Application of Artificial Intelligence Methods in Investigating Homework Cheating Behavior and Its Key Predictive Factors among Elementary School Students

Authors: Zhao Li, Zheng Yi, Zhao Junbang, Zhang Rui, Fang Fang, FU Genyue, Li Kang, Zhao Li, Li Kang

Date: 2022-12-01T00:00:00+00:00

Abstract

Homework cheating among primary school students represents a long-neglected research priority within the field of psychology, while machine learning constitutes an emerging artificial intelligence science in the digital intelligence era. This study administered a questionnaire survey to 2,098 students in grades 2 through 6, employing machine learning methods to investigate the influence of individual cognition, moral judgment, peer behavior, as well as gender, grade level, and academic achievement on homework cheating behavior among primary school students. Results indicate that the ensemble machine learning model achieved a prediction accuracy (mean AUC) of 80.46% for homework cheating among primary school students; the four factors exhibiting the strongest predictive effects on homework cheating, in descending order, are individuals' acceptance of homework cheating, the prevalence and frequency of observed peer cheating, and their own academic achievement.

Full Text

The Application of Artificial Intelligence Methods in Examining Elementary School Students' Homework Cheating Behavior and Its Key Predictors

ZHAO Li[1], ZHENG Yi[2], ZHAO Junbang[3], ZHANG Rui[1,4], FANG Fang[5], FU Genyue[1], LI Kang[6]

[1] Department of Psychology, Hangzhou Normal University, Hangzhou 311121, China
[2] Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou

311121, China
[3] College of Child Development and Education, Zhejiang Normal University, Hangzhou 311231, China
[4] Hangzhou Xiayan Elementary School, Hangzhou 311121, China
[5] School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100871, China
[6] Ontario Institute for Studies in Education, University of Toronto, Ontario M5R 2X2, Canada

Abstract

Academic cheating has been a persistent challenge for educators throughout history, with homework cheating representing a significant yet understudied manifestation of this problem. Despite recent educational reforms, homework remains a central academic task for elementary school students in China. However, research over the past century has predominantly focused on college and secondary school students, neglecting the critical developmental period of elementary school when academic integrity begins to form. Moreover, existing studies have primarily examined exam cheating rather than homework cheating. The present research addresses this significant gap by employing advanced artificial intelligence methods to investigate the development of homework cheating among elementary school children and identify its key contributing factors, thereby providing a scientific foundation for developing early intervention strategies to promote academic integrity.

We surveyed 2,098 students from grades 2 through 6 using a comprehensive questionnaire. Machine learning techniques were employed to examine how individual cognitive factors, moral judgments, peer behaviors, and demographic variables such as gender, grade level, and academic achievement influence homework cheating behavior among elementary school students. The results demonstrated that the ensemble machine learning model achieved a mean prediction accuracy (AUC) of 80.46% for identifying homework cheating. The four strongest predictors were, in descending order: students' own acceptance of homework cheating, the perceived prevalence and frequency of peer cheating, and their own academic achievement.

Keywords: elementary school students, honesty, academic cheating, homework cheating, machine learning, prediction, peer behavior

Introduction

Academic cheating has troubled educators for centuries, representing a secretive and intentional violation of academic integrity standards undertaken to obtain desired grades or rankings. Homework cheating constitutes one of its primary manifestations. Although recent "double reduction" policies have significantly reduced academic burdens for elementary school students in China, homework remains the most important academic task for primary school students compared to Western contexts, and homework cheating remains prevalent among this population.

Elementary school represents a critical period for the formation of moral consciousness and the development of moral behavioral habits. Cheating behaviors during this stage can lead to poor academic performance and subsequent exam cheating, eventually becoming habitual. Moreover, if cheating behaviors during childhood are not addressed with timely guidance and intervention, they may lead to more serious problem behaviors in adulthood, such as lawbreaking or criminal activity (Williams & Williams, 2012). The "double reduction" policy emphasizes not only burden reduction but also improving homework quality and eliminating the "grade-only" mentality. Consequently, homework quality and other routine academic performance may become more important than exam scores in future evaluation systems. Therefore, this study provides an in-depth examination of the development of homework cheating behavior during elementary school and its key influencing factors, aiming to provide scientific evidence for understanding the developmental mechanisms of cheating and developing targeted early intervention methods.

Academic research on cheating has nearly a century of history (Hartshorne & May, 1928). However, the vast majority of existing empirical studies have been conducted abroad and have focused almost exclusively on college and secondary school students (see Anderman & Midgley, 2004; Cizek, 1999; Hrabak et al., 2004), with minimal attention to elementary school populations (Hartshorne & May, 1928). Although Hartshorne and May's 1928 study revealed that children in elementary school spontaneously engage in academic cheating, subsequent educational reforms in Western countries led to the elimination of exams and homework in primary schools, causing research on elementary school cheating to stagnate. Furthermore, existing research has concentrated primarily on exam cheating (Bong, 2008; Freire, 2014), with very few studies addressing homework cheating. To date, no research has specifically examined homework cheating behavior among elementary school students.

Domestic research on academic cheating in China began even later, focusing exclusively on college students (e.g., Shu et al., 2018; Yi, 2021), leaving empirical research on elementary school cheating, particularly homework cheating, completely unaddressed. While findings from college student research may provide some insights, elementary school students have significantly lower levels of moral cognitive development, making it unclear whether these results apply to younger populations. Therefore, this study addresses this gap by examining potential factors influencing current homework cheating behavior among Chinese elementary school students and identifying key influencing factors based on existing empirical research both domestically and internationally. The findings aim to provide important scientific evidence for constructing theoretical models of children's honesty behavior development and for improving and localizing moral behavior development theories.

Specifically, this study examines students from grades 2 through 6 using a large-sample questionnaire survey combined with machine learning methodology (Machine Learning, Pedregosa et al., 2011) to investigate grade-level developmental trends in homework cheating behavior and analyze the influence of various factors and their relative importance.

