The Gender Differences on Physics National Examination (NE) Performance in East Java During the Last 3 Years (2016-2018)

Madlazim[1], Siti Jamiatul Husnaini[2]
[1] Physics Education Department, State University of Surabaya
[2] Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology

Abstract

Physics is a one of the compulsory subject tested in the national examination (NE) for senior high school level. However, since 2017 physics becomes an optional so as not the whole national exam participants took physics for passing a secondary education. Physics its self requires both numeracy skills and analysis and interpretation ability to understand the physical phenomenon. These higher order cognitive skills lead a challenge between genders in particular nature science students. This study explores the genders differences on physics national exam performance in East Java during the last 3 years (2016-2018). A total of 57.561 students (62.67% females) in 2016, 19.441 (66.14% females) in 2017, and 13.143 (50.61% females) were measured. This study revealed that male students outperformed than females students in all tested indicators during the last 3 years. A stereotype threat and computer self-efficacy were discussed along with the suggestion further work was presented.

Keywords:

A. National exam and Natural Sciences

National exam (NE) is held every year by Indonesia ministry of education and culture for students who are going to the next education level. This exam has the important role to map an educational system quality (Retnawati, Arlinwibowo, Wulandari, & Pradani, 2018). The results then are considered by educators, policy maker, and government to formulate the succeeding curriculum. Brookhart and Nitko (2014) stated that the evaluation processes is an attempt to get information in making student-related decisions, learning, curriculum, and educational policies. Therefore, the result of NE can be used to depict the lack and the inherent problem of the existing Indonesia curriculum. According to BSNP (2018), NE results have the passing grade categories: The students who have 85 <NE≤ 100 is very good or (A), 70 <NE≤ 85 is "good" or (B), 55 <NE≤ 70 is "sufficient" or (C), and 0 ≤NE≤ 55 is "less" or (D). Based on this classification, the government expected the students at least to graduate with C category.

In addition to the policy of the passing grade, the NE technical implementation also changes. Since 2015, the NE has utilized computers or known as UNBK. The UNBK makes NE results more reliable because it has a high integrity index. The simulation conducted by Arief and Suryani (2016) revealed that the fraudulent level of UNBK has decreased by an average of 12% each year. In other words, the national integrity index has increased. With a high integrity index, UNBK may provide more accurate information about students' performance. Therefore, the difficulties faced by students

can be mapped properly and the curriculum improvement can be adjusted according to the needs.

This study aims to explore the gender differences on physics national exam performance for secondary school level in East Java. According to Government Regulation of the Republic of Indonesia number 17 of 2010 article 79, secondary school level has some fields to accomodate learning and competencies needed by the students to continue to the higher education level. Such fields are natural sciences, social sciences, and language. Regarding to natural sciences, it examines natural phenomena, including processes (scientific methods), scientific attitudes, products (facts, concepts, theories, laws) and their applications. The 2004 curriculum defined natural science is a systematic way of finding out the nature, facts, concepts, principles, and processes of discovery and having a scientific attitude. Since elementary schools, natural science curriculum is set to students to learn about themselves and their environment (Ministry of National Education, 2015). Physics is a part of the natural sciences. Iskandar (1997) explained that physics is the knowledge about: (1) observing what happened, (2) understanding what is observed, (3) using new knowledge to predict what happened, and (4) testing the forecast under conditions to see whether it is correct. Based on the statements above, it can be concluded that physics is a process of activities to study nature through scientific work to understand the concepts, principles, laws, and scientific attitudes so that it will be useful for everyday life. At the secondary education level, physics NE covered measurement and kinematics (M1), dynamics (M2), work, energy, and collisions (M3), heat (M4), waves and optics (M5) electricity, magnetism, and nucleonics (M6).

B. Participants

The participants of this study were students who took the physics NE during 2016-2018 in East Java (38 regencies). A total of 57.561 students (36.074 females, 21.487 males) in 2016, 19.441 (12.858 females, 6.583 males) in 2017, and 13.143 (6.652 females, 6.491 males) were measured to explore the gender differences.

