

Do these widely believed factors really affect college students' grades?

Rongxuan Ye *

Sichuan Normal University, Sichuan China

* Corresponding Author Email: 757198296@qq.com

Abstract. Presumably everyone knows that the grades of college students are generally not ideal. As far as our daily cognition is concerned, the factors that determine the grades of college students are nothing more than hard work, the quality of the learning environment, and the education of the family. But I think everyone has a certain misunderstanding about the determinants of college students' grades. As a college student, I can clearly feel that university study is completely different from middle school study. Here I have selected a few typical indicators that everyone believes influence college student grades, explored the reasons for this, and tested whether this is the case.

Keywords: education; college; grades.

1. Introduction

We were always encouraged by our teachers in middle school, "You will be relaxed after you go to college". However, the result is not as good as expected, as a college student, we can still find that many of my classmates are bothered by the final exam, and even have difficulty passing the exam. It is not only a hot topic among college students, but also for education institutes. According to a study published in the "Nature" sub-journal "Natural Human Behavior"[1] - based on the results of a test called Supertext which is administered at Stanford University, Moscow State Higher School of Economics, Educational Testing Service (ETS), and domestically. A test co-sponsored by Beijing, Tsinghua University, and partner universities in India. The test is designed to track changes in the abilities of students majoring in computer science and electrical engineering in physics, mathematics, and more, and compare the results. This result surprisingly found that after Chinese college students study in college, their learning ability does not increase but declines. Why do Chinese students rank among the best in the world in middle school, but in college, Chinese students' learning ability and grades decline? It is very important to find out the influencing factors that affect the performance of Chinese college students.

According to early research, it is commonly believed that these factors may have a high effect on students' grades: Preview habit, Parents education level, The study time, The study place, Lunch price.

So, are these widely believed reasons applicable to college students in China?

2. Approach

We have collected 205 pieces of data (After deleting invalid data, it is 198) from a survey on students from the Chinese college student, from the app "WenJuanXin (<https://www.wjx.cn/vm/mLs9DeP.aspx#>), about information about Chinese college students' grades and their condition regarding these factors. The survey contained a variety of validated scales measuring the grades of the students and the degree of the problems they are facing. We will try to find the relationship between the grades and these factors by using the statistical method, like confident intervals, Hypothesis testing, and so on. We use analysis tools like SPSS and EXCEL to evaluate the impact of these factors on college students' performance and conduct an inferential analysis to obtain statistical significance and actual effects. We will try to find out whether these factors have significant effects on the grade of college students.

2.1. Data

We obtained 212 pieces of data through an online questionnaire survey, and after filtering out invalid ones, 198 pieces of data remained. Here, we numbered the data and divided them into two groups according to their parity. Through the probability distribution map (Fig. 1) made in python based on the average GPA score, we observe that the distribution of the two sets of data is roughly the same, so here we temporarily regard these data is reliable.

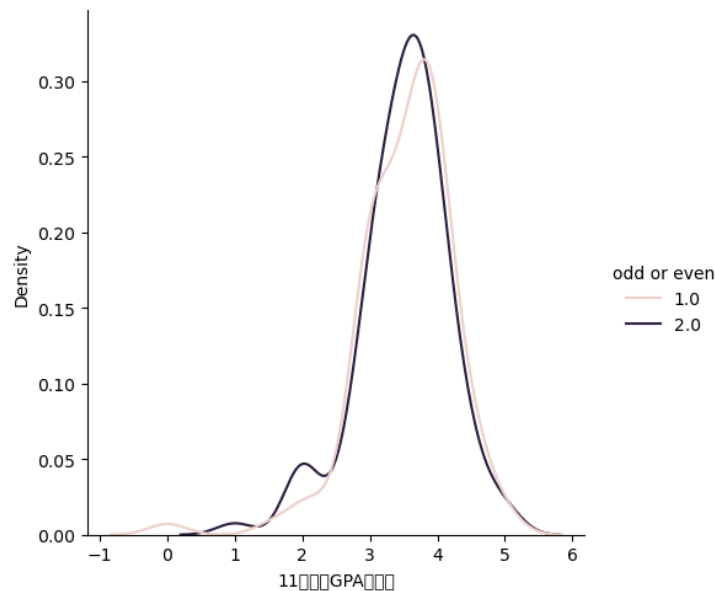


Fig 1. The probability distribution map

2.2. Analysis

First, let's take a look at the results of these students as a whole. There are 198 student samples. After calculation, the average GPA score of these students is: =3.52.

