

ROBUST REGRESSION FOR ESTIMATING THE IMPACT OF STUDENT'S SOCIAL BEHAVIORS ON SCIENTIFIC LITERACY

Umi Mahmudah^{1*}, Muhamad Chamdani², Tarmidzi¹, Siti Fatimah²

¹State Islamic Institute of Pekalongan, Indonesia

²Sebelas Maret University, Indonesia

*e-mail: umi.mahmudah@iainpekalongan.ac.id

Abstract: Scientific literacy is a vital program for current generation of students in the whole world. The purpose of this study is to estimate the impact of students' social behaviors on scientific literacy by using robust regression. A robust approach is applied to overcome the weaknesses of the ordinary least squares (OLS) regression model. The method is believed to produce better accuracy of estimation. This study applied a survey method with a total of 200 respondents who participated in this study. The respondents were randomly selected at a university in Indonesia. There were two instruments used, namely questionnaire sheets using a Likert scale and test sheets. There were four dimensions used in this study to measure the students' social behaviors, which are cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4). The analysis results revealed that all of independent variables had positive and significant impacts on students' scientific literacy (Y). Furthermore, the highest impact was collaboration (X_1) and the lowest impact was pro-social behavior towards peers (X_4), where the robust regression coefficients were .262 and .238, respectively.

Keywords: *robust regression, scientific literacy, social behaviors*

ROBUST REGRESSION UNTUK MENGESTIMASI DAMPAK PERILAKU SOSIAL MAHASISWA TERHADAP LITERASI SAINS

Abstrak: Literasi ilmiah merupakan program yang vital bagi mahasiswa generasi sekarang di seluruh dunia. Tujuan dari penelitian ini adalah untuk mengestimasi dampak perilaku sosial mahasiswa terhadap literasi ilmiah dengan menggunakan robust regression. Pendekatan robust diterapkan untuk mengatasi kelemahan-kelemahan dari model regresi OLS. Metode ini diyakini menghasilkan akurasi estimasi yang lebih baik. Penelitian ini menerapkan metode survei yang dilakukan dengan menggunakan subjek penelitian sebanyak 200 mahasiswa sarjana yang dipilih secara acak di sebuah universitas di Indonesia. Instrumen yang digunakan ada dua, yaitu lembar angket dengan menggunakan skala likert dan lembar tes. Ada empat kategori yang digunakan dalam penelitian ini untuk mengukur perilaku sosial siswa, yaitu kerja sama (X_1), ketegasan (X_2), kontrol diri (X_3) dan perilaku prososial terhadap teman sebaya (X_4). Hasil analisis mengungkapkan bahwa semua variabel independen memiliki dampak positif dan signifikan terhadap literasi sains siswa (Y). Selanjutnya, dampak tertinggi adalah kerjasama (X_1) dan terendah adalah perilaku prososial terhadap teman sebaya (X_4), di mana koefisien robust regression masing-masing adalah 0,262 dan 0,238.

Kata Kunci: *robust regression, literasi sains, perilaku sosial*

INTRODUCTION

Scientific literacy is a very fundamental in increasing human resources, who are able to compete globally. Today is an era where everything is based on science and technology. Their utilization is very massive therefore anyone, whether an individual, a community or a country that is unable to keep up with the development of science and technology will

be left behind. Scientific literacy is a model of developing knowledge and competencies that must be possessed at the same time mastered by the learners. This program has become very well known since it was raised by the Organization for Economic Co-operation and Development (OECD) as the main focus in the 2006 of the Programme for International Student Assessment (PISA), which aims to evaluate the education

system of countries. The definition of scientific literacy based on the OECD (2003) is the ability to use scientific knowledge, to identify questions or problems, and to describe conclusions based on data or evidences.

National Science Education Standards defines the scientific literacy as an understanding of both scientific content and practices. Then, this also includes the ability to use that knowledge to participate in decision-making that may affect the global community (NRC, 1996). Scientific literacy provides convenience in understanding phenomena to be able to help natural decision-making that may change human activities in the future (OECD, 2003). Simply put, good scientific literacy can be known when someone is able to observe, to evaluate, and to interpret scientific knowledge. DeBoer (2000) states that scientific literacy reflects a broad understanding of science for educational purposes, while Bybee, McCrae, & Barry (2009) state it is the ultimate goal of science education. Thus, learning science given to students should be able to produce students who are literate in science.

Based on the results of preliminary observations during science learning, the majority of students (around 95%) stated that science was difficult while students who stated science was easy had a very alarming number, which was only 1%. Then, the preliminary observation also indicate that more than 90% of students are afraid of science materials, even before they know the contents of these materials. Students tend to face difficulties in understanding science concepts and completing assignments given by lecturers, especially questions that require a high level of understanding. These problems have an impact on students' ability to think creatively and their critical thinking skills. Having the ability to think creatively and critically is one of the abilities that must be possessed by students in the 21st century. Thus, Indonesian students can compete internationally.

Implementing scientific literacy in the real world is very important. Scientific literacy may produce scientific literate thinkers to make wise choices based on data and scientific information. Ultimately, scientific literacy can be used to resist racism, sexism, bigotry and social injustice (Maienschein, 1998). Scientifically literate people have skeptical and creative mind habits that are valuable in promoting scientific decision

making. Thus, they tend to be immune to hoaxes because of their ability to properly digest information even though this comes rapidly.