C. Method

Physics NE data was collected through the PAMER application released annually by the Assessment Center, Ministry of Education and Culture. The application provides information about the NE results, including the average grade and the grade of each indicators, hence, it can be seen which indicators must be paid attention more closely by educators and related parties.

D. Results

1. Gender differences on physics NE performance

Figure 1 presented the gender differences on physics NE performance during the last 3 years (2016-2018) in 38 regencies of East Java. In 2016, 2017, and 2018 physics NE, the male performance (63.01; 50.81; and 48.17, respectively) was higher than those of female (52.61; 41.92; and 45.00, respectively). The average of the last 3 years also revealed that male (54.00) performed better than those of female (46.95). Furthermore, male performance decreased from 2016 to 2018, while female performance decreased from 2016 to 2017 and increased from 2017 to 2018.

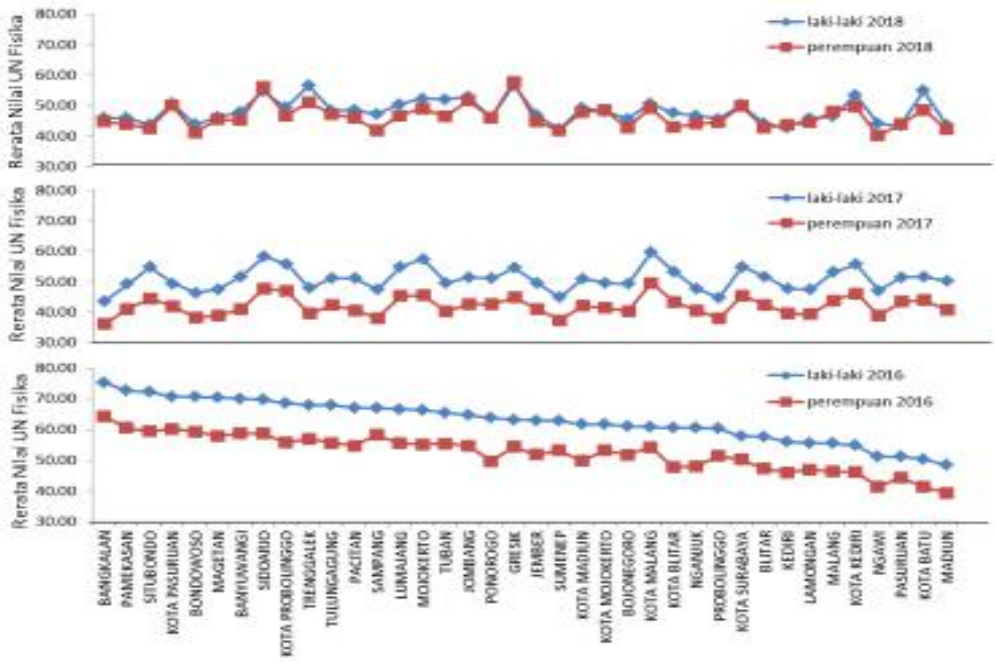

Figure 1. The gender differences on physics NE performance during the last 3 years (2016-2018)

It showed that most regencies have a similar trend. Gender differences may decrease from 2016 to 2018 due to a new policy making physics as an optional examined subjects. In 2016, when physics was a compulsory for all the NE participants, female performance was lower than those of male in all regencies. However, female tended to perform as well as male did since 2017. NE participants could opt physics rather than other natural sciences subjects, such as chemistry and biology, because of perceiving of more self-efficacious and interest. Self-efficacy has been reported as a strong predictor of academic performance (Britner & Pajares, 2006). Students with high self-efficacy are more likely to work harder at a task, to see the difficulty as a challenge (Ketelhut, 2007), and to successfully achieve learning goals (Britner & Pajares, 2006 & Wicaksono et. al., 22017) over less self-efficacious students. Moreover, a longitudinal study found that there is a dynamic and reciprocal relationship between interest and performance over time (Harackiewicz, Durik, Barron, Linnenbrink-Garcia, & Tauer, 2008). It indicated that students' interest and self-efficacy may affects their performance.

