Logistic Regression Analysis of Students' Mathematical Beliefs of University Students

V. SUTHAR1
B. K. SOLANGI
A. A. KHOOHARO
Sindh Agriculture University Tandojam
Pakistan

R. A. TARMIZI
University Putra Malaysia, Serdang, Selangor
Malaysia

Abstract:

Malaysian science and mathematics schooling have currently undergone careful changes. Mathematics has constantly been an utter for many professions; for that reason courses of mathematics provide as a 'filter' that effectively screen out many students out of careers. Observed facts had revealed that belief of students in mathematics is critical in career ambition meeting. While, mathematical beliefs of students are directly interconnected to their mathematics ability and kinds of educational skill among graduate students. In this study students had finalized a self-reported questionnaire related to student mathematical beliefs on three aspects (beliefs about mathematics, beliefs about importance of mathematics and beliefs on one’s ability in mathematics). The Cronbach’s alpha was 0.85, representing a high level of internal consistency recorded by the reliability index. Mathematics ability (GPA) proceedings were achieved from the academic divisions of different Universities. Students were categories into the high and low ability group according to these proceedings. The authors examined mathematical beliefs of the student appropriate for a survey design in this study based on a three dimensional logistic regression model estimation technique. The results showed that, the

1 Corresponding author: E-mail: vsutahar@yahoo.co.uk
relationships between student beliefs about importance of mathematics and beliefs on one's ability in mathematics and with mathematics achievement was significant. Whereas, the Hosmer and Lemeshow results indicated that, there is a good model fit as the figures did not significantly deviate from the model.

Key words: mathematics ability, mathematics beliefs, logistic regression model, likelihood method-of-estimation

Introduction

Educators admit that the students’ beliefs and learning strategies have great impact on students’ academic achievement. Previous studies proved that educators and parents long have been cursed by the problem of students' low achievement in schools. The proficient and experienced teachers are aware that many students complete academic courses without understanding the basic concepts, self-regulations and skills that teachers intend for them to learn. In many countries like Australia, USA, and UK, professional strategies have been build up more than two decades to support mathematics and science prospectus standards (National Research Council, 1996; National Council for Teachers of Mathematics, 2000). These reforms are based on a solid foundation of research on student learning to achieve everlasting success. According to Haigh and Rehfeld (1995), (p.241) nearly all these efforts have been comply upon the postulations that integrations increases the achievement of students in science and mathematics.

Mathematics is a fundamental science subject, which taught all over the world as an individuals’ building foundation from pre schooling age. Mathematics is the only subject which is known either by literate as well as illiterate. Stage, (2001) posed that a mathematics is a barrier to success in college and university for many students. It is viewed as a difficult subject to master due to its symbolic and abstract nature. Stage (2001,
p.203) argued, “A student who is futile in mastering mathematics skills loses opportunities to enroll in a broad range of college courses, thus limiting career choice.”

The immature beliefs and inappropriate self-regulations have strong impact on students’ critical thinking, problem-solving performance, and their approach to learning through innovative skills and engagement with mathematics (Schoenfeld, 1983). At the same time, successful problem solving in mathematics requires students to be able to select and use task-appropriate cognitive strategies for understanding, representing, and solving problems (Schoenfeld, 1985). Numerous researchers (Dahl et al., 2005) mentioned that beliefs of students concern with the nature of knowledge and understanding influence the strategies used by them for several purposes such as learning, developing the confidence ability, humanizing the critical and creative thinking, improving the problem solving methods and more diversify the self-regulated learning, which in turn, affects their academic achievement in mathematics (Hofer, 1999). Pentrich, (1999) examined that the students’ beliefs and self-regulation strategies influence the students’ mathematics achievement.