The Indonesian government and practitioners of science education have given good attention to the importance of scientific literacy for students in Indonesia even though the term of scientific literacy is not explicitly stated in the latest education curriculum. The content of core and basic competencies shows that the development of students' scientific literacy is one of the goals of science education in junior high schools.

Supahar, Rosana, Ramadani, & Dewi (2017) report that science students are required to have skills in observing by using appropriate equipment, carry out experiments based on procedures, recording the results of observations and measurements, making conclusions as well as communicating them verbally and in writing according to the facts and evidences.

It is a real effort from the government in order to improve the level of Indonesian literacy which is categorized as very low. It can be seen very clearly from the assessment results of science learning outcomes at the international level conducted by the OECD on the Program for International Student Assessment (PISA). In fact, Indonesia has always been ranked below, with point's scores always below the OECD average score.

The results of the 2018 PISA survey released by the OECD (2019) also show that students in Indonesia have lower scores than the OECD average scores on the three basic competencies, namely reading, mathematics and science. Further, the report indicates that Indonesian literacy levels on all three basic competencies have fluctuated. Overall, the results of Indonesia's achievements based on the PISA survey are always below the OECD average. The PISA results from 2006 to 2018 show an increasing trend in science scores. Then, there is an increase of 21 points in the 2012 to 2015 of PISA results, which are 382 points to 403 points. However, there is a decrease of 7 points, from 403 in 2015 to 396 in 2018.

The purpose of this study is to analyze the social behavior of students towards scientific literacy. It is important to note that scientific literacy competencies established by the OECD (2003) are used as a reference to measure student

literacy levels. There are three competencies, which are the ability to observe, the ability to evaluate, and the ability to interpret scientific knowledge. The first competency includes three basic abilities; (1) the ability to explain various scientific phenomena related to the development of science and technology, (2) the ability to see natural phenomena, and (3) the ability to describe clearly and systematically. The second competency is the ability to evaluate, which has good knowledge and ability to answer all scientific information asked by users. The last competence is the ability of interpretation, namely understanding scientific knowledge through data sources and scientific evidence. This competency requires accuracy and certainty in explaining the material when interpreting data from tables, diagrams, graphs, prototypes, and other types of data.

There are studies that are conducted in analyzing the relationship between student's behavioral patterns and academic performance. Most of studies are commonly suggested that there is positive and significant relationship between academic and social behaviors and academic performance (see Lechner, Danner, & Rammstedt, 2017; Farrington, Roderick, Allensworth, Nagaoka, Keyes & Johnson, 2012; Borghans, Golsteyn, Heckman, & Humphries, 2011; Truesdell & Abramson, 1992).

Macan, Shahani, Dipboye, & Phillips (1990) associate the academic performance and time management by using as many as 160 students. The result reveals that students who have full control over their time have a greater evaluation of their performance, better work and life satisfaction. Then, the r result also reports a positive and significant correlation between time management and Grade Point Average (GPA). Gašević, Zouaq, & Janzen. (2013) found there is positive and significant relationship between academic performance and social ties, where the study uses as many as 505 students at a university in Canada. Farrington et al. (2012) report that a social behavior is one of the non-cognitive factors that influence student success while deVries (2018) states that social behavior is related to both math and reading grades.

This study uses robust approach on regression model to produce higher accuracy of estimation results. It is well-known that the traditional regression model, which applies

ordinary least square (OLS) method, tends to obtain misleading results due to its sensitivity of the existence of outliers (Fox & Weisberg, 2011). It is important to note that the outliers are extreme values in data set which yield a less accurate estimation result (Zimmerman, 1998; Fox, 2002).

A robust regression is a powerful tool in order to deal with the outliers' problem therefore the method is widely applied in various fields of studies. Most researchers suggest that the method is very useful as well as fundamental in data analysis. The most interesting thing about robust regression is that there is no need for assumption tests that should be done when using the OLS method. This is because the robust approach is a nonparametric method based on sample replication. This method is very useful to restrain the impact of outliers in regression analysis (Chen, 2007). In addition, this method is also powerful for detecting the presence of outliers and at the same time providing estimation results that are resistant to its existence (Alma, 2011).

In order to get the empirical results, this study uses a questionnaire that is distributed to as many as 200 students who are randomly selected. The statistical data is analyzed by using R program. Even though there are studies that analyze the impact of students' social behavior on academic performance. However, its effect on student literacy levels is still very limited, especially on Indonesian undergraduate students. Therefore, this research becomes very important in order to provide an overview and reference for the development of scientific literacy levels of students, especially in Indonesia. As mentioned earlier, the facts show that Indonesian students' scientific literacy is very low. It is necessary to find a real solution immediately to improve the level of student scientific literacy. The purpose of this study is to estimate the impact of students' social behaviors on scientific literacy by using robust regression. This study is expected to help the Indonesian government and education practitioners for this purpose by considering the aspect of social behaviors of students.

