and easiest way for lecturers and professors to evaluate students. Testing is also widely used in low-stakes conditions in which the results of the tests are rather trivial for students. For example the Program for International Student Assessment (PISA) and the Trends in International Mathematics or Science Study (TIMSS) use standardized tests to assess the impact of education quality and to understand what causes differences in achievement across nations. The results, while important to governmental institutions, tend not to be particularly significant to students.

In order to motivate students not willing to participate in low-stakes testing and voluntary academic work, it is important to know their characteristics (i.e. academic motivation and self-evaluation [self-efficacy, self-esteem]). Furthermore, motivating students to be better in their process of acquiring knowledge would greatly improve the quality of higher education.

2. Previous studies

There have been few attempts to distinguish students who are ready to do low-stakes tests and those who are not (Brown and Gaxiola 2010, Eklöf 2010, Wise and DeMars 2005). Unfortunately, there are no studies at all investigating how different these students are in terms of psychological indicators. It is natural to investigate candidates for psychological indicators related to the willingness to take low-stakes tests from the set of psychological factors known to be related to students' academic achievement. So far, research (Deci and Ryan 2000, Marsh and Hau 2003, Rosenberg et al. 1995) has shown positive relationships between academic achievement, academic motivation, and self-evaluation, but the relationships between self-beliefs and the willingness to takes low-stakes tests have not been investigated. Therefore, the main purpose of the current study is to explore whether students' willingness to take low-stakes tests is related to their self-evaluation and academic motivation. Besides self-evaluation and academic motivation, previous academic results could play an important role in students' decision whether to take low-stakes tests or not, so we are interested in how previous academic results influence this decision. As there are reported gender differences in the case of low-stakes testing (Eklöf 2007; Schnipke 1995; Wise, Kingsbury, Thomason, and Kong, 2004)), our purpose is to investigate if there are gender differences in the willingness to take low-stakes tests and if are there differences in low-stakes tests results.

2.1. Motivation to learn and test-taking motivation

In the broad sense, motivation is defined as "to be moved to do something" (Deci and Ryan 1985). One can think about motivation as a unitary construct. It means that motivation (to act) can vary from a little to a great deal. There are quite many motivation theories, but one of the widely used approaches is to divide motivation into intrinsic and extrinsic. Intrinsic motivation refers to doing something because it is inherently interesting or enjoyable, and extrinsic motivation refers to doing something because it leads to a separable outcome.

One of the subcategories of motivation in the educational context is academic motivation – a person's desire regarding academic subjects when competence is judged against a standard of performance or excellence (Eccles and Wigfield 2002). Academic motivation and its relatedness to academic achievement have been widely researched and it is a known fact that motivation is positively related to students´ academic results (Chemers, Hu and Garcia 2001, Deci and Ryan 2000, Phan, 2010, Täht and Must 2009). It could be argued that the more motivated students are, the deeper their information processing, which in turn, is associated with higher academic achievement (Deci and Ryan 2000). It has been previously shown that academic motivation is important in assessment situations.

One specific kind of motivation has been investigated within the framework of low-stakes and high-stakes tests: according to Eklöf (2010), test-taking motivation is the motivation to perform well in a given test or in a given situation. It has been found to depend on whether the specific test is low-stakes or high-stakes for the students taking it (Brown and Gaxiola 2010, Eklöf 2010, Segal 2012, Wise and DeMars 2005). In the case of a high-stakes test, the test has at least some academic or other meaningful consequence for the student (Cole and Osterlind 2008), whereas in the context of a low-stakes assessment, there are typically no consequences (Wise and DeMars 2005). Here, motivation becomes the key element in the performance of the individual (Brown and Gaxiola 2010). Namely, when students feel that there is no consequence for them of the exam, whether positive or negative, they are less likely to be motivated to try their best (Eklöf 2010, Wise and DeMars 2005).

2.2. Self-evaluations and their relatedness to academic achievement

In addition to motivation, self-evaluations have been found to be positively related to students´ academic achievement (Marsh and Hau 2003, Pullmann and Allik 2008, Rosenberg et al. 1995). Self-evaluation could be seen as an umbrella concept that includes several different self-beliefs (Judge et al. 2002). In the current study, we have focused on general and academic self-esteem and self-efficacy. Self-esteem is an overall appraisal of one's self-worth (Rosenberg 1965), whereas academic self-esteem has been described as a self-evaluation in school performance (Rosenberg et al. 1995). Self-efficacy, however, is one's belief in one's ability to succeed in specific situations, and academic self-efficacy is an individual's belief that they can successfully succeed at a designated level on an academic task (Bandura 1977).