Machine learning represents an emerging data analysis method in the digital intelligence era and constitutes the core methodology of artificial intelligence. It relies on computer algorithms to simulate human behavior through data analysis and modeling. In recent years, psychological researchers have applied machine learning to studies of emotion (Just et al., 2017) and psychopathology (Bartlett et al., 2014; Livieris et al., 2018), with a few researchers beginning to apply these methods to child studies (Bruer et al., 2019; Zanette et al., 2016). Compared to traditional statistical modeling, machine learning offers four distinct advantages.

First, machine learning' s data processing approach enhances external validity. Traditional analytical methods (e.g., Generalized Linear Models, GLM; Generalized Estimating Equations, GEE) typically analyze all data simultaneously, often resulting in overfitting and poor generalizability. In contrast, machine learning (which requires larger sample sizes) divides data into three subsets: a training set, a test set, and a holdout set. The training set is used to fit the model, the test set to validate it, and these subsets are then recombined and randomly re-divided multiple times to generate multiple predictive models. Finally, the holdout set, which has not been used in any training or testing, is used to evaluate the models' predictive power on completely new data. This validation process assesses the model' s external validity (Campbell, 1986).

Second, machine learning encompasses multiple algorithms that can flexibly handle complex and variable relationships between variables. This study employs ensemble learning (Ensemble Learning, see Ykhlef & Bouchaffra, 2017), a recently popular approach that first trains different machine learning algorithms on the same sample. These include logistic regression (Logistic Regression, Yarkoni & Westfall, 2017) for linear relationships, and Multilayer Perceptron (MLP), eXtreme Gradient Boosting (XGBoost), and Random Forest for nonlinear relationships (see Golino et al., 2014). The training results from all algorithms are then integrated. The advantage of ensemble methods lies in their ability to combine the strengths of various algorithms to maximize explanation of the relationship between predictors and outcome variables, thereby achieving optimal predictive performance.

Third, machine learning can quantify the relative importance of different influencing factors through Shapley values. Proposed by Nobel laureate Lloyd Shapley in 1953 (Shapley, 1953), Shapley values measure the relative contribution of each predictor to the outcome variable and serve as an important reference indicator in machine learning results (Smith & Alvarez, 2021). Larger values indicate stronger predictive power, while values approaching zero suggest negligible predictive effect.

Finally, machine learning quantifies the predictive effects of all variables to create a predictive model for the outcome variable (e.g., "the probability of elementary school students cheating on homework"). Once established, inputting a student' s scores on the predictor variables yields the probability of that student cheating on homework. This enables preliminary prediction of cheating likelihood, allowing teachers and parents to provide necessary attention and targeted educational interventions.

Based on these advantages, this study examines the influence of selected factors on elementary school students' homework cheating behavior. Given the near-complete absence of research on predictors of homework cheating in elementary school, this study primarily references Murdock and Anderman' s (2006) academic cheating motivation model and previous research on homework cheating among secondary and college students.

Murdock and Anderman' s (2006) academic cheating motivation model is one of the most influential and widely applied models in cheating research, particularly regarding exam cheating. The model integrates findings from studies primarily focused on secondary and college students, categorizing cheating motivation into three aspects: "costs of cheating," "purposes of cheating," and "pre-cheating cognitions about self and outcomes (Can I do it?)." The "costs of cheating" include consequences of being caught, individual moral levels, peer cheating consequences, and integrity-related regulations. The "purposes of cheating" include peer pressure, intelligence, and classroom climate. The "pre-cheating cognitions" primarily involve self-efficacy and outcome expectations (Murdock & Anderman, 2006).

Based on this model and considering the practical applicability and generalizability of findings in elementary schools, this study selected potential influencing factors for elementary school homework cheating, focusing on the following aspects.

First, the severity of cheating consequences. In Murdock and Anderman's (2006) model, consequences represent a primary motivation for academic cheating, and empirical studies have shown that consequence severity influences secondary and college students' cheating (Kam et al., 2017; McCabe & Treviño, 1997; Molnar & Kletke, 2012). Practically, "punishment" has long been considered the "most effective/best measure" against cheating by Chinese educators. However, from ancient criminal penalties for imperial examination fraud to modern institutional penalties (e.g., canceling scores, disqualification, public criticism), regulations target exam cheating, not homework cheating. Homework cheating is less likely to be detected, more frequent, and involves more participants. Because it is perceived as low-risk with no adverse consequences, homework cheating is common and frequent. Therefore, we hypothesize that perceived severity of consequences is a primary factor influencing elementary school homework cheating.

Second, acceptability of cheating. Murdock and Anderman' s (2006) model identifies individual moral level as another important factor. Research on college and secondary students shows that moral awareness influences judgments about cheating acceptability (Cheung et al., 2016; Lee et al., 2020) and perceptions of others' acceptance (Ives & Giukin, 2020). Misjudgments may lead individuals to rationalize or minimize cheating behavior. We hypothesize that individual and perceived peer acceptability of cheating are important factors influencing elementary school homework cheating.

Third, students' evaluations of cheating prevention effectiveness. Murdock and Anderman' s (2006) model suggests that school integrity regulations influence cheating behavior, a finding supported by subsequent research (Ramberg & Modin, 2019). However, surveys indicate that students may not fully understand academic integrity policies (Bretag et al., 2014; Gullifer & Tyson, 2014), potentially explaining why cheating persists—relevant policies may not exert their intended deterrent effect. We hypothesize that students' evaluations of prevention strategies effectively predict their homework cheating behavior.

Fourth, peer cheating behavior. Empirical research on secondary and college students shows that observing peer cheating may lead individuals to rationalize the behavior (Jurdi et al., 2011; McCabe & Abdallah, 2008), thereby increasing their own probability of engaging in homework cheating (Hrabak et al., 2004). Additionally, observing peer cheating may generate feelings of "unfairness," further increasing cheating propensity. We hypothesize that peer cheating behavior effectively predicts elementary school students' homework cheating.

