

## Does Skipping or Repeating a Grade Affect Self-Assessed Intelligence?

Timothy Smith

Georgia College and State University

In the United States, 1.5% of all children repeated a grade in 2010 and grade skipping has come under criticism. When examining the impact of skipping or repeating, the economics literature presents contradictory results often based on flawed methodology. I utilize self-assessed intelligence data, and a strong set of covariates, from the National Longitudinal Study of Adolescent to Adult Health to determine whether grade skipping or grade retention have an impact on school age students. Self-assessed intelligence may be free of some of the problems present in the outcome variables used in previous studies, such as academic performance. I find that grade retention has a significant negative effect, while the results on grade skipping are unclear.

## **I. Introduction**

Grade retention and grade skipping are two options available to students enrolled in public education in the United States. Grade retention involves a student being held back a year and repeating the same course material or grade that the student encountered in the previous school year (Jackson 1975). In the 2009-2010 academic year, 1.5% of all children in the United States were retained (Warren et al. 2014). When a student is deemed sufficiently knowledgeable or advanced in the coursework being taught in a grade, he or she can skip the next grade in order to continue to learn new material. Both of these strategies attempt to alleviate problems at the individual level that arise in the public school system. Grade retention targets students that are struggling academically or socially, while grade skipping hopes to allow high achieving students to avoid boredom and repetition.

Potential concerns have arisen with both of these policies. Grade retention can negatively affect students' social standing and increase drop out rates. Grade retention has been reported by many studies to be ineffective in improving students' performance in future grades measured by typical standardized tests and individual course work grades. These findings are summarized in a meta-analysis conducted by Jimerson (2001). While many such studies have been conducted, a large majority of them have been marked as methodologically flawed upon review, as noted by Jimerson (2001). Studies attempting to evaluate the effectiveness of grade retention on other variables, such as socio-emotional and behavioral outcomes, have had varied results with some in favor of retention and some against it (Jimerson 2001). Grade skipping, on the other hand, has been found to be effective in aiding academically advanced students in their education

and also in their social development (Robeck 1968; Rogers 1992). Other analyses have found that students involved in grade skipping performed better than same age peers, but not better than older-age control groups (Kulik and Kulik 1984). Despite these and other findings, concerns remain regarding knowledge gaps and social problems (Colangelo et al. 2004), and selection bias due to the inherent differences between promoted and non promoted students (Kretschmann and Vock 2014).

Given the methodological problems and lack of consistent results involved in research surrounding grade retention, and the age and persistent concerns present in research on grade skipping, a demand and opportunity for new and diverse research in both areas arises. This study will focus on self-assessed intelligence as compared to other adolescents of the same age as the left hand variable, instead of school-assessed or standardized measures of academic achievement. This change will not only provide an interesting perspective when compared to much of the other research in the field of grade skipping, but it will also create a fix for the methodological problems of much of the research surrounding grade retention by removing the biases towards retention or promotion. It could also shed light on whether or not adolescents who have skipped grades perceive themselves as having knowledge gaps.

## **II. Literature Review**

“The concepts of grade retention and grade skipping and their respective impacts are widely discussed in the academic literature. To begin, a meta-analysis conducted by Steenbergen-Hu and Moon (2011) summarizes previous meta-analyses and provides a detailed summary of current research regarding both socio-emotional outcomes and performance outcomes from studies on acceleration in the last decade. Acceleration

encompasses grade skipping, along with other forms of academic advancement such as early starts, early college entry, AP classes and others. Grade skipping is considered a grade-based acceleration, while AP courses and other non-grade related acceleration would be considered content-based. Their findings present positive effect sizes for both academic performance and socio-emotional variables from 38 studies over the past three decades. While their significance levels do not meet typical criterion levels for academic research ( $p < 0.05$ ), when one study with an effect size dramatically different in sign and magnitude for the other studies is removed, the effect size for academic achievement becomes significant and positive at the 5% level. While their findings represent an encouraging outlook for acceleration in general, grade skipping played a relatively minor role in the study.