Self-evaluations have also been found to be related to learning and academic achievement. Self-efficacy has a positive effect on deep cognitive learning: individuals with a high sense of perceived competence are more likely to spend more time and effort on challenging goals (Prat-Sala and Redford 2010). Another important factor in the learning process is self-esteem. Individuals who feel good about themselves are more likely to succeed in learning (Phan 2010). Students who enter college with confidence in their ability to perform well academically do perform significantly better than less confident students (Chemers, Hu and Garcia 2001).

2.3. Gender differences

It is a well-known fact that there are gender differences in education: grades of girls are higher than those of boys and girls conform more with educational requirements. Mikk, Täht, and Must (2011) found gender differences in educational achievement and argued that one of the likely candidates for gender differences in educational achievement is motivation (Spinath, Freudenthaler and Neubauer 2010). Also, significant differences in test-taking activity between genders have been reported. Based on the data from TIMSS, Eklöf (2007) found that, on average, girls have higher test-taking motivation than boys. Similarly, Kinzie et al. (2007) reported that male examinees' motivation scores were 10% lower than those of female examinees. According to Cole and Osterlind (2008), female college students are more willing to take low-stakes tests than males, as 30.5% of females took the low-stakes test in question, compared to 19.5% of males. Also, Segal (2012) found females to be more likely to invest effort in tests even without incentives.

2.4. Approach with latent variables

Next, in order to avoid correlations between overlapping constructs and to get a more general understanding of how self-evaluations and academic motivation are related to the willingness to take low-stakes test we consider a latent variable approach.

Previous studies have shown that self-evaluations, specifically self-esteem and self-efficacy, are positively correlated to each other (Maccio and Schuler 2012, Dachlbeck and Lightsey 2008, Huang and Liu 2007). There is reason to believe that general and academic self-evaluation are also positively related. Pullmann and Allik (2008) pointed out that there are positive correlations between general and academic self-esteem. Judge et al. (2002) showed that self-esteem and self-efficacy are part of multidimensional construct, they named this Core Self-Evaluations. Based on different self-evaluation research**,** the same tendencies have been seen in the case of more specific self-evaluations, like general and academic self-efficacy and self-esteem (Maccio and Schuler 2012, Dachlbeck and Lightsey 2008, Pullmann and Allik 2008). So, based on previous research, it is natural to assume that self-evaluations are influenced by latent traits that include both general and academic self-esteem and self-efficacy.

Similarly, it has been found that motivation (including academic motivation) could be considered as a singular construct (Stover et al. 2012), since the structures of extrinsic and intrinsic motivation are similar (Deci and Ryan 1985, 2000). Whether intrinsic or extrinsic, motivation leads to action. Therefore a person could be simultaneously extrinsically and intrinsically motivated to do something. For example, Täht and Must (2009) showed that motivation can be seen as one latent variable.

2.5. Purpose of the study

As mentioned above, research has shown that academic motivation and students' self-evaluation are positively related to academic performance, but to our knowledge there are no studies focusing on the correlations between academic motivation, self-evaluation, and low-stakes test results. In addition, to the best of our knowledge, there are no studies assessing differences between low-stakes test-takers and non-takers in self-evaluations, academic motivation, previous academic outcome, and differences across genders in these variables. So, we believe that it is important to address these issues as they need to be taken into account in dealing with low-stakes testing. Therefore, we put forward the following questions:

1. Should the dimensionality of self-evaluative and motivational scales be reconsidered?
2. How do test-takers differ from those students who do not take the test in terms of academic motivation, self-evaluation, and previous academic results?
3. Is it possible to predict test-taking activity and test results based on academic motivation, self-evaluation, and previous academic results?
4. Are there gender differences in test-taking activity, self-evaluation, and motivation?

