

# The Impact of Internet Usage on Students' Cognitive Levels: An Empirical Study Based on CEPS

Zihan Tu

*Jiangxi Normal University, Nanchang, Jiangxi, China*

*\*Corresponding Author*

**Abstract:** The 21st century is an era of rapid information development, where the application of the internet in education has become a trend, or even an educational investment. Many scholars believe that the use of the internet in education can enhance students' learning outcomes, while its impact on students' cognitive abilities has also emerged as a focal point in educational research. This study is based on the follow-up survey of the China Education Panel Survey (CEPS 2014-2015) data provided by Renmin University of China. Using multiple linear regression as the research method, it explores the impact of internet usage time on students' cognitive abilities. The findings reveal a significant and robust negative effect of weekly internet usage time on the development of students' cognitive abilities, suggesting that educators should guide students in reasonably allocating their online time in educational practice. It also highlights the importance of cultivating healthy internet usage habits among students.

**Keywords:** Network; Education; Cognitive Ability; CEPS

## 1. Introduction

The 21st century is an era of rapid development of information technology. The popularity and in-depth application of the Internet are profoundly reshaping the social ecology, especially for the growth and development of young people. In this context, the application of the internet in the field of education has become an irreversible trend and is even seen as an important investment in education<sup>[5]</sup>. The deep integration of digital technology has not only changed the way knowledge is disseminated, but also brought new opportunities and challenges for educational equity and quality improvement. The national level attaches great importance to the coordinated development of digital

technology and education. For example, the report of the 20th National Congress summarized the achievements of the past decade, pointing out that we have faced challenges such as the rampant chaos of online public opinion, highlighting the urgency of network environment governance<sup>[6]</sup>; The "15th Five Year Plan" further clarifies the need to "deepen the construction of digital China" and "guide and regulate the healthy development of online literature, online games, online audio-visual, etc., and strengthen the protection of minors' networks"<sup>[7]</sup>. These policy orientations provide a solid institutional basis for exploring the impact of Internet use on young people, and also highlight the real need to balance technology dividends and potential risks<sup>[8]</sup>.

However, despite the continuous improvement of network access rates, there is still controversy over their actual effects on student development. On the one hand, the Internet can be used as a tool to expand learning resources and stimulate innovative thinking; On the other hand, excessive use or entertainment tendencies may squeeze the time and psychological resources needed for cognitive development. In this context, in-depth analysis of the relationship between internet time, a key behavioral indicator, and cognitive ability has important theoretical value and practical significance<sup>[9]</sup>.

## 2. Problem Statement

With the deep integration of digital technology, teenagers' online time is increasing, and teenagers' Internet use behavior shows new characteristics such as universality, younger age, mobility, etc<sup>[10]</sup>. The Internet access rate of primary and secondary school students in China has exceeded 90%, and the weekly online time has become an important indicator to measure their digital life. However, empirical research on the relationship between online time and cognitive ability shows a complex picture. Some scholars' research has revealed the relationship

between Internet entertainment preferences and students' cognitive ability development, and found that Internet entertainment preferences are negatively correlated with students' cognitive ability<sup>[1]</sup>; Scholars have explored the relationship between information technology and student literacy, and explored the heterogeneous impact of information technology on academic performance<sup>[2]</sup>; Another study has begun to focus on the impact of information technology on the cognitive abilities of middle school students, and has attempted to use multi-layer linear models and other mechanisms for analysis<sup>[3]</sup>. Although scholars have gained some research experience in discussing the impact of information technology utilization on cognitive abilities, research on the influence of internet time, a key behavioral indicator, on cognitive abilities is still worth discussing and studying, providing a more scientific basis for policy implementation and formulation.

Based on existing research, this study proposes the following hypothesis: Internet time has a significant negative impact on students' cognitive abilities.

### 3. Research Design

#### 3.1 Data Source

The data used in this article is from the China Education Panel Studies (CEPS) 2014-2015 follow-up data provided by Renmin University of China. The survey targeted the adolescent population in junior high school and selected 28 counties, 112 schools, 438 classes, and 19487 students from across the country. Information on the growth of all surveyed students was collected, including personal basic information, school learning situation, and internet usage. Additionally, CEPS specially designed test questions to test students' cognitive abilities, providing researchers with rich data basis. This article uses the latest 2014-2015 academic year tracking survey data from the CEPS database as the research data. In addition, samples with missing key variables were excluded, resulting in a remaining sample of 8414.