Objectives

The specific objectives of the present study are:
1. To explore the extent of each construct to predict the mathematics ability of undergraduate students in university Putra Malaysia.
2. To investigate the relationships among the subconstructs of students’ mathematical beliefs.
3. To check the reliability coefficients of the sub-scales used in main construct i.e. students’ mathematical beliefs.
Significance of the study

The rapid changing in daily life by the economical and technological conditions, that individuals have never noticed before that broad range of the difficulty require them to solve it. People must get permanent learning skill and abilities in such challenging task to overcome their goals. Main purpose of this study is to examine how the implementation of relatively simple mathematical beliefs and self-regulated knowledge of students in mathematics affects an undergraduate students' ability in mathematics achievement. Several studies have analyzed the effects of students’ mathematical beliefs and critical thinking on mathematics achievement and self-regulated strategies on motivation, the present study is intended to augment the body of research related to beliefs on mathematics and self-regulated knowledge of students and their effects on student’s ability in mathematics achievement.

Materials and Methods

A research was carried out during the survey using self-response feedback form administered to the undergraduate mathematics students of the university. The questionnaire comprised on two parts in first the demographic information was asked and in second part measuring mathematical beliefs upon 47 objects with 5-positions Likert scale questions (1 and 5 being strongly disagree and agree respectively). Beliefs in mathematics measurement of students were acclimatized from Vallerand et al. (1992). During the lecture session the questionnaires were administered to undergraduate students of university of three different institutions of academic year 2009-2010. Survey feedbacks were tabulated and reported in the shape of frequencies and percentages as advised by Fraenkel and Wallen (1994). The study subjects was also carried out in the different universities of Malaysia in three mathematics
institutes for both male and female undergraduate students enrolled during the second semester of 2009-2010. Schedules were prepared to manage the instruments acquiring upon the authorization and agreement. Students respondent were also repeated that their exact responses would not be allocated with their course instructors and would not affect their course record. They were asked over to answer to the matters properly and sincerely. In current study the consistency coefficients of the subscales were ensured in this study. Assess the ability of the students in the mathematics were observed from grade scores obtained from offices records. And for the prediction of mathematics ability of students from a logistic regression performed between mathematics high- and low-ability in their university studies. Logistic regression model was also used to investigate the influence of demographic and educational aspects mathematical ability of university students. Further, to ensure the accuracy and predictive validity of the logistic regression model the analysis was conducted.

Results

Mathematics beliefs about students’ mathematical high- and low-ability influenced by the key research question for this study, and it is very important from a theoretical and practical point of view as a new area of vital research surrounding the mathematical beliefs is beginning to get significant attention in psychology and mathematics education research (Leder, Pehkonen, and Torner, 2002). Further this study was also to determine the predictive model depicting the relationship between the predictive and criterion variables using both the logistic regression model.
Mathematics Ability

Similarly, consider the students’ mathematics ability is the dependent variable in this study, which measured by GPA and obtained from the different academic divisions of universities in Malaysia. There was no missing data, therefore, n = 473 responses were valid with a minimum value of 1.63, maximum value of 4.00, a mean of 3.35 with standard error is .024, a standard deviation of .56, and skewness and kurtosis have values of -.78 and .297 respectively, which are responsible of the distributional behaviour and shape of the variable. While this research was carried out to determine the affective and cognitive factors of students’ mathematics ability in Malaysian research public universities. It is essential to provide the further information on the mathematics score (GPA). Table 1 illustrates the frequencies and percentages columns are the average of each of the GPA (grades) obtained in the mathematics regular examinations for all the 473 respondents of the sample. In this sample only one student obtained “D+” grade, 7 students obtained “C-” grade, 16 students achieved “C” grades, 23 students received “C+” grades, 46 students obtained “B-” grades, 143 students received “B” grades, 87 students obtained “B+” grades, 78 students received “A-” grades, and 72 students obtained “A” grades. In addition, the mathematics score (GPA) and percent of students, they obtained different grades; are presented in the first and last columns of the Table 1 respectively. For more explanation the mathematics score (GPA) variable, a “D+” grade was equal to “1.50”, “C+” grade equal to “1.75”, in this way the detail is given in the first column of Table 1.