3. Data and method

3.1. Sample

Our research is part of a longitudinal survey focused on students' self-evaluation, learning motivation, and mental abilities. The duration of the survey was from fall 2012 to summer 2013. Undergraduates were informed about the possibility to participate in a longitudinal study via e-mail, web page, and presentations in institutions of higher education. The survey was conducted in a web-based testing environment. In this paper, we concentrate only on the data collected at the beginning of fall 2012 from first-year undergraduate students (N=603) from different institutions of higher education in Estonia. Participants filled out questionnaires about academic motivation, academic and general self-esteem and self-efficacy, and academic self-concept and completed a mental ability test (a shortened version of the scholastic aptitude test). They also reported their national examination results (NER), gender, and age. Filling out all the tests available in

the survey was not obligatory and therefore the number of respondents varies between tests. The only external motivator for answering the questionnaires and taking the mental ability test was the opportunity to find out their test results compared to other participants in the survey. The sample gender distribution was not equal: 410 of the participants were female (67%) and 193 students (33%) male. In Estonian higher education, approximately 60% of students admitted are women (Tõnisson 2011), so our sample approximately represented the student population. The average age for the sample was 20.4 years (sd = 3.2): for females 20.1 years (sd = 3.2) and male students, 21.0 years (sd = 3.1).

3.2. Measures

Five different self-report scales were used in order to measure student self-evaluation. All scales consisted of several attitudinal statements, four of which were rated on a 5-point scale, ranging from 1 (strongly disagree) to 5 (strongly agree). Motivation scale statements were rated on a scale that ranged from 1 (strongly disagree) to 7 (strongly agree). The scale score was calculated as the total sum of the item values. We used the following scales:

- General self-efficacy (GSEf) – 5 items from the Estonian version of Schwarzer and Jerusalem Scale of General Self-Efficacy (Rimm and Jerusalem 1999).
- Academic self-efficacy (ASEf) - 16 items from the Estonian Scale of Academic Self-Efficacy (Üpraus 2009).
- General self-esteem (GSE) - the Estonian version of the Rosenberg Self-Esteem Scale (Rosenberg 1965; Pullmann and Allik 2000)
- Academic self-esteem (ASE) – an 8-item scale developed by Pullmann and Allik (2008) as an analogy to the Marsh (1992) scale for assessment of perception of self-competence in the academic domain.
- Academic motivation – the Academic Motivation Scale (AMS-C 28) College Version (Vallerand et al. 1992, 1993) Estonian version (Mägi 2012). The scale consisted of 24 items in total, of which 12 items measured intrinsic (IM) and 12 extrinsic motivation (EM).

Educational achievement was estimated by the index derived from the results of the national examination result taken when graduating from high school. In order to graduate from high school, students in Estonia have to take 5 examinations, from which at least 3 have to be national examinations. Only one national exam is obligatory – the Estonian language – and students can choose the remaining two. As the number and topics of national examinations taken is different between students, the national examination results in our study do not represent the mean score in examinations, but rather the scores students achieved in their examinations. National examination scores were self-reported on a 100-point scale. For the general national examination result index, some transformations to a 9-point scale were made. Further explanation of the scale is shown in Table 1. For instance, 9 points were received by students whose grades were all higher than 90;

8 points were received by students who had two grades higher than 90, and so on. Students whose grades were all lower than 50 got 0 on our new scale.

Table 1. The encoding of national examination scores

Encoding	Scores in national examinations
0	All grades are lower than 50
1	One grade is between 50–70, other scores are lower
2	Two grades are between 50–70, other scores are lower
3	Three grades are between 50–70, other scores are lower
4	One grade is between 79–90, other scores are lower
5	Two grades are between 79–90, other scores are lower
6	Three grades are between 79–90, other scores are lower
7	One grade is higher than 90, other scores are lower
8	Two grades are higher than 90, other scores are lower
9	All grades are higher than 90

General cognitive ability (GCA) was estimated by the shortened version of the admission test of the University of Tartu (Must and Allik 2001). The test consists of 3 sections (vocabulary, mathematics, and spatial reasoning) – there were 15 items in all subtests, with 45 questions altogether. We consider the GCA test a low-stakes test, as the test-takers had no external rewards and the test results had no consequences for them. The only motivating factor was the abovementioned personal feedback.

Test-taking activity was based on the time participants spent on different sections of the GCA test. A threshold of 6 minutes was determined based on the mean time spent on the vocabulary test, as it was the first subtest in the GCA. Based on their test-taking activity and determined threshold, participants were divided into two groups – test-takers and non-takers.

We used correlation analysis (Pearson correlations), Student t-test, factor analysis, and logistic and linear regression analysis with SPSS 20.0.