#### 3.2 Selection of Variables

The dependent variable selected in this article is the cognitive ability of students, and the standardized cognitive ability score provided by CEPS is used. The cognitive ability test in CEPS involves three dimensions: language, graphics, and quantitative reasoning. Then, a three parameter IRT model is used to standardize the cognitive ability of students and obtain the score. The core independent variable selected in this article is the weekly internet week duration of students. In addition, this article further added baseline cognitive ability as the core variable, and variables such as family background and individual student characteristics as other control variables in the research process.

#### 3.3 Data Processing

For the dependent variable(cognitive\_ability), this study directly used the standardized cognitive ability score from CEPS 2014-2015, which has been calculated using a three parameter IRT model and is reliable and comparable. For the core independent variable of internet time, this study selected the median of the answer time range based on the two questions of internet time from Monday to Friday and internet time on weekends in the questionnaire, and then added them up to obtain the continuous variable of total internet time per week (internet\_weekly). For the core control variable base cognitive ability (base\_cog), the standardized score of cognitive ability from CEPS2013-2014 was directly used. The assignment operations related to other control variables are shown in Table 1.

At the same time, in order to exclude the influence of siblings in the family on students' individual cognitive ability, this study decided to use the propensity score matching method (PSM) proposed by Rosenbaum and Rubin<sup>[4]</sup> to correct errors, and divide the control group and the treatment group according to whether they are the only child. The only child is the treatment group, and the non only child is the control group. The overall data processing is shown in Table 2.

**Table 1. Assignment of Other Control Variables**

| Variable                 | Symbol | Description                                   | Assignment   |
|--------------------------|--------|---|--|
| Only child               | Child  | Are you an only child                         | Yes=1, No=0  |
| Educational expectations | Expect | To what extent do you want your child to read | It doesn't matter=0, don't study now=1, graduating from junior high school=2, vocational school/technical school=3, vocational high school=4, regular high school=5, college diploma=6, undergraduate degree=7, master's degree=8, doctoral degree=9 |

|                             |        |   |   |
|-----------------------------|--------|---|---|
| Family economic status      | Status | The current economic conditions of the family | Very difficult=1, relatively difficult=2, moderate=3, relatively wealthy=4, very wealthy=5  |
| Household registration type | Hukou  | Type of student's household registration      | Rural household registration=1, non rural household registration=2  |
| Gender                      | Male   | Student gender                                | Male=1, Female=2  |
| Father's education status   | Fa_edu | The education level of the student's father   | No education=1, primary school=2, junior high school=3, vocational school/technical school=4, vocational high school=5, high school=6, college diploma=7, undergraduate=8, graduate and above=9 |
| Mother's education status   | Mo_edu | The education level of the student's mother   | No education=1, primary school=2, junior high school=3, vocational school/technical school=4, vocational high school=5, high school=6, college diploma=7, undergraduate=8, graduate and above=9 |

**Table 2. Overall Data Processing**

| Variable name                   | Variable description                                     | Processing Group |                    | Control Group |                    | Bilateral t-test |
|---------------------------------|--|------------------|--------------------|---------------|--------------------|------------------|
|                                 |  | Mean             | Standard deviation | Mean          | Standard deviation |                  |
| Base period cognitive ability   | CEPS2013-2014 Standardized Score for Cognitive Abilities | 0.228            | 0.845              | -0.086        | 0.853              | -16.8781***      |
| Current cognitive ability       | CEPS2014-2015 Standardized Cognitive Ability Score       | 0.507            | 0.169              | 0.198         | 0.819              | -17.7740***      |
| Cognitive Increment             | Current cognitive ability - base cognitive ability       | 0.279            | 0.790              | 0.284         | 0.808              | 0.3023           |
| Weekly internet usage time      | Total online and gaming time per week                    | 15.397           | 8.172              | 14.641        | 8.752              | -4.0844***       |
| Family economic status          | The current economic conditions of the family            | 2.941            | 0.473              | 2.704         | 0.655              | -19.1766***      |
| Educational expectations        | To what extent do parents want their children to read    | 6.761            | 1.802              | 6.319         | 1.992              | -10.6522***      |
| Household registration type     | Types of student household registration                  | 1.682            | 0.466              | 1.313         | 0.464              | -36.1734***      |
| Gender                          | Student gender   | 1.470            | 0.499              | 1.522         | 0.500              | 4.7696***        |
| Father's Education Status       | Student father's level of education                      | 5.061            | 2.116              | 3.685         | 1.671              | -32.5029***      |
| Mother's educational background | The education level of the student's mother              | 4.793            | 2.085              | 3.242         | 1.600              | -37.5344***      |
| Only child                      | Are you an only child                                    | 1.000            | 0.000              | 0.000         | 0.000              | 0.000***         |