METHODS

This study estimates the relationship between students' social behaviors and scientific literacy by using robust regression. Many researchers believe that this method is able

to produce better results than the traditional regression, namely ordinary least square (OLS) method. This study applies a survey method conducted by using 200 research subjects which are randomly selected at a university in Indonesia. There are two instruments used, namely questionnaire sheets using a Likert scale and test sheets. The questionnaire is used to measure the variables of students' social behavior. Meanwhile, scientific literacy variable is measured using 15 questions in the form of written assignments, where these questions refer to the indicators used by the OECD. The questions include three scientific literacy competencies, namely the ability to observe, the ability to evaluate, and the ability to interpret scientific knowledge. It is important to point out that the level of validity and reliability of the instruments used in this study shows good categories. Based on the results of the analysis using SPSS shows that all question items have a significance value below .05, which means that they are statistically valid. Meanwhile, the reliability test using Alpha Cronbach yields a value of .82 for instruments on social behavior and its value is .78 for the instruments on the test sheet. This study uses a quantitative approach to describe the results of the analysis by using robust regression. Statistical analyses are performed by using R program applying a significance level of .05.

For the purpose of this study, data collection is carried out on randomly selected students at State Islamic Institute of Pekalongan University in Indonesia. The dependent variable (Y) used in this study is the level of scientific literacy of students. As mentioned earlier, to measure the level of scientific literacy, this study uses three competencies following the provisions of the OECD (2003). Meanwhile, independent variable in this study is social behavior (X), which is consisted of 4 (four) categories, namely cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4). These four independent variables follow the indicators of social behavior used by El Mallah (2014) in estimating the relationship between the forms of social behavior and aggression to forecast academic outcomes.

OLS regression is the most common method in analyzing the impact of independent variables (X) on dependent variable (Y).

However, it is well-know that this method tends to produce misleading. Alma (2011) state that OLS method should not be applied to regression analysis when the data set is contaminated by the outliers because it may produce inaccurate as well as misleading results. Other than that, OLS method is not appropriate to use if the data set is not normally distributed.

Many researchers believe that robust regression is one of the most powerful tools to overcome the weaknesses of the OLS method (see Chen, 2007; Fox & Weisber, 2011; Alma, 2011). Therefore, this method is believed to yield the estimation results with high accuracy. Basically, robust regression is an iterative procedure for analyzing the coefficient estimates of regression by identifying outliers in the data set as well as minimizing their impacts to the estimation results. This study applies robust regression to obtain the empirical results of the impact of students' social behaviors (X) on scientific literacy (Y).

There are several estimation methods of robust regression, namely M-estimation (maximum likelihood), MM-estimation (method of moment), S-estimation, LTS-estimation (least trimmed square) and LMS-estimation (least mean square). However, this study uses MM-estimator to provide the empirical results of relationship between the students' social behaviors and scientific literacy. This estimator is a special type of M-estimator of robust regression. The most important key in MM-estimation is using S-Estimator of robust regression to estimate the regression parameters. Further, M-estimator is used to determine the regression parameters by minimizing the scale of its residuals. MM-Estimation is defined as follows:

$$\tilde{\beta}_M = \arg \min_{\beta} \sum_{i=1}^n \rho \left(\frac{e_i}{s} \right)$$

Where $\rho(\cdot)$ is objective function, e_i is the residual, s is the scale of the residuals form M-estimation of robust regression. Meanwhile, the general forms of both the M-Estimation and S-Estimation methods are defined as follows:

$$\tilde{\beta}_M = \min_{\beta} \sum_{i=1}^n \rho \left(\frac{y_i - x_i' \tilde{\beta}}{s} \right)$$

and

$$\tilde{\beta}_S = \arg \min_{\beta} s \sum_{i=1}^n (e_1, e_2, \dots, e_k)$$

Where y_i is dependent variable, x_i is independent variable, and β is coefficient regression. Generally, there are six steps of MM-estimator are as follows (Fox & Weisberg, 2011):

Step 1: calculating the residual values of S-Estimation by using $e_i = y_i - \hat{y}_i$.

Step 2: calculating the scale values of S-Estimation by using e_i from step 1 and then calculating the initial weight by using the following formula.

$$w_i = \begin{cases} \left[1 - \left(\frac{e_i}{4.685} \right)^2 \right]^2 & \text{if } |e_i| \leq 4.685 \\ 0 & \text{if } |e_i| > 4.685 \end{cases}$$

Step 3: calculating the coefficient of regression by using e_i with the scale of S-Estimation as the initial iteration of weight least square estimation.

Step 4: calculating the values of new weight by using the scale estimation of the initial iteration of weight least square.

Step 5: calculating the coefficient regression by using $\hat{\beta}^{(m+1)} = (X' W^m X)^{-1} X' W^m y$.

Step 6: repeating step 2 to step 5 to determine a convergent value of $\hat{\beta}^{(m+1)}$.

These steps are used to determine the impact of independent variables, which are cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4) on student's scientific literacy (Y). The empirical results are expected to be free from the influence of outliers so that they have high accuracy. Therefore, the final conclusion does not produce misleading decisions. Simply put, MM-Estimation start by performing S-Estimation then using the results as the initial point for M-Estimation in the second step. As mentioned before, MM-Estimation is a combination of these two methods therefore this method is able to address outliers in both dependent and independent variables.