<table>
<thead>
<tr>
<th>Mathematics Score (GPA)</th>
<th>Grades</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.50</td>
<td>D+</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>1.75</td>
<td>C</td>
<td>7</td>
<td>1.4</td>
</tr>
</tbody>
</table>
Exploratory Data Analysis

Table 2 shows the frequency for two groups of mathematics high- and low-ability of students for binary logistic regression analysis was performed. Peng, Lee, & Ingersoll, (2002) stated that logistic regression is a suitable multivariate method for describing and testing associations between a dichotomous/binary (0/1) response variable and a many categorical variables. A group of mathematics low-ability (0) students was comprised on 143 (30.2%) with 108 (75.5%) females and 35 (24.5%) male respondents. The second group of mathematics high-ability (1) students was comprised on 330 (69.8%) with 238 (72.1%) female and 92 (27.9%) for male respondents (Table 3).

Table 2: Frequency shows a high- and low-ability of undergraduate mathematics students

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Ability (0)</td>
<td>143</td>
<td>30.2</td>
</tr>
<tr>
<td>High Ability (1)</td>
<td>330</td>
<td>69.8</td>
</tr>
<tr>
<td>Total</td>
<td>473</td>
<td>100.0</td>
</tr>
</tbody>
</table>

The assessment of the relationship between response (ability (GPA) in mathematics) and predictor variables is done by using the binary logistic regression analysis in this research. Further the analysis was conducted to achieve estimating effects of the predictor variables on dependent variable of the respondent. Logistic regression was indicated because the dependent variable (high- and low-ability in mathematics) is dichotomous/binary (0/1). Model assumptions were checked and it was determined that there was no any subscale of three main
constructs was not strongly correlated (Cohen, 1988). Therefore, multicollinearity of predictors (3 subscales of students’ beliefs in mathematics) cannot inflate the estimates of variance. After fulfill the assumption of logistic regression model, it is not desirable to minimize these effects before analyzing the results (Howell, 1999; Pedhazur, & Schmelkin, 1991). Supporting data for the present study from the user’s manual also not showed a high correlation between these scales (Weinstein & Palmer, 2002).

The descriptive statistics of the independent variables were separated in mathematics high- and low-ability for the entire sample is provided in Table 3. Two groups (mathematics high- and low-ability) having mean and standard deviation of the age of the respondents were 21.14 and 20.98 years and .88 and .90 years respectively. Table depicted the minimum, maximum, mean, and standard deviation of student ability for both group of mathematics high- and low-ability. The mean scores of the students’ mathematical beliefs were demonstrated on high- and low-ability two student groups. The highest mean score of students’ beliefs about mathematics was 3.797 for beliefs about importance of mathematics in mathematics high-ability group and 3.575 for beliefs about ones mathematics ability in mathematics low ability group of the respondents. The lowest mean score was 3.730 for beliefs about mathematics in high ability group and 3.452 also for beliefs about mathematics in low ability of two groups.

Table 3: Descriptive statistics of high- and low-ability in mathematics for logistic regression analysis

<table>
<thead>
<tr>
<th>Variables</th>
<th>High ability in mathematics</th>
<th>Low ability in mathematics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minum</td>
<td>Max</td>
</tr>
<tr>
<td>Ability in Mathematics</td>
<td>3.13</td>
<td>4.00</td>
</tr>
<tr>
<td>Beliefs about mathematics</td>
<td>2.00</td>
<td>5.00</td>
</tr>
</tbody>
</table>
The analysis was carried out on the entire sample (n = 473) was used. In this case forward linear regression was employed using an alpha (α) value of .05 as the cut-off in. Enter logistic regression analysis method was chosen because this method is widely used in educational research (George & Mallery, 2003). In binary logistic regression analysis, age, race and the 3 students’ beliefs in mathematics subscales were used as predictor variables. General descriptive data is offered first. This data is followed by the logistic regression results.