We use the confidence interval to estimate the average of the average GPA \bar{x} scores of all Chinese college students. Here we choose 95% confidence interval.

According to the central limit the theory, with the sample size is large enough, we can assume the distribution of the mean of average GPA score of college students in China as a normal distribution.

With the confidence interval is 95%, the significant level is $\alpha=0.05$, $\alpha/2=0.025$, with the degree of freedom: $df =n-1=197$ (n is the sample size ,198), so the critical t-value is: $t_{\alpha/2} =1.972$.

We can get the 95% confidence interval of the mean of average GPA score of all Chinese college students with this formal below:

$$\bar{x} \pm t_{\alpha/2} \frac{S}{\sqrt{n}} \quad (1)$$

So, after calculation, we can get the 95% confidence interval of the mean is (3.539,3.621).

So, we have 95% confidence to say that the mean of the average GPA score is between 3.539 to 3.621.

2.2.1 Parents education level

In the eyes of most people, parents' academic qualifications are always linked with their children's achievements unconsciously. There is an old saying in China that "like father, like son". Moreover, an academic report has shown that the mother's education level will be linked to the child's final education level [2]. Here we do a simple analysis based on the samples of these students on their parents' education level and their GPA scores.

Here we divide these students Into 3 groups according to their parents' education level, namely "middle school", "high school", "undergraduate and above".

Here is a boxplot of the GPA scores (3.1) of the different groups of students, we can see that each group has slightly different grades and observe that the average GPA scores of children of highly educated parents do seem to be higher. Then we will use statistical methods to verify whether the parents of college students have different education backgrounds, and their GPA scores are also different?

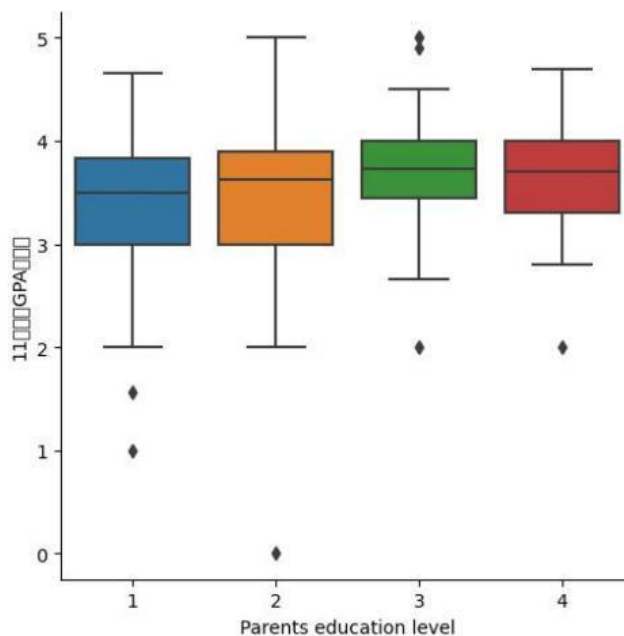


Fig 2. Block diagram of average GPA scores of students in different groups

Note: 1 represents students whose parents education level is “middle school”; 2 represents students whose parents education level is “high school”; 3 represents students whose parents education level is “undergraduate”; 4 represents students whose parents education level is “above undergraduate”; The y-axis is the average GPA score of these students)

In order to find out the relationship between parents’ highest education and grades. We decided to use the ANOVA test (The function of the ANOVA (Analysis of Variance) test is to compare whether there are significant differences in the mean values between multiple groups. It is a statistical method used to analyze the differences between numerical data observed under different conditions.) to test whether the college students of parents with different educational levels will have different grades.

