ACT WORKING PAPER 2017 (1)

Can Psychosocial Factors Predict First-to-Second Year College Retention Above and Beyond Standard Variables?

A Mixed Effects Multinomial Regression Analysis

David R. King, Georgia Institute of Technology
Edwin Ndum, PhD, ACT

February, 2017

ACT Working Paper Series

ACT working papers document preliminary research. The papers are intended to promote discussion and feedback before formal publication. The research does not necessarily reflect the views of ACT.

David R. King is a graduate student in the School of Psychology at Georgia Institute of Technology. His research interests include multidimensional item response theory and Bayesian parameter estimation.

Edwin Ndum is a research scientist in Statistical and Applied Research. He does research on test reliability and validity of test scores, trends, and differential effects of psychosocial assessments.

The authors thank Jeff Allen, Krista Mattern, Alex Casillas, Edgar Sanchez, Kurt Burkum, Teri Fisher, and Emily Neff for their helpful comments and suggestions on earlier drafts of this report.

Abstract

We examined the validity of 10 psychosocial factors for predicting retention status (stay, transfer, or drop out) at the start of second year of college by fitting a mixed-effects multinomial logistic regression model. Data consisted of retention records of 9,364 students from 31 four-year institutions. Predictors included commonly used variables for predicting college retention (i.e., high school grade point average [HSGPA], ACT Composite score, gender, family income, and race/ethnicity) and the 10 psychosocial factors. We examined the differential effects of psychosocial factors on retention status, namely which psychosocial attributes differentiated students who stayed from those who transferred or dropped out, and which psychosocial factors differentiated students who transferred from students who dropped out. Our results indicated that *Academic Discipline, Commitment to College, Social Connection*, and *Academic Self-Confidence* incrementally predicted first-to-second year retention status above and beyond the commonly used variables. Specifically, students with higher *Academic Discipline*, *Commitment to College*, or *Social Connection* were *more* likely to stay at their institutions than to drop out or transfer; whereas students with higher *Academic Self-Confidence* were *less* likely to stay than to drop out or transfer, and also *less* likely to transfer than to drop out. These results were confirmed using a follow-up cross-validation analysis on a holdout sample of 4,804 students. These findings may have implications for interventions aimed at identifying students at risk of dropping out of college and for improving retention to the second year of college.

Keywords: psychosocial factors, college retention, validity research, multilevel modeling

Can Psychosocial Factors Predict First-to-Second Year College Retention Above and Beyond Standard Variables? A Mixed Effects Multinomial Regression Analysis

A recent national survey on student retention (Habley, Valiga, McClanahan, & Burkum, 2010) observed that approximately 26% of students at four-year institutions left their original institutions for some reason (e.g., dropped out, transferred) between the first and second year (ACT, 2016, Figure 2). As a result of this considerable decline in second-year enrollment, college retention research often focuses on measuring variables related to first-to-second year retention (e.g., Herzog, 2005; Belloc, Maruotti, Petrella, 2011).[1]

Students leave college for a number of reasons including academic issues (McGrath & Braunstein, 1997), personality characteristics (Tross, Harper, Osher, & Kneidinger, 2000), attitudes toward college (Rivas, Sauer, Glynn, & Miller, 2007), and financial hardship (Ishitani & DesJardins, 2002). One line of research has focused on psychosocial factors (PSFs) and study skills as predictors of college retention (Robbins, Lauver, Le, Davis, Langley, & Carlstrom, 2004; Le, Casillas, Robbins, & Langley, 2005).

PSFs are important predictors of college retention because they can be used to identify students at risk of leaving college (Allen, Robbins, & Sawyer, 2010). PSFs have been found to significantly predict third-year college retention, even after accounting for standard variables such as high school grade point average (HSGPA), ACT Composite score, gender, family income, and race/ethnicity (Allen, Robbins, Casillas, & Oh, 2008). Considering the substantial decline in second-year enrollment, the current study focuses on the relationship between PSFs and first-to-second year college retention.

[1] First-to-second year retention is defined as the percentage of students who enrolled as first-time, full-time students who return one year later as either full- time or part-time students at the same institution.

Although retention is often operationalized as a binary outcome (i.e. student stays vs. student does not stay), the dichotomy may confound important within-group differences among students who do not stay for various reasons. Of particular interest in the retention literature are differences between students who drop out and students who transfer (Herzog, 2005). In a study examining the relationship between PSFs and retention to the third-year of college, students with higher first-year academic performance were *more* likely to transfer than drop out (Allen et al., 2008). This finding was corroborated in a case study on retention at Italian universities that found students with a better educational background to be *more* likely to transfer and *less* likely to drop out (Belloc, Maruotti, & Petrella, 2011). Given these differences, the current study also examined the differential effects of PSFs on students who transferred and students who dropped out.

Objectives of the current study

The primary goal of the current study was to examine the validity of 10 PSFs for predicting first-to-second year college retention. In that regard, we investigated the following research questions:

1. Are PSFs related to first-to-second year retention?
2. Do PSFs incrementally predict college retention beyond academic and demographic variables?
3. Which PSFs differentiate students who stay at their institution from students who transfer or drop out, and/or differentiate students who transfer to other institutions from those who drop out?
4. How do estimates from the current study compare with estimates from an earlier study examining these PSFs?

5. Can the PSFs identify students at risk of leaving their institutions? (6) Do the PSFs provide useful information for creating effective interventions and increasing first-to-second year retention rates?

Method

Operationalizing retention

A limitation in modeling an outcome with three nominal categories (stay, transfer, or drop out) is that comparisons are typically performed between a reference level (e.g., drop out) and all other levels of the outcome variable (e.g., drop out vs. transfer, drop out vs. stay). These comparisons may preclude a comparison of substantive interest, such as *stay* versus *transfer* and *drop out*. As a solution, Hedeker (2003) suggested specifying $C-1$ contrasts, where C is the number of levels of the outcome. For the current study, the probability that student i in school j has first-to-second year retention status c, conditional on random effects $\boldsymbol{\beta}$, is given by:

$$\Pr(Y_{ij} = c \mid \boldsymbol{\beta}) = \frac{\exp(z_{ijc})}{\sum_{h=1}^{C} \exp(z_{ijh})} \text{ for } c = 1, 2, ..., C$$

where z_{ijc} is a multinomial logit equal to the sum of the weighted predictors.