RESULTS AND DISCUSSION

Results

Descriptive statistics are presented to obtain the general view of all respondents used in this study. According to the data set, the majority of respondents are known to be male, which are as many as 110 respondents (55%) compare to female respondents, i.e., 90 respondents (45%).

Meanwhile, the average of respondents' age is about 23 years old with standard deviation is 5 years. The youngest and the older respondents are 17 and 35 years old, respectively. The level understanding of scientific literacy where respondents are asked about their confidence in mathematics and science knowledge and its applications is as follows. Approximately half of those surveyed (46%) say that they are not confident about their knowledge in mathematics and science. A minority of the respondents (21%) indicates that they slightly confident about their ability in mathematics and science. Meanwhile, there are 33% of the respondents have good confidence in their knowledge.

A closer look at the data set reveals that the majority respondents who have good confidence level of scientific literacy; they come from science major in high school therefore already have enough background in mathematics and science. Further, the data set also indicates that most of the social science based respondents are not at all confident or have slightly confidence in their knowledge of mathematics and science. Furthermore, the data set also reveal the respondents with science major in high school is about 44.5% compare to those who are social major students, which is 55.5%.

A further statistical analysis reveal that the original data are not normally distributed (or close to normal distribution) therefore parametric method such as OLS method is not applicable in determining the impact of students' social behavior on scientific literacy. This study uses Kolmogorov-Smirnov test for analyzing the normality of the original data from questionnaire distribution. Table 1 presents the results of normality test that show all independent variables, namely cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4) are not normally distributed.

It is important to note that the data set are considered to be normally distributed based on Kolmogorov-Smirnov test when the value of Kolmogorov-Smirnov Z is small. Other than that, asymptotic significance level is able to be used to determine whether the data is normally distributed. The value of asymptotic significance level for two-tailed test is greater than .05 then the variable is considered normally distributed. Contrary, when the value of asymptotic significance level is less than .05 then the

Table 1. Normality Test

		Y	X_1	X_2	X_3	X_4
Normal Parameters	Mean	23.89	23.24	20.88	17.84	20.53
	Std. Deviation	3.695	3.403	4.230	6.342	4.750
Most Extreme Differences	Absolute	.075	.412	.235	.236	.247
	Positive	.075	.298	.182	.198	.168
	Negative	-.057	-.412	-.235	-.236	-.247
Kolmogorov-Smirnov Z		1.067	5.833	3.328	3.331	3.488
Asymp. Sig. (2-tailed)		.205	.000	.000	.000	.000

variable is not normally distributed.

From Table 1 it is clearly seen that the only variable is normally distributed is students' scientific literacy (Y) where the asymptotic significance level is .205, which is greater than .05. Since the values of asymptotic significance level for the four independent variables, which are cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4) are less than .05 then it is safe to assume that the independent variables used in this study are not normally distributed.

A further analysis also reveals that all variables are contaminated by outliers as shown in Figure 1. The outliers are defined as the extreme values that are far away from the data set. It is apparent that both of dependent and independent variables that are used in this study are contaminated by the outliers. Therefore, in order to estimate the impact of students' social behaviors on scientific is more appropriate to use robust regression than the traditional OLS method of regression. Table 2 indicates the empirical results from MM-estimation of robust regression.

The results of the analyses produce several important findings as follows. The robust residual standard error is about 1.444, which is considered quite small. Equally important, the convergent value of the robust coefficient regression is achieved at 13 IRWLS iterations. A further novel finding is that the analysis produces multiple *R-squared* by .784 while the *Adjusted R-squared* is .780.

Furthermore, the results of the analysis provide clear support for the positive impact of social behaviors of the students on their scientific literacy. Importantly, the results not only produce positive impacts but also provide clear evidence that these impacts are statistically significant at the significance level of .05. From

Table 2 it can be seen clearly that all independent variables used in this study, namely cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4) have positive and significant impact on the dependent variable (Y), which is students' scientific literacy. Overall, the estimates of the coefficient vector of robust regression by using MM-Estimation for the four independent variables are .262; .240; .249; and .238, respectively. Table 2 also indicates that all independent variables produce *p-value* by .000, which is less than .05.

Discussions

This study analyzes the impact of students' social behavior on scientific literacy by using robust regression. Overall, these empirical results indicate that the four categories in social behaviors used in this study provide quite similar number of coefficient regression, where the numbers are in the range .24 to .26. All things considered, the empirical results produce clear support that the students' social behaviors influence the scientific literacy positively. The results are in accordance with research by deVries (2018) states that social behavior is related to both mathematics and reading grades. This means that social behavior has a strong influence on student achievement, especially mathematics and reading grades. Repolschi (2017) states social behavior has a significant influence on critical thinking skills where all students should have this skill to have good scientific literacy. Through critical thinking skills and scientific literacy, students are able to take actions/decisions responsibly towards both of the context and conditions of life in the social environment (Suwono, Pratiwi, Susanto, & Susilo, 2017). Turiman, Omar, Daud, & Osman (2012) state that scientific literacy is basically combining both insights related to science and