Step1: The null hypothesis: H_0 : There is no significant difference in the average GPA score of college students whose parents have different levels of education.

The alternative hypothesis: H_1 : The college students whose parents have different education level have different GPA score.

Step2: The whole sample size is: $n=198$, the total degree of freedom is: $df=n-1=197$. With the number of groups is: $c=3$, so the degree of freedom among groups is: $df_1=2$. The degree of freedom within groups is: $df_2=n-c-1=195$

Step3: Here we assume the average GPA score distribution is normal distribution, and the population of each groups have equal variance. With we need to compare 3 groups, we decide to use ANOVA test

Step4: here is the result of ANOVA test by EXCEL.

Table 1. The parents education level in relation to the GPA score of the students ANOVA test

Source of variance	SS	df	MS	F	P-value	F crit
Among groups	3.560432	3	1.186811	2.579884	0.05483	2.651153
Within groups	89.24484	194	0.460025			
Total		197				

We can find the the P-value is 0.05483

Step5: If we choose the significant level: $\alpha=0.1$, we can get that the P-value is lower than α of 0.1, so, we have enough evidence to reject null hypothesis.

From the results, we have a 90% probability that the average GPA score of college students has a certain relationship with the education level of their parents. And it has literature shown that there is a significant positive relationship between the educational level of the parents and the educational level of the children [4]. But if we set the significant level: $\alpha=0.05$, we cannot reject the null hypothesis. So this statement is not absolute.

2.2.2 The study time

In life, we can often hear that hard work brings a harvest, perhaps this is also the factor that most people think affects grades, so the time for learning is also within our consideration.

We all generally think that the longer the study time, the better the grades, but this is often not the case in our own study life (our own feelings). So We analyzed the academic performance of these students, and the following are the results of our analysis.

We divided the study time of the students into groups, which were 1-3h, 4-6h, 7-9h, and above 9h. four groups. Define the average GPA score of students in each group as μ_i ($i=1,2,3,4$). In order to study whether the study time will have an impact on academic performance, we decided to first use a kdeplot to roughly compare the groups with average GPA. This is a plot comparison of the average GPA scores of different groups we made with python(5.1). Through intuitive experience, this graph is not as generally expected that the group that wants to study for a longer time has a higher average GPA score. Instead, the average GPA score of each group presents a similar shape. From here it is not seem as the more time you spend on study, the better grade you will have. Just looking at the chart, it seems that there is no relationship between the two, so let's test it in a statistical way to see if we can find evidence of a relationship.

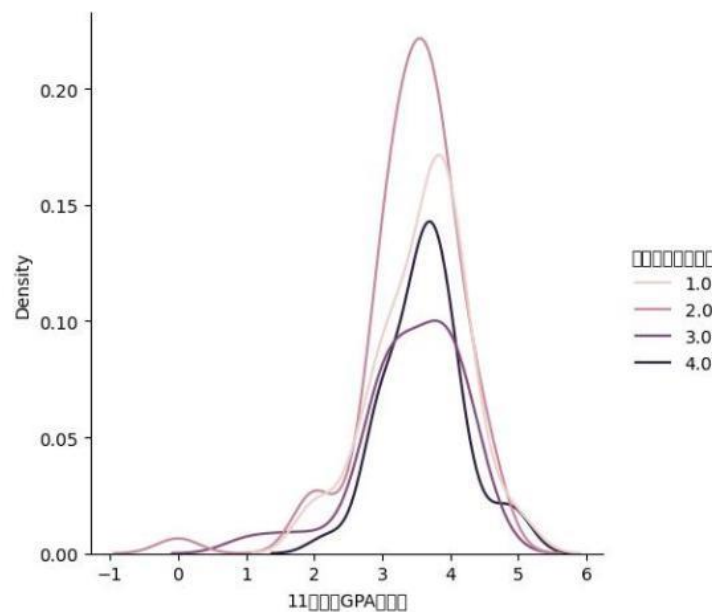


Fig 3. Comparison Chart of Average GPA of Different Groups

Note: 1 represents students whose study time is “1-3h”; 2 represents students whose study time is “4-6h”; 3 represents students whose study time is “7-9h”; 4 represents the students whose study time is “>9h”; The x-axis is the average GPA score of these students.