Customized contrasts between the three levels of retention (stay, transfer, or drop out) can be specified by multiplying the multinomial logits by the desired contrast coefficients. For example, dummy coding gives the usual outcome probabilities for the multinomial logistic regression model (the red numbers were used to emphasize the combination of the contrast coefficients):

$$\Pr(Y_{ij} = drop\,out \mid \boldsymbol{\beta}) = \frac{\exp(0z_{ij2} + 0z_{ij3})}{\exp(0z_{ij2} + 0z_{ij3}) + \exp(1z_{ij2} + 0z_{ij3}) + \exp(0z_{ij2} + 1z_{ij3})} = \frac{1}{1 + \sum_{h=2}^{C} \exp(z_{ijh})},$$

$$\Pr(Y_{ij} = transfer \mid \boldsymbol{\beta}) = \frac{\exp(1z_{ij2} + 0z_{ij3})}{\exp(0z_{ij2} + 0z_{ij3}) + \exp(1z_{ij2} + 0z_{ij3}) + \exp(0z_{ij2} + 1z_{ij3})} = \frac{\exp(z_{ij2})}{1 + \sum_{h=2}^{C} \exp(z_{ijh})}, \text{ and}$$

$$\Pr(Y_{ij} = stay \mid \boldsymbol{\beta}) = \frac{\exp(0z_{ij2} + 1z_{ij3})}{\exp(0z_{ij2} + 0z_{ij3}) + \exp(1z_{ij2} + 0z_{ij3}) + \exp(0z_{ij2} + 1z_{ij3})} = \frac{\exp(z_{ij3})}{1 + \sum_{h=2}^{C} \exp(z_{ijh})}.$$

In the current study, orthogonal contrasts were specified for the three categories of retention status. The first contrast compared students who stayed against the combined average of students who transferred and dropped out. The second contrast compared students who transferred (coefficient = 1/2) against students who dropped out (coefficient = –1/2). Note that these two sets of contrasts are orthogonal, and therefore, fully account for non-overlapping variability in the retention outcome. By weighting the probability functions with these contrast coefficients, the multinomial logits, z_{ijc}, become interpretable as the log-odds of "stay" versus "transfer/drop out" and the log-odds of "transfer" versus "drop out." Namely, multiplying the multinomial logits by the contrast coefficients associated with each outcome-type resulted in:

$$\Pr(Y_{ij} = drop\,out \mid \boldsymbol{\beta}) = \frac{\exp[(-1/3)z_{ij2} + (-1/2)z_{ij3}]}{\exp[(-1/3)z_{ij2} + (-1/2)z_{ij3}] + \exp[(-1/3)z_{ij2} + (1/2)z_{ij3}] + \exp[(2/3)z_{ij2} + (0)z_{ij3}]},$$

$$\Pr(Y_{ij} = transfer \mid \boldsymbol{\beta}) = \frac{\exp[(-1/3)z_{ij2} + (1/2)z_{ij3}]}{\exp[(-1/3)z_{ij2} + (-1/2)z_{ij3}] + \exp[(-1/3)z_{ij2} + (1/2)z_{ij3}] + \exp[(2/3)z_{ij2} + (0)z_{ij3}]},$$

and

$$\Pr(Y_{ij} = stay \mid \boldsymbol{\beta}) = \frac{\exp[(2/3)z_{ij2} + (0)z_{ij3}]}{\exp[(-1/3)z_{ij2} + (-1/2)z_{ij3}] + \exp[(-1/3)z_{ij2} + (1/2)z_{ij3}] + \exp[(2/3)z_{ij2} + (0)z_{ij3}]}.$$

Measuring psychosocial factors

This study examined 10 PSFs (*Academic Discipline*, *Academic Self-Confidence*, *Commitment to College*, *Communication Skills*, *General Determination*, *Goal Striving*, *Social Activity*, *Social Connection*, *Steadiness*, and *Study Skills*) measured by ACT Engage® College (ACT, 2012), an instrument formerly known as the Student Readiness Inventory (Le et al., 2005). The psychosocial constructs were developed based on previous research on the effects of nonacademic-related skills on academic outcomes. Following item-development, the survey was administered to high school and college students and an exploratory factor analysis was conducted to determine the number of interpretable factors to extract. Items that loaded on these factors were selected. To cross-validate the observed factor structure, a confirmatory factor analysis was conducted on a holdout sample with items reselected based on the magnitudes of the factor loadings. The first-order factors were then factored again to determine the structure of the second-order factors. The final model consisted of 10 first-order latent constructs, each measured by 10 to 12 graded response items, and three second-order constructs. The model provided a good-fit for the data (χ^2 (10,250) = 10,486.72, CFI = .99, RMSEA = .012) and the factor structures were consistent with the *a priori* operational definitions of the 10 PSFs and the three higher-order domains.

A validation study showed three PSFs – *Academic Discipline*, *Commitment to College*, and *Social Connection* – were statistically significant predictors of second-to-third-year college-retention, even after accounting for the effects of academic achievement and demographic variables on retention (Allen et al., 2008). *Commitment to College* and *Social*

Connection had direct effects on third-year retention, whereas *Academic Discipline* indirectly affected retention by means of first-year academic performance.

Data preparation

The data used for this study consisted of 14,411 students from 45 four-year institutions who completed the ACT Engage College assessment. First-to-second year of college retention status was determined from enrollment data obtained from the National Student Clearinghouse.[2] First-to-second year college retention status by gender, family income, and race is shown in Table 1 (Appendix A). Gender was defined as a binary predictor, with male coded as "1" and female coded as "0." Family income was defined as an ordinal predictor, with family income less than $36,000 coded as "1," income between $36,000 and $100,000 coded as "2," and income greater than $100,000 coded as "3." The 10 PSFs were each rescaled to have a mean of zero and a standard deviation of one.

Students with full information on required variables were retained for the analysis (n =14,411; starting sample = 21,031). We randomly sampled and set aside one-third of the cases (n = 4,804) for the cross-validation analysis. We dropped institutions (level two units) with fewer than 10 students to improve the accuracy of the parameter estimates (Austin, 2010). This resulted in a training set of 9,364 students nested in 31 four-year institutions. The data did not contain any variable that could be used to determine the destination-institution for students who transferred from first to second year of college. However, we assumed that those who stayed for the second year enrolled in one of the 31 institutions, whereas those who dropped out did not register in any of the institutions.

[2]www.studentclearinghouse.org

Estimation procedure

We estimated model parameters using adaptive Gaussian quadrature (Rabe-Hesketh, Skrondal, & Pickles, 2002) with the SAS Nonlinear Mixed Models procedure (PROC NLMIXED). The NLMIXED procedure allowed for direct implementation of the log likelihood function and the specification of custom contrasts in the multinomial logistic model. We specified both fixed effects (the student-level predictors) and random effects (the school-level intercepts). The random intercept for each contrast accounted for the clustering of students within schools. These intercepts were assumed to be multivariate normal, with the mean vector set to zero and the covariance matrix specified as an identity matrix.

Results

Demographic characteristics and retention status

We obtained descriptive statistics to assess the relationship between the PSFs and first-to-second year college retention. In general, PSF scores were higher for students who stayed at their institutions than for students who transferred or dropped out; and higher for students who transferred than for students who dropped out. For example, students who stayed at their institutions had higher standardized *Academic Discipline* scores (average= .09) than students who transferred (*average* = –0.14) or dropped out (*average* = –0.23).

Point-biserial correlations (r_{pb}) were calculated between the standardized PSF scales and the "stay vs. drop out/transfer" and "transfer vs. stay" contrasts. All correlations between PSFs and "stay vs. drop out/transfer" were positive (r_{pb} = .01 to .12), indicating that the PSF scores of students who stayed at their institutions tended to be higher than the combined average PSF scores of students who transferred and students who dropped out. Furthermore, most of the correlations between the PSFs and "transfer vs. drop out" were positive (r_{pb} = –.03 to .06),

indicating that in general, students who transferred had higher PSF scores than students who dropped out. The means of the standardized PSF scores for each retention status category are shown in Table 2 (Appendix A), and the point-biserial correlations between standardized PSF scales and retention status are shown in Table 3 (Appendix A).

We examined the relationship between PSFs and the "stay vs. transfer/drop out" in greater detail by splitting each PSF into deciles (10 equal groups), and then, calculated the retention rate (i.e. the percentage of students who stay at their institutions) at each decile. We found that the largest increases in retention occurred between the first and third deciles for each scale (i.e. bottom 20% of scores).

