

Using the Regression Discontinuity Design (RDD) to Study the Impact of Changes in the Learning Environment on Students' Mental Health

The advancement of academic stages, such as transitioning from primary school to junior high school or from senior high school to university, is a crucial phase in the growth of teenagers. However, it is also a period when their social environment undergoes drastic changes. Students need to adapt to new peer groups, unfamiliar teachers, and different school systems, often losing their original friendships and familiar social support systems. These changes may trigger anxiety, loneliness, and psychological stress, especially having a more significant impact on children with weaker social adaptation abilities. If these psychological issues are not intervened in a timely manner, they may further develop into long-term psychological disorders, affecting students' academic performance, interpersonal relationships, and even their overall well-being. Addressing these issues not only helps to promote the healthy growth of children but also contributes to reducing the social costs associated with adolescent mental health problems. From a social science perspective, understanding how the structural changes in the education system affect the individual's psychological trajectory can help reveal the complex relationship between institutional transitions and human behavior.

Despite the great significance of this issue, causal research surrounding the "psychological problems caused by school transitions" remains relatively scarce, mainly due to difficulties in data acquisition. Mental health data often suffer from low reporting rates or inconsistent measurements, and longitudinal tracking of students

during the school transition process requires high time and economic costs. These obstacles limit the application of traditional empirical methods in identifying the impact of social environment shocks.

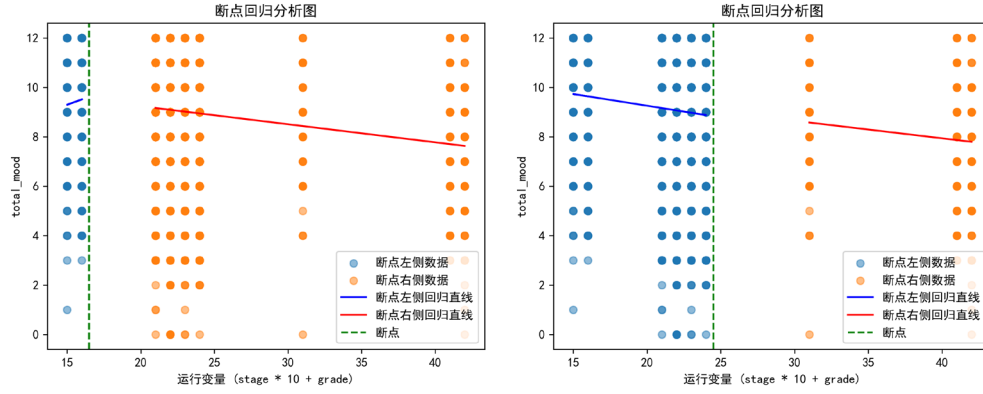
To overcome these limitations, this study will obtain data from reliable and trustworthy data sources and adopt the Regression Discontinuity Design (RDD) to identify the causal impact of school transition time points on students' mental health. By utilizing the "artificial cut-off points" in grade boundaries and school transition policies, RDD can provide robust causal inferences in the absence of randomized experiments.

The main question this study aims to answer is: To what extent does the change in the learning environment affect students' mental health? This question is helpful in further understanding the relationship between the learning environment and students' psychological development in the field of social science, refining the theoretical framework regarding the interaction between learning and psychological development in educational psychology, and providing a theoretical basis for subsequent research. In addition, this research question is also of great significance for machine learning, especially in applications such as predictive modeling of vulnerable students and the development of early warning systems. Therefore, understanding the social and emotional impact of school transitions not only promotes the development of education and psychology but also provides theoretical support for constructing machine learning systems with social awareness and ethical foundations.

This dataset is sourced from the social science database on Nature. The data

collectors obtained officially recognized student data through cooperation with the Spanish government. It collected data from a total of 5,890 students of various age groups in Spain, including basic socio-demographic statistical data and test questions for various abilities. This study mainly focuses on three questions related to mental health among them. Through the application of RDD, we can understand whether the change in the learning environment has an impact on students' mental health.

By sorting the academic stages (stage) and grades (grade) in the dataset, we mainly use (1,6)/(2,1) (transition from primary school to secondary school) and (2,4)/(3,1) (transition from secondary school to university/vocational school) as two cut-off points for separate RDD explorations. In terms of data preprocessing, since using the birth year to represent school transitions is not rigorous, the abscissa is replaced by stage*10 + grade. Thus, 16 and 21 are taken as the first cut-off point, and 24 and 31 are taken as the second cut-off point for regression discontinuity analysis. Calculate the total score total_mood (ranging from 0 to 12, with a higher score indicating a more positive mood) of the three emotion-related questions as the response variable for regression analysis. Use a local linear regression model. For the data on the left side of the cut-off point, construct the regression equation $Y_i = \beta_{0L} + \beta_{1L}X_i + \epsilon_i$; for the data on the right side of the cut-off point, construct the regression equation $Y_i = \beta_{0R} + \beta_{1R}X_i + \epsilon_i$, where Y_i is the outcome variable, X_i is the cut-off variable, β is the regression coefficient, and ϵ_i is the error term.