While the numbers of empirical studies that look at grade skipping as their primary variable of concern are few and far between, there are several studies of note that present findings that inform this examination. One such study, conducted by Park, Lubinski and Benbow (2012) examined the longitudinal effects of grade skipping in STEM (Science, Technology, Engineering and Mathematics) fields. Using a sample size of 3,467, they use exact and propensity score matching to create comparison groups and examine the effects of grade skipping on time to first publication and completion of first degree. They found that students who skipped a grade outperformed their same-age peers in time to first publication, time to first degree, total number of citations, and likelihood of entering a STEM or other research related field. The authors did not find any significant differences when comparing the grade skipping students to same-grade peers,

meaning students who are in the promoted students new grade and are a year older than the promoted students.

Another study used German students to examine the effects of grade skipping on a multitude of academic performance variables, while attempting to control for several types of bias associated with other literature. Their goal was to attempt to verify results found by older studies (Rogers 1992; Kulik 1984) by recreating the same experiments with more controls. They used same-grade comparison. This study also used propensity score matching to create groups and evaluated academic performance. They found that, after controlling for more factors and creating more balanced groups, only a small significant effect on spelling performance was found, with no differences in mathematics or overall language scores (Kretshmann et al. 2014). This supports other literature such as Park et al.(2013), and Kulik (2004).

These three studies present a unified result. Students who skip a grade are able keep up with same-grade peers, while outperforming same-age peers. However, this study builds upon previous literature with an innovative model which includes a psychological variable with specific regard to grade skipping.

Grade retention has a larger body of literature available, although much of the older literature has been discounted by newer studies, as mentioned earlier (Jimerson 2001). Jimerson's meta-analysis provides a concise analysis of much of the older and some more recent literature, and concludes that the current research provides a sufficient body of support against grade retention that the focus of any future research should be towards other workable solutions and child-by-child based interventions (2001). His analysis finds negative effects in both the short term and long term, comparing both

performance and socio-emotional variables (2001). Jimerson has also published a paper referencing his and other academic papers analyzing the effects of grade retention, and has consistently found no positive outcomes associated with retention (2012).

In contrast to the conclusion drawn by Jimerson, several newer meta-analyses have revisited the use of grade retention, and have called into question the methods used by older meta-analyses. Allen et al. addresses Jimerson specifically and challenges the “score card” approach used to compare studies. Conducting a comprehensive comparison of the studies used prior, they find that many of the studies do not use adequate statistical methods and controls, and that those who do fail to find benefits, but do not find consistent harmful short term or long-term effects. While the two analyses (Jimerson and Allen et al.) agree that no benefits are found from grade retention, Allen et al challenges the accepted notion that grade retention has harmful effects. They also suggest that future research should focus on finding situations where grade retention is beneficial, and narrowing the use of grade retention to children that will benefit from it. The comparison of these two papers provides a summary of the two common viewpoints on retention.

### **III. Data**

The data utilized are from the National Longitudinal Study of Adolescent to Adult Health conducted by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill. This study began in 1994 with a national representative survey of students in seventh through twelfth grade. It was created in response to a mandate from the U.S. Congress and is funded by several grants from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The data set is nationally representative and covers a wide variety of topics spanning

from sexual and general health to in-home life and friendships. There are two versions of the data set, a public-use data set and a private-use set; this paper uses the public-use data set which has 6,504 observations included within it. The summary statistics of the data are available in table 4.

The data had to be cleaned in several ways before use. The BIO\_SEX variable in the ADD Health data set is coded with responses as 2 or 1 instead of 1 or 0. I created a new variable “female” such that when an observation is female the variable is equal to 1; otherwise it is equal to 0. The original survey is also coded in a way such that an individual could mark himself or herself as Hispanic/Latino but then also mark him or herself as another race. I recoded the set of race variables such that if an individual answered that they were Hispanic/Latino they would not appear as any other race. If there were missing observations or responses (such as legitimate skips) that didn’t have statistical meaning for any of the variables, I removed these observations entirely. This resulted in a decrease in sample size from 6504 to 4671 observations. I also created several other dummy variables for ease of interpretation, including: HAPPY, SKIPSCHOOL, WANTSCOLLEGE, SMOKES, PARENT1ED and SLEEP variables. These variables were categorical in some way and were converted to a ‘yes-no’ condition. I also recoded my dependent variable as a dummy variable instead of a categorical variable for my initial regression. I utilize the original ordinal variable in a robustness check. I did this by creating a dummy (named: above) that was equal to 1 if an individual rated his or herself as any of the options that represented above average intelligence.