Logistic Regression Analysis

The probability for students’ ability in mathematics can be written by the logistic regression equation as:

\[ P(\text{ability}) = \frac{1}{1 + e^{-Z}} \]

The binary logistic regression analysis model was carried out on the data for the entire sample with student ability in mathematics was dependent variable and 3 affective and one demographic categorical (race) predictor variables. In the current study the sample data set, dependent variable was dichotomous (0/1) high- and low-ability in mathematics of respondent. Binary logistic regression analysis model was chosen because this method is widely used in education research (George and Mallery, 2003). In this case the logistic model was more accurate at predicting student mathematics ability than the null model.

The binary logistic regression analysis was done included a set of 4 parameters that studied, if the two groups of students (students’ high- and low-ability in mathematics) differed in affected and demographic (race) variables towards mathematics ability. For the entire sample of data set (n = 473)
for student mathematics high- and low-ability a binary logistic regression model was applied. The significant differences of affective variables were observed among the three sub constructs were those of “Beliefs about mathematics”, “Beliefs about importance of mathematics”, and “Beliefs about ones’ ability in mathematics” variables (p < .00) are shown in Table 4.

As well, on the scale for each unit increase in the respondents’ amount of Beliefs about mathematics and beliefs about ones’ ability in mathematics subscale for same construct, the respondents’ probability of being in the high–ability group would likely increase by .82.7% with .203 and 100.6% with .246 standard error respectively. Also all three subscales of students’ beliefs about mathematics variable were contributed statistically significant (p < .001) for classifying the correctly in the two student groups (high- and low-ability in mathematics). The odds ratio of this subscale is 2.287 and 2.736, which indicates that it is more likely to increase 173.6% and 128.7% with 95% respectively. The confidence interval of Beliefs about mathematics was .754 and 2.453 lower and upper respectively, when student change the category from 0 to 1 (low- to high-ability) group.

Table 4: Coefficient of predictors

<table>
<thead>
<tr>
<th>Variables</th>
<th>B</th>
<th>S.E.</th>
<th>Wald</th>
<th>df</th>
<th>Sig.</th>
<th>Exp(B)</th>
<th>95.0% C.I. for EXP(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beliefs about mathematics</td>
<td>.827</td>
<td>.203</td>
<td>16.675</td>
<td>1</td>
<td>.000</td>
<td>2.287</td>
<td>.754  2.453</td>
</tr>
<tr>
<td>Beliefs about importance of</td>
<td>.563</td>
<td>.168</td>
<td>11.271</td>
<td>1</td>
<td>.001</td>
<td>1.756</td>
<td>1.031 2.411</td>
</tr>
<tr>
<td>mathematics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beliefs about ones’ mathematics</td>
<td>1.006</td>
<td>.246</td>
<td>16.680</td>
<td>1</td>
<td>.000</td>
<td>2.736</td>
<td>1.333 4.942</td>
</tr>
<tr>
<td>ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>race(1)</td>
<td>-.473</td>
<td>1.249</td>
<td>.146</td>
<td>1</td>
<td>.703</td>
<td>.621</td>
<td>.054 7.179</td>
</tr>
<tr>
<td>race(2)</td>
<td>.453</td>
<td>1.278</td>
<td>.126</td>
<td>1</td>
<td>.723</td>
<td>1.574</td>
<td>.129 19.250</td>
</tr>
</tbody>
</table>
Binary logistic regression analysis was conducted with racial/ethnic groups of respondents as predictor variables. The race was used as predictor variable in the general estimating model, it was contributing statistically significant (p = .018). Furthermore, the race/ethnicity was recoded in four dummy variables for the current study for sample size, n = 473 (Table 4). These dummy variables were created, presenting Malay or not, Chinese or not, Indian or not, and others or not respectively. The Malay category was considered as the reference category.

