# Research on the psychological gap, personality and achievement of in-school youth based on regression analysis

Lei Zhou*

*Student Work Place, Huazhong Agricultural University, Wuhan, Hubei, China*

ABSTRACT: Although our society is in rapid development, the psychological problems among in-school university students are increasingly obvious. According to this problem, this thesis applied the Psychological Gap Scale made by Caixia Ma with EPQ and AMS, and made random questionnaire survey among 400 students in a comprehensive university. The survey found out that the average scores of all psychological gap dimensions exceeded the critical value 3, showing most students in that university have psychological gap. Their personality stability, introversion and extroversion are all above the national norm level while their stubbornness is lower than it. Besides, the students' motivation in pursuing success is stronger than their motivation in avoiding failure. In the last part, this thesis reached the conclusion that personality leaves a great impact in the students' psychology through regression analysis model and study of the quantitative relations among personality, achievement and psychological gap.

*Keywords*: regression analysis; psychological gap; personality; achievement motivation; university students

## 1 INTRODUCTION

As an important part of society, university students can affect the future of a country or even a nation. Their university life is a period for them to become psychologically mature and get ready for bearing part social responsibilities. It is also a period for university students to realize diverse development including learning, relationship, adaptation, and self-evaluation. Therefore, university life can definitely leave psychological gap on most students. At present, university students with hidden psychological trouble are increasing. It is hard to effectively predict and control the trend. In addition, as employment situation is becoming more and more serious while the number of university student is becoming bigger and bigger, university student's social status drops sharply. They are no longer representatives of knowledge. Instead, they are now laborers with university education. Hence, significant psychological gap can form in this way. Besides, different students have different motivation in achievement and personality. Therefore, differences also exist in the expression form of psy-

chological gap. Based on the above, it can be concluded that there's significant social importance in studying in-school youth's motivation in achievement, personality, and psychological gap.

## 2 REGRESSION ANALYSIS MODEL

Simple linear regression model:

$$y = \beta_0 + \beta_i x + \varepsilon \tag{1}$$

In which, $\beta_0$, $\beta_1$ refer to regression coefficient and $\varepsilon$ refers to random error. Always assume $\varepsilon \sim N(0, \sigma^2)$, random variable is $y \sim N(\beta_0 + \beta_1 x, \sigma^2)$. Parameter $\beta_0$, $\beta_1$ in equation (1) still takes the least square method for estimation. Thus, select estimated value $\hat{\beta}_j$ and make $\hat{\beta}_j = \beta_j$. When $j=0,1$, the residual sum of square $Q = \sum_{i=1}^{m} \varepsilon_i^2 = \sum_{i=1}^{m} (y_i - \beta_0 - \beta_1 x_{i1})^2$ will reach its lowest value. Therefore, set $\dfrac{\partial Q}{\partial \beta_j} = 0, j = 0,1,2,\cdots,n)$ and can obtain

*Corresponding author: 623270182@qq.com

$$\begin{cases} \dfrac{\partial Q}{\partial \beta_0} = -2\sum_{i=1}^{m}(y_i - \beta_0 - \beta_1 x_1) = 0 \\ \dfrac{\partial Q}{\partial \beta_j} = -2\sum_{i=1}^{m}(y_i - \beta_0 - \beta_1 x_1) = 0, j = 0,1 \end{cases}$$

After arrangement, it can be converted into the following normal equations system:

$$\begin{cases} \beta_0 n + \beta_1 \sum_{i=1}^{n} x_{i1} + \beta_2 \sum_{i=1}^{n} x_{i2} + \cdots \beta_m \sum_{i=1}^{n} x_{im} = \sum_{i=1}^{n} y_i \\ \beta_0 \sum_{i=1}^{n} x_{i1} + \beta_1 \sum_{i=1}^{n} x^2_{i1} + \beta_2 \sum_{i=1}^{n} x_{i1}x_{i2} + \cdots \beta_m \sum_{i=1}^{n} x_{i1}x_{im} = \sum_{i=1}^{n} x_{i1}y_i, m = 0,1 \\ \beta_0 \sum_{i=1}^{n} x_{im} + \beta_1 \sum_{i=1}^{n} x_{im}x_{i1} + \beta_2 \sum_{i=1}^{n} x_{im}x_{i2} + \cdots \beta_m \sum_{i=1}^{n} x^2_{im} = \sum_{i=1}^{n} x_{im}y_i \end{cases}$$
(2)

The matrix form is $X^T X \beta = X^T Y$. When matrix $X$ is in full rank, $X^T X$ can be invertible square matrix and obtain $\hat{\beta} = (X^T X)^{-1} X^T Y$. Bring $\hat{\beta}$ back to the original model, the estimated value of y can be obtained as $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$. However, the fitted value of this data group is $\hat{Y} = X \hat{\beta}$. The fitting error $e = Y - \hat{Y}$ can be called residual error and can be the estimation of random error $\varepsilon$. $Q = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ is residual sum of square which is $Q(\beta)$.