My dependent variable had a mean of .566; indicating that slightly above half of the sample believed that they were above average intelligence. My first key independent variable, SKIPPED, represents only two percent of my sample, while my other key variable, HELDBACK, represents almost twenty percent of my sample. This means that many more students have been held back than have skipped a grade. Twenty-two percent of my population is black, which is above the national average. This is because the ADD Health survey intentionally over-included African Americans. Over ninety-four percent of my sample stated that they were interested in going to college. Of the parents of students in the survey, only twenty-six percent of them finished at least a bachelor's degree. The average income of a household was forty-eight thousand dollars a year.

#### **IV. Methodology**

##### **Theory**

A person's perceived intelligence is a factor in his or her future education decisions, which, in turn, will affect his or her economic well-being and economic knowledge. This in itself is cause for evaluation of factors that could potentially cause harmful or beneficial effects to a person's thoughts on their intelligence.

Grade retention and grade skipping both have the potential to cause change in a person's perceived intelligence. If an adolescent, especially a child during the first few years of school, is told that he or she is going to have to repeat the same material that he or she has already learned once, leave behind all of his or her friends, and deal with the burden of explaining that he or she was held back to current friends that could definitely be cause for psychological distress, particularly with respect to intelligence level, because grades and academic progress are so tightly linked to intelligence in the United States.

The same, but opposite effect, can be said for grade skipping. If a child is told that he or she is going to not only complete their current grade, but also skip a grade to the grade after that it could be a huge boost to the confidence and perceived intelligence, for the reason described above.

### **Model**

I used a probit model to estimate the effects of skipping or repeating a grade on an individual's self-assessed intelligence. A probit differs from OLS in that it forces coefficients to be between negative one and one. This is necessary because if the dependent variable of a model is a dummy variable, OLS can create uninterpretable results that fall outside of negative one to one. Running a probit and computing marginal effects causes all results to be interpretable. I estimate an OLS linear probability model as well, for comparison. The model for both is displayed below:

$$\begin{aligned}
 \text{ABOVE}_i = & \beta_0 + \beta_1 \text{SKIPPED}_i + \beta_2 \text{HELDBACK}_i + \beta_3 \text{HISPANIC}_i + \beta_4 \text{BLACK}_i \\
 & + \beta_5 \text{NATIVE\_AMERICAN}_i + \beta_6 \text{ASIAN}_i + \beta_7 \text{ELECTRONICSUSE}_i + \beta_8 \text{SPORTS}_i \\
 & + \beta_9 \text{FRIENDS}_i + \beta_{10} \text{BODYIMAGE}_i + \beta_{11} \text{SLEEP}_i + \beta_{12} \text{HAPPY}_i \quad (1) \\
 & + \beta_{13} \text{PARENTALINVOLVE}_i + \beta_{14} \text{SMOKES}_i + \beta_{15} \text{PARENTINCOME}_i + \beta_{16} \text{FEMALE}_i \\
 & + \beta_{17} \text{SKIPSCHOOL}_i + \beta_{18} \text{WANTSCOLLEGE}_i + \beta_{19} \text{PARENT1ED}_i + e_i
 \end{aligned}$$

ABOVE is a dummy variable equal to 1 if a person rated themselves as any of the choices that represented above average intelligence on the scale. SKIPPED and HELDBACK are dummy variables equal to 1 if an individual skipped a grade or was held back a grade. The variables that begin with HISPANIC and end with ASIAN are a collection of dummy variables representing different races. ELECTRONICSUSE through SMOKES are a collection of variables representing different lifestyles choices the individual made ranked either on some scale or as dummy variable. PARENTINCOME is the income of the individual's parent in thousands. FEMALE is a dummy variable

equal to 1 if the individual is a female. SKIPSCHOOL and WANTSCOLLEGE are dummy variables equal to 1 if the individual had ever skipped school, and if the person rated themselves as wanting to go to college. PARENT1ED is a dummy variable equal to 1 if the child's primary parent had a bachelor's degree. After estimating the probit regression I compute marginal effects.