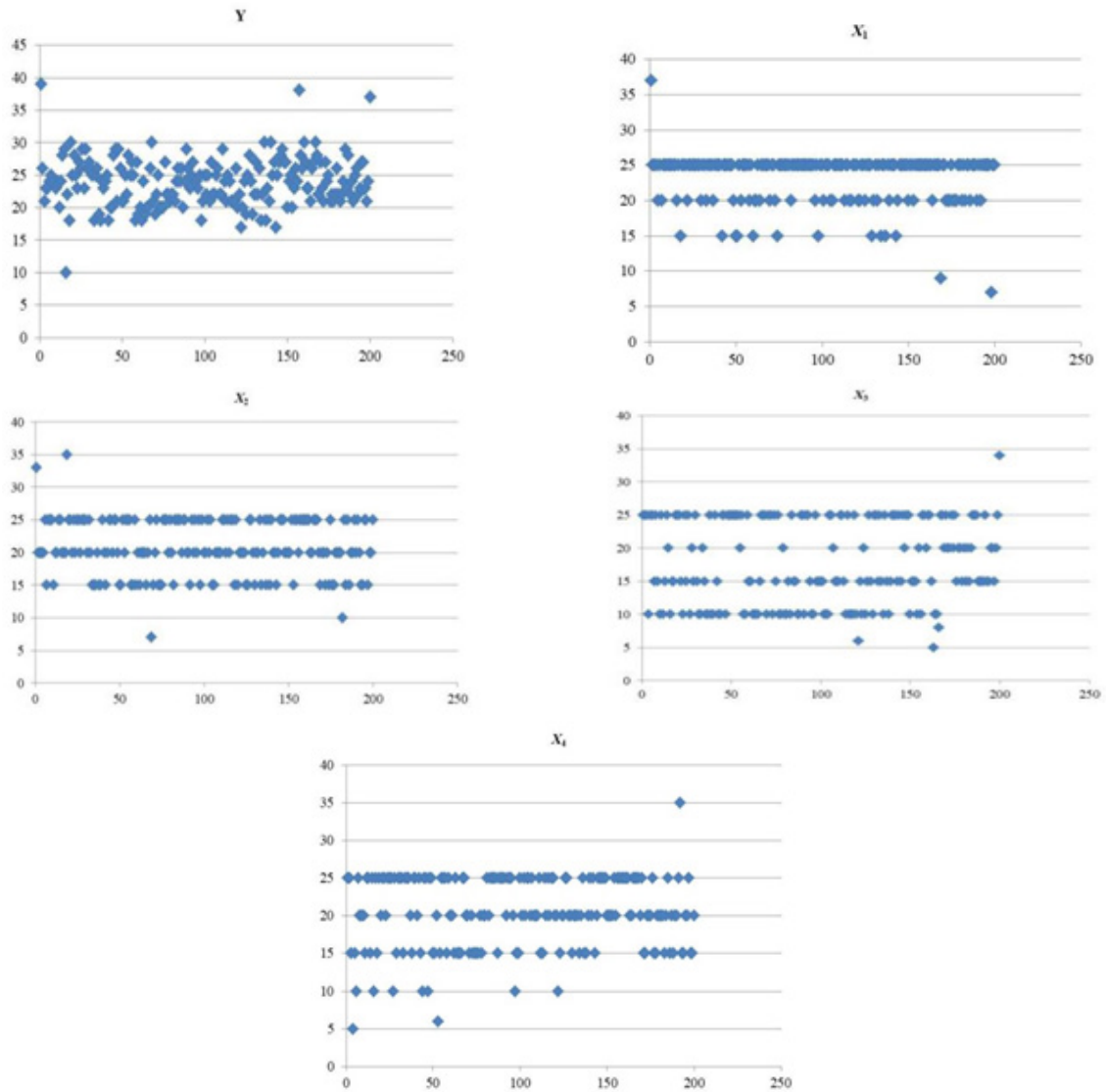


Figure 1. Outliers Detection

Table 2. Estimation Results of MM-Estimation

Variables	Estimate	Std. Error	<i>t</i> -value	<i>p</i> -value
(Intercept)	3.325	1.033	3.218	.002
Cooperation (X_1)	.262	.057	4.624	.000
Assertiveness (X_2)	.240	.030	8.027	.000
Self-control (X_3)	.249	.018	13.425	.000
Pro-social behaviors toward peers (X_4)	.238	.033	7.152	.000

scientific knowledge, as well as skills such as inquiry, critical thinking, problem solving, and decision making.

Furthermore, the results also indicate that the residual standard error of robust regression using MM-Estimation is about 1.444, which is quite small. Another key to remember, this

value is the scale of the residuals used in the estimate standard error is related to the accuracy of the estimation results statistically. And, small standard error implies better accuracy of the estimation. This result is in accordance with previous researches (see Zimmeraman, 1998; Chen, 2007). All of these studies report small

values of standard error in robust regression analysis. Overall, the analysis results provide clear evidences that robust regression by using MM-Estimation yields increasingly good results so that there is no possibility to misleading results.

As mentioned before, MM-Estimator of robust regression also produces the multiple R-squared by .784. This means there are about 78.4 % contributions of all the independent variables in interpreting students' scientific literacy while the remaining numbers are explained by other variables. Basically, this value is used to measure how far the ability of the proposed model in Figure 2 in explaining the variation in students' scientific literacy. It is also important to point out that when the value of the multiple R-squared close to one then all the independent variables used in this study, which are cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4) are expected to be able to provide all information that are needed in predicting the dependent variable.

