https://doi.org/10.29013/EJHSS-22-3-86-92

Yuxuan Yang,
Miller School of Albemarle, VA

EXAM PRESSURE AND MOTIVATION AND LIFE SATISFACTION AMONG ADOLESCENTS

Abstract

Objective: This study aims to compare the exam-related pressure and motivation and life satisfaction among adolescents between China and the United States.

Methods: This study used data from the 2015 Programme for International Student Assessment (PISA). PISA is the survey of adolescent students as well as their parents and schools around the world. Logistic regression analysis and artificial neural network are applied. ROC curves and Area Under the Curve (AUC) are estimated.

Results: logistic regression showed that U.S. students are 1.137 times more likely to feel satisfied than Chinese students; exam-related pressure variables are associated with lower life satisfaction (OR < 1), while exam-related motivation variables are associated with students' higher life satisfaction (OR > 1). Artificial neural network also showed that the variables of country, exam-related pressure and motivations affect students' life satisfaction.

Conclusion: Data from this international survey revealed that students in the U.S. have higher life satisfaction than students from China. Exam-related pressure decrease life satisfaction while exam-related motivation increases life satisfaction.

Keywords: exam-related pressure and motivation; logistic regression analysis; artificial neural network.

1. Study objectives

This study aims to compare the exam-related pressure and motivation and life satisfaction among adolescents between China and the United States.

2. Materials and Methods

2.1. Data source

This study used data from the 2015 Programme for International Student Assessment (PISA, website: http://www.oecd.org/pisa/) [1]. PISA is the survey of adolescent students as well as their parents and schools around the world, conducted by the Organization for Economic Co-operation and Development (OECD). It is conducted every three years to tests 15-year-old students in reading, mathematics and science. PISA choose the age of 15 because it is believed that students at this age can decide whether or not they want to continue their education.

The 2015 data is the most recent available PISA data by the time of this study. It (http://www.oecd.org/pisa/data/2015database/) includes five main data files: a student-questionnaire data file, a school-questionnaire data file, a teacher-questionnaire data file, a cognitive item data file and a file with questionnaire timing data. We used the student data.

2.2 Variables

2.2.1 outcome: life satisfaction

Students were asked "*Overall, how satisfied are you with your life as a whole these days?*"

Responses ranged from 0 to 10.

2.2.2 Exam-related pressure and motivation

Table 1.

	question	variable name
Exam-related pressure	I often worry that it will be difficult for me taking a test.	worry_difficult
	I worry that I will get poor <grades> at school	vorry_poor_perforn
	Even if I am well prepared for a test I feel very anxious	anxious
	I get very tense when I study for a test.	tense
	I get nervous when I don't know how to solve a task at school.	nervous
Exam-related motivation	I want top grades in most or all of my courses.	want_top_grade
	I want to be able to select from among the best opportunities available when I graduate.	want_opportunity
	I want to be the best, whatever I do.	want_best
	I see myself as an ambitious person.	ambitious
	I want to be one of the best students in my class	want_best

Response options were: strongly disagree, disagree, agree, strongly agree.

2.2.3 Other variables:

Students' age and gender are included in the logistic model. These are variables that may affect the outcome, therefore, by including them in the model, their potential confounding effect can be controlled.

2.3 data analysis

Logistic regression analysis and artificial neural network are applied.

Logistic Regression Modeling is a popular analytic technique to analyze the association between a set of predictors and a binary outcome. In this study, the outcome is "having higher life satisfaction".

The general formula of logistic regression is: $ln(\text{odds of an event occurring}) = ln\left(\frac{P}{P-1}\right) = \beta + \beta_1 \cdot \cdot X_1 + \beta_2 \cdot X_2 + \ldots + \beta_n \cdot X_n \cdot P$ is the probability of an event, which is convertible with odds. X_n is a predictor variable, and β_n is a regression coefficient. The relationship between the odds ratio and the coefficients is $OR = e^{\beta}$.

- If the coefficient β of a variable X_n is larger than 0, is related to a higher odds/probability of the event. The odds ratio related to X_n is above 1 in this case.
- If the coefficient of a variable X_n is equal to 0, X_n is not related to the event. The odds ratio related to is equal to 1 in this case.
- If the coefficient of a variable X_n is smaller than 0, X_n is related to a lower odds/probability of the event. The odds ratio related to X_n is below 1 in this case.