In addition to the probit, I also estimate an ordered probit in order to gain back some of the variation lost when converting my categorical dependent variable into a dummy variable. This equation is a modified version of equation 1 and is restated below in equation 2. An ordered probit allows for analysis of categorical variables by assuming that the distance between each answer choice is equal which then creates an opportunity to analysis outcomes for each specific answer.

$$\begin{aligned}
 \text{INTELLIGENCE}_i = & \beta_0 + \beta_1 \text{SKIPPED}_i + \beta_2 \text{HELDBACK}_i + \beta_3 \text{HISPANIC}_i + \beta_4 \text{BLACK}_i \\
 & + \beta_5 \text{NATIVE\_AMERICAN}_i + \beta_6 \text{ASIAN}_i + \beta_7 \text{ELECTRONICSUSE}_i + \beta_8 \text{SPORTS}_i \\
 & + \beta_9 \text{FRIENDS}_i + \beta_{10} \text{BODYIMAGE}_i + \beta_{11} \text{SLEEP}_i + \beta_{12} \text{HAPPY}_i \quad (2) \\
 & + \beta_{13} \text{PARENTALINVOLVE}_i + \beta_{14} \text{SMOKES}_i + \beta_{15} \text{PARENTINCOME}_i + \beta_{16} \text{FEMALE}_i \\
 & + \beta_{17} \text{SKIPSCHOOL}_i + \beta_{18} \text{WANTSCOLLEGE}_i + \beta_{19} \text{PARENT1ED}_i + e_i
 \end{aligned}$$

My original dependent variable has six answers: moderately below average, slightly below average, average, slightly above average, moderately above average, and highly above average. The ordered probit computes likelihoods for each potential answer for the dependent variable, which are then understood through computing marginal effects for each outcome. What follows is a more specific analysis of how skipping or repeating a grade affects individuals at each answer choice. Estimating an ordered probit provides more specific results, and will hopefully be a confirmation and increase the robustness of my original findings.

## V. Results and Discussion

My two key variables in all regressions were SKIPPED and HELDBACK. The results of both the OLS and probit are presented in tables 1 and 2. The coefficient of SKIPPED in both the OLS and probit models is not statistically different from zero and cannot be interpreted. My conjecture on why this result occurred is that a myriad of results can occur when an individual skips a grade, ranging from continued success to higher difficulty completing the new grade's coursework. In addition to these reasons, only a miniscule portion of my sample skipped a grade (~2%). This makes it more difficult to find significant results. The coefficient on HELDBACK was statistically significant and had a meaningful magnitude ( $p < .001$ ). The coefficient was -0.197 meaning that an individual was 19.7 percentage points more likely to report him or herself as below average intelligence if he or she was held back. This represents a 35% decrease from the mean of the ABOVE variable. This is the most important finding of results and has potentially significant policy implications that will be discussed later.

While being Hispanic lowers the likelihood of marking yourself as above average intelligence by 8 percentage points, significant at the 95% level, being Black increases this likelihood by 8 percentage points - also significant at the 95% level. The variables SPORTS, SLEEP and HAPPY all have small marginal significant effects. Playing after-school sports or rating oneself as happy with one's life overall results in a small increase in the likelihood of rating oneself as above average intelligence, while getting enough sleep actually decreases the likelihood of rating oneself as above average intelligence. Being female reduces the likelihood of rating oneself as above average intelligence by 5

percentage points, while having a parent with a bachelor's degree increases this likelihood by 12 percentage points.

The ordered probit provides significant support for my original findings concerning grade retention, and actually provides some significant results concerning grade skipping as well. Table 3 shows the results for SKIPPED and HELDBACK at each of the answer choices. At the outcomes that represent the answers for moderately below average, slightly below average and average (1, 2, and 3 in the regression orders) being held back caused an individual to be 1, 3.5 and 10 percentage points more likely to choose that response, all else constant ( $p < 0.001$  for all three). For the outcomes that represent above average responses (4, 5, and 6) the sign of the coefficient flipped, which is consistent with my probit results. The coefficients for each were -2, -9.5, and -3 ( $p < 0.001$ ). This means that an individual that was held back was that many percentage points less likely to report any of the above average responses, respectfully.

The results for skipping a grade are also interesting for several of the outcomes mentioned above. For the moderately and slightly below average responses, individuals who were skipped a grade were less likely than others to report these responses ( $p < 0.05$  and  $p < 0.1$  respectively). Also, for the moderately above average response, responders who had skipped a grade were 6 percentage points more likely to choose this response than others, all held constant. ( $p < 0.1$ ) The others results were statistically insignificant.