4. Results

4.1. Dimensionality of self-evaluation and motivation scales

The correlations between self-evaluation and motivation scales are low to modest – correlations are in the range of r = .132 to r = .524 (Table 2). With the aim to get a more general estimate of students' self-evaluation, factor analysis was used (principal axis factoring, Varimax rotation). The number of factors that could emerge from the analysis was not determined beforehand.

Two different dimensions emerged – one that influences academic and general self-esteem and self-efficacy (explained 35% of the common variance of variables) and another that influences intrinsic and extrinsic motivation (explained 21% of variance) (Table 3). The first factor was named self-evaluation (SEF) and the second one academic motivation (AMF).

Table 2. Correlations between self-evaluation scales and low-stakes test results (GCA)

	ASE	GSEf	ASEf	IM	EM	GCA
General self-esteem (GSE)	**.524****	**.627****	**.472****	**.157****	.066	.016
Academic self-esteem (ASE)	1	**.475****	**.577****	**.193****	–.014	**.251****
General self-efficacy (GSEf)		1	**.497****	**.243****	**.132****	.051
Academic self-efficacy (ASEf)			1	**.454****	**.247****	.018
Intrinsic motivation (IM)				1	**.508****	–.017
Extrinsic motivation (EM)					1	**–.156****

**p<0.01

Table 3. Factor structure of students' self-evaluation and motivation scales (factor loadings)

	Self-evaluation	Academic motivation
GSE	.763	
ASE	.728	
GSEf	.715	
ASEf	.662	
EM		.836
IM		.834

GSE – general self-esteem, ASE – academic self-esteem, GSEf – general self-efficacy, ASEf – academic self-efficacy, EM – extrinsic motivation, IM – intrinsic motivation.

4.2. Differences between genders

Of students who participated in our study, 46.3% (N = 279) did not take the low-stakes general cognitive ability test (non-takers). Female students were more willing to take a low-stakes test, as 58% of female students took the test compared to 44% of male students.

We used a t-test to clarify if there were any gender differences in national exam results, test results, SEF, and AMF. Female and male students were different in terms of AMF ($t(443) = 5.78$, $p<.000$) and low-stakes test results ($t(322) = –5.36$, $p<.000$) . Female students reported higher levels of academic motivation and had lower results in the low-stakes test (Table 4).

Table 4. Gender differences in national examination and test results, self-evaluation, and motivation factors

	Female		Male		t-test			Effect size
	Average	St. dev.	Average	St. dev.	t-value	df	p-value*	Cohen's d
Motivation factor	.16	.79	–.33	.94	5.78	443	**.000**	.564
Self-evaluation factor	–.05	.89	.10	.91	–1.59	443	.113	–.167
National examination result	5.98	1.78	5.78	1.86	1.24	541	.214	.110
Low-stakes test result (GCA)	22.68	10.23	29.58	10.06	–5.36	322	**.000**	–.680

4.3. Differences between test-takers and non-takers

A t-test was conducted in order to check if there were mean differences in national examination results, SEF, and AMF in the two groups that differed in terms of test-taking activity.

Test-takers, compared to non-takers, had lower levels of SEF (t (443) =2.10, p= 0.036) and higher levels of national examination results (t (541) = 4.83, p<.000) (Table 5).

Table 5. Undergraduates' SEF and national examination results. Differences between the scores of low-stakes test-takers and non-takers

	Non-takers		Test takers		t-test			Effect size
	Avg.	St. dev.	Avg.	St. dev.	t-value	df	p-value*	Cohen's d
Motivation factor	–.09	.92	.04	.86	–1.42	443	.157	–.146
Self-evaluation factor	.12	.75	–.05	.95	2.10	443	**.036**	.199
National examination result	5.49	1.81	6.23	1.74	–4.83	541	**.000**	–.417

4.4. Predicting test-taking activity

We used logistic regression in order to predict test-taking activity based on gender, SEF, AMF, and national examination results.