An artificial neural network (ANN), often just called a "neural network" (NN), is a mathematical model or computational model based on biological neural networks, in other words, is an emulation of biological neural system. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In more practical terms neural networks are non-linear statistical data modeling tools. ANN is widely used these days to model complex

relationships between inputs and outputs or to find patterns in data. This model was done using R software 'neuralnet' package.

In order to assess how well a model performs, a ROC curve is generated which is a measure of classifier performance [2]. It is a graphic that shows the diagnostic ability of a model in predicting binary outcome as its discrimination threshold is varied. Ultimately, the focus is the area under the ROC curve, or AUROC. AUROC ranges from 0.50 to 1.00, and the higher AUROC, the better the model is. It should be noted, however, that in this study, the purpose is to compare which model has a higher AUROC, instead of the absolute values of AUROC.

4. Results

There are 11.065 participants with complete information of the variables, including 7946 from China and 3119 from the USA.

Student gender distribution	
boys	49.8%
girls	50.2%

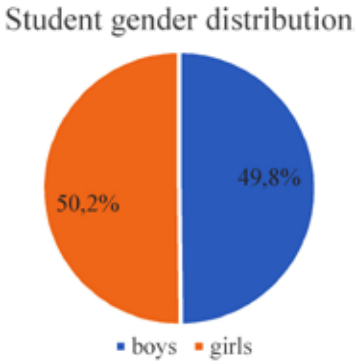


Figure 1.

Boys are more likely to have higher life satisfaction.

% of high life satisfaction	
boys	52.1%
girls	48.0%

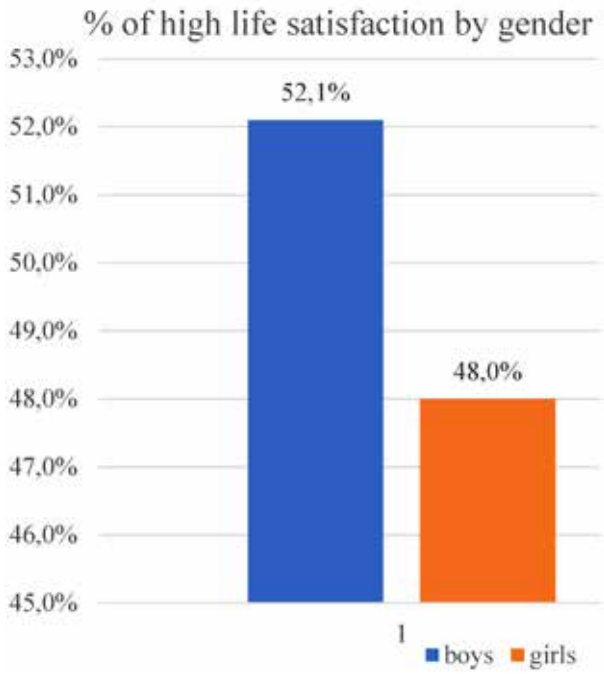


Figure 2.

Students in the U.S. have higher proportion of high life satisfaction.

% of high life satisfaction	
China	48.0%
USA	55.9%

Figure 3.

Results from Logistic Regression

Table 2.

Odds Ratios from Logistic Regression modeling					
	P value		**Odds Ratio**	**Lower CI**	**Upper CI**
1	***2***	***3***	***4***	***5***	***6***
student gender: male vs female	0.163		1.057	0.977	1.143

1	*2*	*3*	*4*	*5*	*6*
age	0.491		0.95	0.821	1.099
worry_difficult	0.003	**	0.907	0.849	0.969
worry_poor_perform	<0.001	***	0.822	0.771	0.875
anxious	0.292		0.966	0.906	1.03
tense	0.207		0.959	0.9	1.022
nervous	<0.001	***	0.833	0.787	0.881
want_top_grade	0.011	*	1.091	1.019	1.167
want_opportunity	0.532		0.973	0.893	1.059
want_best	0.001	**	1.119	1.046	1.198
ambitious	<0.001	***	1.342	1.261	1.428
USA	0.036	*	1.137	1.008	1.282

From the table, exam pressure variables are associated with lower life satisfaction (OR < 1), while

Exam motivation variables are significantly associated with students' higher life satisfaction (OR > 1).

e.g., a student who worries about difficulty of the exam is 90.7% times as likely as a student not reporting worrying to have high life satisfaction.