Six of the PSF scales were monotonically related to retention status. For example, retention rates increased at each decile of *Academic Discipline*, as shown in Figure 1 (in Appendix B). In contrast, four of the PSFs, namely *Goal Striving, Social Activity, Study Skills,* and *Steadiness*, were *not* monotonically related to retention. For these scales, retention rates peaked in the 4^{th} to 6^{th} decile range and then decreased in the 8^{th} to 10^{th} decile range. For example, retention rates peaked at the sixth decile of *Social Activity*, and then decreased at higher deciles as shown in Figure 2 (Appendix B). The relationships between retention and the other PSFs at each decile are shown in Figure 3 through Figure 10 in Appendix B.

Incremental prediction of PSFs

We assessed incremental prediction of PSFs beyond standard variables by specifying full and reduced regression models. The standard variables in the reduced model included academic achievement variables (HSGPA & ACT Composite score) and demographic variables (gender, race, & family income). The full model included these standard variables and the 10 psychosocial scales. We standardized all continuous predictors to facilitate the comparison of the

estimated parameters, and used effect coding for the categorical predictors for the purpose of interpreting the intercept as the overall estimated log-odds for a given contrast.

A likelihood ratio test[3] was conducted to assess relative fit between the full and reduced models. The full model fitted the data significantly better than the reduced model (χ^2 (df = 20) = 147, $p < .01$), indicating that the 10 PSFs from ACT Engage College incrementally predicted college retention beyond the standard variables.

Two additional likelihood ratio tests were conducted to determine the necessity of including the random effects in the model. The first test compared the full model with independent random intercepts to the full model without random intercepts (i.e. fixed effects only). The model with independent random intercepts fit the data significantly better than the full model without random intercepts (χ^2 (df = 2) = 97, $p < .01$). The second test compared the full model with correlated random intercepts (i.e., an additional parameter estimate measuring the correlation between random intercepts) to the full model with independent random intercepts. The full model with correlated random intercepts did not fit the data significantly better than the full model with independent random intercepts, with χ^2 (df = 1) = 3 and p = .051. Therefore, the model with independent random intercepts was retained.

Relationship between PSFs and first-to-second year of college retention

Students with higher *Academic Discipline* ($\widehat{\beta}$ = 0.284), *Commitment to College* ($\widehat{\beta}$ = 0.127), and *Social Connection* ($\hat{\beta}$ = 0.135) were significantly *more* likely to stay at their college than to transfer or drop out. Meanwhile, students with higher *Academic Self-Confidence* ($\hat{\beta}$ = –0.216) were significantly *less* likely to stay than to transfer or drop out. *Academic Self-*

[3] A likelihood ratio test was appropriate because the reduced model was nested within the full model and we estimated both models using maximum likelihood.

Confidence was the only scale that significantly differentiated students who transferred from students who dropped out. Students with higher *Academic Self-Confidence* ($\hat{\beta}$ = – 0.216) were significantly *less* likely to transfer than to drop out. Parameter estimates and standard errors for the full model are shown in Table 4 (Appendix A).

Comparison to results from Allen et al. (2008)

The Allen et al. (2008) study[4] included first-year academic performance, a measure not included in the current study. Furthermore, Allen et al. (2008) used only three of the 10 PSFs in the current study to model second-to-third year retention, whereas this study used all 10 PSFs to model first-to-second year retention.

A separate set of contrasts between retention status outcomes was specified to compare model estimates obtained from the current sample with estimates obtained from an earlier validation sample (Allen et al., 2008). In Allen et al. (2008), contrasts were specified between "stay vs. drop out" and "transfer vs. drop out;" all categorical variables were dummy-coded; "Male" was specified as the reference group for *Gender*; "White" for *Race*; and "Moderate Income" for *Family Income.*

Although the predictors used in the current study and the Allen et al. (2008) study were not the same, the model estimates were still comparable. The estimates for demographic predictors were in the same direction and achieved the same level of statistical significance in both studies. In the current study, after controlling for family income and race/ethnicity, males were *more* likely than females to stay at their institutions than to drop out ($\hat{\beta}$ = 0.036). Controlling for family income and gender, African Americans were *more* likely than Caucasians

[4] As a cautionary note, we should mention that the predictors in the two studies were not the same and, therefore, cannot be directly compared.

to stay than to drop out ($\hat{\beta}$ = 0.321), Hispanics were *more* likely than Caucasians to stay than drop out ($\hat{\beta}$ = 0.159), and students of other ethnicities were *less* likely than Caucasians to stay than drop out ($\hat{\beta}$ = –0.289).

The family income predictor in the current study was operationalized differently than the socioeconomic status (SES) predictor in the Allen et al. (2008) study, but the results were still comparable. Namely, after controlling for gender and race/ethnicity, students from lower-income families were *less* likely than students from moderate-income families to stay than drop out ($\hat{\beta}$ = –0.467), whereas students from high-income families were *more* likely than students from moderate-income families to stay than drop out ($\hat{\beta}$ = 0.276).

Three PSFs common to both studies were *Academic Discipline, Commitment to College,* and *Social Connection*. Most notably, students with higher *Academic Discipline* were *more* likely to stay than drop out in the current study ($\hat{\beta}$ = 0.341), but were *less* likely to stay than drop out in the Allen et al. (2008) study ($\hat{\beta}$ = –0.179). An explanation for this discrepancy is that *Academic Discipline* was mediated by first-year academic performance in the Allen et al. (2008) study – this study did not control for first-year academic performance. Indeed, the total effect of *Academic Discipline* was positive and statistically significant in the Allen et al. (2008) study.

Furthermore, *Commitment to College* and *Social Connection* were statistically significant and in the same direction in both studies. Namely, students with higher *Commitment to College* were *more* likely to stay than drop out ($\hat{\beta}$ = 0.120), and students with higher *Social Connection* were also *more* likely to stay than drop out ($\hat{\beta}$ = 0.137). Full model estimates for the "stay vs. drop out" and "transfer vs. drop out" contrasts are shown in Table 5 (Appendix A).

Can psychosocial factors identify students at risk of leaving their institutions?

We used parameter estimates from the full and reduced models to calculate the estimated probability that each student in the cross validation (holdout) sample stayed at his or her institution from the first-to-second year. To examine whether the estimated probabilities could be used to accurately identify students who transferred or dropped out, we examined thresholds at the 5th, 10th, 15th, 20th, and 25th percentiles.[5] A student was classified as "low risk" (of transfer/drop out) if the estimated probability was greater than a specified threshold and classified as "at risk" if the estimated probability was less than or equal to the threshold. Following each classification, we recorded the number of times the student was correctly classified as staying (i.e., true positives), incorrectly classified as staying (false positives), correctly classified as not staying (true negatives), and incorrectly classified as not staying (false negatives).

The full and reduced models correctly identified students at risk of transfer/drop out with negative predictive values[6] ranging from 9% to 16% higher than random chance. Negative predictive values were 0 to 2% higher for the full model than for the reduced model. Detailed results from the classification analysis are shown in Table 6 (Appendix A).

In practice, school administrators and/or guidance counselors may not be familiar with parameter estimates obtained from a logistic regression and may prefer a more straightforward method for identifying students who drop out or transfer. We examined the incremental

[5] Percentiles are based on students estimated probabilities of retention. For example, a threshold at the 5th percentile differentiates the 5% of students with the lowest probability of retention from the 95% of students with the highest probability of retention.