Furthermore, the results of the analysis also show that the competency of the ability to observe produces the highest results while the competency of ability to evaluate yields the lowest results. Based on interviews, students tend to be afraid when facing questions in the form of developed concept. In order to develop evaluation skills must be supported by studying scientific literature. Large and credible literature studies can be used as a basis for strengthening and developing scientific concepts and at the same time being able to identify the strengths and weaknesses of a literature (Colthorpe,

Abraha, Zimbardi, Ainscough, Spiers, Chen, & Lavidis, 2017).

According to Table 2, the value of Y intercept is 3.325 which mean the height of the regression line when it crosses the Y axis. In other words, the students' scientific literacy is expected to be about 3.325 when all other independent variables are zero. Table 2 also indicates the robust regression coefficient of the cooperation (X_1) is .262 so that for an increasing unit in the variable of cooperation is expected to increase a .262 unit in scientific literacy, holding all other variables are constant. It is clear that cooperation variable (X_1) influences students' scientific literacy positively. Based on table 2, the p-value of this independent variable is .000 so that the positive impact of cooperation variable on scientific literacy is considered statistically significant at the level of significance .05. The aspect of cooperation shows the highest results compared to other aspects. The result indicates that the aspect has a strong influence on scientific literacy. This is in line with research conducted by Dries (2019) who reports that there is an effective advantage between teamwork and scientific literacy. Collaboration between students helps them to improve literacy, metacognition, communication and social skills (Ramadhanti & Yanda, 2018).

Turning now to the empirical evidence of the impact of assertiveness variable (X_2) on students' scientific literacy where the impact is both positive and significant. According to Table 2, the coefficient estimate and p-value of robust regression are .240 and .000, respectively. Therefore, for every point increase in assertiveness variable (X_2), it is predicted a

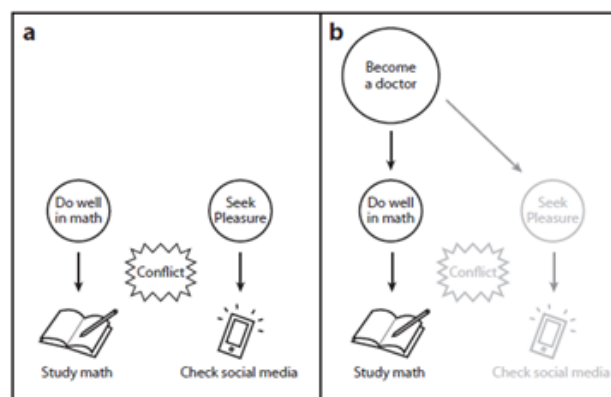


Figure 2. Illustration of Self-Control

Source: Duckworth *et al.* (2019)

.240 point increase in students' scientific literacy when all other variables are constant. The assertiveness aspect has a positive influence on scientific literacy. This is in line with research from Belim & Almedia (2018) which found that assertiveness, clear communication and positivity are important and strategic elements for optimizing literacy.

Furthermore, based on Table 2 it can be seen clearly there is positive and significant impact self-control variable (X_3) on students' scientific literacy. The robust regression coefficient of the self-control (X_3) is .249 so that for a point increase in the variable of self-control is predicted to increase students' scientific literacy by .249 when all other variables are constant. Based on Table 2, the *p-value* of this independent variable is .000 so that the positive impact of self-control variable (X_3) is considered statistically significant at the level of significance .05. Self-control refers to the harmony of thoughts, feelings and actions with a goal that is valued constantly in facing more attractive alternatives (Duckworth, Taxer, Eskreis-Winkler, Galla, & Gross, 2019). Self-control is believed to influence student academic achievement. Ribeiro (2018) illustrates the interaction between high literacy and low self-control, which has a significant impact on academic readiness as shown in Figure 2. Figure 2 explains the following important points, namely (a) studying mathematics and checking social media are mutually exclusive responses. Learning is an academic goal for success in mathematics, while examining social media is a goal for momentary pleasure; (b) doing good in mathematics is related to the student's goal of becoming a doctor, while pleasure is its goal. Thus, self-control imposes impulse to study mathematics while refraining from wanting to check social media (Duckworth *et al.*, 2019).

Furthermore, the impact of the last independent variable, which is pro-social behaviors toward peers (X_4) on students' scientific literacy is also considered statistically significant where the *p-value* is less than .05. Meanwhile, the coefficient regression is .238 which mean that for every increase point in pro-social behaviors toward peers variable (X_4) is expected to increase students' scientific literacy by .238 point. This is in line with research conducted by Llorent, Gomez, Farrington, &

Zych (2020) which states that social emotional competence and empathy have interrelated relationships with literacy skills.

Taken together, these results suggest that there is positive as well as significant relationship between students' social behaviors, which are measured by four categories cooperation (X_1), assertiveness (X_2), self-control (X_3) and pro-social behaviors toward peers (X_4) and their scientific literacy. Comparing these results, from Table 2 it can be seen clearly that the highest and the lowest coefficient estimates are occurred on cooperation (X_1) and pro-social behaviors toward peers (X_4) variables, which are .262 and .238, respectively. Therefore, cooperation variable (X_1) has the highest impact on students' scientific literacy while the variable of pro-social behavior toward peers (X_4) has the lowest impact.