In addition to these predictors, this study examines demographic and personal background variables including age, gender, and self-reported academic achievement. Previous research on secondary and college students finds that lower-achieving students are more likely to cheat to obtain better grades (Newstead et al., 1996; Özcan et al., 2019), and students with low self-efficacy (poor self-perceived ability) are more likely to cheat (see Murdock & Anderman, 2006). Demographic variables such as age and gender also correlate with cheating (Błachnio, 2019; Cizek, 1999; Freire, 2014; Jurdi et al., 2011). We hypothesize these variables significantly predict elementary school homework cheating.

In summary, this study employs questionnaire surveys and machine learning to examine: (1) perceptions of potential consequences and their severity, (2) individual and perceived peer acceptability of cheating, (3) evaluations of cheating prevention strategies, (4) observed peer cheating behavior, and (5) demographic variables. Based on these factors, we construct machine learning models using ensemble methods to analyze and compare their predictive effects on homework cheating behavior.

Method

2.1 Participants This study received ethical approval from the Hangzhou Normal University Academic Ethics Committee and obtained informed consent from participating schools and parents. We selected three different types of elementary schools in a prefecture-level city in Zhejiang Province: one regular public school, one public school with many children of migrant workers, and one private school. Students from grades 2 through 6 were selected, with several classes randomly chosen from each grade for whole-class survey administration (Grade 1 students were excluded due to limited literacy and reading comprehension skills and minimal homework; Zhang, 2019). A total of 2,300 elementary school students participated. Questionnaires with completion rates below 70% (157) and those missing responses to outcome variable items (45) were excluded, yielding a final valid sample of 2,098 students. The mean age was 10.04 ± 1.40 years (53% male). All students were of Han ethnicity. Grade, gender, and school distributions are shown in .

2.2.1 Questionnaire Development and Administration Given the absence of existing questionnaires specifically for elementary school homework cheating, we developed a self-report questionnaire based on previous research on secondary and college students. Development occurred in three stages.

Stage 1: Interviews. Based on Lim and See' s (2001) survey of college students' cheating attitudes and perceived consequences, we developed interview protocols for both students and teachers. Student interviews focused on current cheating situations (especially homework cheating), manifestations, and attitudes toward cheating. Teacher interviews additionally covered school and family approaches to addressing cheating and their effectiveness. Thirty-nine students (who did not participate in subsequent surveys) and nine teachers from the three schools were interviewed. Interviews were recorded and transcribed by two psychology graduate students blind to the research purpose. Results indicated that homework cheating existed as early as grade 2, primarily manifesting as copying answer keys and classmates'work. Students perceived the most severe consequences as teacher or parent criticism and peer ridicule. They considered effective cheating-reduction strategies to include mastering knowledge through diligent study, increasing punishment severity, and informing parents.

Stage 2: Pilot testing. Based on interview results and Bucciol et al.' s (2017) college cheating questionnaire, we developed an initial version. One hundred fifty-eight students from grades 2, 4, and 5 (who did not participate in the formal survey) completed pilot testing. Based on responses and student feedback, we revised unclear or problematic items. Two psychometrics experts reviewed the revised questionnaire, resulting in the final version.

Stage 3: Formal survey. Students from grades 2 through 6 were selected proportionally from each class across the three schools (ensuring broad academic achievement distribution for representativeness). Surveys were administered in classrooms in a one-to-many format, distributed and collected on-site. To minimize concerns, surveys were anonymous, and the entire process was organized by research assistants without teachers present.

2.3 Formal Questionnaire Structure

The questionnaire measured two main components: (1) the key outcome variable —homework cheating behavior, and (2) predictor variables influencing homework cheating, including individual psychological and demographic variables.

2.3.1 Outcome Variable: Homework Cheating Behavior Students responded to the question, "Have you ever engaged in behaviors such as copying others' homework, copying answer keys, or having others do your homework?" on a 5-point Likert scale from 1 (never) to 5 (very frequently). Analysis revealed a clear positive skew (high proportion of "1" responses, low and similar proportions for "2"–"5"). To avoid statistical bias, the outcome was recoded as a binary variable: scores of "1" coded as "no cheating behavior" (0) and scores of "2"–"5" coded as "cheating behavior" (1).

2.3.2 Predictor Variables (1) Perceived severity of consequences. Based on interview results, five potential consequences of homework cheating were listed (Cronbach's $\alpha = 0.787$), such as "teacher criticism." Students rated each consequence's severity from 1 (not at all serious) to 5 (very serious).

(2) Acceptability of cheating. Two items measured: (a) students' own acceptance of cheating ("self-acceptability") and (b) perceived peer acceptance ("peer acceptability"), rated from 1 (completely unacceptable) to 5 (completely acceptable).

(3) Perceived effectiveness of cheating-reduction strategies. Based on interviews, nine common strategies were listed (Cronbach's $\alpha = 0.781$), such as "teachers grading homework more carefully to catch cheaters." Students rated each strategy's effectiveness from 1 (not at all useful) to 5 (very useful).

(4) Observed peer cheating behavior. Three aspects were measured: (a) prevalence ("peer cheating-prevalence")—how common cheating is among classmates, rated 1 (never seen) to 5 (almost everyone does it); (b) overall frequency ("peer cheating-overall frequency")—how frequently classmates cheat, rated 1 (never) to 5 (frequently); and (c) specific frequency of three common cheating forms identified in interviews (copying answer keys, copying others' work, having others do homework), each rated 1 (never) to 5 (frequently).

(5) Demographic information. Including school type (regular public, migrant worker public, private), age, gender, grade (2-6), only-child status, and self-reported academic achievement (above, equal to, or below class average).

2.4 Machine Learning Model Construction

Data were analyzed using SPSS 24.0. Following descriptive statistics, machine learning modeling was conducted. Given the diversity of algorithms and unknown relationships between predictors and outcomes, we employed ensemble learning, fitting prediction models for homework cheating using four algorithms and integrating them via stacking. The four algorithms were: logistic regression, Multilayer Perceptron (MLP), eXtreme Gradient Boosting (XGBoost), and Random Forest.