[6] Negative predictive values are the proportion of students who were correctly classified as "at risk" out of the total number of students who transferred or dropped out.

predictive ability of the most salient PSFs (i.e. *Academic Discipline, Commitment to College,* and *Social Connection*) in identifying students who dropped out or transferred.

The probability that a randomly selected student transferred or dropped out was approximately 25%. Using standardized ACT Composite score to identify students as "at risk" of transfer/drop out, the probability of selecting a student who actually transferred or dropped out increased to approximately 33%. Further, using an unweighted composite of standardized *Academic Discipline, Commitment to College,* and *Social Connection,* the probability of selecting a student who actually transferred or dropped out increased to approximately 36%. Lastly, when using both standardized ACT Composite score and the three standardized PSF scales, the probability of selecting a student who actually transferred or dropped out increased to approximately 37%.

The results indicate *Academic Discipline, Commitment to College,* and *Social Connection* provide increased accuracy in predicting "transfer/drop out" over ACT Composite score. Further, combining ACT Composite score with the three significant PSFs best predicted "transfer/drop out." Full results are shown in Table 7 (Appendix A) and corroborate similar findings reported in the ACT Engage College User's Guide (ACT, 2012).

Do the psychosocial factors provide useful information for creating effective interventions and increasing first-to-second year college retention rates?

Identification of students at risk of leaving their institutions only improves retention if the students experience effective interventions. An advantage of using PSF scores to identify students at risk of leaving an institution is that the PSF scores can also be used to design effective interventions. Further information on interpreting PSF scores and designing appropriate intervention can be found in the ACT Engage College User's Guide (ACT, 2012).

Students with predicted probabilities of retention in the bottom 5% reported much lower *Academic Discipline, Commitment to College,* and *Social Connection* than students with higher predicted probabilities. Interventions that focus on improving these particular psychosocial skills can increase a student's probability of first-to-second year retention. For example, our study showed that if a student in the 5th percentile increased his or her *Academic Discipline* by one standard deviation, the student's estimated probability of staying at his/her institution increased from .65 to .71. If the student increased his or her *Academic Discipline, Commitment to College,* or *Social Connection* by one standard deviation, the student's probability of first-to-second year retention increased from .65 to .76. Estimated increases for the 5th, 10th, 15th, 20th, and 25th percentiles are shown in Table 8 (in Appendix A).

Conclusion

This study examined the differential effects of 10 PSFs on retention characterized by "stay vs. transfer/drop out" and "transfer vs. drop out." Among the 10 PSFs examined, *Academic Discipline, Commitment to College, Social Connection,* and *Academic Self-Confidence* incrementally predicted first-to-second year retention status beyond standard variables. That is, students with higher *Academic Discipline, Commitment to College,* and *Social Connection* were more likely to stay at their institutions than to drop out or transfer, whereas students with higher *Academic Self-Confidence* were less likely to stay than to drop out or transfer. *Academic Self-Confidence* also differentiated students who transferred from students who dropped out: namely, students with higher *Academic Self-Confidence* were less likely to transfer than to drop out – suggesting that students may not return to college for reasons that may be implicitly related to the belief in their ability to perform well in college. Among the 10 PSFs examined, the effect of *Academic Self-Confidence* on retention seemed counterintuitive to the expectations that higher

self-confidence would result in improved retention (e.g., DeWitz, Woolsey, & Walsh; 2009). As such, the impact of *Academic Self-Confidence* warranted further examination.

It is plausible to expect that students with higher academic self-confidence stay in college rather than transfer or drop out. However, it is also logical to reason that the more academically-self-confident-students believe they can succeed at any institution, making it more likely for these students to transfer rather than to stay. Also, the effect of *Academic Self-Confidence* on retention status could be an artifact of the mix of constraints in the model. In this light, the authors suspected that the effect of *Academic Self-Confidence* on retention status could have been impacted by the effects of other predictors in the model – for instance, high school GPA, ACT Composite score, and *Academic Discipline,* significant predictors of retention status that were also highly correlated with *Academic Self-Confidence* (Table 9 in Appendix A). *Academic Self-Confidence* was positively associated with retention status (stay vs. transfer/drop out as shown in Table 3 in Appendix A), meaning higher *Academic Self-Confidence* scores were associated with a higher probability of a student staying (versus not staying) at his or her institution. Notwithstanding, the odds of a student staying in the same institution in the second year reduced with increased *Academic Self-Confidence*, after controlling for other constraints in the model (Table 4 in Appendix A). When high school GPA, ACT Composite, and *Academic Discipline* variables were excluded from the model, we found that the more-academically-self-confident students were more likely to stay than transfer or drop out ($\hat{\beta}$ = 0.072, *SE* = 0.030, *p* = 0.021), unlike the prior conclusion derived from Table 4 (Appendix A). Thus, it could be argued that the higher academically self-confident students were not different from those who obtained better high school GPAs, scored higher on ACT Composite test, and/or were the most academically disciplined. Beaton and Eaton (2006) proposed that students who have well-

established and strong academic self-concepts have more confidence in their ability to succeed, portending a composite meaning of academic self-confidence.

Academic self-confidence has often been likened to academic self-efficacy, which has also been shown to be a non-significant determinant of student's persistence in college (Zajacova, Lynch, & Espenshade, 2005; Becker & Gable, 2009) or negatively impact first-to-second-year retention (Devonport & Lane, 2006). As such, our findings on the impact of *Academic self-confidence* on retention status are consistent with prior researches on the subject.

Utility of Results

The stability of the parameter estimates was examined through a cross-validation analysis on a holdout sample. Students were classified as "at risk" or "low risk" of drop out/transfer. The 10 PSFs increased identification accuracy of "at risk" students by a small margin over the model without the PSFs, and by a large margin over random identification. The results from this analysis suggest that the 10 PSFs are useful for identifying students "at risk" of leaving their institutions, beyond the traditionally used academic factors. Hence, the prediction equation developed in this study might be used for identifying "at risk" students for targeted interventions. A simpler approach for identifying "at risk" students involves using overall ACT score in combination with the three most significant PSFs (*Academic Discipline, Commitment to College,* and *Social Connection).* Although all students benefit from intervention programs, our findings suggest the programs that might be most rewarding to students with the requisite academic abilities to succeed in college (but who might drop out because of psychosocial and/or other non-academic skills deficiencies), should be geared toward improving both academic and nonacademic factors – like *Academic Discipline, Commitment to College,* and *Social Connection.* An assessment of *Academic Self-Confidence* may be used as a tool to identify

students at risk of dropping out of college (Devonport & Lane, 2006). Overall, these findings provide tools and rationales to format holistic-based targeted-interventions for improving college retention.

Reference

ACT. (2012). *ACT Engage® College user's guide*. Iowa City, IA: ACT. Retrieved from http://www.act.org/engage/pdf/engage_college_users_guide.pdf.

ACT. (2016). *2016 retention/completion summary tables*. Iowa City, IA: ACT. Retrieved from http://www.act.org/content/dam/act/unsecured/documents/Retention-Summary-Tables-2016.pdf.

Allen, J., Robbins, S. B., Casillas, A., & Oh, I. (2008). Third-year college retention and transfer: Effects of academic performance, motivation, and social connectedness. *Research in Higher Education, 49*(7), 647-664.