Because there has three variables here, so we decided to use ANOVA test (Analysis of variance (ANOVA) is a collection of statistical models and their associated estimation procedures used to analyze the differences among mean) test to check the results. The following are the analysis results we got through SPSS. (We assume the distribution of their grade in each study groups as normal distribution)

Step1: We set two Hypothesis:

Null hypothesis: H0: The study time have no significant effect on the grade of college students, the average GPA score in each groups is appremixly the same

Alternative hypothesis:H1: The study time truly have some effect on college students grade, at least two proportion mean are different.

The number of groups: $c=4$, sum of sample size from all populations: $n=199$

Step2: $df_1=c-1=3$, $df_2=n-c=194$ (df_1 , and df_2 is the degree of freedom of among groups and degree of freedom within groups)

we choose the significance level, $\alpha=0.05$

Step3: Since our sample size is 3, which is larger than two variables, we use ANOVA test (we assumed the population is normal distribution and are average GPA score in each groups is independent, and the three population have equal variance.)

Step4:

Table 2. The study time in relation to the GPA score of the students in AVOVA test

Source of variance	SS	df	MS	F	P-value	F crit
Between groups	1.178656	3	0.392885	0.831851	0.477825	2.651153
Within groups	91.62662	194	0.472302			
Total	92.80527	197				

Step5: We can find that the p-value of the samples is 0.478.

Step6: The p-value is larger than the significance level of 0.05, We do not have enough evidence to reject H0.

This result shows that base on this data, we have no enough evidence to say that there is significant relationship between the length of study and the GPA score of college students too.

In fact, this result is not surprising. As early as 2004, some scholar concluded that the regularity of learning is an important factor in academic performance, while learning time is only a secondary factor.[5]. And also, this study experiment also found that people with high grades spent the most time studying, followed by those with low grades, and those with medium grades were the least, similar to the results of the histogram we made there. So although we also can't prove there has no relationship between the study time and grade, we can at least think that the common view that the length of study is the key to determining the grades of college students is not always suit in all situations such as for college students' grade.

2.2.3 Satisfaction with the major

There is a proverb "Interest is the greatest teacher". In most people's opinion, the more interested you are, the easier it is for you to achieve success in this area. In a research experiment on non-intellectual factors, 100 sophomores were investigated through questionnaires on non-intellectual factors, teacher introductions, and examination results. It was found that among non-intellectual factors, the score of interest is 76, ranked 11th [6], which effectively shows that interest is a significant factor in determining grades. Satisfaction with your major is similar to interest in your major in some respects, so is it the same as interest, people who are satisfied with their major are more likely to get good grades?

But here we have to consider this issue, that is, if the student's satisfaction with the major is related to the student's willingness to fill out the questionnaire, this may lead to deviations in the results of the analysis, so this analysis is only for reference.

In the statistical process, the professional satisfaction is divided into 5 categories, very satisfied, satisfied, unsatisfied, unsatisfied, and dissatisfied. For simplicity, and to make sample size in each group in large enough, we classify very satisfied and satisfied into the "satisfied" category, and the rest are classified into "other" categories.