2.4.1 Four Machine Learning Algorithms and Procedures Logistic regression is a generalized linear model for predicting categorical (primarily binary) variables using a logistic function. MLP is a feedforward artificial neural network with input, output, and one or more hidden layers, adjusting connection weights between neurons to fit the model. XGBoost is an ensemble method that continuously trains and optimizes decision trees, summing outputs from each training iteration to obtain final predictions. Random Forest is similar to XGBoost but differs in that its output is simply the result of majority voting without further optimization of decision tree training results.

For MLP, we constructed networks with hidden layers and hyperbolic tangent activation functions. Covariates were standardized before training. Conjugate Gradient Descent ($\lambda = 0.0000005$, $\sigma = 0.00005$, interval center 0, interval offset $\pm$0.5) adjusted connection weights to minimize prediction errors for training set samples. Test set prediction errors were calculated after each training iteration to ensure error reduction did not result from overfitting. XGBoost used default parameters with a tree-based method (gbtree), 100 decision trees, and (boosting learning rate) default of 0.3 as a shrinkage technique to prevent overfitting. Random Forest used bootstrap sampling when constructing decision trees, with 100 trees and the number of features to consider at each split set to the square root of total features, both to improve prediction accuracy and prevent overfitting. All covariates for training and test sets were standardized.

All four algorithms followed five steps: (1) Randomly split all data into three independent sets: training (64%), test (16%), and holdout (20%). (2) Train on the training set with 32 feature inputs: 22 from nine questionnaire items (eight questions from plus options and self-reported achievement) and 10 dummy variables converted from categorical variables (school, grade, gender, only-child status; reference levels: regular public school, grade 2, female, only child). (3) Test the model on the test set to obtain performance metrics. (4) Combine training and test sets, randomly re-divide into new training and test sets, and repeat steps 2–3 to obtain a second model. Repeating this process 100 times ("split-train-test-recombine-split") yields 100 models, ensuring stability regardless of sample allocation. (5) Validate models using the holdout set (never used in training or testing) to assess external validity.

2.4.2 Integration of Machine Learning Results Following analysis with the four algorithms, we used stacking to integrate results. Specifically, we set logistic regression, MLP, XGBoost, and Random Forest to train on raw data, then integrated their training results (stacking with five-fold cross-validation), and finally validated the integrated results on test and holdout sets (validation algorithm set to logistic regression). This integration leverages the strengths of each algorithm to obtain the final optimal model.

2.4.3 Key Metrics for Machine Learning Models Validation using the holdout set yields two key metrics for each algorithm and the ensemble. First, the Area Under the Curve (AUC) of the Receiver Operating Characteristic (ROC) curve. ROC and AUC are common machine learning performance metrics characterizing classification models. ROC plots True Positive Rate (correctly predicted positives/all positives) against False Positive Rate (negatives incorrectly predicted as positives/all negatives) (see [Figure 1: see original paper]). AUC represents the area under this curve; larger AUC indicates better classifier performance, assessing both sensitivity (true positive rate, i.e., accurately predicting "cheating behavior") and specificity (1 -false positive rate, i.e., accurately predicting "no cheating behavior"). Sensitivity and specificity have an inverse functional relationship, with optimal combination at the ROC curve' s diagonal. Specific sensitivity/specificity thresholds depend on application context.

[Figure 1: see original paper] illustrates the ROC curve showing how model sensitivity (true positive rate) changes with specificity (false positive rate). The area under the curve (AUC) indicates overall model performance; greater distance from the identity line (dashed diagonal) represents better classification ability.

Second, Shapley values for each predictor. Shapley values address complex allocation problems (Shapley, 1953). For example, if A, B, and C complete a task together, determining A' s fair bonus requires calculating: work completed by A alone; work completed by B and C with A minus work by B/C alone; and work by all three minus work by B and C together. The mean of these values represents A' s marginal contribution. This method comprehensively considers all cooperative and individual relationships, providing fair and scientific calculation of marginal contributions. Consequently, researchers increasingly use Shapley values to assess predictors' relative contributions in models (Ghorbani & Zou, 2019; Lundberg & Lee, 2017).

In this study, Shapley values for predictors follow these rules, where "A, B, C" are different predictors and "work" is contribution size (prediction accuracy). A predictor' s Shapley value represents its actual marginal contribution to overall model prediction accuracy, with magnitude indicating relative importance.

Results

3.1.1 Cheating Behavior [Figure 2: see original paper] shows the percentage of students in grades 2–6 who self-reported cheating. Cheating rates increased noticeably in grade 3 and stabilized from grade 4 onward. Binary logistic regression with grade as predictor (grade 2 as reference) and cheating behavior as outcome revealed significant grade differences ($p < 0.001$). Grade 2 had significantly lower cheating rates than other grades (grade 2 vs. 3: $p = 0.002$, B =

0.51, OR = 1.67, 95% CI = 1.21-2.29; grade 2 vs. grades 4/5/6: $ps < 0.001$, B = 0.70, 0.57, 0.87; OR = 2.01, 1.77, 2.39, 95% CI = 1.46-2.75, 1.30-2.40, 1.77-3.22). However, grades 3-6 did not differ significantly from each other ($ps > 0.05$).

3.1.2 Factors Influencing Cheating Behavior shows means and standard deviations for each predictor. Students generally considered "being punished by teacher" (M = 3.65) the most serious consequence and viewed cheating as unacceptable (M = 2.02, significantly different from neutral point 3, $p < 0.001$). They also perceived low peer acceptance of cheating (M = 2.08, $p < 0.001$). Additionally, students considered "strengthening classroom practice to master knowledge in class" (M = 3.28) the most effective cheating-reduction strategy, and identified "copying others' homework" (M = 2.01) as peers' most common cheating method.

3.2 Machine Learning Analysis To accurately calculate predictor effects, only participants with 100% response rates on predictor items were included, yielding a final machine learning sample of 1,637. As described, we used four machine learning algorithms and integrated their results via ensemble learning.

3.2.1 AUC Analysis of Four Algorithms and Ensemble Learning shows mean AUC values for 100 models simulated by each of the four algorithms and the ensemble method. All AUC means significantly exceeded chance level (50%, $ps < 0.001$). The ensemble learning model achieved a mean AUC of 80.46%, indicating an 80.46% probability of accurately predicting whether elementary school students would cheat on homework.