Allen, J., Robbins, S. B., & Sawyer, R. (2010). Can measuring psychosocial factors promote college success? *Applied Measurement in Education*, *23*(1), 1-22.

Austin, P. C. (2010). Estimating multilevel logistic regression models when the number of clusters is low: A comparison of different statistical software procedures. *The International Journal of Biostatistics 6*(1), Art. 16.

Bean, J. P., & Eaton, S. (2000). *A Psychological Model of College Student Retention. In Reworking the Departure Puzzle: New Theory and Research on College Student Retention,* edited by J. M. Braxton, 73-89. Nashville, TN: University of Vanderbilt Press.

Becker, S. P., & Gable, R. K. (2009). The relationship of self-efficacy and GPA, attendance, and college Student retention. *NERA Conference Proceedings 2009*, 26. Retrieved from http://digitalcommons.uconn.edu/nera_2009/26.

Belloc, F., Maruotti, A., & Petrella, L. (2011). How individual characteristics affect university students drop-out: A semiparametric mixed-effects model for an Italian case study. *Journal of Applied Statistics, 38*(10), 2225-2239.

Devonport, T. J., & Lane, A. M. (2006). Relationships between self-efficacy, coping and student retention. *Social Behavior & Personality: An International Journal, 34*(2), 127-138.

DeWitz, S. J., Woolsey, M. L., & Walsh, W. B. (2009). College student retention: An exploration of the relationship between self-efficacy beliefs and purpose in life among college students. *Journal of College Student Development, 50*(1), 19-34. doi:10.1353/csd.0.0049

Habley, W., Valiga, M., McClanahan, R., & Burkum, K. (2010). *What works in student retention? Fourth national survey.* Iowa City, IA: ACT.

Hedeker, D. (2003). A mixed-effects multinomial logistic regression model. *Statistics in Medicine, 22*(9), 1433–1446.

Herzog, S. (2005). Measuring determinants of student return vs. drop/stopout vs. transfer: A first-to-second year analysis of new freshman. *Research in Higher Education, 46*(8), 883-928.

Ishitani, T. T., & DesJardins, S. L. (2002). A longitudinal investigation of dropout from college in the United States. *Journal of College Student Retention: Research, Theory & Practice*, *4*(2), 173-201.

Le, H., Casillas, A., Robbins, S. B., & Langley, R. (2005). Motivational and skills, social, and self-management predictors of college outcomes: Constructing the student readiness inventory. *Educational and Psychological Measurement, 65*(3), 482-508.

McGrath, M., & Braunstein, A. (1997). The prediction of freshmen attrition: An examination of the importance of certain demographic, academic, financial and social factors. *College Student Journal, 31*(3), 396-408.

Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal, 2*(1), 1–21.

Rivas, R. M., Sauer, P. L., Glynn, J. G., & Miller, T. E. (2007). Persist/dropout differences in pre-matriculation attitudes of freshman towards college attrition: A longitudinal multiple group structural equations model. *College Teaching Methods & Styles Journal (CTMS), 3*(3), 55-68.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, *130*(2), 261.

Tross, S. A., Harper, J. P., Osher, L. W., & Kneidinger, L. M. (2000). Not just the usual cast of characteristics: Using personality to predict college performance and retention. *Journal of College Student Development, 41*(3), 323-334.

Zajacova, A., Lynch, S. M., & Espenshade, T. J. (2005). Self-efficacy, stress, and academic success in college. *Research In Higher Education, 46*(6), 677-706. doi:10.1007/s11162-004-4139-z.

Appendix A

Tables showing results of analyses

Table 1. Percentage First-to-Second Year College Retention Status by Gender, Family Income, and Race

			Enrollment status		
Group		*n*	Drop (%)	Transfer (%)	Return (%)
Gender	Female	8,348	10.4	13.4	76.3
	Male	6,063	14.3	12.7	73.0
Income	Low	3,829	15.8	12.2	72.0
	Moderate	7,567	11.0	13.2	75.8
	High	3,015	10.0	13.7	76.4
Race	Asian	470	10.4	9.2	80.4
	Black	1,595	13.3	12.9	73.8
	Hispanic	1,627	15.2	11.6	73.2
	White	10,145	11.0	13.5	75.5
	Other	574	20.0	12.7	67.2
Overall		14,411	12.1	13.0	74.9

Table 2. Means of Standardized PSF Scales by Retention Status

	Drop Out (n = 1,737)	Transfer (n = 1,882)	Stay (n = 10,792)
Academic Discipline	-0.23	-0.14	0.09
Academic Self-Confidence	-0.04	-0.07	0.05
Commitment to College	-0.12	-0.06	0.06
Communication Skills	-0.07	-0.03	0.03
General Determination	-0.06	-0.04	0.06
Goal Striving	-0.03	-0.02	0.04
Social Activity	-0.08	0.04	0.00
Social Connection	-0.15	-0.02	0.04
Steadiness	-0.04	-0.07	0.01
Study Skills	0.01	-0.06	0.04

Table 3. Point-Biserial Correlations Between Standardized PSF Scales and Retention Status Contrasts

	Stay vs. Transfer/Drop Out (*n*=14,411)	Transfer vs. Drop Out (*n*=3,619)
Academic Discipline	0.12**	0.04**
Academic Self-Confidence	0.05**	-0.02
Commitment to College	0.07**	0.03
Communication Skills	0.03**	0.02
General Determination	0.05**	0.01
Goal Striving	0.03**	0.01
Social Activity	0.01	0.06**
Social Connection	0.05**	0.06**
Steadiness	0.03**	-0.01
Study Skills	0.03**	-0.03*

Note. ** $p < .01$, * $p < .05$

Table 4. Parameter Estimates and Standard Errors (SE) for Full Model

	Stay vs. Transfer/Drop Out		Transfer vs. Drop Out	
Predictor	Estimate	SE	Estimate	SE
Intercept	2.231**	(0.101)	0.115	(0.137)
High School GPA	0.312**	(0.030)	0.175**	(0.049)
ACT Composite score	0.310**	(0.038)	0.056	(0.064)
Gender (Female)	-0.063*	(0.029)	0.164**	(0.049)
Race (Asian)	0.174	(0.146)	0.344	(0.256)
Race (Black)	0.249**	(0.082)	-0.140	(0.137)
Race (Hispanic)	-0.050	(0.101)	0.141	(0.167)
Race (Other)	-0.165	(0.104)	-0.495**	(0.169)
Family income (low)	-0.212**	(0.043)	-0.387**	(0.073)
Family income (high)	0.183**	(0.048)	0.316**	(0.083)
Academic Discipline	0.284**	(0.041)	0.107	(0.067)
Academic Self-Confidence	-0.216**	(0.036)	-0.189**	(0.061)
Commitment to College	0.127**	(0.030)	-0.012	(0.049)
Communication Skills	-0.046	(0.037)	-0.035	(0.061)
General Determination	-0.055	(0.049)	-0.081	(0.082)
Goal Striving	-0.026	(0.049)	0.144	(0.082)
Social Activity	-0.019	(0.035)	0.061	(0.059)
Social Connection	0.135**	(0.037)	0.004	(0.062)
Steadiness	0.012	(0.031)	0.012	(0.052)
Study Skills	0.004	(0.035)	-0.086	(0.058)
Institutional variation	0.144*	(0.062)	0.134	(0.071)