Same as the previous module of counting parents' education, classify average GPA score above 3 as "good", below as "normal", according to the previous theoretically speaking, the proportion of

good in the “satisfied” group should be higher than that of the “other” group. Through the result (Tab 3, Fig. 4), we found that the “satisfied” group good proportion is 81.8%, and the “other” group is 73.7%. We draw a preliminary conclusion that the degree of satisfaction (interest) to the major has an positive impact on the grades of college students. In order to prove the rigorousness of the conclusion, we verify our conclusions through hypothesis text:

Table 3. The frequency table and proportion chart of satisfaction with the major and whether the student’s grades are good

Frequency	Good	Normal	Total
Satisfied	81	18	99
Others	73	26	99
Total	154	44	198

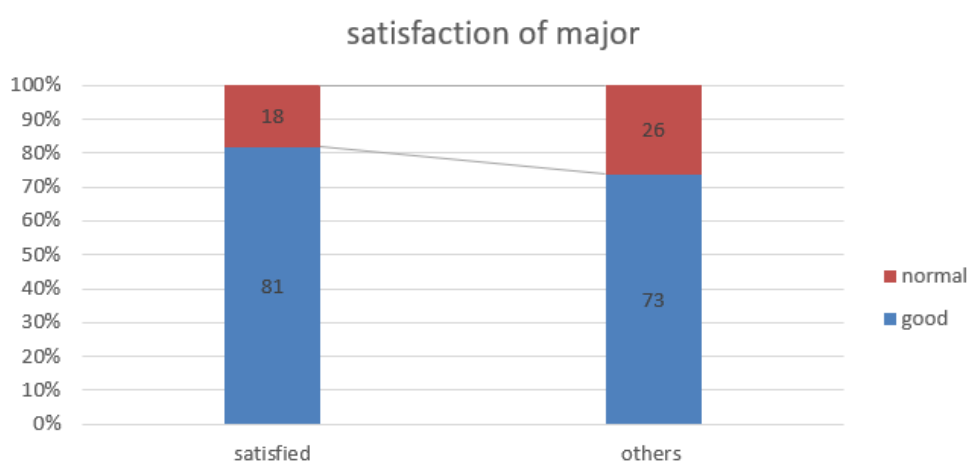


Fig 4. Comparison of satisfaction between different groups

Hypothesis text (with significant level: $\alpha=0.1$)

Step 1: set hypothesis:

Null hypothesis (H0): The “good” proportion of the “satisfied” group is no more than the good of the “others” group. $\pi_1 - \pi_2 \leq 0$ (π_1, π_2 is the proportion of “good” in “satisfied” group and “others” group respectively)

Alternative Hypothesis (H1): The good rate of the “satisfied” group is higher than that of the “others” group. $\pi_1 - \pi_2 > 0$

Step 2: Refer to the test result (1.4), we can get the “good” rate of the “satisfied” group: $p_1 = 81/99 = 81.8\%$, the “good” rate of the “others” group: $p_2 = 73/99 = 73.7\%$, and we calculate the good rate of the two groups: $p = 154/198 = 77.8\%$. We can also get the sample size of the “satisfied” group, $n_1 = 99$, and the sample size of the “others” group, $n_2 = 99$. The whole sample size, $n = n_1 + n_2 = 198$.

Step3: Because we need to test the proportion different between two groups, and both np and $n(1-p)$ is bigger than 5, we decide to use two sample z test.

Step4: With the significant level is 0.05, we can get the critical z value: $z\alpha = 1.28$

Step5: We compute the statistic z value:

$$z\text{-stat} = \frac{p_1 - p_2 - (\pi_1 - \pi_2)}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 1.371 \tag{2}$$

Step6: we can find that the z-stat value 1.371 is greater than the critical z value $z\alpha = 1.28$, so here we have 90% confidence to say that satisfaction of the major have significant influence on the college students grade.

This hypothesis test shows that the satisfaction with the major probably will affect the academic performance. So, It is reasonable to think that the performance of college students and the degree of college students' satisfaction have some relationship.

2.2.4 Lunch price

It seems that there will be such a phenomenon in life. There will always be people who link family conditions with students' grades. Here we will explore whether there is a direct relationship between the two. Since it is inconvenient to obtain the family conditions of the students, we use the lunch price of the students to indirectly measure the family conditions of the students. There will inevitably be some errors in such an operation, so here we just roughly analyze the relationship between the two.