[Figure 3: see original paper] displays AUC values for 100 models from each algorithm and the ensemble method. Ensemble learning models showed high overall sensitivity and specificity (1 -false positive rate). Subsequent analyses focus on ensemble learning results due to its integration of four algorithms' strengths. Converting the ensemble AUC mean to Cohen' s d yielded 1.214 (95% CI: 1.205-1.222), indicating a very large effect size (Cohen, 1988; Cohen' s d > 1.2 indicates very large effect).

Note: a, b, c, d, e represent AUC values for logistic regression, XGBoost, MLP, Random Forest, and ensemble learning, respectively.

3.2.2 Shapley Value Analysis of Predictors in the Ensemble Model Ensemble learning with holdout set validation yielded significant Shapley values for all predictors ($ps < 0.05$), indicating significant marginal contributions to model prediction accuracy. Ranking predictors by Shapley value magnitude reveals their relative importance. [Figure 4: see original paper] shows major predictors with marginal contributions $\geq$ 1%, which differed substantially in importance and could be divided into four groups.

[Figure 4: see original paper] Shapley values and 95% confidence intervals for major predictors of homework cheating. "⊖" indicates negative prediction; all others show positive prediction.

Group 1: Students' own acceptance of cheating (higher acceptance predicts cheating). This predictor's Shapley value reached 10.49%, representing a 10.49% marginal contribution to overall prediction accuracy. Paired t-tests showed this value significantly exceeded the second-ranked predictor' s 3.83% ($t = 23.88$, $df = 327$, $p < 0.001$), indicating far superior predictive power.

Group 2: Shapley values of 2%-4%. Peer cheating prevalence (3.83%) significantly contributed more than remaining variables (vs. 3.26%, $t = 1.98$, $df = 327$, $p = 0.048$). Also included: students' relative achievement level (reverse-scored), peer cheating prevalence, peer cheating frequency, and frequency of peers "copying others' homework." These three variables did not differ significantly (see [Figure 4: see original paper] for values; t = 0.57, 1.78, 1.23, dfs = 327, p = 0.57, 0.22, 0.08) and were significantly correlated ($r = 0.21$, $p < 0.001$), indicating equivalent predictive power. Specifically, lower self-rated achievement and more prevalent/frequent peer cheating positively predicted cheating behavior.

Group 3: Shapley values of 1%-2% with smaller marginal contributions, significantly lower than 2.9% ($t = -6.99$, $df = 327$, $p < 0.001$). Included: grade level (grades 6 and 4 > grade 2), perceived peer acceptance of cheating, frequency of peers "copying answer keys" (higher perceived acceptance/frequency predicts cheating), and severity ratings for three consequences ("parental punishment," "parental criticism," "teacher punishment" ; higher perceived severity predicts less cheating; r = 0.44, 0.34, 0.36, ps < 0.001), and school type (migrant worker schools < regular public schools).

Group 4: Remaining variables with significant but very weak contributions (Shapley values < 0.01). These minimally important predictors included: ratings of all nine cheating-reduction strategies, severity ratings for some consequences ("teacher criticism,""peer ridicule"), and some demographic information (see).

shows means, standard deviations, and 95% confidence intervals for Shapley values of secondary predictors.

Discussion

This study systematically examined elementary school students'homework cheating behavior and its relationships with individual cognitive and contextual variables (perceived consequence severity, self and peer acceptability, intervention strategy effectiveness), peer cheating prevalence and frequency, and demographic variables (gender, grade, school type) using questionnaire surveys and machine learning, an artificial intelligence core methodology.

First, approximately 33% of elementary school students self-reported homework cheating, with rates showing grade-level trends. Specifically, grade 2 represents the emergence stage, grade 3 a surge stage, followed by a plateau; by grade 6, the rate reached 40.5%. This indicates homework cheating emerges in grade 2 and becomes relatively common by grade 6. The rapid increase in grade 3 may result from increased homework volume and frequency. The plateau from grade 4 suggests moral development level is not a key factor in cheating behavior during elementary school (a hypothesis supported by research on children' s lying; Lee, 2013). Thus, cheating as a habitual behavior, once formed, may persist or increase without effective intervention, highlighting the importance of early honesty education and intervention.

Second, all examined predictors significantly predicted homework cheating. Machine learning results showed that integrating different algorithms via ensemble learning produced a model with high sensitivity and specificity (mean AUC = 80.46%), meaning the model has an 80.46% probability of correctly predicting cheating based on these predictors. The Cohen' s d derived from AUC also showed a very large effect size (Cohen' s d > 1.2).

Third, Shapley value analysis revealed that while all predictors contributed significantly to the model' s high accuracy, their importance varied substantially. Specifically:

- Students' own acceptance of homework cheating was the most critical predictor. Greater acceptance strongly predicted self-reported cheating, consistent with findings from secondary and college students (Abaraogu et al., 2016; Ives & Giukin, 2020). Murdock and Anderman (2006) proposed that individual differences in cheating acceptability judgments may stem from two sources: (1) not considering cheating immoral, or (2) knowing it' s immoral but rationalizing one's own behavior to reduce cognitive conflict. Elementary school students' judgments may be influenced by these same factors.
- Peer cheating prevalence and frequency were also important predictors, consistent with research showing peer cheating leads individuals to rationalize their own behavior (Hrabak et al., 2004; Ghanem & Mozahem, 2019; McCabe & Treviño, 1993). A recent meta-analysis of multinational studies identified peer cheating as one of the most important factors in academic dishonesty, with this "contagion effect" more pronounced in collectivist than individualist cultures (Zhao & Mao et al., 2022).
- Self-rated academic achievement showed similar importance. Lower-achieving students were more likely to report homework cheating, consistent with secondary/college findings (Newstead et al., 1996; Özcan et al., 2019), reflecting a potential motivation to avoid failure (Oran et al., 2016). Additionally, lower-achieving students may have weaker self-efficacy, which predicts cheating (see Murdock & Anderman, 2006).