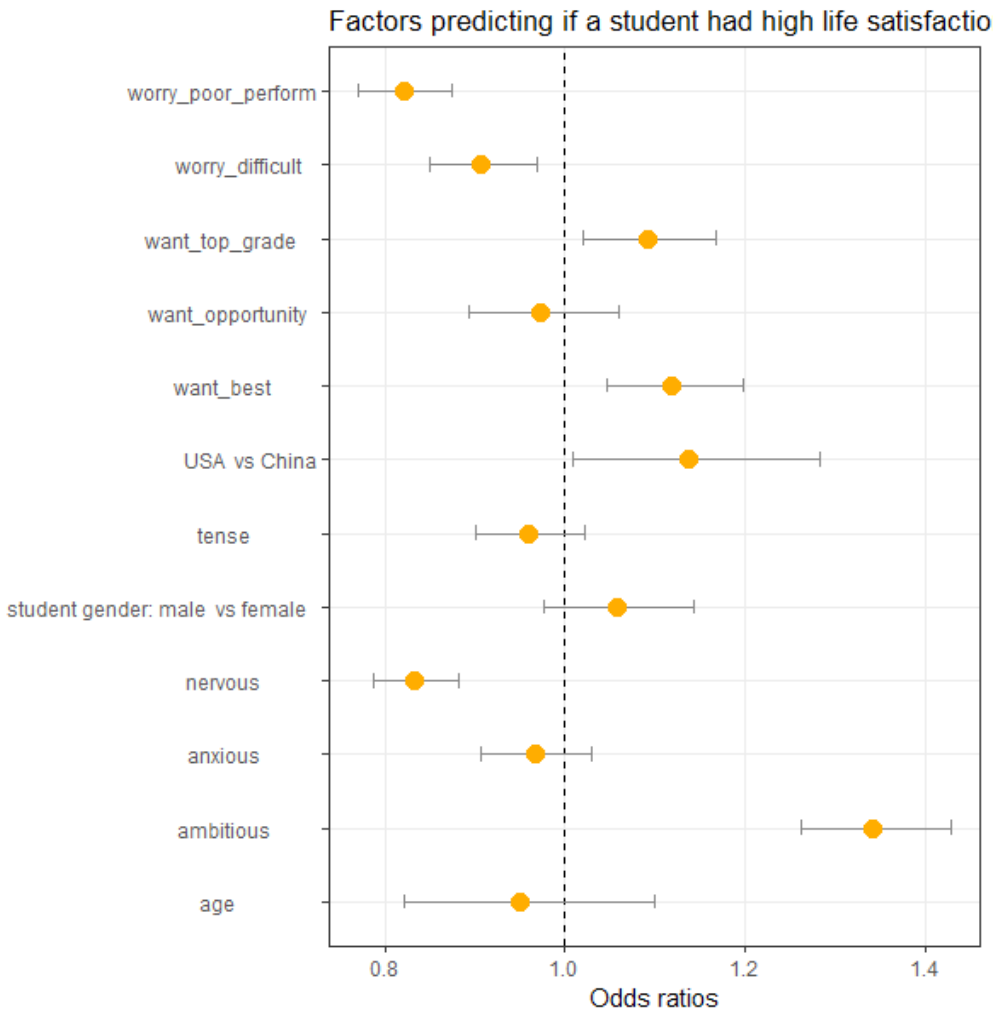


Figure 4.

The Odds Ratio also confirmed that students in the U.S. have higher life satisfaction than students from China. OR = 1.137, meaning that U.S. students are 1.137 times more likely to feel satisfied than Chinese students.

This model has an AUC of 63%. Given that I am only using the exam-related factors to predict students' life satisfaction, this is a pretty good number.