In addition, there is a significant high correlation between male and female performance, $r = .95$, $p < .001$ (Table 1). It means that female performance increases when male performance increases, and vice versa. (Fig. 1). A two-way test was conducted to measure the significant difference between genders for 3 years exam (the average of 2016, 2017, and 2018 physics NE result). The result revealed there is a significant gender difference on physics NE performance, $t = 41.26$, $p < .001$. It rejected H_0, where H_0: $\mu_1 = \mu_2$ and accepted H_1: $\mu_1 \neq \mu_2$. Gender gap is still an unclear issue among researchers. Some studies indicated females do not surpass males in general intelligence (Halpern, 2000), while other studies found that females performs better than males at school (Deary, Strand, Smith, & Fernandes, 2007; Downey &

Vogt Yuan, 2005; Ellis et al. 2008). In the current study, the gender gap was found on physics NE performance. A potential explanation for this persistent gap is that stereotype threat, the fear of confirming a stereotype about one self, is inhibiting females' performance. Kost-Smith et al. (2010) testified in their work that female are more worried and nervous (showing a stereotype threat) in taking exam.

The stereotype threat means a fear of doing something that will accidentally confirm such stereotype (Steele, 1997). This stereotypical threats could have a negative impact on members of stereotyped groups and result in worse outcomes. Spencer, Steele, and Quinn (1999) tested the effects of stereotype threats by looking at the females' achievement on difficult mathematical tests and found that females performed worse than male after telling them a stereotypical threats on math tests between genders. However, after the researchers presented that the female students have a chance to do as well as male did with an effort and a hard work, females' performance increased. Previous studies confirmed that this threats also can be reduced through self-affirmation (Cohen, Garcia, Apfel, & Master, 2006; Martens, Johns, Greenberg, & Schimel, 2006). Self-affirmation is a process in which a person asserts his overall sense of self-worth and integrity (Steele, 1997). Self- affirmation assumes that people are motivated to maintain a positive sense of overall their integrity, identity and value. When someone's integrity or identity is threatened, someone looks for ways to deal with such threat. Martens et al. (2006) found that female given the opportunity to write their self-affirmation 15 minutes before the science test showed the similar results to male's achievement.

Moreover, the use of computer-based NE may affect the gender's performance. Imhof, Vollmeyer, and Beierlein (2007) stated that male students outperformed female students at a computer-based task. For the domain of computer use, it has been shown repeatedly that higher levels of computer self-efficacy are correlated with higher levels of computer use, more efficient user strategies, more positive affect when using computers, and lower levels of computer anxiety (Dickha¨user & Stiensmeier-Pelster, 2002, 2003; Shapka & Ferrari, 2003). Since 2015, computer-based NE was implemented, so it may be one of the gender gap's reason. The further work should be conducted to verify it.

Table 1. Correlation and comparison of male and female performance

	Male	*Female*
	53.99956	46.94877
Mean	14	193
	10.67350	8.887927
Variance	604	83
Observations (n)	38	38
	0.947224	
Pearson Correlation	05	
Hypothesized Mean Difference	0	
df	37	
	41.25924	
t Stat	37	

	7.63818
P(T<=t) one-tail	E-33
	1.687093
t Critical one-tail	62
	1.52764
P(T<=t) two-tail	E-32
	2.026192
t Critical two-tail	463

2. Genders differences of each tested indicator

This section explained the results of male and female's performance on each indicator for 3 years in 38 regencies of East Java, that is kinematics (M1), dynamics (M2), work, energy, and collisions (M3), heat (M4), waves and optics (M5) electricity, magnetism, and nucleonics (M6). Figure 3 and 4 explained the male and female performance, respectively.