Whether there's any linear relation between dependent variable $y$ and independent variable $x_1,\ldots,x_m$ can accord with the linear relation shown in equation (1) needs check. Take the original assumption as $H_0$: $\beta_j = 0$. When $H_0$ is established,

$$F = \frac{U/m}{Q/(n-m-1)} \sim F(m, n-m-1)$$.

Under significance level $\alpha$, there's $\alpha$ quantile $F_\alpha(m, n-m-1)$. If $F < F_\alpha(m, n-m-1)$, accept $H_0$; or, refuse $H_0$.

Multiple Linear Regression (MLR) model:

$$\begin{cases} y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_m x_m + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases}$$
(3)

In which $\beta_0, \beta_1,\ldots, \beta_m$, $\sigma^2$ are all irrelevant unknown parameters of $x_1,\ldots,x_m$. $\beta_0, \beta_1,\ldots, \beta_m$, are called regression coefficients. Parameters $\beta_0, \beta_1,\ldots, \beta_m$ in equation (3) still take the least square method for estimation. Thus, estimated value $\hat{\beta}_j$ shall be taken. Make $\hat{\beta}_j = \beta_j$ and $j = 0,1,\ldots,m$, the error sum of squares $Q = \sum_{i=1}^{m} \varepsilon_i^2 = \sum_{i=1}^{m}(y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_m x_{im})^2$ will reach its lowest value and obtain normal equations system $X^T X \beta = X^T Y$. When matrix $X$ is in full column rank, it can obtain $\hat{\beta} = (X^T X)^{-1} X^T Y$. Bring $\hat{\beta}$ back to the original model. The estimated value of $y$ can be ob-

tained as $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots \hat{\beta}_m x_m$. Then, residual sum of squares $Q(\beta)$ can be obtained. Whether there's any linear relation between dependent variable $y$ and independent variable $x_1,\ldots,x_m$ can accord with the linear relation shown in equation (3) needs the same check: take the original assumption as $H_0$: $\beta_j = 0$. When $H_0$ is established, $F = \frac{U/m}{Q/(n-m-1)} \sim F(m, n-m-1)$. Under significance level, there's $\alpha$ quantile $F_\alpha(m, n-m-1)$. If $F < F_\alpha(m, n-m-1)$, accept $H_0$; or, refuse $H_0$.

## 3 REGRESSION ANALYSIS OF IN-SCHOOL YOUTH'S PSYCHOLOGICAL GAP, PERSONALITY, AND ACHIEVEMENT

This thesis took undergraduates of a comprehensive university as research object and randomly selected four hundred students from four majors including arts, history, sciences and engineering, one hundred students from each major. The students are freshmen, sophomores, juniors and seniors. No special requirement was asked for gender or family background. Each student was randomly selected. (See Table 1)

Table 1. Survey conditions

| Item | Sub-item | Number of students participating in the questionnaire survey | Proportion |
|---|---|---|---|
| Major | Arts | 90 | 22.6% |
| | History | 70 | 17.4% |
| | Sciences | 128 | 31.9% |
| | Engineering | 112 | 28.1% |
| Grade | Freshman | 100 | 25% |
| | Sophomore | 100 | 25% |
| | Junior | 100 | 25% |
| | Senior | 100 | 25% |
| Gender | Female | 200 | 50% |
| | Male | 200 | 50% |
| Family background | Urban | 200 | 50% |
| | Rural | 200 | 50% |

The questionnaire applied the Psychological Gap Scale made by Caixia Ma, EPQ and AMS. See Table 2 given below for the general situation of the three scale dimensions.