Furthermore, another analysis results reveal that the students' gender and age influence positively their scientific literacy. The coefficient regression of age is .210 which mean that the older the respondents then their scientific literacy are expected to be higher. Meanwhile, the coefficient estimate of students' gender is about .260 therefore for female students, the predicted scientific literacy score is expected to be .260 point higher than for male students. This is very interesting finding because generally, male students are commonly have better scientific literacy (OECD, 2019). But, this result also consistent with the report published by OECD (2019) which states that female students in Indonesia tend to have better scientific literacy compare to male students. Further, level of scientific literacy confidence of students also has positive impact on scientific literacy where the coefficient estimate is 2.181.

Statistics descriptive of the data set indicate almost half of the respondents do not have good confidence about their knowledge as well as ability in mathematics and science. A research conducted by House & Telese (2017) reveals that self-confidence about science is significantly related to science achievement. Students who have high levels of achievement in science can learn faster in science and usually succeed in science. There are only 33% of the participants say they have high level confidence regarding mathematics and sciences knowledge and its applications. The data set also reveal that from the respondents who have a high level of

confidence, the majority are female students. Furthermore, the statistical analysis indicates that the original data from questionnaires distribution are not normally distributed therefore the traditional OLS method is not appropriate to obtain the accurate regression coefficient (see Table 1). Other than that, figure 4 also reveals that the original data used in this study are contaminated by outliers. These are sufficient reasons to rule out the OLS method and then using robust regression to obtain better estimation results.

As mentioned before that IRWLS approximation is used to determine the iteration in MM-estimation of robust regression for analyzing the impact of students' social behaviors on scientific literacy. The analysis results indicate that the number of IRWLS is 13 iterations to produce the convergent value of the coefficient regression. The convergent value means that there are no changes or very few changes to the regression coefficient from the next iteration. Because of that, the values of standard errors, estimated regression coefficients, and t-test scores and p-values are calculated based on its residuals. This result is in accordance with research conducted by Fox & Weisberg (2011) which states that robust regression produces estimates with a higher level of accuracy compared to the OLS regression.

Summing it up, scientific literacy is fundamental for students in current era because of the importance of science and technology and its application. This study provides empirical evidence in relation between social behaviors and scientific literacy. This study presents an overview of scientific literacy of Indonesian undergraduate students. This is important results that are expected to be used as a reference or a basis for improving science literacy of students in Indonesia by considering their social behaviors.

CONCLUSION

This study analyzes the impact of social behaviors on students' scientific literacy by applying MM-Estimator of robust regression. However, this method is used to overcome the weakness of the OLS regression such as the existence of outliers and non-normal distribution of the data set. The total number of undergraduate students is 200, which are selected randomly at a university in Indonesia.

To obtain the empirical results, this study uses as many as four independent variables, which are cooperation, assertiveness, self-control and pro-social behaviors toward peers.

The results reveal that students' social behavior has positive and significant impact on scientific literacy. Furthermore, the four categories of social behavior also indicate to influence both positively and significantly the students' scientific literacy. The results also show that cooperation variable has the highest impact on students' scientific literacy while pro-social behavior toward peers' variable has the lowest impact.

ACKNOWLEDGEMENTS

The author would like to thank the Editors and anonymous reviewers for their generous comments and support during the review process.

REFERENCES

- Alma, O. G. (2011). Comparison of robust regression methods in linear regression. *International Journal Contemporary Mathematical Sciences*, 6(9), 409-421. <http://www.m-hikari.com/ijcms-2011/9-12-2011/almaIJCMS9-12-2011>.
- Belim, C., & Almeida, C. V. (2018). Healthy thanks to communication: A model of communication competence to optimize health literacy - assertiveness, clear language, and positivity. In V. E. Papalois & M. Theodosopoulou (Eds.). *Optimizing health literacy for improved clinical practice*. Hershey, PA: IGI Global, pp. 124-152. doi:10.4018/978-1-5225-4074-8.ch008.
- Borghans, L., Golsteyn, B. H., Heckman, J., & Humphries, J. E. (2011). Identification problems in personality psychology. *Personal Individual Differences*, 51(3), 315-320. doi:10.1016/j.paid.2011.03.029.
- Bybee, R., McCrae, B., & Barry, R. (2009). PISA 2006: An assessment of scientific literacy. *Journal of Research in Science Teaching*, 48(6), 862-883. doi:10.1002/tea.20333.
- Colthorpe, K., Abraha, H. M., Zimbardi K.,