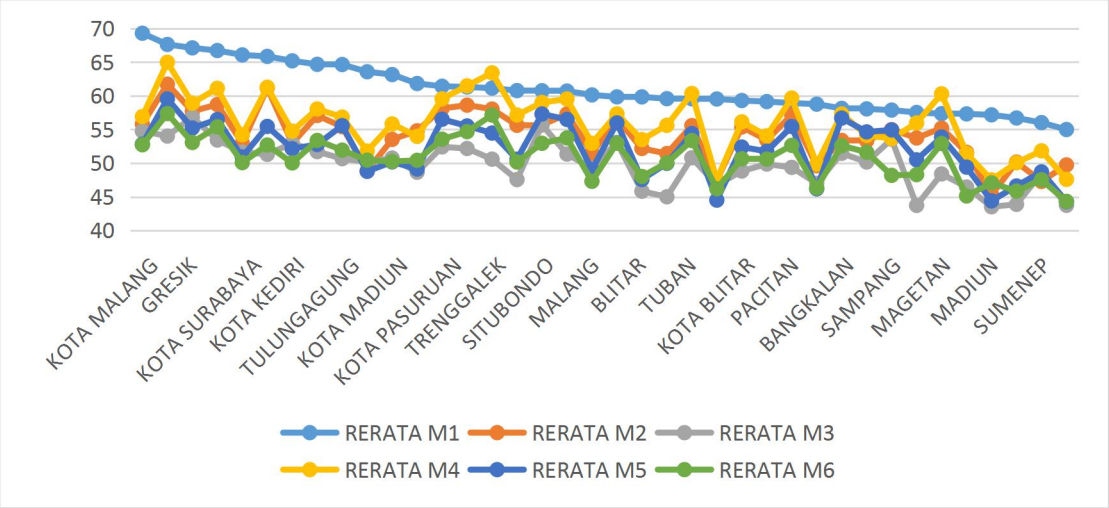

Figure 3. Male's physics NE performance in all regencies, East Java.

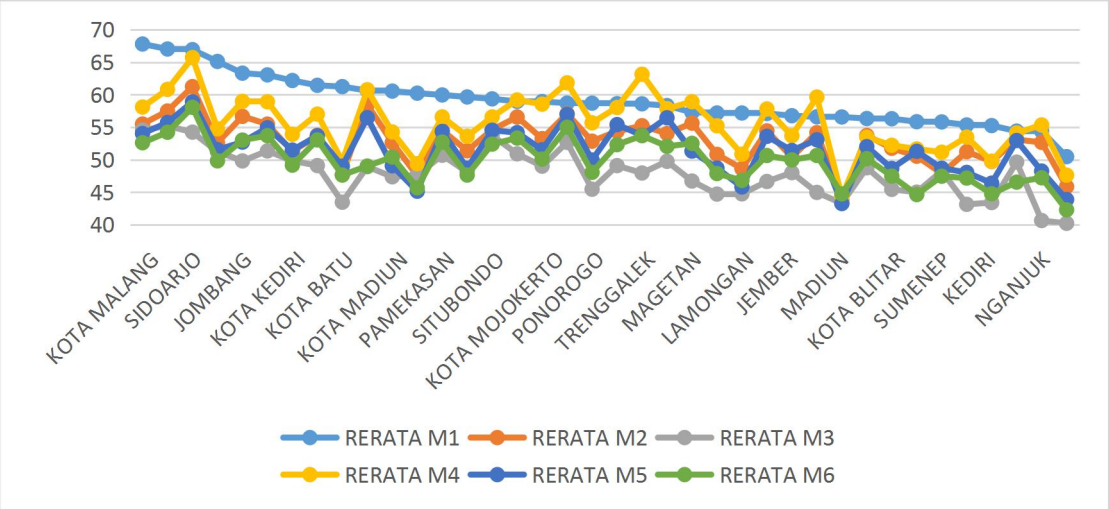

Figure 4. Female's physics NE performance in all regencies, East Java.