We used three models in logistic regression – first with genders together, second with only the male student sample, and third with only the female student sample. At first, we ran logistic regression with four variables: national examination results, SEF, AMF, and gender. AMF was not statistically significant in the model (Table 6). Models with other variables (Model 1) explained 8 to 12% of test-taking activity variability. A test of the full model against a constant-only model was statistically significant, indicating that the predictors as a set reliably distinguished between test-takers and non-takers ($\chi^2 = 37.70$, p <.000 with df = 3). Odds ratio (Exp(B)) values for gender, SEF, and national examination results are 0.41, 0.76, and 1.25, respectively. Male students and students with higher

Table 6. Predicting test-taking activity (logistic regression)

		B	Wald	Sig.	Exp(B)
Model 1	Gender	–.902	16.70	**.000**	**.406**
	Self-evaluation factor	–.272	4.74	**.030**	**.762**
	National examination result	.225	13.45	**.000**	**1.252**
Model 2	Self-evaluation factor	–.275	2.04	.154	.760
	National examination result	.583	24.62	**.000**	**1.791**
Model 3	Self-evaluation factor	–.252	2.95	.086	.777
	National examination result	–.009	.01	.915	.991

Model 1 – genders together (0 = female; 1 = male), Model 2 – only male students. Model 3 – only female students.

levels of self-evaluation are less likely to take the test whereas students with higher national examination results are more likely to take the test. Results of the logistic regression show that gender is an important variable predicting low-stakes test-taking activity, even when national examination results and self-evaluation are controlled for.

The second model (Table 6) considered only the male student sample. One statistically significant variable remained in the model: national examination results ($p < .000$). The model explained 20 – 27% of test-taking activity. A test that compared the full model against a constant only model was statistically significant ($\chi^2(2) = 33.44$, $p < .000$). Just as in the case of the general model, students who had higher levels of national examination results were more likely to take the low-stakes test (Exp(B) accordingly 1.68). In the case of the female student sample (Model 3), no statistically significant predictors occurred ($\chi^2(2) = 3.18$, $p = 204$).

4.5. Predicting test results (GCA)

We used linear regression to predict low-stakes test results. More specifically, we used gender, national examination results, SEF, and AMF as independent variables. Only gender ($t = 4.34$, $p<.000$) and national examination results ($t = 11.11$, $p<.000$) were statistically significant predictors for test results (Table 7). The model explains 36% of the variation in test results ($F(2) = 86.67$, $p<.000$). As gender is a statistically significant predictor of test results, and the previous findings bring out a number of differences between genders, we conducted a linear regression for genders separately. In the male sample, only national examination results had predictive value ($t = 6.21$, $p<.000$); the model predicted 31% of low-stakes test result variation ($F(1) = 36.23$, $p<.000$). In the female sample, both AMF ($t = –2.26$, $p = .025$) and national examination result ($t = 9.64$, $p<.000$) were significant predictors for low-stakes test results. While the relationship between test results and national examination results is positive, the relationship with motivation is negative – female students with higher motivation have lower results in the low-stakes test. The model conducted on the female

Table 7. Predicting low-stakes test results

Gender		Standardized B	t	Sig.
Total sample	NER*	.530	11.11	.000
	SEF*	.005	.107	.915
	AMF*	–.053	–1.11	.270
	Gender	.211	4.34	.000
Female	NER*	.547	9.64	.000
	SEF*	.066	1.17	.243
	AMF*	–.127	–2.26	.025
Male	NER*	.126	6.21	.000
	SEF*	–.158	–1.62	.110
	AMF*	.126	1.34	.185

Dependent variable: low-stakes test result
*NER – national examination result; SEF – self-evaluation factor; AMF – motivational factor

sample explained 32% of the variation in test results ($F(1) = 51.29$, $p<.000$). So, national examination results are an important predictor of low-stakes test results for both genders, but AMF is a significant predictor only in the female student sample. Correlations in Table 2 also illustrate the negative relationships between motivation and test results and give reason to believe that the negative predictive value is due to extrinsic motivation.

5. Discussion

Our purpose in the current study was to find out if students who are more motivated and have higher self-evaluations are more willing to participate in low-stakes tests and get better results. We were also interested to see whether their gender and previous academic results have an impact on the results of low-stakes tests. The latter would help clarify whether tests taken in low-stakes conditions are influenced by any motivational or self-evaluation factors.

Based on the results of exploratory factor analysis, we showed that four self-evaluation scales – general and academic self-esteem and self-efficacy – are influenced by one latent variable we named the self-evaluation factor (SEF). Similarly, two motivational scales – extrinsic and intrinsic motivation – were influenced by the academic motivation factor (AMF). Previous studies also support our findings (Stover et al. 2012, Täht and Must 2009).

5.1. Gender differences

Women were more willing to participate in the survey in general (70% of participants were female students) and in the low-stakes test (58% of female students took the test, whereas the percentage for male students was 44). Similar tendencies have also been reported in previous studies. Kinzie et al. (2007) found that female students devoted more time and effort to academic activities such as studying and also participated more often in a learning community. Male undergraduates, on the other hand, engage less frequently in academically challenging activities, and are systematically less engaged than their female counterparts.