- Ainscough, L., Spiers, J. G., Chen, H. C., & Lavidis, N. A. (2017). Assessing students' ability to critically evaluate evidence in an inquiry-based undergraduate laboratory course. *Advances in Physiology Education, 41*(1), 154-162. doi:10.1152/advan.00118.2016.
- DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationship to science education reform. *Journal of Research in Science Teaching, 37*(6), 582-601. doi:10.1002/1098-2736(200008)37:6<582::AID-TEA5>3.0.CO;2-L.
- DeVries, J. M., Rathmann, K., & Gebhardt, M. (2018). How does social behavior relate to both grades and achievement scores? *Frontier Psychology, 9*(857), 1-8. doi:10.3389/fpsyg.2018.00857.
- Dries, D. R. (2019). Gains in affect and in scientific literacy from peer-led team learning in a literature-based first-semester Biochemistry course. *Biochemistry and Molecular Biology, 33*(1). doi:10.1096/fasebj.2019.33.1_supplement.617.9.
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M., & Gross, J. J. (2019). Self-control and academic achievement. *Annual Review of Psychology, 70*(1), 373-399. doi:10.1146/annurev-psych-010418-103230.
- El Mallah, S. (2014). *Social behavior and academic performance: Examining relations between forms of pro-social behavior and aggression in predicting academic outcomes*. (Thesis, Virginia Polytechnic Institute and State University). <http://hdl.handle.net/10919/76942>.
- Farrington, C. A., Roderick, M., Allensworth, E., Nagaoka, J., Keyes, T. S., & Johnson, D. W. (2012). *Teaching adolescents to become learners. the role of noncognitive factors in shaping school performance: A critical literature review*. Chicago: University of Chicago Consortium on Chicago School Research.
- Fox, J. (2002). *Bootstrapping robust regression models: Appendix to an R and S-PLUS companion to applied regression*. <http://cran.r-project.org/doc/contrib/FoxCompanion/appendix-bootstrapping.pdf>.
- Fox, J. & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed). Thousand Oaks, CA, Sage.
- Gašević, D., Zouaq, A., & Janzen, R. (2013). "Choose your classmates, your GPA is at stake!" The association of cross-class social ties and academic performance. *American Behavioral Scientist, 57*(10), 1460-1479. doi:10.1177/0002764213479362.
- House, J. D., & Telese, J. A. (2017). Confidence in science and achievement outcomes of fourth-grade students in Korea: Results from the TIMSS 2011 assessment. *Project Innovation, 137*(4), 389-392. <https://www.ingentaconnect.com/content/prin/ed/2017/00000137/00000004/art00004#Refs>.
- Lechner, C., Danner, D., & Rammstedt, B. (2017). How is personality related to intelligence and achievement? A replication and extension of Borghans et al. and Salkever. *Personality and Individual Differences, 111*(June), 86-91. doi:10.1016/j.paid.2017.01.040.
- Llorent, V. J., Gomez, A. L., Farrington, D. P., & Zych, I. (2020). Social and emotional competencies and empathy as predictors of literacy competence. *Psicothema, 32*(1), 47-53. doi:10.7334/psicothema2019.106.
- Macan, T. H., Shahani, C., Dipboye, R. L., & Phillips, A. P. (1990). College students' time management: Correlations with academic performance and stress. *Journal of Education Psychology, 82*(4), 760-768. doi:10.1037/0022-0663.82.4.760.
- Maienschein, J. (1998). Scientific literacy. *Science, 281*(5379), 917. doi:10.1126/science.281.5379.917.
- NRC (National Research Council). (1996). *National science education standards*.

- Washington, DC: The National Academies Press. doi:10.17226/4962.
- OECD (Organization for Economic Co-operation and Development). (2003). *Scientific literacy*. <http://www.oecd.org>.
- OECD. (2019). *Programme for international student assessment (PISA) results from PISA 2018*. https://www.oecd.org/pisa/publications/PISA2018_CN_IDN.pdf.
- Ramadhanti, D., & Yanda, D. P. (2018). Understanding poetry through the use of cooperative learning model. *Cakrawala Pendidikan*, 32(3), 436-446. doi:10.21831/cp.v38i3.20675.
- Ribeiro, A. (2018). *The effects of child literacy and self-control on school readiness based on teacher report*. (Master Thesis, University of Rhode Island). <https://digitalcommons.uri.edu/theses/1201>.
- Repolschi, O. (2017). Beliefs and behaviors in learning critical thinking skills. *Philosophy, Social and Human Disciplines, II*, 43-56. <http://www.apshus.usv.ro/arhiva/2015II/004.%20pp.%2043-56.pdf>.
- Supahar, Rosana, D., Ramadani, M., & Dewi, D. K. (2017). Performance assessment instrument of science process skills conform the nature of science. *Cakrawala Pendidikan*, 36(3), 435-445. doi:10.21831/cp.v36i3.14731.
- Suwono, H., Pratiwi, H. E., Susanto, H., & Susilo, H. (2017). Enhancement of students' biological literacy and critical thinking of biology through socio-biological case-based learning. *Jurnal Pendidikan IPA Indonesia*, 6(2), 213-220. doi:10.15294/jpii.v6i2.9622.
- Truesdell, L. A. & Abramson, T. (1992). Academic behavior and grades of mainstreamed students with mild disabilities. *Exceptional Children*, 58(5), 392-398. doi:10.1177/001440299205800503.
- Turiman, P., Omar, J., Daud, A. M., & Osman, K. (2012). Fostering the 21st Century skills through scientific literacy and science process skills. *Procedia - Social and Behavioral Sciences*, 59(October), 110-116. doi:10.1016/j.sbspro.2012.09.253.
- Zimmerman, D. W. (1998). Invalidation of parametric and nonparametric statistical tests by concurrent violation of two assumptions. *Journal of Experimental Education*, 67(1), 55-68. doi:10.1080/00220979809598344.