Figures 3 and 4 showed that male students (61.00; 54.19; 55.94; 52.14; 50.83, respectively) performed better than female (59.00; 49.90; 47.99; 55.59; 51.47; and 49.86, respectively) on M1, M2, M3, M4, M5, and M6, respectively in 38 regencies, East Java (Madlazim et al, 2015). Among regencies, Sidoarjo and Trenggalek had performance average better for all materials than other regencies, while Ngawi and Batu City are the lowest.

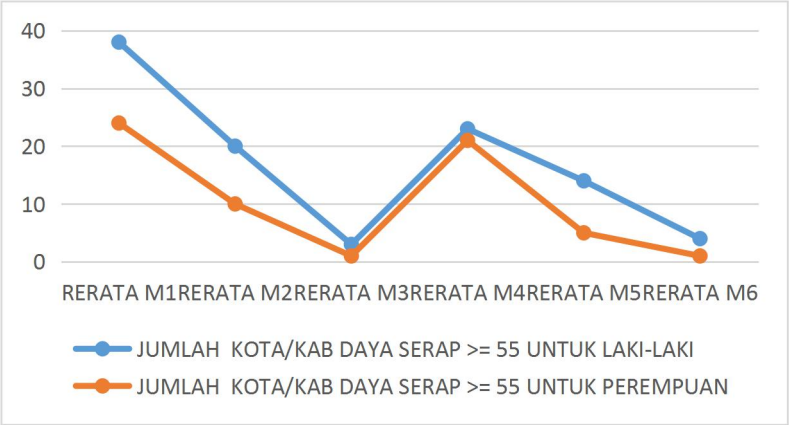

Figure 5. A number of regencies gaining sufficient or (C) category, 55 <NE≤ 70.

Figure 5 showed that most students are capable in the kinematics (M1) than other indicators. Moreover, indicator of work, energy, and collisions (M3) and electricity, magnetism, and nucleonics (M6) appear to be the most difficult topics among students. Furthermore, according to figure 5, gender differences also occurred even after mapping the result into each indicator. It indicated that the stereotype threat is in all tested physics competencies.

E. Conclusion and recommendation

Based on the analysis of physics NE performance during the last 3 years (2016-2018) in East Java, male students consistently achieved better than female students. It may indicate the existing of stereotype threat among genders. Moreover, computer self-efficacy and interest may influence the gender performance as well. However, this result should be used to guide further research that follows more controlled research design meant to test such causal hypotheses.

Self-affirmation is one of the solution to reduce stereotype threat. It is proven by previous studies such as (Cohen et al., 2006; Martens et al., 2006). The empirical studies to testify this issue in the different educational culture is a promising further research. Moreover, the use of computer-based learning should be popularized among students in term of preparing them in national exam. The computer self-efficacy becomes an own challenge for the students. Performance is positively correlated with self-efficacy, $r = .38$ (Multon, Brown, & Lent, 1991). Male students tend to show higher levels of self-efficacy compared to their female counterparts (Hinz, Schumacher, Albani, Schmid, & Brahler, 2006), even in the computer literacy (Durndell, Siann, & Glissov, 1995; Mitra, Lenzmeier, Avon, Qu, & Hazen, 2000). By familiarizing females with computer-based tests/ tasks may certainly increase their computer self-efficacy. The educational experts, practitioners, and instructors are expected to pay attention to the

issue alike this to solve the gender gap's issue. Future curriculum should prepare all Indonesian students to the industry 4.0 era.

Acknowledgments

Thank you to the Ministry of Education and Culture of the Republic of Indonesia for providing data and allowing it to be used in this study.

References

Arief, R., & Suryani, E. (2016). Sistem Dinamik Ujian Nasional Berbasis Komputer Untuk Meminimalkan Resiko Kecurangan Serta Meningkatkan Efektifitas Dan Efisiensi Anggaran. *INTEGER: Journal of Information Technology*, *1*(2).

Britner, S. L., & Pajares, F. (2006). Sources of science self‐efficacy beliefs of middle school students. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *43*(5), 485–499.