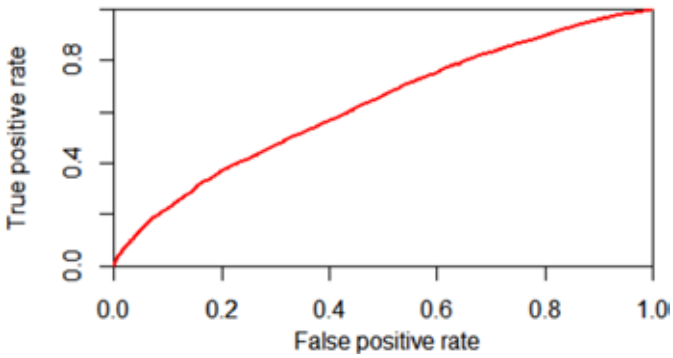


Figure 5. Logistic regression: ROC Curve: AUC = 63%

Results from NN

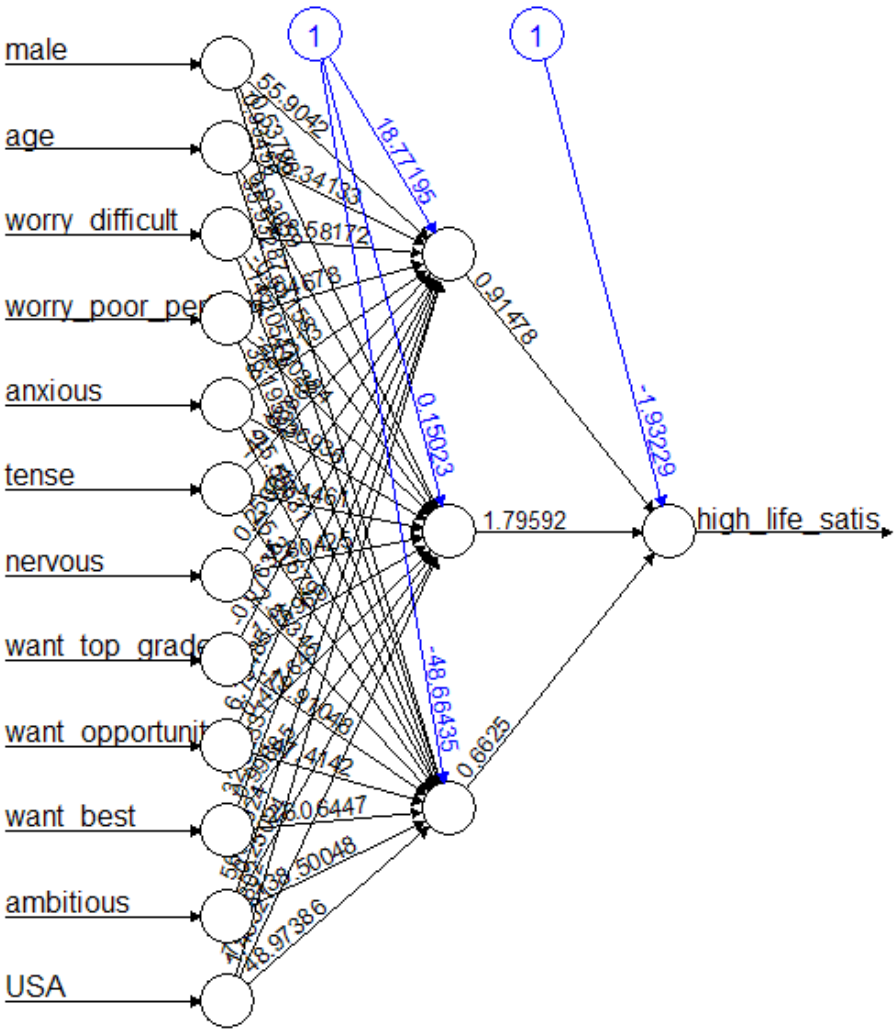
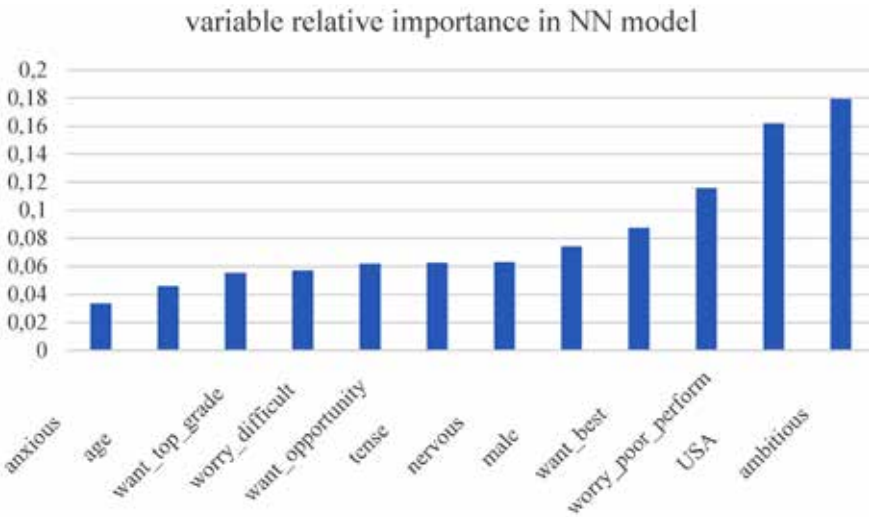