Here, we divide the daily lunch expenses of college students collected into two groups. The first group is those who spend on lunch more than 20 yuan, and the second group is those who spend less than 20 yuan. We also divide base on the average GPA scores into two groups too, some of which are higher or equal to the average grades, and some of them are lower than the average grades. Through the table of statistics (Tab 4), the results show dramatically that those who spend more than 20 yuan seem to have a higher proportion of the grades are higher than the average. So can we conclude that the price of lunch is linked to the grades of college students?

Table 4. Frequency table of lunch prices and whether grades are above average

Price of lunch/Yuan	Above average	Lower	Total
<20	86	84	170
>20	18	10	28
Total	104	94	198

For this reason, we decided to use the hypothesis test method to test our conjecture

Step1, we set the hypothesis: Null Hypothesis (H0): The proportion of "above average" in the ">20" group is equal to the proportion of "above average" in the whole sample.

Alternative Hypothesis(H1): In ">20" group has a higher proportion of "above average".

Step 2, we choose the significant level, $\alpha=0.05$ and we can get the critical Z-value is 1.96 and we can set the not reject area of $[-1.96, 1.96]$. We calculate the proportion of "above average" in the ">20" group, $p=18/28=0.643$. The proportion of "above average" in the whole sample space, $p_0=104/198=0.525$. The size of ">20" group sample is $n_1=28$, the size of the sample space is $n_2=198$.

Step3, we do the calculation of the test for p:

$$Z_{\text{stat}} = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = 1.244 \quad (3)$$

We compare Z_{stat} 1.244 with critical Zvalue 1.96, which is lower than critical Z-value, so in no-reject area.

Step 4, so, we can conclude that we do not enough evidence to reject the null hypothesis, the proportion of "above average" in the "good group" is the same as in the whole sample space.

According to the test results, we cannot trust that the price of lunch is linked to the grades of college students base on the data. Stories like this are not uncommon in life. The poor have achieved excellent results through hard work, so being rich in life is not an absolute condition for determining success in learning. Here we believe that there is no basis for linking living conditions with college students' success in learning.

2.2.5 Preview habit

From childhood to adulthood, the preview is always assigned to us by the teacher as homework. But in college, since the preview is no longer mandatory, this habit seems to be dropped by us. So will the pre-study habits we have left behind have something to do with the decline in the grades of our college students?

According to these 198 samples, through the list of average scores (Tab 5), we can clearly see that the scores of students who have the habit of previewing are slightly higher than those of students who do not have the habit of previewing. And according the the confidence interval we calculate in the

beginning (3.539,3.621). Which is out of the range ,so we initially think that this is a factor affecting the performance.

We define students whose average GPA score is higher than 3 as "good", and students whose grades are lower than 3 as "normal". Here we can feel that in this sample (Tab 6), the "good" rate of the students who have previewed is obviously higher than that of the students who have not previewed.

Table 5. Table of preview habits and average GPA scores

Preview habit	Average GPA score
No	3.401555556
Have	3.767777778
Total	3.518080808

Table 6. stock columns for preview habit and "good" rate

Porportion of "good" preview habit	"good" number		
	Good	Normal	Total
No	97	38	136
Have	57	6	63
Total	154	44	198

Here, we use two sample hypothesis texts to prove our inference: We choose the significant level 0.05, and we assume the standard deviation in each is the same.

Step1: we set the null hypothesis: H0: There have no significant different for mean of their GPA average score in two groups (Whether the students have the habit of the preview.)

the alternative hypothesis:H1: There have significant different of these students in two groups

Step2: The mean of the students average GPA score who have the habit of preview is: $x_1 = 3.768$, the sample size is $n_1 = 63$, degree of freedom is: $df_1 = 62$. The mean of the student average GPA score who have no habit of preview: $x_2 = 3.402$, the sample size is $n_2 = 135$, degree of freedom: $df_2 = 134$. Of

all, we can also calculate the Pooled variation $Sp^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 - 1 + n_2 - 1} = 0.381$

Step3: With the standard deviation in two groups in unknown, we decide to use the test.