Note: \* \* \* < 0.01

### 3.4 Research Model

#### 3.4.1 Benchmark OLS model

In order to estimate the average impact of internet time on cognitive ability, the following multiple linear regression model was established (without controlling for baseline cognitive ability):

$$Cognitive\_ability_i = \alpha + \beta_1 Internet_i + \beta_2 Controls_i + \mu_i \quad (1)$$

In equation(1), subscripts 'i' and 't' respectively represent families and individual students;  $Cognitive\_ability_i$  is the dependent variable, representing the standardized score of cognitive ability for the i-th adolescent in the current period;  $Internet_i$  is the core independent variable of this study, which represents the average weekly total time (hour) spent online by the adolescent during the observation period; The core coefficient is  $\beta_1$ , which measures the 'net impact' of weekly internet time on cognitive

ability after controlling for other factors;  $Controls_i$  is a set of control variables, including: whether one is an only child, family economic status, parental education expectations, household registration type, father's education status, mother's education status, and gender;  $\mu_i$  is the random error term of the equation.

#### 3.4.2 OLS regression model after propensity score matching

To more accurately estimate the treatment effect of internet usage and simultaneously assess the causal effect of only-child status on cognitive ability, the following PSM-matched OLS regression model was established:

$$Cognitive\_ability_i = \alpha + \beta_1 Child_i + \beta_2 bas\_cog_i + \beta_3 Controls_i + \mu_i \quad (2)$$

In equation(2),  $Cognitive\_ability_i$  is the dependent variable, representing the standardized score of the i-th adolescent's current cognitive ability;  $Child_i$  is the core

processing variable in this study and a dummy variable;  $\beta_1$  is the core coefficient, which measures the average processing effect (ATT) of being an only child on cognitive ability after balancing sample selection bias through propensity score matching (PSM) and controlling for other factors;  $Bas\_cog_i$  is the key control variable, representing the standardized score of cognitive ability in the  $i$ -th adolescent base period. Controlling for this variable aims to strip away the influence of initial ability level and make  $\beta_1$  closer to the causal estimation of cognitive ability "growth";  $Controls_i$  is a set of covariates used in the matching process, which are ultimately entered into the regression equation through sample weights, including weekly internet time, family economic status, parental education expectations, household registration type, father's education status, mother's education status, and gender.

### 3.4.3 Value added model for controlling base period capability

In order to separate the influence of students' initial ability level, this study directly estimated the effect of internet time on cognitive ability growth, and then constructed the following educational value-added model:

$$Cognitive\_ability_i = \alpha + \beta_1 Internet_i + \beta_2 base\_cog_i + \beta_3 Controls_i + \epsilon_i \quad (3)$$

In equation(3),  $base\_cog_i$  is the key control variable, representing the standardized score of student  $i$ 's baseline cognitive ability; After introducing this variable, the estimation of coefficient  $\beta_1$  is closer to the impact of internet time on the increase or appreciation of students' cognitive abilities.

### 3.4.4 A robust model with capability increment as the dependent variable

As a robustness test for the core conclusion, the dependent variable is replaced with a clear cognitive increment, and the model is set as follows:

$$\Delta Cognitive\_ability_i = \alpha + \gamma_1 Internet_i + \gamma_2 Controls_i + v_i \quad (4)$$

In equation(4),  $\Delta Cognitive\_ability_i$  representing the cognitive increment of student  $i$ ;  $\gamma_1$  is the core coefficient of concern in this model, which directly measures the impact of internet time on the growth rate of cognitive ability.