Females appeared to have higher academic motivation compared to male students. This finding was supported by Eklöf (2007), according to whom female students reported a higher level of test-taking motivation than males. However, male students had higher results in the low-stakes test, which could be expected, as it was a GCA consisting of three subtests: vocabulary, mathematics, and spatial reasoning. As previous studies have shown, males tend to score higher in maths and spatial reasoning, whereas females get better results in vocabulary tests (Lynn and Mikk 2008, Mikk, Täht and Must 2012). Therefore, the overall better results of males are not surprising.

5.2. Differences between test-takers and non-takers and predicting test-taking activity

Our main purpose was to investigate if test-takers differed from non-takers in terms of motivation, self-evaluation factor, and previous academic results. Students who took the low-stakes test had lower levels of self-evaluation and higher levels of national examination results, effect sizes respectively d = .20 and d = –.42. This is congruent with the findings of Pullmann and Allik (2008), who reported that students with lower self-evaluation have been found to have higher academic achievements. We also found that gender, self-evaluation factor, and national examination results predicted 8–12% of test-taking activity.

When looking at genders separately, only male students had statistically significant differences between test-takers and non-takers in academic achievement. For male students, national examination results alone were a significant predictor, which explained 20–27% of the test-taking activity. The result is quite surprising as there were no differences between genders in national examination results, but, among males, those with higher national examination results were more willing to participate in low-stakes tests.

Our study gives reason to believe that previous success in academic work may also be a motivator to participate in low-stakes tests. Therefore, it could be that male students with higher previous results are more willing to put themselves to the test.

5.3. Predicting results of low-stakes tests

Gender and national examination result were significant predictors that predicted 36% of variance in test results. The models here were quite different when looking at genders separately. Being a male student and having higher results in national examinations also predicted higher results in the low-stakes test. For female students, both national examination results and the motivation factor predicted test results, while students with higher levels of motivation had lower results in tests. In the case of male students, only national examination results had any predictive value. This is quite unexpected and controversial as previous studies have shown that academic motivation is positively correlated with academic achievement (Deci and Ryan 2000, Eccles and Wigfield 2002, Täht and Must 2009). These new findings give reason to believe that students who are more oriented toward high academic achievement are less willing to engage in low-stakes tests. It could be speculated that, since the test results are unimportant, they cannot see any reason to waste time on taking it.

5.4. The importance for higher education

The current study could help differentiate students who are rather willing to participate in academic work for which there is no immediate gratification from those who are not. It could be argued that more willing students are easier to work with in universities as they have more motivation to study. It is often assumed by

university teaching staff that, since students have voluntarily come to obtain higher education, they are naturally oriented toward high academic achievement. However, based on the results of our research, it is important to note that if the test scores of female and male students are not looked at separately, significant differences could be overlooked. This may lead to biased results in predicting test-taking activity and outcomes in low-stakes tests. It could be speculated that since male students who have lower scores in national examinations are less motivated, they can also potentially be less willing to participate in non-obligatory academic work, i.e. attending lectures or reading additional literature. This could become a problem for them later as they may fall behind on their studies and eventually drop out of university. If higher education institutions were to address this, it could potentially help solve the problem of males having a higher dropout rate.

Furthermore, while being low-stakes for students, the tests might be highly important for institutions conducting the research. Therefore, researchers who use questionnaires, surveys, and tests should consider the possible factors influencing the validity of their studies while using student samples. An important question here is how to motivate students to take low-stakes tests as well as participate in non-obligatory academic work. One way is to raise the stakes by making participation more externally motivated. For example, to give students extra credits for attending lectures or reading additional literature recommended by lecturers.

6. Conclusion

In the current study, we found that the best predictors for test-taking activity appeared to be the results of national examinations and motivation, which is congruent with previous studies. Also, it is important to take into account that when interpreting low-stakes tests, significant differences could be overlooked when genders are considered together.

Finally, we have some ideas about how future studies on the same subject could be improved. Our sample was biased as there were significantly more female students, so we think it would be important to try to reduce this bias in future studies. Also, we had no knowledge about the background of the students, so it could be that the participating students´ level of academic motivation was higher to begin with. Moreover, the students´ emotional and physical conditions and the environment in which the test was taken could influence the results.