From Table 2, it can be seen that the average scores of all dimensions used to measure in-school youth's psychological gap exceed the critical value 3. To rank the scores from higher to lower, we can get learning, interpersonal communication, extracurricular activities, relationship, and self-assessment. A higher score can manifest more significant gap. Moreover, the total can also show that most in-school youth in that university have psychological gap. Their emotional stability and extroversion and introversion are all higher than normal national level while their stubbornness is lower than the standard. Their motivation in pursuing suc-

cess is stronger than the motivation in avoiding failure.

Table 2. General situation of in-school youth's psychological gap, personality, and achievement

| Performance | Population | Average score | Standard deviation |
|---|---|---|---|
| Self-assessment | 400 | 3.057 | 0.741 |
| Study | 400 | 3.145 | 0.576 |
| Extracurricular activities | 400 | 3.098 | 0.566 |
| Interpersonal communication | 400 | 3.137 | 0.456 |
| Relationship | 400 | 3.079 | 0.531 |
| Total score of psychological gap | 400 | 3.115 | 0.435 |
| Stubbornness | 400 | 2.663 | 1.814 |
| Extroversion and introversion | 400 | 7.731 | 2.841 |
| Emotional stability | 400 | 5.52 | 3.036 |
| Motivation in pursuing success | 400 | 23.899 | 6.351 |
| Motivation in avoiding failure | 400 | 21.665 | 6.991 |
| Total score of motivation in achievement | 400 | 2.301 | 9.388 |

In the next part, this thesis will apply regression analysis model to study the relations among in-school youth's psychological gap, motivation in achievement, and personality.

### 3.1 Relations between motivations in personality and achievement

Take $x_1$, $x_2$, $x_3$ to respectively represent stubbornness, extroversion & introversion, and emotional stability of personality. $y_1$, $y_2$ respectively represent motivation in pursuing success and motivation in avoiding failure. Use regression analysis model can obtain the relations between $y_1$, $y_2$ and $x_1$, $x_2$, $x_3$ as follows:

$$y_1 = 4.1662 - 3.34594\,x_1 + 2.12156\,x_2 + 1.9603\,x_3$$

$$y_2 = 24.9066 - 2.89979\,x_1 - 2.9502\,x_2 + 4.80568\,x_3$$

From the above, it can be seen that stubbornness and extroversion & introversion have great influence on motivation in pursuing success while extroversion & introversion and emotional stability have great influence on motivation in avoiding failure.

### 3.2 Relations between personality and psychological gap

Make $x_1$, $x_2$, $x_3$ respectively represent stubbornness, extroversion & introversion, and emotional stability of personality. $y_1$, $y_2$, $y_3$, $y_4$, $y_5$ respectively represent self-assessment, learning, extracurricular activities, interpersonal communication, and relationship. Use regression analysis model to analyze the corresponding relations between $y_1$, $y_2$, $y_3$, $y_4$, $y_5$ and $x_1$, $x_2$, $x_3$ as follows:

$$y_1 = 21.2314 + 2.33641\,x_1 - 5.75462\,x_2 + 3.63634\,x_3$$

$$y_2 = 14.771 + 1.79261\,x_1 - 3.76246\,x_2 + 2.31026\,x_3$$

$$y_3 = 2.36054 - 0.0475\,x_1 + 0.0555\,x_2 + 0.0821\,x_3$$

$$y_4 = 2.02587 + 0.62442\,x_1 + 0.01392\,x_2 - 0.11454\,x_3$$

$$y_5 = 1.7228 + 0.46435\,x_1 - 0.04175\,x_2 + 0.09808\,x_3$$

From the above, it can be seen that emotional stability and extroversion & introversion have great influence on self-assessment, learning, and curricular activities while stubbornness and emotional stability have great influence on interpersonal communication and relationship.

### 3.3 Relations between motivation in achievement and psychological gap

Make $x_1$, $x_2$ respectively represent motivation in pursuing success and motivation in avoiding failure. $y_1$, $y_2$, $y_3$, $y_4$, $y_5$ respectively represent self-assessment, learning, extracurricular activities, interpersonal communication, and relationship. Use regression analysis model to analyze the corresponding relations between $y_1$, $y_2$, $y_3$, $y_4$, $y_5$ and $x_1$, $x_2$ as follows:

$$y_1 = 0.43771 - 0.04739x_1 + 0.17345x_2$$

$$y_2 = -2.1108 + 0.0165x_1 + 0.2267x_2$$

$$y_3 = -1.2292 - 0.0879\,x_1 + 0.2962\,x_2$$

$$y_4 = -1.2686 - 0.0739x_1 + 0.2833x_2$$

$$y_5 = -1.6315 - 0.0973\,x_1 + 0.3255\,x_2$$

From the above, it can be seen that motivation in avoiding failure have great influence on self-assessment, learning, extracurricular activities, interpersonal communication, and relationship.