As observed from the charts, whether transitioning from primary school to secondary school or from secondary school to a higher-level learning environment, there is no obvious impact on students' emotions. However, as students advance in their learning stages, their overall emotional level shows a downward trend. In the correlation test, the calculated R^2 value is -0.1, indicating a negative correlation between the advancement of students' learning stages and their emotional levels, but it is not statistically significant. Therefore, the conclusion is that there is no need to be overly concerned that the changes in interpersonal relationships brought about by the change in the learning environment will affect students' mental health.

This study has made certain contributions to the existing literature in the fields of machine learning and social science. In the field of social science, it has broken through the dilemma of difficult data acquisition in traditional empirical methods when studying psychological problems caused by school transitions. By adopting the Regression Discontinuity Design (RDD) to identify the causal impact of school transition time points on students' mental health, it provides a new empirical approach for the research on the relationship between educational system transitions and individual psychological trajectories in educational sociology. However, despite the achievements of this study,

there are still limitations. The data only comes from Spain, and the geographical limitations of the sample may affect the generalizability of the conclusions. Future research can attempt to conduct similar studies in different countries and cultural backgrounds to compare the impact of school transitions on students' mental health under different educational systems and social and cultural environments. In addition, this study only focuses on the impact of changes in interpersonal relationships during the advancement of academic stages on students' mental health. Subsequent research can incorporate more variables, such as changes in the family environment and learning pressure, to comprehensively analyze the comprehensive impact mechanism of changes in the learning environment on students' mental health and expand the research boundaries of the interdisciplinary field of educational psychology and machine learning.

The results of this study have important implications for the real world. At the level of educational practice, it enables educators to understand that there is no need to be overly worried about the impact of changes in interpersonal relationships caused by school transitions on students' mental health. This helps them allocate educational resources more rationally and shift their focus more to aspects such as the adjustment of students' long-term learning pressure and the cultivation of learning abilities. From a social perspective, reducing excessive attention to changes in interpersonal relationships during the school transition stage can guide social resources to be more precisely invested in the key issues of adolescent mental health, such as intervening in the impact of factors like learning pressure and family environment on students' mental

health, promoting the overall improvement of adolescent mental health levels, and thus reducing the social costs associated with adolescent mental health problems and creating a more favorable social environment for the growth of adolescents.

The machine learning methods adopted in this study strictly adhere to the principles of AI governance during the implementation process. In the data acquisition stage, cooperation with the Spanish government to obtain officially recognized data ensures the legal and compliant source of the data, protects students' privacy, and practices the principles of ethical integrity. In addition, the results of this study can promote the inclusiveness of AI development and deployment, provide more equitable mental health monitoring and support services for students with different learning backgrounds and abilities, contribute to achieving the sustainable development goals in the field of education, and promote long-term prosperity and social well-being.

Appendix

In this study, we explored the impact of changes in the learning environment on students' mental health through the analysis of relevant data. However, there are still some aspects of the research methods that can be optimized. During the data collection phase, we used rating questions to measure students' mental health status. Although this method is convenient for quantitative processing, it has certain limitations, that is, it may lead to respondents not answering the questions seriously. For example, some students may randomly select answers due to boredom with the questionnaire, misunderstandings, or a casual attitude. This will undoubtedly distort the authenticity of the data and further affect the accuracy and reliability of the research results.

To overcome this potential problem, future research can try to introduce innovative methods such as natural language processing. Specifically, open-ended questions can be set in the questionnaire, allowing students to freely elaborate on their feelings and experiences during the process of changes in the learning environment. Subsequently, natural language processing technology can be used to quantitatively evaluate the emotional level of each respondent by identifying keywords, sentiment words, and semantic associations in the students' answers. For instance, if negative words such as "anxiety" and "loneliness" frequently appear in the students' answers, the system can assign a relatively low emotional score accordingly; if positive words such as "excitement" and "expectation" appear, a higher score can be given. In this way, not only can we obtain more authentic and in-depth information about students' psychological states, but we can also analyze students' mental health status from

multiple perspectives, making up for the deficiencies of relying solely on rating questions and providing a more solid data foundation and research perspective for studying the relationship between changes in the learning environment and students' mental health.