Figure 6.

In the above plot, the net is essentially a black box so we cannot say that much about the fitting or the weights. However, it is sufficient to say that the model algorithm has converged and therefore the model is ready to be used.



	rel_imp
male	0.07421744
age	0.04604342
worry_difficult	0.05726405
worry_poor_perform	0.11610010
anxious	0.03373606
tense	0.06272287
nervous	0.06310223
want_top_grade	0.05547345
want_opportunity	0.06202663
want_best	0.08774197
ambitious	0.17947600
USA	0.16209578

Figure 7.

Variable importance from the NN model is listed. The relative importance of a specific explanatory variable can be determined by identifying all weighted connections between the nodes of interest. It can be interpreted as the strength of association between an explanatory variable and the response variable. The number indicates relative importance with the absolute magnitude from zero to one [3].

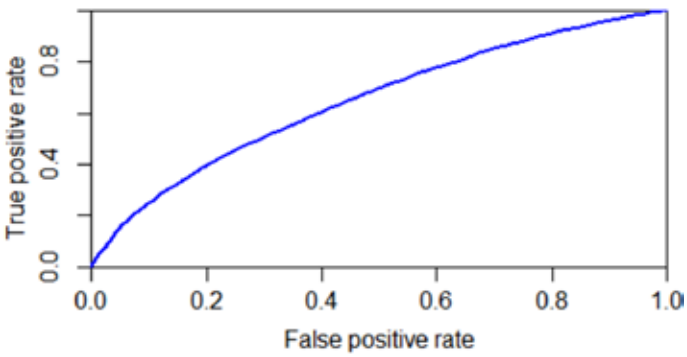


Figure 8. Neural Network model: ROC Curve: AUC = 65%

From the model, it can be seen that "ambition", "USA (country)" and "worry_poor_perform" are some of the most important factors associated with students' interest in science study. AUC is 65% for the NN model.

4. Interpretation of results

Boys have higher life satisfaction than girls.

Students in the U. S. A have higher life satisfaction than students in China.

Both Logistic regression and Neural Network modeling revealed that exam-related pressure and motivation are associated with students' life satisfaction.

Exam-related pressure decreases life satisfaction, while exam-related motivation increases life satisfaction.

5. Conclusion

Data from this international survey revealed that students in the U.S.A have higher life satisfaction than students in China. Both Logistic regression and Neural Network modeling revealed that exam-related pressure and motivation are associated with students' life satisfaction.

References:

1. OECD. Programme for International Student Assessment (PISA) 2015. Data base: Student Questionnaire data file, Organization for Economic Co-operation and Development. Paris, France. 2016.
2. Evaluation of Predictive Models. Decision Systems Group, Brigham and Women's Hospital Harvard Medical School.

3. Garson G. D. Interpreting neural network connection weights. Artif. Intell. Expert 6, 1991. – P. 46–51.
4. Peng C. J., Lee K. L., Ingersoll G. M. An Introduction to Logistic Regression Analysis and Reporting. The Journal of Educational Research, – 96(1). – P. 3–14.
5. Tabachnick B., and Fidell L. Using Multivariate Statistics (4th Ed.). Needham Heights, MA: Allyn & Bacon. 2001.
6. Stat Soft. Electronic Statistics Textbook, URL: http://www.statsoft.com/textbook/stathome.html
7. Stokes M., Davis, C. S. Categorical Data Analysis Using the SAS System, SAS Institute Inc., 1995.