The binary prediction model indicates that the estimated odds ($e^{\beta_i}$, $i = 1, 2, 3$) of student mathematics ability for Chinese equal .62 times the estimated odds for Malay with 95% confidence of interval .054 and 7.18 lower and upper level respectively, when student change the category from 0 to 1 (low- to high-ability) group. Secondly, the estimated odds of student ability for Indian are 1.57 times the odds for Malay with 95% confidence of interval .129 and 19.25 lower and upper level respectively, when respondent change the category from 0 to 1 (low- to high-ability) group in Table 4.

Third and finally, the estimated odds of student ability for others are .19 times the odds for Malay with 95% confidence of interval .007 and 2.689 lower and upper level respectively, when respondent change the category from 0 to 1 (low- to high-ability) group. These results show that the estimated likelihood of student ability in mathematics for Chinese and Indian are lesser than that for Malay (reference category). In the contrarily, others (last category of the race) are more likely to be able than that for reference (see Table 4).

After distinguish between two groups of students, who perform high-ability 330 (69.8%) cases and those who perform...
low-ability in mathematics 143 (30.2%) cases of the university mathematics undergraduate students in Malaysia, a binary regression model was applied. The effects of binary logistic regression model presented an affective association between the three sub construct variables and the response variable in this research study. In order to determine the accuracy of logistic regression model that predicts students’ mathematical ability based on 3 subscales of students’ mathematical beliefs and race/ethnicity of respondents among the university mathematics undergraduate students in Malaysia.

Likelihood is the probability of observed results; assumed the construct estimates and -2 times the log likelihood (-2LL) is a measure of how well the estimated model fits the categorical data. For the logistic regression model that contains only the constant, -2LL is 579.730, as shown in Table 5. Goodness-of-fit statistic for the model with significant factors, the value of -2LL for the proposed model is $-2 \log \text{likelihood} = 477.329$; $\chi^2 (6) = 102.408$, $p < 0.000$, which is smaller than the -2LL for the model containing only a constant. The Cox & Snell $R^2$ and Nagelkerke $\tilde{R}^2$ are statistics that endeavor to quantify the explained proportion “variation” in the binary regression model. The value of Cox & Snell $R^2$ is .20 and reveals that about 20% of the variation in the outcomes the model explains variable. Likewise, the Nagelkerke $\tilde{R}^2$ is shown in the same table as .28, which specifies that about 28% of the variation in the outcome variable is elucidated by the logistic regression model.

Table 5: Omnibus tests of model coefficients and model statistics

<table>
<thead>
<tr>
<th>Omnibus Tests</th>
<th>Model Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-square</td>
<td>$\log$ likelihood</td>
</tr>
<tr>
<td></td>
<td>only with constant</td>
</tr>
<tr>
<td></td>
<td>with all predictor</td>
</tr>
<tr>
<td>Sig</td>
<td>$\log$ likelihood</td>
</tr>
<tr>
<td>f</td>
<td>Cox &amp; Snell $R^2$</td>
</tr>
<tr>
<td></td>
<td>Nagelkerke $\tilde{R}^2$</td>
</tr>
<tr>
<td>Model</td>
<td>102.408</td>
</tr>
</tbody>
</table>

(a) and (b) Estimation terminated at number 3 and 7 iterations respectively.
The omnibus test disclosed a comprehensive significant model \( \chi^2 (6) = 102.408, (p < 0.000) \) in Table 5. Additionally, the regression analysis exhibited that nine predictors were significant at the 5% level to the response variable was high- and low-ability in mathematics (see Table 4). Furthermore, Table 5 presents the 0.456 and .645 values for Cox and Snell \( R^2 \) and the Nagelkerke \( \bar{R}^2 \) respectively. These pseudo-\( R^2 \) procedures can be treated as “somewhat analogous to \( R^2 \) in linear regression” (McCoach & Siegle, 2003). Furthermore, the Hosmer and Lemeshow test was non-significant with a \( \chi^2 (8) \) of 7.033, \( p = 0.533 \) which showed a goodness-fit fit of model, since the data did not significantly deviate from the model.