Note. ** $p < .01$, * $p < .05$

Table 5. Parameter Estimates and Standard Errors (SE) for Model with Dummy Contrasts Between Outcomes for Comparison with Allen et al. (2008)

	Stay vs. Drop Out		Transfer vs. Drop Out	
Predictor	Estimate	SE	Estimate	SE
Intercept	2.193**	(0.111)	0.082	(0.155)
High School GPA	0.397**	(0.041)	0.174**	(0.049)
ACT Composite score	0.340**	(0.053)	0.055	(0.064)
Gender (Female)	0.036	(0.081)	0.325**	(0.099)
Race (Asian)	0.481	(0.263)	0.198	(0.317)
Race (Black)	0.321**	(0.123)	-0.279	(0.155)
Race (Hispanic)	0.159	(0.165)	0.010	(0.201)
Race (Other)	-0.289	(0.150)	-0.663**	(0.199)
Family income (low)	-0.467**	(0.085)	-0.446**	(0.106)
Family income (high)	0.276*	(0.109)	0.244	(0.126)
Academic Discipline	0.341**	(0.056)	0.111	(0.067)
Academic Self-Confidence	-0.312**	(0.051)	-0.189**	(0.061)
Commitment to College	0.120**	(0.042)	-0.013	(0.049)
Communication Skills	-0.063	(0.051)	-0.033	(0.061)
General Determination	-0.097	(0.068)	-0.084	(0.082)
Goal Striving	0.045	(0.068)	0.143	(0.082)
Social Activity	0.012	(0.048)	0.062	(0.059)
Social Connection	0.137**	(0.051)	-0.000	(0.062)
Steadiness	0.020	(0.043)	0.010	(0.051)
Study Skills	-0.038	(0.049)	-0.083	(0.058)
Institutional variation	0.114*	(0.052)	0.203	(0.110)

Note. ** $p < .01$, * $p < .05$

Table 6. Identification of Students at Risk of Leaving Their Institutions at Various Thresholds

Model	Percentile	Threshold	Correct	Incorrect	Accuracy
Full	5	0.74	103	137	0.43
	10	0.79	192	288	0.40
	15	0.81	277	446	0.38
	20	0.83	364	593	0.38
	25	0.85	432	775	0.36
Reduced	5	0.76	99	144	0.41
	10	0.80	189	293	0.39
	15	0.83	261	447	0.37
	20	0.84	355	606	0.37
	25	0.85	426	774	0.36
Random	5	0.05	65	179	0.27
	10	0.10	128	343	0.27
	15	0.15	192	548	0.26
	20	0.20	261	720	0.27
	25	0.25	327	883	0.27

Note. Correct = correctly identified as "at risk" (true negatives); Incorrect = incorrectly identified as "at risk" (false negatives); Accuracy = identification accuracy (negative predictive value)

Table 7. Percent of 4-Year Students Accurately Identified as At-Risk

Selection method	Drop Out (%)	Transfer (%)	Overall (%)
Random selection	12.05	13.05	25.10
ACT Composite Score only*	19.20	13.95	33.15
ACT Engage only*	18.51	17.27	35.78
ACT Composite Score + ACT Engage	20.99	16.16	37.15

Note. ACT Engage stands for the PSFs

*Students scoring in the bottom 5% of these populations were flagged

Table 8. Average ACT Engage Estimates by Percentile and Estimated Probabilities Based on ACT Engage Improvements

	ACT Engage estimates			Probability of retention				
Percentile	AD	CC	SC	Baseline	AD +1SD	CC +1SD	SC +1SD	Overall
5	-1.39	-0.99	-0.75	0.65	0.71	0.68	0.68	0.76
10	-1.05	-0.73	-0.61	0.70	0.76	0.73	0.73	0.80
15	-0.92	-0.60	-0.53	0.73	0.78	0.75	0.76	0.82
20	-0.75	-0.49	-0.43	0.75	0.80	0.77	0.77	0.84
25	-0.65	-0.43	-0.38	0.76	0.81	0.78	0.78	0.85

Note. AD = *Academic Discipline*; CC = *Commitment to College*; SC = *Social Connection*; +1SD = estimated probability of retention if the student increases his or her ACT Engage score by one standard deviation; Overall = estimated probability of retention if the student increases all three AD, CC, and SC scores by one standard deviation.

Table 9. Correlation Between the Academic and Nonacademic (Psychosocial Factors) Variables in the Model

	1	2	*3*	4	5	6	7	8	9	10	11	12
1. High School GPA	1.00											
2. ACT Composite	0.38**	1.00										
3. Academic Discipline	0.39**	0.00	*1.00*									
4. *Academic Self-Confidence*	*0.30***	*0.43***	*0.38***	*1.00*								
5. Commitment to College	0.14**	0.01	*0.44***	0.31**	1.00							
6. Communication Skills	0.10**	-0.01	*0.38***	0.20**	0.36**	1.00						
7. General Determination	0.20**	-0.08**	*0.70***	0.35**	0.50**	0.56**	1.00					
8. Goal Striving	0.12**	-0.07**	*0.59***	0.47**	0.51**	0.50**	0.77**	1.00				
9. Social Activity	-0.01	-0.03*	*0.15***	0.25**	0.28**	0.33**	0.27**	0.43**	1.00			
10. Social Connection	0.05**	0.00	*0.27***	0.19**	0.35**	0.53**	0.38**	0.47**	0.61**	1.00		
11. Steadiness	0.10**	0.03**	*0.37***	0.35**	0.27**	0.42**	0.39**	0.44**	0.28**	0.24**	1.00	
12. Study Skills	0.10**	-0.06**	*0.49***	0.30**	0.30**	0.52**	0.58**	0.58**	0.21**	0.33**	0.36**	1.00