The results of the ordered probit support my original findings and provide further support in that not only does being held back make an individual less likely to report themselves as one of the above average answers, but it also shows that these same individuals were more likely to report themselves as below average. It also provides some

significant results concerning grade skipping, and these results are consistent with the theory of the model. The consistency of these results strengthens my original findings.

Overall, my results present interesting findings that have real world implications concerning policy for middle and high school education. The increased likelihood of individuals who are held back to rate themselves as below average intelligence provides an interesting con-current point to other research against grade retention. My results also suggest that appropriate grade skipping can improve students' thoughts on their own intelligence, although these results are not as strong as the results concerning retention. Students who are retained believe that they are less intelligent than their peers; this could have long-term effects on their future education level, employment efforts, workforce participation and other labor market outcomes. These results could be used to support legislature or regulations that limit or create new guidelines for grade retention, or as a citation for future papers that use performance variables as their dependent variables. By providing these results, those papers can address some of the assertions of bias by citing that this paper shows lower perceived intelligence in those students.

Other less important but interesting points for discussion include some of the significant results on race variables, and the results on variables such as 'sports' and 'sleep'. The race variables for black individuals and Hispanic individuals both raise interesting questions regarding the cause of these coefficients. Perhaps future studies could investigate the relationship between this finding of lower intelligence and check for correlation between this and performance variables such as test scores.

## **VI. Conclusions**

Given my results, the policy of grade retention should be revisited and examined further in order to decide whether the potential adverse effects to self-assessed intelligence are worth any potential advantage provided by forcing the student through the same material. In addition, consistent with other current literature, I would suggest that a study be conducted that provides additional support to retained students and examines if this provides a greater benefit.

Although my original results regarding grade skipping were not significant, my ordered probit results do suggest a moderate correlation between skipping a grade and reporting a higher perceived intelligence rating, although the strength of these results is modest. Overall this research supports other current research and provides an interesting perspective when discussing the problems of same-grade versus same-peer groups. To address the original research question pointedly, my findings suggest repeating a grade negatively affects self-assessed intelligence, while skipping a grade positively affects self-assessed intelligence.

## Bibliography

- Allen, C. S., Chen, Q., Willson, V. L., & Hughes, J. N. 2009. Quality of Research Design Moderates Effects of Grade Retention on Achievement: A Meta-analytic, Multi-level Analysis. *Educational Evaluation and Policy Analysis*, 31(4), 480–499. <http://doi.org/10.3102/0162373709352239>
- Colangelo, N., S. Assouline, M. Gross. 2004. A Nation Deceived: How Schools Hold Back America's Brightest Students. The Connie Belin & Jacqueline N. Blank International Center for Gifted Education and Talent Development. The University of Iowa.
- Jackson, B. 1975. The research evidence on the effects of grade retention. *Review of Education Research* 45(4): 613-635.
- Jimerson, S. 2001. Meta-analysis of grade retention research: Implications for practice in the 21st Century. *School Psychology Review* 30(3): 420-437.
- Kretschmann, J., Vock, M., & Lüdtke, O. 2014. Acceleration in elementary school: Using propensity score matching to estimate the effects on academic achievement. *Journal Of Educational Psychology*, 106(4), 1080-1095. doi:10.1037/a0036631
- Kulik, J., and C. Kulik. 1984. Effects of accelerated instruction on students. *Review of Education Research*. 54(3): 409-425.
- Park, G., D. Lubinski & C. Benbow. 2012. When Less Is More: Effects of Grade Skipping on Adult STEM Productivity Among Mathematically Precocious Adolescents. *Journal of Education Psychology*, 105(1): 176-198.
- Robeck, M. 1968. California Project Talent: Acceleration Programs for Intellectually Gifted Pupils. Office of Education, Washington, DC.
- Rogers, K. 1992. A best-evidence synthesis of research on acceleration options for gifted students. Trillium Press.
- Warren, J., E. Hoffman, M. Andrew. 2014. Patterns and trends in grade retention rates in the United States, 1995-2010. *Educational Researcher*.