### 3.4 Relations among personality, motivation in achievement, and psychological gap

Make $x_1$, $x_2$, $x_3$, $x_4$, $x_5$ respectively represent stubbornness, extroversion & introversion, emotional stability, motivation in pursuing success, and motivation in avoiding failure. $y_1$, $y_2$, $y_3$, $y_4$, $y_5$ respectively represent self-assessment, learning, extracurricular activities, interpersonal communication, and relationship. Use regression analysis model to analyze the corresponding relations between $y_1$, $y_2$, $y_3$, $y_4$, $y_5$ and $x_1$, $x_2$, $x_3$, $x_4$, $x_5$ as follows:

$$y_1 = -0.2728 - 0.5231x_1 + 0.0562x_2 + 0.4661x_3 + 0.0714x_4 + 0.00092x_5$$

$$y_2 = 0.3842 + 0.0209x_1 + 0.1764x_2 - 0.0132x_3 + 0.0231x_4 + 0.0407x_5$$

$$y_3 = 0.20003 - 0.0865x_1 + 0.3345x_2 - 0.08901x_3 + 0.0005x_4 + 0.0474x_5$$

$$y_4 = 1.5757 + 0.5465x_1 + 0.2355x_2 - 0.3903x_3 - 0.0533x_4 + 0.0789x_5$$

$$y_5 = 0.483 + 0.0647x_1 + 0.3802x_2 - 0.1837x_3 - 0.0411x_4 + 0.0686x_5$$

From the above, it can be concluded that stubborn-

ness and emotional stability have great influence on self-assessment. Stubbornness, extroversion & introversion, and motivation in avoiding failure have great influence on learning. Stubbornness, extroversion &introversion, and emotional stability have great influence on extracurricular activities and interpersonal communication. Motivation in avoiding failure has great influence on relationship. Therefore, personality contains more influence on psychological gap.

At last, make $x_1$, $x_2$, $x_3$, $x_4$, $x_5$ respectively represent stubbornness, extroversion & introversion, emotional stability motivation in pursuing success, and motivation in avoiding failure. $y$ refers to the total score of psychological gap. By using regression analysis, we can obtain

$$y = -49.7248 - 7.5849\,x_1 + 4.1163\,x_2 + 7.1093_3 + 0.8075\,x_4 - 0.8307\,x_5$$

It can also be proved that personality has great influence on psychological gap.

## 4 CONCLUSIONS

According to current in-school youth's psychological problems, this thesis applied regression analysis model and mathematical software to conduct quantitative analysis of university student's personality, motivation in achievement, and psychological gap. It obtained the specific relations among in-school youth's personality, motivation in achievement, and psychological gap and reached the conclusion that personality leaves a great influence on in-school youth's psychology. Hence, more attention shall be paid to the cultivation on in-school youth's personality.

## REFERENCES

[1] Gui, Y.P. 2009. *Study of Relations between University Student's Personality Motivation in Achievement and Psychological Gap*. Fuzhou: Fujian Normal University.

[2] Zhang, W.Y., Liu, G.F. & Wu, Y.H. 2012. Regression analysis on correlative factors between university student's psychological health and physical exercise. *Preventive Medicine*, 39(5): 1162-1165.

[3] Fan, C.X., Ma, S.B., Lin, H.S. & Wang, H.S. 2001. Gradual regression analysis of university student's psychological health condition and influential factors. *Chinese Journal of School Health*, 22(3): 215-216.

[4] Zhong Wenjuan, Huang Wanqi, Zhang Yinyou & He Zun. 2007. Logistic regression analysis of the influencing factors of psychological problems faced by college students. *Journal of Clinical Rehabilitative Tissue Engineering Research*, 11(17): 3454-3456.

[5] Li, J. 2012. Regression analysis of the personality characteristics and coping modes of Mongol and Han university students. *Journal of Inner Mongolia Normal University (Education Sciences Edition)*, 25(3): 47-50.

[6] Huang, Z.P. & Li, S.Z. 2012. Application of Mathematica in MLR Analysis. *Journal of Qufu Normal University*, 38(4): 28-31.