Other factors with weaker but significant predictive effects included grade level, school type, peer acceptability, and consequence severity. For example, regular public school students showed higher cheating rates than migrant worker school students, possibly due to differences in school climate, policies, and teacher-student relationships that shape learning atmosphere (Ramberg & Modin, 2019; McCabe et al., 2012). Perceived peer acceptability had weaker predictive power than students' own acceptability and observed peer cheating, suggesting peer influence operates primarily through observable behavior rather than others' moral awareness. The minimal effect of "peer ridicule" as a consequence supports this interpretation.

Perceived consequence severity was a weaker predictor than hypothesized, possibly because China lacks substantive homework cheating punishment systems. Students' limited understanding of consequences derives mainly from exam cheating contexts. Notably, despite strict exam cheating penalties, exam cheating persists, suggesting reliance on external enforcement without cultivating students' moral values is ineffective.

Students' evaluations of cheating-reduction strategies and some demographic variables were the weakest predictors (though still significantly > 0). The weak predictive power of strategy effectiveness ratings likely reflects students' generally low evaluations of common anti-cheating strategies.

This study offers theoretical innovation as the first comprehensive model of elementary school homework cheating predictors, quantifying and ranking their importance. Notably, elementary school cheating factors differ from those affecting secondary/college students. For example, consequence severity is crucial for predicting secondary/college exam cheating in Murdock and Anderman's (2006) model but had minimal impact on elementary homework cheating, highlighting developmental specificity. Some factors affecting college students (e.g., only-child status, gender) contributed minimally to predicting elementary cheating, while others (e.g., grade level) showed stronger effects not found in older students (Ives et al., 2017).

However, some key elements from Murdock and Anderman's (2006) model—students' own cheating acceptability, peer cheating prevalence/frequency, and achievement level—also predicted elementary homework cheating, showing some continuity across developmental stages (Abaraogu et al., 2016; Ghanem & Mozahem, 2019).

Additionally, situational cognitions (e.g., self-acceptability) played important roles, providing new insights into the long-standing debate about whether cheating is situationally or dispositionally driven. While past research emphasized situational over personal factors (Hartshorne & May, 1928), our findings suggest interaction between situation and personal traits may be important.

Methodologically, this study innovates by applying machine learning to children's moral development for the first time. Results demonstrate machine learning is feasible for analyzing child behavioral data, offering a new research approach in the digital intelligence era.

Practically, we created a predictive model (80%+ accuracy) that could be developed into an app or web-based test. Students would complete the questionnaire, and the model would output individual cheating probabilities (0%-100%) to help teachers and parents identify at-risk students. However, results should not be used to label students, requiring careful consideration of research ethics and educational approaches. Since some predictors contributed minimally, future research could develop a shorter questionnaire by removing low-contribution items, facilitating broader application in schools and families.

This study also provides concrete, actionable approaches for honesty education. Since cheating acceptability (moral awareness) rather than consequence severity predicts cheating, educators should focus on building correct academic integrity cognitions rather than emphasizing punishments. Given strong peer influence, teachers and parents should minimize negative peer effects and promote positive ones, such as praising students who complete work independently regardless of accuracy, emphasizing that independent completion matters more than correctness (Misselbrook, 2014; Siev & Kliger, 2019). Schools should also respond to the "double reduction" policy by emphasizing quality over quantity and establishing healthy learning climates that eliminate "grade-only" mentalities, helping students view homework as knowledge consolidation rather than competition (Misselbrook, 2014; Siev & Kliger, 2019). Finally, since cheating is habitual (Davy et al., 2007) and elementary school cheating rates are relatively low across ages (Cochran, 2015), early intervention during elementary or even preschool years can "nip cheating in the bud."

Limitations include: (1) The 80.46% prediction accuracy still allows nearly 20% error, requiring future model optimization through theoretical exploration of additional predictors and larger samples for parameter tuning and external validation. (2) Self-report measures may underestimate actual cheating due to social desirability. Future research could combine self-reports with behavioral experiments, though existing paradigms suit older students better and require adaptation for children. (3) This study examined traditional cheating methods (copying) but not digital-age methods (e.g., using homework-grading apps or other online methods), warranting future investigation.

In conclusion, this first study of elementary school homework cheating using machine learning ensemble algorithms systematically identified key factors and their relative importance. Results show 33% of elementary students self-report homework cheating, with rates increasing across grades. The predictive model achieved 80.46% AUC, accurately predicting cheating behavior. Elementary school homework cheating depends heavily on students' cheating acceptability, peer cheating behavior, and their own achievement levels. These findings advance theoretical understanding of early academic integrity development and provide scientific evidence for early intervention. Moreover, machine learning proves effective for analyzing developmental data.

ChinaRxiv 中X

References

Abaraogu, U. O., Henning, M. A., Okpare, M. C., & Rajput, V. (2016). Disclosing academic dishonesty: perspectives from Nigerian and New Zealand health professional students. *Ethics & Behavior, 26*(5), 431-447. https://doi.org/10.1080/10508422.2015.1055494

Anderman, E. M., & Midgley, C. (2004). Changes in self-reported academic cheating across the transition from middle school to high school. *Contemporary Educational Psychology, 29*(4), 499-517. https://doi.org/10.1016/j.cedpsych.2004.02.002

Bartlett, M. S., Littlewort, G., Frank, M., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. *Current Biology, 24*(7), 738-743. https://doi.org/10.1016/j.cub.2014.02.009

Błachnio, A. (2019). Don' t cheat, be happy. Self-control, self-beliefs, and satisfaction with life in academic honesty: A cross-sectional study in Poland. *Scandinavian Journal of Psychology, 60*, 261-266. https://doi.org/10.1111/sjop.12534

Bong, M. (2008). Effects of parent-child relationships and classroom goal structures and motivation, help-seeking avoidance, and cheating. *Journal of Experimental Education, 76*(2), 191-217. https://doi.org/10.3200/JEXE.76.2.191-217

Bretag, T., Mahmud, S., Wallace, M., Walker, R., McGowan, U., East, J., ···& James, C. (2014). 'Teach us how to do it properly!' An Australian academic integrity student survey. *Studies in higher education, 39*(7), 1150-1169. https://doi.org/10.1080/03075079.2013.777406

Bruer, K. C., Zanette, S., Ding, X. P., Lyon, T. D., & Lee, K. (2019). Identifying liars through automatic decoding of children' s facial expressions. *Child Development, 91*, e995-e1011. https://doi.org/10.1111/cdev.13336

Bucciol, A., Cicognani, S., & Montinari, N. (2017). Cheating in academia: The relevance of social factors. Working Papers 15/2017, University of Verona, Department of Economics.