Note. N = 9,364 students. ** $p < .01$, * $p < .05$

Appendix B

Charts illustrating retention rates at deciles of each psychosocial factor

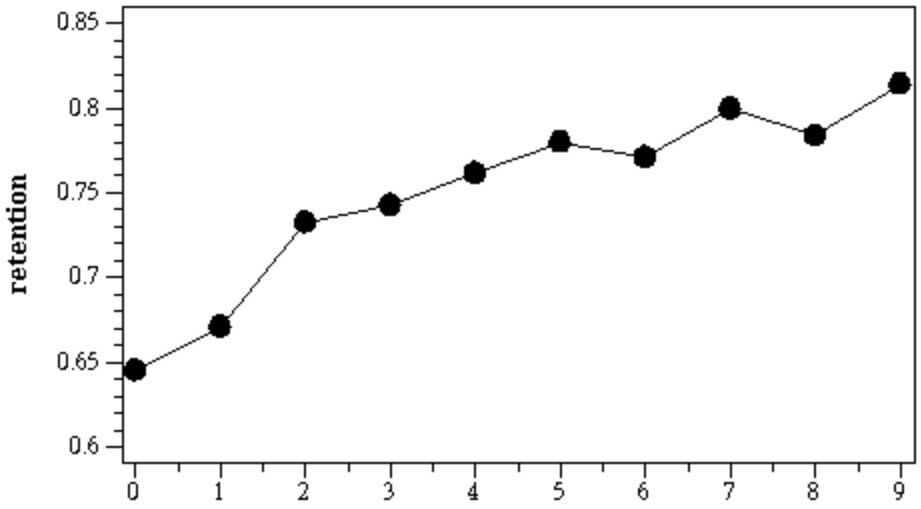

Figure 1. Average retention at each decile of *Academic Discipline*

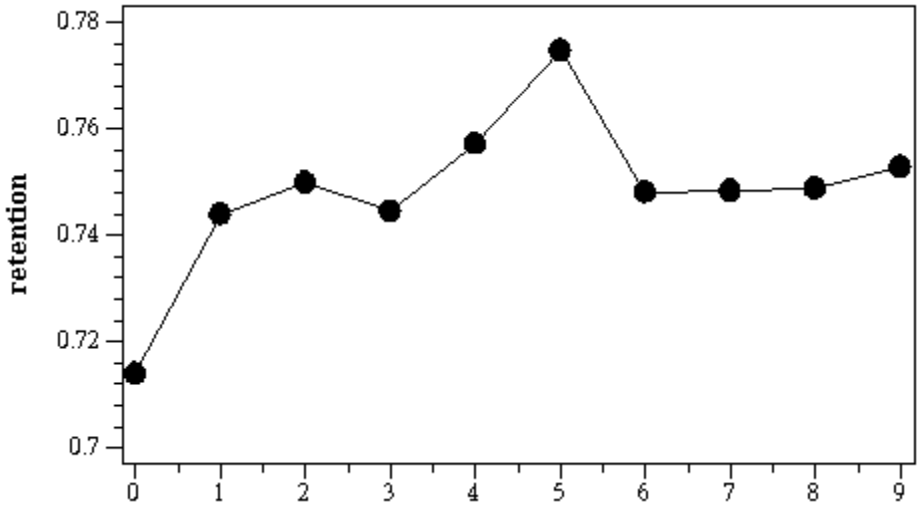

Figure 2. Mean retention at each decile of *Social Activity*

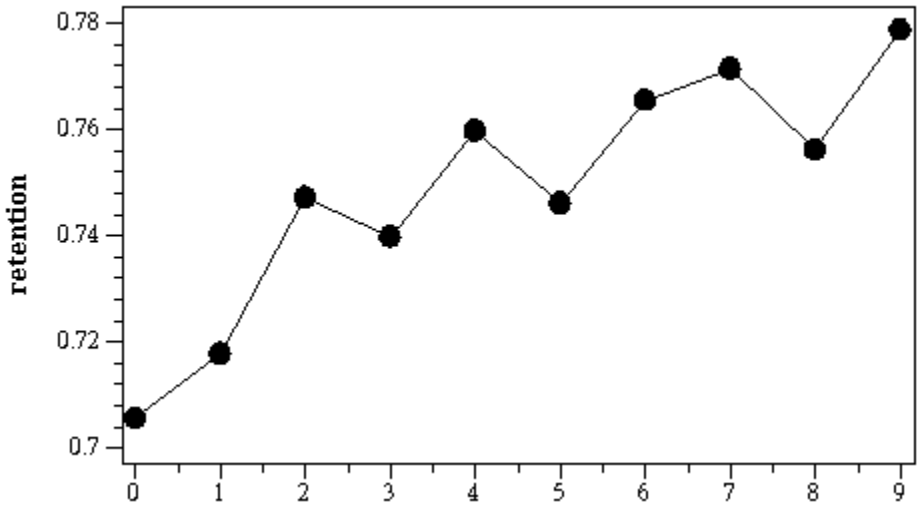

Figure 3. Mean retention at each decile of *Academic Self-Confidence*

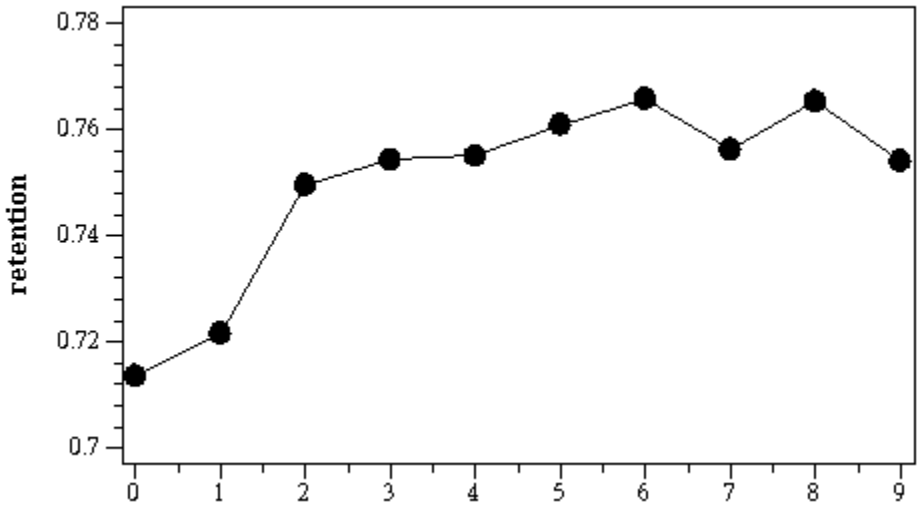

Figure 4. Mean retention at each decile of *Communication Skills*

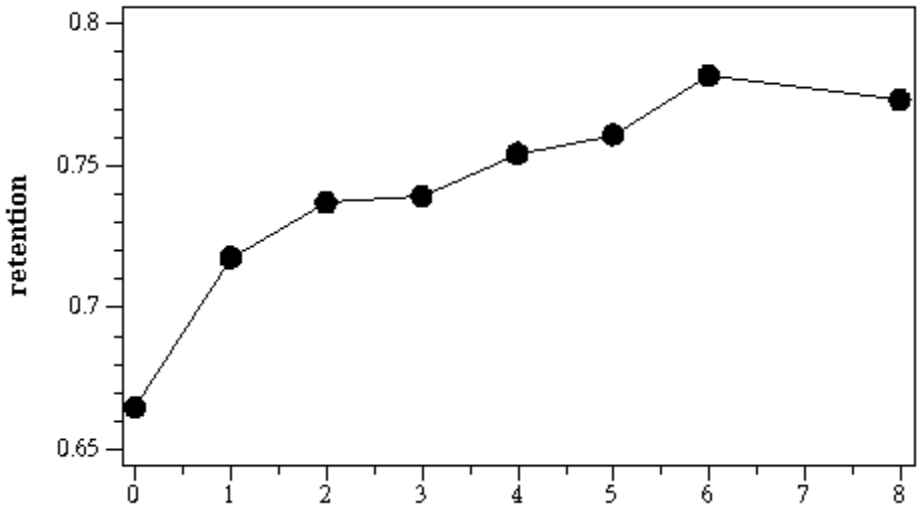