### 3.5 Heterogeneity Test

In order to examine whether the impact of internet time on students' cognitive abilities varies with other variables, this study also conducted heterogeneity research, studying the heterogeneity of variables such as gender, family economic status, education expectations, household registration type, and parental education level.

### 4. Research Findings

This study empirically tested the impact of weekly internet time on the cognitive development of middle school students using a multiple linear regression model. At the same time, in order to obtain more robust estimates, the analysis process used sampling weights for CEPS data to ensure the representativeness of the results, and adopted a strategy of gradually adding control variables. The estimated results of the final benchmark regression are shown in Table 3.

**Table 3. Benchmark Regression Results**

| Variable                    | Equation(1)      | Equation(2)      | Equation(3)      | Equation(4)     |
|-----------------------------|------------------|------------------|------------------|-----------------|
| Constant term               | -0.991*(0.053)   | -0.443(0.100)    | -0.479**(0.047)  | 0.176*(0.052)   |
| Only child                  | 0.130**(0.019)   | 0.086**(0.023)   | 0.069**(0.017)   | -0.008**(0.020) |
| Weekly internet usage time  | -0.004***(0.001) | -0.005***(0.002) | -0.003***(0.001) | -0.002**(0.001) |
| Family economic status      | 0.110**(0.015)   | 0.041**(0.025)   | 0.055**(0.013)   | -0.016**(0.016) |
| Educational expectations    | 0.107***(0.006)  | 0.068***(0.007)  | 0.071***(0.005)  | 0.025***(0.005) |
| Household registration type | -0.008**(0.020)  | 0.034**(0.024)   | -0.013**(0.017)  | -0.019**(0.020) |
| Basic cognitive ability     | —                | 0.420**(0.016)   | 0.439**(0.010)   | —               |
| Gender                      | 0.020**(0.016)   | 0.034**(0.023)   | 0.022**(0.014)   | 0.024**(0.018)  |
| Father's education status   | 0.042***(0.006)  | 0.022***(0.008)  | 0.025***(0.005)  | 0.003***(0.006) |
| Mother's education status   | 0.036***(0.006)  | 0.016***(0.007)  | 0.021***(0.005)  | 0.001***(0.006) |
| R <sup>2</sup>              | 0.157            | 0.326            | 0.350            | 0.005           |
| Observed value              | 8414             | 5508             | 8414             | 8414            |

Note: The parentheses indicate the robust standard deviation; \*\*\*<0.01, \*\*<0.05, \*<0.1

### 4.1 Benchmark Regression Analysis

As shown in Table 3, equations (1) to (3) sequentially demonstrate the net impact of

weekly internet time (internet\_weekly) on standardized scores of students' cognitive abilities under the condition of continuously increasing control variables. Firstly, in the

benchmark OLS model of equation (1), this study preliminarily explored the relationship between internet time and cognitive ability. The estimation results showed that internet time had a significant negative impact on cognitive ability, with a coefficient of -0.004 and significant at the 1% statistical level. This preliminary result suggests that an increase in online time is associated with a decrease in cognitive ability scores, with an average decrease of approximately 0.004 standard deviations in students' standardized cognitive ability scores for every 1 hour increase in online time per week.

Secondly, equation (2) presents the estimation results for the Average Treatment effect on the Treated (ATT) of being an only child, based on the PSM-matched sample. After controlling for baseline cognitive ability, family economic status, parental education expectations, household registration type, parental education level, gender, and weekly internet time, the status of being an only child shows a significant positive effect on cognitive ability (coefficient = 0.086,  $p < 0.05$ ). This result indicates that, after balancing selection bias related to family background, the standardized cognitive ability score of only children is, on average, approximately 0.086 standard deviations higher than that of non-only children. The model also includes weekly internet usage time as a control variable, but its coefficient is not the focus of the analysis under this model specification.