**Classification of Cases**

This method of assessing the successive accuracy of a model is to evaluate its ability to predict correctly the outcome category for cases for which outcome is known. If a case has diagnosed diseased, for instance, it can be seen if the case is correctly classified as diseased on the basis of other predictor variables. While the overall percent correctly predicted seems exceptionally good at 74.6%, the researcher must note that blindly estimating the most frequent category (high-ability) for all cases would yield an evidence higher percent correct (74.6%), as noted in Table 6. This implies high-ability status can be differentiated on the basis of students’ mathematical beliefs and race/ethnicity for this survey data.

**Table 6: Classification Table with all predictor variables**

<table>
<thead>
<tr>
<th>Predicted Ability</th>
<th>Observed Low Ability</th>
<th>Observed High Ability</th>
<th>Percentage Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1 Ability</td>
<td>Low Ability</td>
<td>High Ability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>88</td>
<td>38.5</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>298</td>
<td>90.3</td>
</tr>
<tr>
<td>Overall Percentage</td>
<td></td>
<td></td>
<td>74.6</td>
</tr>
</tbody>
</table>
Precisely, this set of variables was explained the variance by the Cox and Snell $R^2$ corresponded 20%, on the other hand the Nagelkerke $\tilde{R}^2$ corresponded 28%. This predicted model demonstrated 90.3% were correctly classified of the students group in the high-ability in mathematics, while 38.5% of the students were correctly classified to be in the low-ability group. As a result in the overall model, 74.6% of the sample was classified correctly. Logistic model was well fitted with 3 predictor subscale variables of the construct: students’ mathematical beliefs and race/ethnicity for the current survey data set of this study.

**Conclusion**

The recent study is that the students’ mathematical beliefs of three mathematical institutes of various Universities in Malaysia are significantly different in high-ability in mathematics from the students who are low-ability in mathematics. Hence the high-ability student in study had higher mathematical beliefs compared to the low-ability student. The students’ mathematical beliefs perform manipulate to a certain degree on high- or low-ability in mathematics. Though, cause-effect relationships cannot be claimed since this is not an experimental or longitudinal study (Papanastasiou and Zembylas, 2006). Therefore, no fundamental effects due to mathematical beliefs of the students can be characterized to the choice of students either to high- or low-ability in mathematics. These results also showed that other surroundings variables may be prominent such as, these secondary or post-secondary educations of students or attendance to the developmental course during these levels.

A univariate logistic regression modeling was revealed in this study. In the current study, maximum likelihood examination of the various beliefs of student also studied between high- and low-ability in mathematics of students.
showed that mathematics high-ability than low-ability classified “students’ beliefs about mathematics” on priority of a knowledge and understanding.

If mathematics and science instructors were merely follow the effects of the statistical sophisticated logistic regression analysis from the current study to object groups for the endorsement of sound knowledge, insight, cognizance and understanding. A program to enhance the awareness and thoughtfulness of the significance of mathematics may have been extended for all students of Malaysian science at secondary and post-secondary levels, by entering in tertiary education at college and University. These outcomes from the logistic regression modeling on the beliefs of students about mathematics pointed out that the university education method was well-organized and differences on the mathematical beliefs of the students at high- and low-ability mathematics groups can be clearly differentiated.

ACKNOWLEDGEMENT

We amusingly wish to thanks all the students who voluntarily participated in this study for their attention in giving feedback the self-reported questionnaire in facilitating the investigators to understand their beliefs about mathematics. We are also grateful to the course instructors. Finally, we state our appreciation to Deputy Registrars (Academic Divisions) of three various universities in Malaysia for their precious support and assistance in providing record (grades) of students.

REFERENCES