Figure 5. Mean retention at each decile of *Commitment to College*

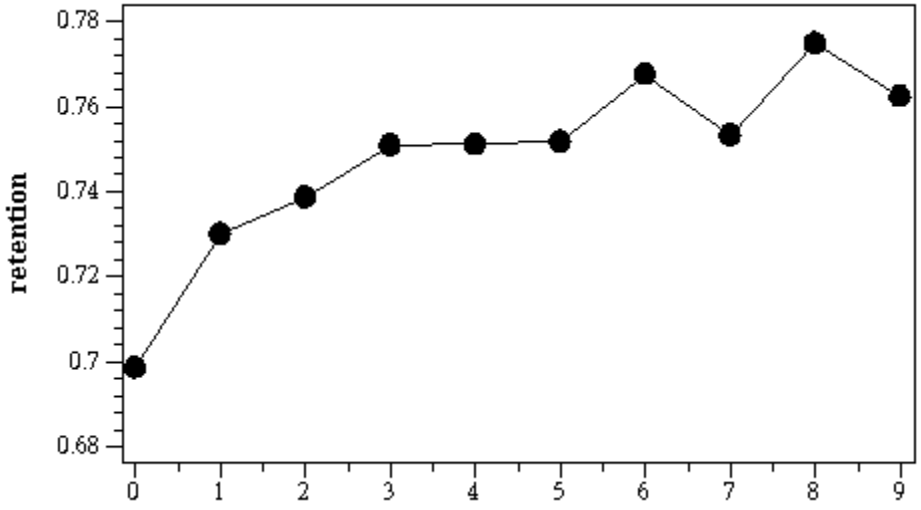

Figure 6. Mean retention at each decile of *General Determination*

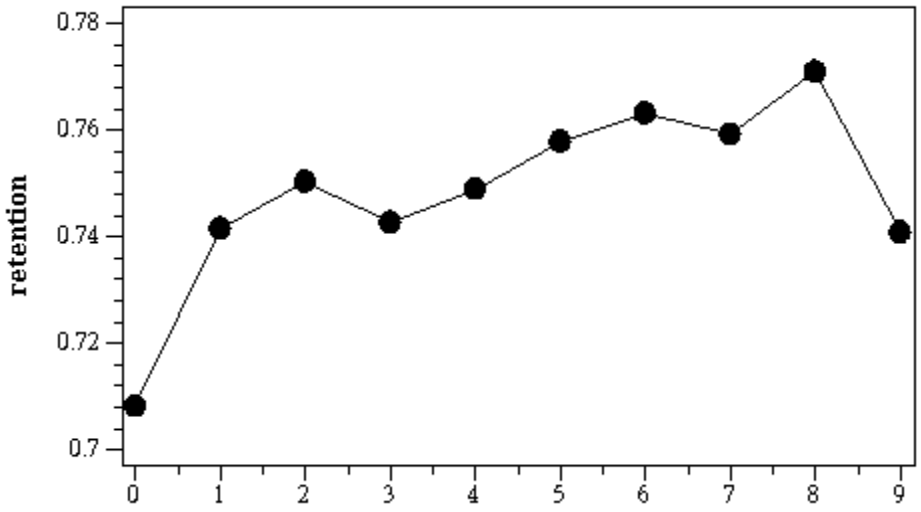

Figure 7. Mean retention at each decile of *Goal Striving*

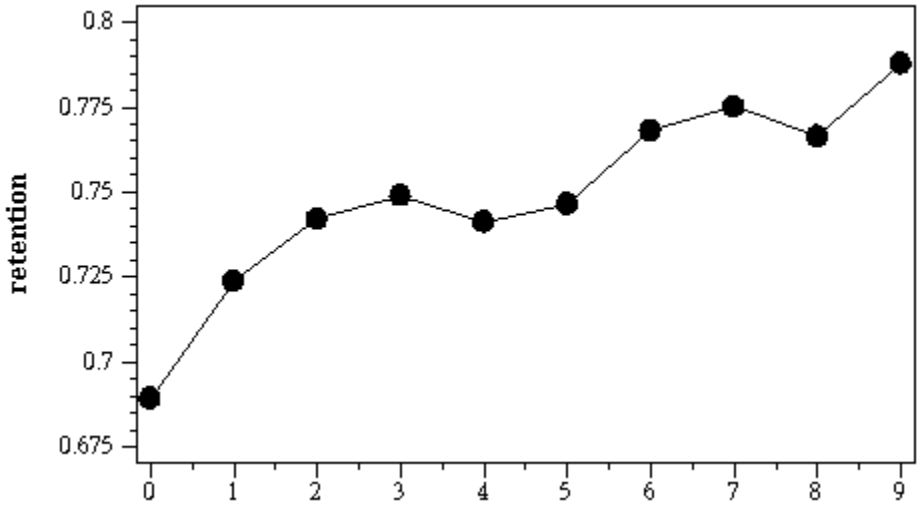

Figure 8. Mean retention at each decile of *Social Connection*

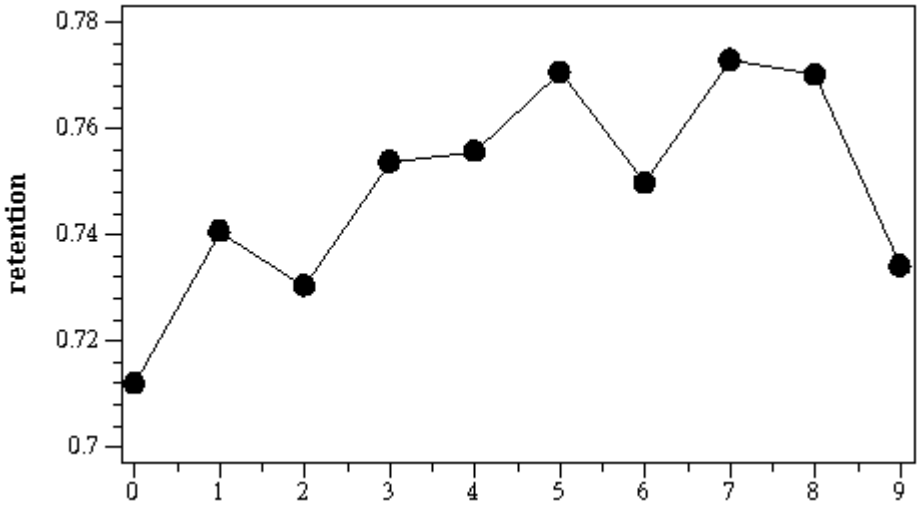

Figure 9. Mean retention at each decile of *Study Skills*

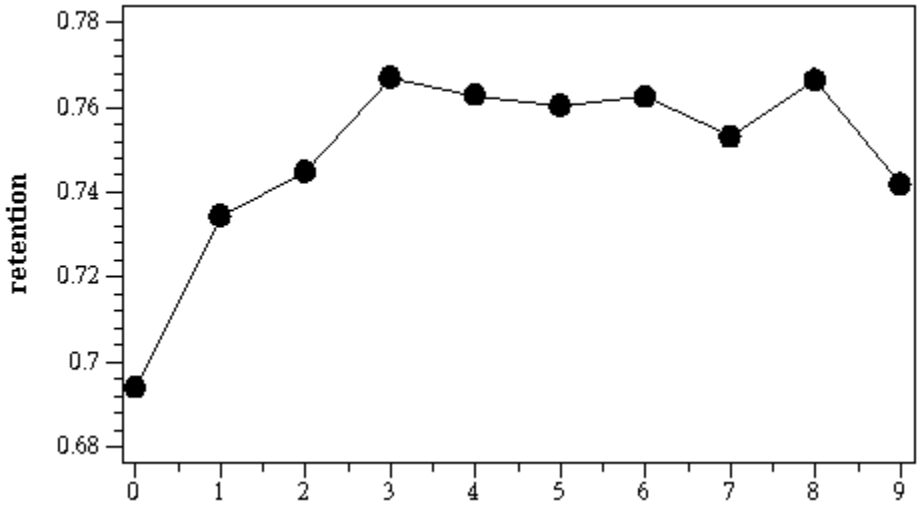

Figure 10. Mean retention at each decile of *Steadiness*