Acknowledgements

A Primus grant (#3-8.2/60) from European Social Fund to Anu Realo and by the Estonian Centre of Behavioral and Health Sciences supported the preparation of this article. We are very grateful for the suggestions made by the anonymous reviewers.

Address:
Mari-Liis Mägi
Institute of Psychology
University of Tartu
Näituse 2
50409 Tartu, Estonia
Tel: +372 5556 9610
E-mail: mari-liis.magi@ut.ee

References

Bandura, Albert (1977) *Social learning theory*. New York: General Learning Press.

Brown, Jeanne M. and Carrie Gaxiola (2010) "Why would they try? Motivation and motivating in low-stakes information skills testing". *Journal of Information Literacy* 4, 23–36.

Chemers, Martin M., Li-tze Hu, and Ben F. Garcia (2001) "Academic self-efficacy and first-year college student performance and adjustment". *Journal of Educational Psychology* 93, 55–64.

Cole, James S. and Steven J. Osterlind (2008) "Investigating differences between low- and high-stakes test performance on a general education exam". *The Journal of General Education* 57, 119–130.

Cole, James S. and David A. Bergin (2005) "Association between motivation and general education standardized test scores". Paper presented at the Association for Institutional Research Annual Convention.

Dachlbeck , David T. and Owen Richard Lightsey, Jr. (2008) "Generalized self-efficacy, coping, and self-esteem as predictors of psychological adjustment among children with disabilities or chronic illnesses". *Children's Health Care* 37, 293–315.

Deci, E.L. and R.M. Ryan (1985) *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, Edward L. and Richard M. Ryan (2000) "Intrinsic and extrinsic motivations: classic definitions and new directions". *Contemporary Educational Psychology* 25, 54–67.

Eklöf, Hanna (2007) "Test-taking motivation on low-stakes tests: a Swedish TIMSS 2003 example. Issues and methodologies in large-scale assessments". *International Journal of Testing* 7, 311–326.

Eklöf, Hanna (2010) "Skill and will: test-taking motivation and assessment quality". *Assessment in Education: Principles, Policy and Practice* 17, 345–356.

Huang, Xishan, Ming Liu, and Kunio Shiomi (2007) "An analysis of the relationships between teacher efficacy, teacher self-esteem and orientations to seek help". *Social Behaviour and Personality* 35, 707–716.

Judge, Timothy A., Amir Erez, Joyde E. Bono, and Carl J. Thoresen (2002) "Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct?". *Journal of Personality and Social Psychology* 83, 693–710.

Judge, Timothy A., Edwin A. Locke, and Cathy C. Durham (1998) "Dispositional effects on job and life satisfaction: the role of core evaluations". *Journal of Applied Psychology* 83, 17–34.

Kinzie, Jillian, Robert Gonyea, George D. Kuh, Paul Umbach, Charlie Blaich, and Ali Korkmaz (2007) "The Relationship between gender and student engagement in college". Paper presented at the 32nd annual conference of the Association for the Study of Higher Education, Louisville, KY.

Lau, Abigail R. and Dena A. Pastor (2007) "A hierarchical linear model of variability in test motivation across students and within students across tests". Paper presented at the October 2007 meeting of the Northeastern Educational Research Association.

Lynn, Richard and Jaan Mikk (2008) "Science: sex differences in attainment". *The Journal of Social, Political and Economic Studies* 33, 101–124.

Maccio, Elaine M. and Jamie T. Schuler (2012) "Substance use, self-esteem, and self-efficacy among homeless and runaway youth in New Orleans". *Child Adolescent Social Work Journal* 29, 123–136.

Marsh, Herbert W. and Kit-Tai Hau (2003) "Big-fish-little-pond effect on academic self-concept: a cross-cultural (26-country) test of the negative effects of academically selective schools". *American Psychologist* 58, 364–376.

Mikk, Jaan, Karin Täht, and Olev Must (2011) "Sex differences in educational attainment". *Personality and Individual Differences* 53*,* 132–136.

Must, Olev and Jüri Allik (2001) "Tunne oma võimeid. Abivahend eneseanalüüsiks". [Know your abilities: a tool for self-analysis.] Tartu: Tartu University Press.

Mägi, Mari-Liis (2012) "Tartu Ülikooli bakalaureuseastme üliõpilaste akadeemilise motivatsiooni võrdlus kahe teaduskonna näitel". [The comparison of bachelor students´ motivation in the example of two faculties in the University of Tartu.] Seminar paper. University of Tartu.