Finally, equation (3) adds the most critical baseline cognitive ability variable on the basis of equation (2). This model is the most complete setting of this study, aimed at controlling students' initial ability level, so as to more accurately estimate the impact of internet time on cognitive enhancement. The final research results showed that the regression results of the value-added model were in line with theoretical expectations: base period cognitive ability showed a strong positive predictive effect on current period ability (coefficient=0.439,  $p < 0.01$ ). More importantly, after removing the interference of differences in students' initial

ability levels, the coefficient of weekly internet time remains significantly negative at the 1% level (coefficient=-0.003). This means that after excluding the influence of differences in students' original cognitive abilities, for every additional hour of online time per week, the standardized score of students' cognitive abilities will still decrease by an average of about 0.003 standard deviations. The explanatory power of this model ( $R^2=0.350$ ) has been significantly improved compared to the benchmark model, indicating that the model is more fully designed. In addition, as a robustness test, equation (4) replaces the dependent variable with the increase in students' cognitive ability ( $\Delta$  Cognitive\_ability). Although the coefficient of online time is still negative, its statistical significance has weakened. This may indicate that internet usage time has a certain impact on the short-term (one-year) growth rate.

#### 4.2 The Impact of Control Variables

In addition to the core independent variable, the results of other control variables also provide valuable findings. In addition, parents' educational expectations show a highly significant positive impact in all models, and are the most stable and strongest promoting factor for predicting students' cognitive abilities. This indicates that parents' expectations and investment in their children's academic performance can effectively translate into students' intrinsic motivation and academic performance. The educational level of parents also shows a significant positive effect in most models, jointly confirming the core role of family human capital in intergenerational transmission. In contrast, the influence of family economic status and household registration type became insignificant after controlling for baseline cognitive ability, suggesting that their effects may be more achieved by influencing students' initial abilities.

#### 4.3 Heterogeneity Test Results

**Table 4. Benchmark Regression Results**

| Variable                     | Coefficient | Standard error | t-value | p-value | 95% confidence interval |
|------------------------------|-------------|----------------|---------|---------|-------------------------|
| Weekly internet usage time   | -0.004      | 0.001          | -4.56   | 0.000   | [-0.0063,-0.0025]       |
| Only child                   | 0.130       | 0.018          | 7.01    | 0.000   | [0.0933,0.1658]         |
| Educational expectations     | 0.107       | 0.004          | 24.69   | 0.000   | [0.0989,0.1160]         |
| Family economic expectations | 0.110       | 0.014          | 7.58    | 0.000   | [0.0814,0.1381]         |
| Household registration type  | -0.008      | 0.019          | -0.42   | 0.672   | [-0.0449,0.089]         |

|                            |        |       |        |       |                  |
|----------------------------|--------|-------|--------|-------|------------------|
| Gender                     | 0.020  | 0.016 | 1.24   | 0.215 | [-0.0118,0.0525] |
| Father's educational level | 0.042  | 0.006 | 7.31   | 0.000 | [0.0305,0.0528]  |
| Mother's educational level | 0.036  | 0.006 | 6.20   | 0.000 | [0.0246,0.0474]  |
| Constant term              | -1.003 | 0.055 | -18.15 | 0.000 | [-1.118,-08951]  |
| Number of observations     | 8414   |       |        |       |                  |
| R <sup>2</sup>             | 0.1567 |       |        |       |                  |

**Table 5. Heterogeneity test results**

| Model Type                  | Absorption of internet time | Coefficient of interaction term | Interaction item p-value | R <sup>2</sup> | Is there heterogeneity present |
|-----------------------------|-----------------------------|---------------------------------|--------------------------|----------------|--------------------------------|
| Baseline model              | -0.0044***                  | —                               | —                        | 0.1567         | —                              |
| Gender                      | -0.0061***                  | 0.0037                          | 0.057                    | 0.1578         | Significant edge               |
| Household registration type | 0.0029***                   | -0.0034                         | 0.076                    | 0.1570         | Significant edge               |
| Educational expectations    | -0.0021***                  | -0.0004                         | 0.423                    | 0.1567         | not significant                |
| Only child                  | -0.0039***                  | -0.0012                         | 0.527                    | 0.1567         | not significant                |
| Family economic status      | 0.0058***                   | -0.0037                         | 0.018                    | 0.1572         | significant                    |
| Father's educational level  | -0.0009***                  | -0.0009                         | 0.086                    | 0.1570         | Significant edge               |
| Mother's educational level  | -0.0038***                  | -0.0002                         | 0.723                    | 0.1567         | not significant                |