Brookhart, S. M., & Nitko, A. J. (2014). *Educational assessment of students*. New York, NY: Pearson Education Inc.

BSNP. 2018. Buku Saku UN. Retrieved from https://bsnp-indonesia.org/wp-content/uploads/2018/03/Buku-Saku-UN-6-Des-2017-Edit-21-Maret-2018.pdf.

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*(5791), 1307–1310.

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21.

Dickha¨user, O., & Stiensmeier-Pelster, J. (2002). Gender differences in computer work: evidence for the model of achievement-related choices. *Contemporary Educational Psychology, 27*, 486–496.

Dickha¨user, O., & Stiensmeier-Pelster, J. (2003). Gender differences in the choice of computer courses: applying an expectancy-value model. *Social Psychology of Education, 6*, 173–189.

Downey, D. B., & Vogt Yuan, A. S. (2005). Sex differences in school performance during high school: Puzzling patterns and possible explanations. *The Sociological Quarterly*, *46*(2), 299–321.

Durndell, A., Siann, G., & Glissov, P. (1995). Gender and computing: Persisting differences. *Educational Research, 37*, 219–227.

Ellis, L., Hershberger, S., Field, E., Wersinger, S., Pellis, S., Geary, D., et al. (2008). *Sex differences: Summarizing more than a century of scientific research*. New York: Psychology Press.

Halpern, D. F. (2000). *Sex differences in cognitive abilities (3rd Ed.)*. Mahwah: Erlbaum.

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal

relations between achievement goals, interest, and performance. *Journal of Educational Psychology, 100*(1), 105–122.

Hinz, A., Schumacher, J., Albani, C., Schmid, G., & Brähler, E. (2006). Bevölkerungsrepräsentative Normier-ung der Skala zur Allgemeinen Selbstwirksamkeitserwartung. [Standardization of the General Self-Efficacy Scale in the German population]. *Diagnostica, 52*(1), 26–32.

Imhof, M., Vollmeyer, R., & Beierlein, C. (2007). Computer use and the gender gap: The issue of access, use, motivation, and performance. *Computers in Human Behavior, 23*(6), 2823–2837.

Ketelhut, D. J. (2007). The impact of student self-efficacy on scientific inquiry skills: An exploratory investigation in River City, a multi-user virtual environment. *Journal of Science Education and Technology*, *16*(1), 99–111.

Kost‐Smith, L. E., Pollock, S. J., Finkelstein, N. D., Cohen, G. L., Ito, T. A., & Miyake, A. (2010, October). Gender differences in physics 1: The impact of a self‐affirmation intervention. In AIP Conference Proceedings (Vol. 1289, No. 1, pp. 197–200). AIP.

Madlazim, Supriyono and Mukhayyarotin Niswati Rodliyatul Jauhariyah (2015). Student's scientific abilities improvement by using guided inquiry laboratory*, Journal of Science Education*, 16(2):58-62.

Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology, 42*(2), 236–243.

Mitra, A., Lenzmeier, S., Avon, R., Qu, N., & Hazen, M. (2000). Gender and computer-use in an academic institution: Report from a longitudinal study. Journal of Educational *Computing Research, 23*(1), 67-84.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30–38.

Retnawati, H., Arlinwibowo, J., Wulandari, N. F., and Pradani, R. G. 2018. Teachers' Difficulties and Strategies in Physics Teaching and Learning That Applying Mathematics. *Journal of Baltic Science Education*, *17*(1), 120–135.

Shapka, J. D., & Ferrari, M. (2003). Computer-related attitudes and actions of teacher candidates. *Computers in Human Behavior, 19*, 319–334.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4–28.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613–629.

Wicaksono, I., Wasis, Madlazim (2017). The effectiveness of virtual science teaching model (VS-TM) to improve student's scientific creativity and concept mastery on senior high school physics subject. *Journal of Baltic Science Education*, 16(4), 549-561.