O'Neil, Harold F., Jamal Abedi, Charlotte Lee, Judy Miyoshi, and Ann Mastergeorge (2005) "Monetary incentives for low-stakes tests". *Educational Assessment* 3, 185–208.

Phan, Huy Phuong (2010) "Students' academic performance and various cognitive processes of learning: an integrative framework and empirical analysis". *Educational Psychology* 30, 297–322.

Prat-Sala*,* Merce and Paul Redford (2010) "The interplay between motivation*,* self-efficacy and approaches to studying". *British Journal of Educational Psychology* 80*,* 283–305.

Pullmann, Helle and Jüri Allik (2000) "The Rosenberg self-esteem scale: its dimensionality, stability and personality correlates in Estonian". *Personality and Individual Differences* 28, 701–715.

Pullmann, Helle and Jüri Allik (2008) "Relations of academic general self-esteem to school achievement". *Personality and Individual Difference* 45, 559–564.

Rimm, Halliki and Matthias Jerusalem (1999). "Adaptation and validation of an Estonian version of the general self-efficacy scale". *Anxiety, Stress and Coping* 12, 329–345.

Rosenberg, Morris (1965) *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rosenberg, Morris, Carmi Schooler, Carrie Schoenbach, and Florence Rosenberg (1995) "Global self-esteem and specific self-esteem: different concepts, different outcomes". *American Sociological Review* 60, 141–156.

Schnipke, Dale L. (1995) *Assessing speededness in computer-based tests using item response times*. (Doctoral dissertation, Johns Hopkins University, 1995.) Dissertation Abstracts International, 57, B759.

Segal, Carmit (2012) "Working when no one is watching: motivation, test scores, and economic success". *Management Science* 58, 1438–1457.

Spinath, C., H.H. Feudenthaler, and A.C. Neubauer (2010) "Domain-specific school achievement in boys and girls as predicted by untelligence, personality and motivation". *Personality and Individual Differences* 48, 481–486.

Stover, Juliana Beatriz, Guadalupe de la Iglesia, Antonio Rial Boubeta, and Mercedes Fernández Liporace (2012) "Academic motivation scale: adaptation and psychometric analyses for high school and college students". *Psychology Research and Behavior Management* 5, 71–83.

Täht, Karin and Olev Must (2009) "Relationship between the educational performance and attitudes of Estonian students. In Jaan Mikk, ed. *Teenagers in Estonia: values and behaviour*, 45–63. Frankfurt am Main: Peter Lang.

Tõnisson, Eve (2011) "Kõrghariduse valdkonna statistiline ülevaade". [Statistical survey of higher education institutions in Estonia]". Available online at <http://www.hm.ee/index.php?048183>. Accessed on 24.09.2013.

Unt, Marge, Karin Täht, Ellu Saar, and Jelena Helemäe (2013) "The expansion of higher education: devaluation or differentation? The Estonian case". In *Higher education at a crossroad: the case of Estonia,* 367–391. Ellu Saar and Rene Mõttus, ed. Frankfurt am Main: Peter Lang.

Vallerand, Robert J., Luc G. Pelletier, Marc R. Blais, Nathalie M. Brière, Caroline B. Senécal, and Évelyne F. Vallières (1992) "The academic motivation scale: a measure of intrinsic,

extrinsic, and amotivation in education". *Educational and Psychological Measurement* 52, 1003–1017.

Vallerand, Robert J., Luc G. Pelletier, Marc R. Blais, Nathalie M. Brière, Caroline B. Senécal, and Évelyne F. Vallières (1993) "On the assessment of intrinsic, extrinsic and amotivation in education: evidence on the concurrent and construct validity of the academic motivation scale". Educational and *Psychological Measurement* 53, 159–172.

Wise, Steven L. and Christine E. DeMars (2005) "Low examinee effort in low-stakes assessment: problems and potential solutions". *Educational Assessment* 10, 1–17.

Wise, Steven L. and Christine E. DeMars (2004) "An investigation of motivation filtering in a statewide achievement testing program". Paper presented at the annual meeting of the National Council on Measurement in Education, San Diego.

Üpraus, Reili (2009) "Akadeemiline edukus: Akadeemilise enese-tõhususe mõõtmine". [Academic achievement: measurement of academic self-efficacy.] Seminar paper. University of Tartu.