Campbell, D. T. (1986). Relabeling internal and external validity for applied social scientists. *New Directions for Program Evaluation, 31*, 67-77. https://doi.org/10.1002/ev.1434

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.), Hillsdale, NJ: Erlbaum.

Cochran, J. K. (2015). The effects of life domains, constraints, and motivations on academic dishonesty: a partial test and extension of Agnew' s general theory. *International Journal of Offender Therapy and Comparative Criminology, 61*(11), 1288. https://doi.org/10.1177/0306624X15618689

Cizek, G.J. (1999), *Handbook of educational policy*, Academic Press, San Diego, CA.

Davy, J. A., Kincaid, J. F., Smith, K. J., & Trawick, M. A. (2007). An examination of the role of attitudinal characteristics and motivation on the cheating behavior of business students. *Ethics & Behavior, 17*(3), 281-302. https://doi.org/10.1080/10508420701519304

Freire, C. (2014). Academic misconduct among Portuguese economics and business undergraduate students–a comparative analysis with other major students. *Journal of Academic Ethics, 12*(1), 43-63. https://doi.org/10.1007/s10805-013-9199-2

Galloway, M. K. (2012). Cheating in advantaged high schools: prevalence, justifications, and possibilities for change. *Ethics & Behavior, 22*(5), 378-399. https://doi.org/10.1080/10508422.2012.679143

Ghanem, C. M., & Mozahem, N. A. (2019). A study of cheating beliefs, engagement, and perception–the case of business and engineering students. *Journal of Academic Ethics, 17*, 291-312. https://doi.org/10.1007/s10805-019-9325-x

Cheung, H. Y., Wu, J., & Huang, Y. (2016). Why do Chinese students cheat? Initial findings based on the self-reports of high school students in China. *The Australian Educational Researcher, 43*(2), 245-271. https://doi.org/10.1007/s13384-016-0201-z

Ghorbani, A., & Zou, J. (2019). Data shapley: Equitable valuation of data for machine learning. Working Paper

Golino, H. F., Gomes, C. M. A., & Andrade, D. (2014). Predicting academic achievement of high-school students using machine learning. *Psychology, 5*(5), 2046-2057. https://doi.org/10.4236/psych.2014.518207

Gullifer, J. M., & Tyson, G. A. (2014). Who has read the policy on plagiarism? Unpacking students' understanding of plagiarism. *Studies in Higher Education, 39*(7), 1202-1218. http://dx.doi.org/10.1080/03075079.2013.777412

Hartshorne, H. & May, M.A. (1928). *Studies in deceit.* New York: McMillan.

Hrabak, M., Vujaklija, A., Vodopivec, I., Hren, D., Marušić, M., & Marušić, A. (2004). Academic misconduct among medical students in a post-communist country. *Medical Education, 38*, 276-285. https://doi.org/10.1111/j.1365-2923.2004.01766.x

Ives, B., Alama, M., Mosora, L. C., Mosora, M., Grosu-Radulescu, L., & Clinciu, A. I., et al. (2017). Patterns and predictors of academic dishonesty in Romanian university students. *Higher Education, 74*, 815-831. https://doi.org/10.1007/s10734-016-0079-8

Ives, B., & Giukin, L. (2020). Patterns and predictors of academic dishonesty in Moldovan university students. *Journal of Academic Ethics, 18*(1), 71-88. https://doi.org/10.1007/s10805-019-09347-z

Jurdi, R., Hage, H. S., & Chow, H. P. H. (2011). Academic dishonesty in the Canadian classroom: behaviours of a sample of university students. *Canadian Journal of Higher Education, 41*(3), 1-35. https://doi.org/10.47678/cjhe.v41i3.2488

Just, M. A., Pan, L., Cherkassky, V. L., McMakin, D. L., Cha, C., Nock, M. K., & Brent, D. (2017). Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth. *Nature Human Behaviour, 1*(12), 911. https://doi.org/10.1038/s41562-017-0234-y

Kam, C. C. S., Hue, M., & Cheung H. Y. (2017). Plagiarism of Chinese secondary school students in Hong Kong. *Ethics & Behavior, 28*(4), 316-335. https://doi.org/10.1080/10508422.2017.1333909

Lee, K. (2013). Little liars: development of verbal deception in children. *Child Development perspectives, 7*(2), 91-96. https://doi.org/10.1111/cdep.12023

Lee, S. D., Kuncel, N. R., & Gau, J. (2020). Personality, attitude, and demographic correlates of academic dishonesty: a meta-analysis. *Psychological Bulletin, 146*(11). https://doi.org/10.1037/bul0000300

Lim, V. K. G., & See, S. K. B. (2001). Attitudes toward, and intentions to report, academic cheating among students in Singapore. *Ethics & Behavior, 11*(3), 261-274. https://doi.org/10.1207/S15327019EB1103_5

Livieris, I. E., Kotsilieris, T., Dimopoulos, I., & Pintelas, P. (2018). Decision support software for forecasting patient’ s length of stay. *Algorithms, 11*(12), 199. https://doi.org/10.3390/a11120199

Lundberg, S. M., & Lee, S. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 4765-4774. https://doi.org/10.48550/arXiv.1705.07874

McCabe, D. L., & Abdallah, F. H. (2008). Academic dishonesty in the Middle East: individual and contextual factors. *Research in Higher Education, 49*(5), 451-467.