Table 1 OLS Results	N=4761	
Variable	Coefficients (Standard Errors)	
SKIPPED	0.0441739	(.04524)
HELDBACK	-0.18235***	(.01836)
HAPPY	0.0568824***	(.00914)
HISPANIC	-0.1098628	(.04597)
WHITE	-0.030693	(.03987)
BLACK	0.0420315	(.04016)
NATIVEAMERICAN	0.0187288	(.04139)
ASIAN	-0.0876099*	(.04982)
ELECTRONICSUSE	-0.0064527	(.00821)
SPORTS	0.0194956***	(.00618)
FRIENDS	-0.0004462	(.00721)
BODYIMAGE	-0.0218811**	(.00891)
SLEEP	-0.0385729**	(.01617)
PARENTINVOLVE	-0.0041724	(.00299)
SMOKES	-0.0776799***	(.01928)
PARENTINCOME	0.0007595***	(.00011)
PARENT1ED	0.126161***	(.01594)
SKIPSCHOOL	-0.0140506	(.01674)
WANTSCOLLEGE	0.1720835***	(.02938)
	***= Significant at 99% level	
	**= Significant at 95% level	
	*= Significant at the 90% level	

Table 2 Probit Results		N=4761
Variable	Marginal Effects	(Standard Errors)
SKIPPED	0.0542011	(.04924)
HELDBACK	-0.1905126***	(.01966)
HAPPY	0.0619408***	(.00988)
HISPANIC	-0.1131253**	(.04943)
WHITE	-0.0344735	(.04239)
BLACK	0.0493833	(.04226)
NATIVEAMERICAN	0.0216102	(.04433)
ASIAN	-0.0984198*	(.05403)
ELECTRONICSUSE	-0.0072212	(.00897)
SPORTS	0.0207389***	(.00676)
FRIENDS	-0.0010425	(.00785)
BODYIMAGE	-0.0238772***	(.00971)
SLEEP	-0.0410172***	(.0176)
PARENTINVOLVE	-0.0043121	(.00328)
SMOKES	-0.0832659***	(.02091)
PARENTINCOME	0.0012954***	(.00026)
PARENTIED	0.1275833***	(.01788)
SKIPSCHOOL	-0.0138816	(.01815)
WANTSCOLLEGE	0.185992***	(.03354)
	***= Significant at 99% level	
	**= Significant at 95% level	
	*= Significant at the 90% level	

Variable (N=4761)	Moderately Below Average	Slightly Below Average	Average
HELDBACK	0.0103(.00194)***	0.0349(.0049)***	0.1035(.0112)***
SKIPPED	-0.0031(.0016)**	-0.0134(.0071)*	-0.0585(.0363)
	Slightly Above Average	Moderately Above Average	Highly Above Average
HELDBACK	-0.0202(.0037)***	-0.0959(.0108)***	-0.0326(.0033)***
SKIPPED	0.0015(.0019)	0.0497(.0298)*	0.0237(.0168)
	***=Significant at the 99% level		
	**=Significant at the 95% level		
	*=Significant at the 90% level		

Table 4 Summary Statistics					
Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
ABOVE	4761	0.5662599	0.4956432	0	1
SKIPPED	4761	0.0211946	0.144048	0	1
HELDBACK	4761	0.1928923	0.3946115	0	1
HISPANIC	4761	0.1031899	0.3042393	0	1
WHITE	4761	0.6471848	0.477897	0	1
BLACK	4761	0.224149	0.4170653	0	1
NATIVEAMERICAN	4761	0.0299722	0.1705288	0	1
ASIAN	4761	0.0338257	0.1807998	0	1
ELECTRONICSUSE	4761	2.385999	0.849114	0	3
SPORTS	4761	1.435239	1.144595	0	3
FRIENDS	4761	1.983943	0.9926708	0	3
BODYIMAGE	4761	3.179191	0.7747373	1	5
SLEEP	4761	0.7332477	0.4423092	0	1
HAPPY	4761	2.146007	0.7967309	0	3
PARENTINVOLVE	4761	4.729822	2.438056	0	7
SMOKES	4761	0.1901092	0.3924292	0	1
PARENTINCOME	4761	48.36609	57.35542	0	999
FEMALE	4761	0.5101691	0.4999501	0	1
PARENTIED	4761	0.2654678	0.4416292	0	1
SKIPSCHOOL	4761	0.265896	0.4418564	0	1
WANTSCOLLEGE	4761	0.9447656	0.2284617	0	1