Step4: With the significant level we choose is 0.05, we can get the critical t value is: 1.972

Step5: We compute the t statistic value is:

$$t\text{-stat} = \frac{x_1 - x_2}{\sqrt{S_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} = 3.882 \quad (4)$$

Step6: We can see that the t-stat is higher than the critical t value of 1.972, so we have enough evidence to say that there have significant differences in the two groups for whether have the habit of the preview.

We can also use the same approach to prove the "good" rate in have preview habit groups is higher.

So, Therefore, it is very reasonable to say that preview is a factor in determining grades. Students who have the habit of preview generally have better grades. Therefore, if you want to improve your grades, it is very important to develop good preview habits.

3. Conclusion and discussion

In this essay, we verified the influencing factors of college students' academic performance. Since we cannot guarantee the reliability of these data 100%, our analysis here is only for reference.

First of all, we found that the GPA scores of most college students are concentrated between 3.539 and 3.621.

As for the family's education level, we found that the family's education level has a certain impact on the performance of college students, which may reflect the impact of family background on students' learning and development. It may be that families with higher educational backgrounds have more advantages in these students' learning attitudes, learning resources, and learning environment, which makes the final grades of college students different.

We did not demonstrate a significant relationship between study time and academic performance in our experiments, which may imply that study time is not the only determinant. However, through observation, we can see that the students who study the longest still have the highest academic performance, so it is very likely that the study time is also a factor of academic excellence.

Regarding professional satisfaction, we found that it also has a certain impact on student achievement. Students with high levels of satisfaction are likely to be more engaged in learning, more interested in learning content, and thus achieve better grades. Therefore, choosing a major of interest, or increasing the interest of the major, may also be used as one of the ways to improve the grades of college students.

The price of lunch has not been found to have a significant impact on grades in the analysis, which may mean that we cannot think that family conditions have no significant impact on college students' grades. But we cannot be sure that this is the case, because the price of lunch is not a direct measure of household conditions, and the analytical tool we used, the hypothesis test, does not prove that there is no relationship, it only proves that we cannot be sure that there is a relationship between the two.

We found that the habit of previewing has a significant impact on student performance. The confidence level in the hypothesis test is over 95%, which means that getting into the habit of previewing is a very good way to improve your grades. Previewing can improve students' understanding of the course content, and can help students keep up with the teacher's progress in class, and can complete the course content better and more efficiently in class. Therefore, developing the habit of previewing is a very good method to improve grade.

This set of analysis is not perfect, for example, it does not consider the relationship between various factors, the reasons for the analysis are only a small part of the many reasons, and the credibility of the data is also questionable. This paper hopes to be used as a reference for college students to improve their grades.

References

- [1] Prashant Loyalka, Ou Lydia Liu, Guirong Li, Elena Kardanova, Skill levels and gains in university STEM education in China, India, Russia and the United States, July 2021
- [2] Zhou Shijun, Li Qingyao, Cui Lizhi, parents 'Education and Children's Education —— Empirical investigation based on CGSS micro data [J]. *Education and Economics*, 2018 (3): 46-53
- [3] (Hao Huijuan, Students' Mathematics Learning Motivation and its Influence on Mathematics Performance, Primary and Secondary Education, No.20,2020,2020 / 10 / 19
- [4] Sun Yongqiang, Yan Yan, "Research on the urban-rural differences in intergenerational transmission of education in China: An empirical analysis based on the China Family Tracking Survey (CFPS) [J]. *Journal of Beijing Normal University (Social Sciences Edition)*, 2015 (6): 59-67
- [5] Li Guanghai, Cui Qunfa. The correlation between academic performance and study time among college students [J]. *Journal of Anyang Normal University*, 2004 (2)
- [6] Li Jin, Analysis on the Factors Affecting College Students' School Record, "Journal of Tianjin Adult Higher Education Union, Issue 3, 2005, 55-59