Note: \*\*\*<0.01, \*\*<0.05, \*<0.1

According to the table 4 and 5, it can be concluded that there is significant heterogeneity in family economic status, indicating that the higher the socioeconomic status, the stronger the negative impact of online time. This may be because high status families originally had more educational resources, and online time occupied these resources; Household registration has significant marginal heterogeneity, and the influence of agricultural household registration groups is relatively small, which may be related to their limited educational resources; Gender and father's education level have marginal significance; Educational expectations, only child status, and maternal education level did not show significant heterogeneity effects. This result suggests that in families that originally had more educational resources, the internet may have more effectively squeezed out more beneficial cognitive development activities; Gender roles and knowledge transmission within the family may play a protective role, to some extent regulating the impact of internet time on cognitive abilities; The negative impact of internet time on cognitive ability is more related to family structural and resource factors rather than personal choices.

## 5. Conclusion

This study is based on the latest education tracking data provided by the China Survey and Data Center of Renmin University of China (CEPS2014-2015). Through multiple regression analysis, the study found that weekly internet time has a significant negative impact on cognitive ability, and this impact exhibits

obvious heterogeneity characteristics. Specifically, students from families with lower socioeconomic status and insufficient parental supervision are more significantly affected. This discovery provides empirical evidence based on a nationally representative sample to understand the relationship between information technology use and students' cognitive development, and demonstrates the importance of cultivating healthy internet usage habits. Firstly, it is suggested that teachers should guide students to allocate their online time reasonably in educational practice, with a focus on managing the internet usage of disadvantaged students; Secondly, schools should strengthen the positive guidance of information technology education; Finally, families need to establish scientific rules for internet usage, and parents should improve their digital literacy to better guide their children in using the internet reasonably. However, it should be pointed out that this study has certain limitations in terms of data, methods, and variable manipulation. Future research can improve the quality of the study by extending the tracking period, adopting precise experimental designs, enriching measurement tools, and testing nonlinear relationships. At the same time, research boundaries can be expanded from the perspectives of heterogeneity of online content, multidimensional development outcomes, and multi-level regulatory factors, in order to comprehensively and deeply understand the complex relationship between online behavior and development of young people in the digital age, and provide more scientific basis for educational practice and policy formulation.

**References**

- [1] FANG Chao , WANG Guxue , HUANG Bin.(2019)Can Information Technology Promote the Development of Students' Cognitive Ability? Net Effect Estimation Based on CEPS.Open Education Research , 25(04):100-110.
- [2] CHEN Chunjin , GU Xiaoqing.(2017).The Impact of Information Technology on Students' Literacy:A Perspective on Education Informatization.Open Education Research , 23(03):37-49.
- [3] SHI Huan-huan , XUE Hai-ping.(2022)Can Reducing Homework Time Exacerbate Inequality in Educational Outcomes? Evidence from the Empirical Study of CEPS.Journal of Educational Studies , 18(06):151-168.
- [4] Paul R. Rosenbaum , Donald B. Rubin , D.B.(1985).Constructing a control group using multivariate matched sampling methods that incorporate the propensity score.American Statistician, 24(4): 521-544.
- [5] Cao DANDan , Luo Shengquan , Yang Xiaoping , Wang Wentao.(2018)Effect of Internet on the Development of Cognitive Ability of Urban and Rural Teenagers.China Educational Technology , (11):9-17.
- [6] Guo Xiuqi.(2020)Research on the Influence of Educational Informatization on the Cognitive Ability Gap Between Urban and Rural Students.Huazhong Normal University.
- [7] CHEN Chunjin , GU Xiaoqing(2017).Does the Internet Expand the Inequality of Education Outcome:An Empirical Study Based on the Programme for International Student Assessment.Peking University Education Review , 15(01):140-153+191-192.
- [8] Liu Yujie.(2021)Research on the influence of online learning on the learning effect of primary and secondary school students.Jinan University.
- [9] HU Yongmei , FAN Wenfeng , DING Weili.(2015)Does Shadow Education Expand the Inequality of Educational Outcomes-An Empirical Study Based on PISA 2012 Shanghai Data.Peking University Education Review , 13(03):29-46+188.
- [10] Heck,R.H.(2007).Examining the relationship between teacher quality as an organizational property of schools and students' achievement and growth rates.Educational Administration Quarterly, 43(4): 399-432.