McCabe, D. L., Butterfield, K. D., & Trevino, L. K. (2012). *Cheating in college: why students do it and what educators can do about it.* Baltimore, MD: John Hopkins University Press.

McCabe, D. L., & Treviño, L. K. (1993). Academic dishonesty: honor codes and other contextual influences. *The Journal of Higher Education, 64*, 522-538.

McCabe, D. L., & Treviño, L. K. (1997). Individual and contextual influences on academic dishonesty: a multicampus investigation. *Research in Higher Education, 38*(3), 379-396. https://doi.org/10.1023/A:1024954224675

Misselbrook, D. (2014). Is honesty the best policy? why trustworthiness is no easy answer. *British Journal of General Practice the Journal of the Royal College of General Practitioners, 64*(628), 558-559. https://doi.org/10.3399/bjgp14X682129

Molnar, K. K., & Kletke, M. G. (2012). Does the type of cheating influence undergraduate students' perceptions of cheating? *Journal of Academic Ethics, 10*(3), 201-212. https://doi.org/10.1007/s10805-012-9164-5

Murdock, T. B., & Anderman, E. M. (2006). Motivational perspectives on student cheating: toward an integrated model of academic dishonesty. *Educational Psychologist, 41*, 129-145. http://dx.doi.org/10.1207/s15 326985ep4103_1

Newstead, S. E., Franklyn-Stokes, A., & Armstead, P. (1996). Individual differences in student cheating. *Journal of Educational Psychology, 88*(2), 229-241.

Oran, N. T., Can, H. O., Senol, S., & Hadimli, A. P. (2016). Academic dishonesty among health science school students. *Nursing Ethics, 23*(8), 919-931. https://doi.org/10.1177/0969733015583929

Özcan, M., Yeniçeri, N., & Çekiç, E. G. (2019). The impact of gender and academic achievement on the violation of academic integrity for medical faculty students, a descriptive cross-sectional survey study. *BMC Medical Education, 19*:427. https://doi.org/10.1186/s12909-019-1865-7

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., & Grisel, O., et al. (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830. https://doi.org/10.1524/auto.2011.0951

Putarek, V., & Pavlin-Bernardic, N. (2020). The role of self-efficacy for self-regulated learning, achievement goals, and engagement in academic cheating. *European Journal of Psychology of Education, 35*(3), 647-671. https://doi.org/10.1007/s10212-019-00443-7

Ramberg, J., & Modin, B. (2019). School effectiveness and student cheating: do students' grades and moral standards matter for this relationship? *Social Psychology of Education, 22*(3), 517-538. https://doi.org/10.1007/s11218-019-09486-6

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games, 2*, 307-317.

Shu, S., Sang, Q., Guo, Y., Huang, X. (2018). Why cheating Is contagious: the mediating effect of social loss and attitude toward cheating. *Psychological Development and Education, 34*(6), 664-671. https://doi.org/10.16187/j.cnki.issn1001-4918.2018.06.04

Siev, S., & Kliger, D. (2019). Cheating in academic exams: a field study. *Dishonesty in Behavioral Economics*, 111-140. https://doi.org/10.1016/B978-0-12-815857-9.00008-X

Smith, M., & Alvarez, F. (2021). Identifying mortality factors from machine learning using Shapley values -a case of COVID19. *Expert Systems with Applications, 176*, 114832. https://doi.org/10.1016/j.eswa.2021.114832

Tsui, A. P. Y., & Ngo, H. Y. (2016). Social predictors of business student cheating behaviour in Chinese societies. *Journal of Academic Ethics, 14*(4), 1-16. https://doi.org/10.1007/s10805-016-9266-6

Williams, M. W. M, & Williams, M. N. (2012). Academic dishonesty, self-control, and general criminality: a prospective and retrospective study of academic dishonesty in a New Zealand university. *Ethics & Behavior, 22*(2), 89-112.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: lessons from machine learning. *Perspectives on Psychological Science: A Journal of the Association for Psychological science, 12*(6), 1100-1122. https://doi.org/10.1177/1745691617693393

Yi, X. M. (2021). University new students' prognosis of being punished (PBP) for violation and the degree of fear of consequences (DFC) of violation and their influence on the possibility of truancy and cheating. *Chinese Journal of Clinical Psychology, 29*(3), 572-576. https://doi.org/10.16128/j.cnki.1005-3611.2021.03.026

Ykhlef, H., & Bouchaffra, D. (2017). Induced subgraph game for ensemble selection. *International Journal on Artificial Intelligence Tools, 26*(1), 1760003. https://doi.org/10.1109/ICTAI.2015.97

Zanette, S., Gao, X. Q., Brunet, M., Bartlett, M. S., & Lee, K. (2016). Automated decoding of facial expressions reveals marked differences in children when telling antisocial versus prosocial lies. *Journal of Experimental Child Psychology, 150*, 165-179. https://doi.org/10.1016/j.jecp.2016.05.007

Zhang, R. (2019). *Academic dishonesty among primary school students and its correlates*. Hangzhou Normal University. https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2019

Zhao, L., Mao, H. Y., Compton, B. J., Peng, J. J., Fu, G., Fang, F., Heyman, G. D., & Lee, K. (2022). Academic dishonesty and its relations to peer cheating and culture: a meta-analysis of the perceived peer cheating effect. *Educational Research Review, 36*, 100455. https://doi.org/10.1016/j.edurev.2022.100455

Zhao, L., Peng, J., Dong, L. D., Li, Y., Mao, H., Compton, B. J., ···& Lee, K. (2022). Effects of test difficulty messaging on academic cheating among middle school children. *Journal of Experimental Child Psychology, 220*, 105417. https://doi.org/10.1016/j.jecp.2022.105417

Zhao, L., Zheng, Y., Mao, H. Y., Compton, B. J., Fu, G., Heyman, G. D., & Lee, K. (2021). Using environmental nudges to reduce academic cheating in young children. *Developmental Science*, e13108. https://doi.org/10.1111/desc.13108

Note: Figure translations are in progress. See original paper for figures.